On the Power of Quantum Fourier Sampling

Bill Fefferman[*1] **and Christopher Umans**[†2]

1 **Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD, USA**
wjf@umd.edu

2 **California Institute of Technology, Pasadena, CA, USA**
umans@cms.caltech.edu

Abstract

A line of work initiated by Terhal and DiVincenzo [19] and Bremner, Jozsa, and Shepherd [6], shows that restricted classes of quantum computation can efficiently sample from probability distributions that cannot be exactly sampled efficiently on a classical computer, unless the **PH** collapses. Aaronson and Arkhipov [3] take this further by considering a distribution that can be sampled efficiently by linear optical quantum computation, that under two feasible conjectures, cannot even be approximately sampled within bounded total variation distance, unless the **PH** collapses.

In this work we use Quantum Fourier Sampling to construct a class of distributions that can be sampled exactly by a quantum computer. We then argue that these distributions cannot be approximately sampled classically, unless the **PH** collapses, under variants of the Aaronson-Arkhipov conjectures.

In particular, we show a general class of quantumly sampleable distributions each of which is based on an "Efficiently Specifiable" polynomial, for which a classical approximate sampler implies an average-case approximation. This class of polynomials contains the Permanent but also includes, for example, the Hamiltonian Cycle polynomial, as well as many other familiar #**P**-hard polynomials.

Since our distribution likely requires the full power of universal quantum computation, while the Aaronson-Arkhipov distribution uses only linear optical quantum computation with noninteracting bosons, why is our result interesting? We can think of at least three reasons:

1. Since the conjectures required in [3] have not yet been proven, it seems worthwhile to weaken them as much as possible. We do this in two ways, by weakening both conjectures to apply to any "Efficiently Specifiable" polynomial, and by weakening the so-called Anti-Concentration conjecture so that it need only hold for one distribution in a broad class of distributions.
2. Our construction can be understood without any knowledge of linear optics. While this may be a disadvantage for experimentalists, in our opinion it results in a very clean and simple exposition that may be more immediately accessible to computer scientists.
3. It is extremely common for quantum computations to employ "Quantum Fourier Sampling" in the following way: first apply a classically efficient function to a uniform superposition of inputs, then apply a Quantum Fourier Transform followed by a measurement. Our distributions are obtained in exactly this way, where the classically efficient function is related to a (presumed) hard polynomial. Establishing rigorously a robust sense in which the central primitive of Quantum Fourier Sampling is classically hard seems a worthwhile goal in itself.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes

Keywords and phrases Quantum Complexity Theory, Sampling Complexity

Digital Object Identifier 10.4230/LIPIcs.TQC.2016.1

[*] BF was supported by NSF CCF-1423544, BSF grant 2010120 and the Department of Defense.
[†] CU was supported by NSF CCF-1423544 and BSF grant 2010120.

1 Introduction

It is a major goal of computational complexity theory to establish "quantum superiority", obtaining provable settings in which quantum algorithms attain speedups over classical algorithms. Despite the importance of this endeavor, the best evidence that quantum computers can efficiently solve *decision* problems outside **NP** comes from oracle results, see, e.g., [1, 11, 10]. A line of work initiated by DiVincenzo and Terhal [19] and Bremner, Jozsa and Shepherd [6] asks whether we can provide a theoretical basis for quantum superiority by studying *distribution sampling problems*. Since then, there have been many other *exact* sampling results, giving examples of distributions with quantum samplers, which cannot be sampled *exactly* by classical randomized algorithms, see, e.g., [9, 12, 15]. These hardness results are restrictive in that they do not hold in the *approximate* setting, whereby the classical algorithm is allowed to sample from any distribution close in total variation distance to the idealized quantum distribution.

Aaronson and Arkhipov took this a step further, by giving a distribution that can be sampled efficiently by a restrictive form of quantum computation, that assuming the validity of two feasible conjectures, cannot be approximately sampled classically[1], unless the **PH** collapses [3]. The equivalent result for decision problems, establishing $\mathbf{BQP} \not\subset \mathbf{BPP}$ unless the **PH** collapses, would be a crowning achievement in quantum complexity theory. In addition, this research has been very popular with experimentalists who hope to perform this task, "Boson Sampling", in their labs. Experimentally, it seems more relevent to analyze the hardness of approximate quantum sampling, since it is unreasonable to expect that any physical realization of a quantum computer can *itself* exactly sample from its idealized distribution.

In addition to experimental motivation, it is also known that if we can find such a quantumly sampleable distribution for which no classical approximate sampler exists, there exists a "search" problem that can be solved by a quantum computer that cannot be solved classically [2]. In a search problem we are given an input $x \in \{0,1\}^n$, and our goal is to output an element in a nonempty set, $A_x \subseteq \{0,1\}^{\mathrm{poly}(n)}$ with high probability. Establishing this separation, which is not known to follow from exact sampling hardness results, would certainly be one of the strongest pieces of evidence to date that quantum computers can outperform their classical counterparts.

In this work we use the same general algorithmic framework used in many quantum algorithms, which we refer to as "Quantum Fourier Sampling", to demonstrate the existence of a general class of distributions that can be sampled exactly by a quantum computer. We then argue that these distributions cannot be approximately sampled classically, unless the **PH** collapses. Perhaps surprisingly, we obtain and generalize many of the same conclusions as Aaronson and Arkhipov [3] with a completely different class of distributions.

Additionally, concurrently, and independent of us, an exciting result by Bremner, Montanaro and Shepherd [7] obtains similar quantum "approximate sampling" results under related but different conjectures. While our construction has the advantage of a broader class of hardness conjectures, their distribution can be sampled by a class of commuting quantum computations known as Instantaneous Quantum Polynomial time, or **IQP**. This is an advantage of their result, since our quantum sampler likely requires the full power of universal quantum computation.

[1] Indeed, this argument and ours hold even if the classical sampler is a randomized algorithm with access to a **PH** oracle. Therefore it can be interpreted as further evidence that quantum computers can solve problems outside the **PH**.

2 Overview

Our goal is to find a class of distributions that can be sampled efficiently on a quantum computer that cannot be approximately sampled classically. A natural methodology toward showing this is to prove that the existence of a classical approximate sampler implies that a #**P**-hard function can be computed in the **PH**. By Toda's Theorem [20], this would imply a collapse of the **PH**.

In this work, we demonstrate a class of distributions that can, at least in principle, be sampled exactly on a quantum computer. We prove that the existence of an approximate sampler for these distributions implies the existence of a procedure that approximates an "Efficiently Specifiable" polynomial on average. Informally, an Efficiently Specifiable polynomial is a sum of multilinear monomials in which the variables in each monomial can be computed efficiently from the index of the monomial. This includes, among others, the Permanent and Hamiltonian Cycle polynomial.

Computing a multiplicative approximation to the Permanent (or the square of Permanent) with integer entries in the worst-case is #**P**-hard, and computing the Permanent on average is #**P**-hard (see e.g., [3] for more details). The challenge to proving our conjectures is to put these together to prove that an average-case multiplicative approximation to the Permanent (or for that matter, any Efficiently Specifiable polynomial) is still a #**P**-hard problem. Since we can't prove these conjectures, and we don't know the ingredients such a proof will require, it seems worthwhile to attempt to generalize the class of distributions that can be sampled quantumly.

The conjectures we need to prove hardness of approximate sampling are weakened analogues of the conjectures in Aaronson and Arkhipov's results [3]. They conjecture that an *additive approximate average-case solution* to the Permanent with respect to the Gaussian distribution with mean 0 and variance 1 is #**P**-hard. They further propose an "Anti-concentration" conjecture which allows them to reduce the hardness of *multiplicative approximate average-case solutions* to the Permanent over the Gaussian distribution to the hardness of *additive average case solutions* to the Permanent over the Gaussian distribution. The parameters of our conjectures match the parameters of theirs, but our conjectures are broader, so that they need only hold for one such Efficiently Specifiable polynomial, (one of which is the Permanent), and any one of a wider class of distributions.

3 Quantum Preliminaries

In this section we cover a few basic priciples of quantum computing needed to understand the content in the paper. For a complete overview there are many references available, e.g., [13, 16].

We first recall the concept of quantum evaluation of an efficiently classically computable function $f : \{0,1\}^n \to \{0,1\}^m$, which in one quantum query to f maps:

$$\sum_{x\in\{0,1\}^n} |x\rangle|z\rangle \to \sum_{x\in\{0,1\}^n} |x\rangle|z \oplus f(x)\rangle \,.$$

Note that this is a unitary map and can be implemented efficiently as long as f is efficiently computable.

We need the following lemma, which will be useful for our quantum sampler.

▶ **Lemma 1.** *Let $h : [m] \to \{0,1\}^n$ be an efficiently computable one-to-one function, and suppose its inverse can also be efficiently computed. Then the superposition $\frac{1}{\sqrt{m}} \sum_{x\in[m]} |h(x)\rangle$ can be efficiently prepared by a quantum algorithm.*

Proof. Our quantum procedure with two quantum registers proceeds as follows:

1. Prepare $\frac{1}{\sqrt{m}} \sum_{x\in[m]} |x\rangle|00...0\rangle$
2. Query h using the first register as input and the second as output:

$$\frac{1}{\sqrt{m}} \sum_{x\in[m]} |x\rangle|h(x)\rangle$$

3. Query h^{-1} using the second register as input and the first as output:

$$\frac{1}{\sqrt{m}} \sum_{x\in[m]} |x \oplus h^{-1}(h(x))\rangle|h(x)\rangle = \frac{1}{\sqrt{m}} \sum_{x\in[m]} |00...0\rangle|h(x)\rangle$$

4. Discard first register ◀

Finally, we will frequently be dealing with the uniform distribution over $\{\pm 1\}^n$ strings, and a natural generalization:

▶ **Definition 2** ($\mathbb{T}_\ell$)**.** Given $\ell > 0$, we define the set $\mathbb{T}_\ell = \{\omega_\ell^0, \omega_\ell^1 ..., \omega_\ell^{\ell-1}\}$ where ω_ℓ is a primitive ℓ-th root of unity.

We note that $\mathbb{T}_\ell$ is just ℓ evenly spaced points on the unit circle, and $\mathbb{T}_2 = \{\pm 1\}$.

4 Efficiently Specifiable Polynomial Sampling on a Quantum Computer

In this section we describe a general class of distributions that can be sampled efficiently on a Quantum Computer.

▶ **Definition 3** (Efficiently Specifiable Polynomial)**.** We say a multilinear homogenous n-variate polynomial Q with coefficients in $\{0, 1\}$ and m monomials is *Efficiently Specifiable* via an efficiently computable, one-to-one function $h : [m] \to \{0, 1\}^n$, with an efficiently computable inverse, if:

$$Q(X_1, X_2..., X_n) = \sum_{z\in[m]} {X_1}^{h(z)_1} {X_2}^{h(z)_2} ... {X_n}^{h(z)_n}.$$

▶ **Definition 4** ($\mathcal{D}_{Q,\ell}$)**.** Suppose Q is an Efficiently Specifiable polynomial with n variables and m monomials. For fixed Q and ℓ, we define the class of distributions $\mathcal{D}_{Q,\ell}$ over ℓ-ary strings $y \in [0, \ell - 1]^n$ given by:

$$\Pr_{\mathcal{D}_{Q,\ell}}[y] = \frac{|Q(Z_y)|^2}{\ell^n m}$$

where $Z_y \in \mathbb{T}_\ell^n$ is a vector of complex values encoded by the string y.

The encoding works by assigning each value $j \in [0, \ell - 1]$ to ω_ℓ^j. For example, notice that when $\ell = 2$ then $y \in \{0, 1\}^n$ and Z_y is simply the corresponding $\{\pm 1\}^n$ assignment with each entry set to 1 if the corresponding entry in y is 0 and -1 if the corresponding entry in y is 1.

▶ **Theorem 5** (Quantum Sampling Theorem)**.** *Given an Efficiently Specifiable polynomial, Q with n variables, m monomials, relative to a function h, and $\ell \leqslant exp(n)$, the resulting $\mathcal{D}_{Q,\ell}$ can be sampled in $poly(n)$ time on a Quantum Computer.*

Proof.

1. We start in a uniform superposition $\frac{1}{\sqrt{m}} \sum_{z \in [m]} |z\rangle$.
2. We then apply Lemma 1 to prepare $\frac{1}{\sqrt{m}} \sum_{z \in [m]} |h(z)\rangle$.
3. Apply Quantum Fourier Transform over $\mathbb{Z}_\ell^n$ to attain $\frac{1}{\sqrt{\ell^n m}} \sum_{y \in [0,\ell-1]^n} \sum_{z \in [m]} \omega_\ell^{<y,h(z)>} |y\rangle$.

Notice that the amplitude of each y basis state in the final state after Step 3 is proportional to the value of $Q(Z_y)$. A measurement in the computational basis will amount to sampling from the distribution $D_{Q,\ell}$ as desired.

Why is each evaluation appearing in the amplitudes of this quantum state? To see this, let's analyze the simple case of $D_{Q,2}$, in which we claim each amplitude of the state after Step 3 is proportional to Q evaluated at a particular $\{\pm 1\}^n$ assignment. Note that in this case the Quantum Fourier Transform we apply in Step 2 is simply $H^{\otimes n}$, where H is the 2×2 Hadamard matrix.

We can think of the Hadamard transform as having columns indexed by all 2^n multilinear monomials $M_1, M_2, ..., M_{2^n}$ on n variables $x_1, x_2, ..., x_n$, and the 2^n rows of the transform as indexed by all possible $\{\pm 1\}^n$ assignments to the n variables. Then the unnormalized (i, j)-th element of the matrix is $M_j(y_i)$, the evaluation of the j-th monomial on the i-th assignment. To prove this, we first observe that the one qubit Hadamard matrix can be seen in this way, where $M_1 = 1$, the "empty monomial" that always evaluates to 1 irrespective of the assignment, and $M_2 = x_1$. The rows of the transform can be indexed by assignments -1 and $+1$, and the unnormalized matrix entries simply correspond to the evaluations of each monomial on the respective assignment, as mentioned earlier. Further, it is easy to see that the tensor product respects this structure, giving rise to our claimed interpretation.

The state we prepare in Step 2, $\frac{1}{\sqrt{m}} \sum_{z \in [m]} |h(z)\rangle$, is simply the quantum state that is uniformly supported over each of the m monomials in Q, and so after applying the Hadamard transform in Step 3, we obtain a state with amplitudes equal to suitably normalized evaluations of Q at each $\{\pm 1\}^n$ assignment. It is not hard to further generalize this argument to the case of $\mathcal{D}_{Q,\ell}$, in which case we apply a similar interpretation to the Quantum Fourier Transform over $\mathbb{Z}_\ell^n$. ◀

5 Classical Hardness of Efficiently Specifiable Polynomial Sampling

We are interested in demonstrating the existence of some distribution that can be sampled exactly by a uniform family of quantum circuits, that cannot be sampled approximately classically. Approximate here means close in Total Variation distance, where we denote the Total Variation distance between two distributions X and Y by $\|X - Y\|$. Thus we define the notion of a Sampler to be a classical randomized algorithm that approximately samples from a given class of distributions:

▶ **Definition 6** (Sampler). Let $\{D_n\}_{n>0}$ be a class of distributions where each D_n is distributed over $\mathbb{C}^n$. Let $r(n) \in \text{poly}(n), \epsilon(n) \in 1/poly(n)$. We say S is a *Sampler* with respect to $\{D_n\}$ if $\|S(0^n, x \sim U_{\{0,1\}^{r(n)}}, 0^{1/\epsilon(n)}) - D_n\| \leqslant \epsilon(n)$ in (classical) polynomial time.

We first recall a theorem due to Stockmeyer [17] on the ability to "approximate count" in the **PH**.

▶ **Theorem 7** (Stockmeyer). *Given as input an efficiently computable function $f : \{0,1\}^n \to \{0,1\}^m$ and $y \in \{0,1\}^m$, there is a procedure that outputs α such that:*

$$(1-\epsilon) \Pr_{x \sim U_{\{0,1\}^n}}[f(x) = y] \leqslant \alpha \leqslant (1+\epsilon) \Pr_{x \sim U_{\{0,1\}^n}}[f(x) = y] .$$

In randomized time $\mathrm{poly}(n, 1/\epsilon)$ *with access to an* **NP** *oracle.*

In this section we use Theorem 7, together with the assumed existence of a Sampler for $\mathcal{D}_{Q,\ell}$ to obtain hardness consequences.

In particular, we show that a Sampler would imply the existence of an efficient approximation to an Efficiently Specifiable polynomial in the following two contexts:

▶ **Definition 8** ($\epsilon-$additive δ-approximate solution). Given a distribution D over $\mathbb{C}^n$ and $P : \mathbb{C}^n \to \mathbb{C}$ we say $T : \mathbb{C}^n \to \mathbb{C}$ is an $\epsilon-$additive approximate $\delta-$average case solution with respect to D, to P, if $\Pr_{x \sim D}[|T(x) - P(x)| \leqslant \epsilon] \geqslant 1 - \delta$.

▶ **Definition 9** ($\epsilon-$multiplicative δ-approximate solution). Given a distribution D over $\mathbb{C}^n$ and a function $P : \mathbb{C}^n \to \mathbb{C}$ we say $T : \mathbb{C}^n \to \mathbb{C}$ is an $\epsilon-$multiplicative approximate $\delta-$average case solution with respect to D, to P, if $\Pr_{x \sim D}[|T(x) - P(x)| \leqslant \epsilon |P(x)|] \geqslant 1 - \delta$.

These definitions formalize a notion that we will need, in which an efficient algorithm computes a particular hard function approximately only on most inputs, and can act arbitrarily on a small fraction of remaining inputs.

Now we prove our main theorem, which informally states that the existence of a Sampler for $\mathcal{D}_{Q,\ell}$ would imply a solution to Q^2 in the following sense: the solution gives a good additive error approximation to $Q^2(X)$ with probability $1 - \delta$ over the choice of assignments X. That is, on a δ-fraction of assignments the output of the solution may not even be additively-close to the desired value of Q^2.

The proof of this theorem is somewhat technical, but the intuition is very clear. If we have access to a classical randomized algorithm that samples from a distribution close in Total Variation distance to $\mathcal{D}_{Q,\ell}$, we would like to use Stockmeyer's Algorithm (Theorem 7) to get a multiplicative estimate to the probability of a particular outcome of the Sampler. After accounting for normalization, this would amount to a multiplicative estimate to the desired evaluation of the Efficiently Specifiable polynomial. Of course, if the Sampler sampled from exactly the distribution $\mathcal{D}_{Q,\ell}$ we'd be able to do this. Unfortunately though, we only know that the distribution sampled by our Sampler is *close* to the ideal distribution $\mathcal{D}_{Q,\ell}$. Therefore, we can't trust that the probability of any particular outcome of the Sampler is exactly the same as the probability of this outcome according to $\mathcal{D}_{Q,\ell}$. One thing we do know, however, is that *most* of the probabilities of the distribution sampled by the Sampler must be additively close to the probabilities of $\mathcal{D}_{Q,\ell}$, since the two distributions are close in Total Variation distance. This will be enough to guarantee that if we use Stockmeyer's algorithm to estimate the probability of a uniformly chosen outcome, with high probability over choice of assignment, we get a decent additive estimate to the evaluation of the Efficiently Specifiable polynomial. Note that our analysis can be thought of as a simplified version of the analysis in [3].

▶ **Theorem 10** (Complexity consequences of Sampler). *Given an Efficiently Specifiable polynomial Q with n variables and m monomials, and a Sampler S with respect to $\mathcal{D}_{Q,\ell}$, there is a randomized procedure computing an $(\epsilon \cdot m)-$additive approximate $\delta-$average case solution with respect to the uniform distribution over $\mathbb{T}_\ell^n$, to the Q^2 function, in randomized time* $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ *with access to an* **NP** *oracle.*

Proof. We need to give a procedure that outputs an ϵm-additive estimate to the Q^2 function evaluated at a uniform setting of the variables, with probability $1-\delta$ over choice of setting. Setting $\nu = \frac{\epsilon\delta}{16}$, suppose S samples from a distribution $\mathcal{D}'$ such that $\|\mathcal{D}_{Q,\ell} - \mathcal{D}'\| \leqslant \nu$. We let p_y be $\Pr_{\mathcal{D}_{Q,\ell}}[y]$ and q_y be $\Pr_{\mathcal{D}'}[y]$.

Our procedure picks a uniformly chosen encoding of an assignment $y \in [0, \ell-1]^n$, and outputs an estimate $\tilde{q}_y$. Note that $p_y = \frac{|Q(Z_y)|^2}{\ell^n m}$. Thus our goal will be to output a $\tilde{q}_y$ that approximates p_y within additive error $\epsilon\frac{m}{\ell^n m} = \frac{\epsilon}{\ell^n}$, in time polynomial in n, $\frac{1}{\epsilon}$, and $\frac{1}{\delta}$.

We need:

$$\Pr_y[|\tilde{q}_y - p_y| > \frac{\epsilon}{\ell^n}] \leqslant \delta\,.$$

First, define for each y, $\Delta_y = |p_y - q_y|$, which by definition gives us $\|\mathcal{D}_{Q,\ell} - \mathcal{D}'\| = \frac{1}{2}\sum_y[\Delta_y]$. Now:

$$E_y[\Delta_y] = \frac{\sum_y[\Delta_y]}{\ell^n} = \frac{2\nu}{\ell^n}\,.$$

And applying Markov's inequality, $\forall k > 1$,

$$\Pr_y[\Delta_y > \frac{k2\nu}{\ell^n}] < \frac{1}{k}\,.$$

Setting $k = \frac{4}{\delta}$ and recalling that $\nu = \frac{\epsilon\delta}{16}$, we have:

$$\Pr_y[\Delta_y > \frac{\epsilon}{2}\cdot\frac{1}{\ell^n}] < \frac{\delta}{4}\,.$$

Then use approximate counting (with an **NP** oracle), using Theorem 7 on the randomness of S to obtain an output $\tilde{q}_y$ so that, for all $\gamma > 0$, in time polynomial in n and $\frac{1}{\gamma}$:

$$\Pr[|\tilde{q}_y - q_y| > \gamma\cdot q_y] < \frac{1}{2^n}\,.$$

Because we can amplify the failure probability of Stockmeyer's algorithm to be inverse exponential. Now because the q_y's are probabilities that sum to 1:

$$E_y[q_y] = \frac{\sum_y q_y}{\ell^n} = \frac{1}{\ell^n} \Rightarrow \Pr_y[q_y > \frac{k}{\ell^n}] < \frac{1}{k}\,.$$

Now, applying the union bound with γ set to $\frac{\epsilon\delta}{8}$:

$$\begin{aligned}\Pr_y[|\tilde{q}_y - p_y| > \frac{\epsilon}{\ell^n}] &\leqslant \Pr_y[|\tilde{q}_y - q_y| > \frac{\epsilon}{2}\cdot\frac{1}{\ell^n}] + \Pr_y[|q_y - p_y| > \frac{\epsilon}{2}\cdot\frac{1}{\ell^n}]\\ &\leqslant \Pr_y[q_y > \frac{k}{\ell^n}] + \Pr[|\tilde{q}_y - q_y| > \gamma\cdot q_y] + \Pr_y[\Delta_y > \frac{\epsilon}{2}\cdot\frac{1}{\ell^n}]\\ &\leqslant \frac{1}{k} + \frac{1}{2^n} + \frac{\delta}{4} = \frac{\delta}{2} + \frac{1}{2^n} \leqslant \delta.\end{aligned}$$

◀

Now, as will be proven in Appendix A, the variance, $\mathrm{Var}\,[Q(X)]$, of the distribution over $\mathbb{C}$ induced by an Efficiently Specifiable Q with m monomials, evaluated at uniformly distributed entries over $\mathbb{T}_\ell^n$ is m, and so the preceeding Theorem 10 promised us we can achieve an $\epsilon\mathrm{Var}\,[Q(X)]$-additive approximation to Q^2, given a Sampler. We now show that, under a conjecture, this approximation can be used to obtain a good multiplicative estimate to Q^2. This conjecture effectively states that the Chebyshev inequality for this random variable is tight.

▶ **Conjecture 11** (Anti-Concentration Conjecture relative to an n-variate polynomial Q and distribution $\mathcal{D}$ over $\mathbb{C}^n$). *There exists a polynomial p such that for all n and $\delta > 0$,*

$$\Pr_{X\sim\mathcal{D}}\left[|Q(X)|^2 < \frac{\mathrm{Var}\left[Q(X)\right]}{p(n,1/\delta)}\right] < \delta\,.$$

▶ **Theorem 12.** *Assuming Conjecture 11, relative to an Efficiently Specifiable polynomial Q and a distribution $\mathcal{D}$, an $\epsilon\mathrm{Var}\left[Q(X)\right]$-additive approximate δ-average case solution with respect to D, to the Q^2 function can be used to obtain an $\epsilon' \leqslant poly(n)\epsilon$-multiplicative approximate $\delta' = 2\delta$-average case solution with respect to $\mathcal{D}$ to Q^2.*

Proof. Suppose λ is, with high probability, an $\epsilon\mathrm{Var}\left[Q(X)\right]$-additive approximation to $|Q(X)|^2$, as guaranteed in the statement of the Theorem. This means:

$$\Pr_{X\sim\mathcal{D}}\left[\left|\lambda - |Q(X)|^2\right| > \epsilon\mathrm{Var}\left[Q(X)\right]\right] < \delta\,.$$

Now assuming Conjecture 11 with polynomial p, we will show that λ is also a multiplicative estimate to $|Q(X)|^2$ with high probability. By the union bound,

$$\begin{aligned}\Pr_{X\sim\mathcal{D}}\left[\frac{\left|\lambda - |Q(X)|^2\right|}{\epsilon p(n,1/\delta)} > |Q(X)|^2\right] &\leqslant \Pr_{X\sim D}\left[\left|\lambda - |Q(X)|^2\right| > \epsilon\mathrm{Var}\left[Q(X)\right]\right] + \\ &\quad \Pr_{X\sim\mathcal{D}}\left[\frac{\epsilon\mathrm{Var}\left[Q(X)\right]}{\epsilon p(n,1/\delta)} > |Q(X)|^2\right] \\ &\leqslant 2\delta\end{aligned}$$

where the second line comes from Conjecture 11. To attain multiplicative error bounds ϵ' and δ' we can set $\delta = \delta'/2$ and $\epsilon = \epsilon'/p(n,1/\delta)$. ◀

For the results in this section to be meaningful, we simply need the Anti-Concentration conjecture to hold for some Efficiently Specifiable polynomial that is #**P**-hard to compute, relative to any distribution we can sample from (either $U_{\{\pm1\}^n}$, or $\mathcal{B}(0,k)^n$). We note that Aaronson and Arkhipov [3] conjectures the same statement as Conjecture 11 for the special case of the **Permanent** function relative to matrices with entries distributed independently from the complex Gaussian distribution of mean 0 and variance 1.

Additionally, we acknowledge a result of Tao and Vu [18] who show:

▶ **Theorem 13** (Tao & Vu). *For all $\epsilon > 0$ and sufficiently large n,*

$$\Pr_{X\sim U_{\{\pm1\}^{n\times n}}}\left[|\mathbf{Permanent}[X]| < \frac{\sqrt{n!}}{n^{\epsilon n}}\right] < \frac{1}{n^{0.1}}\,.$$

Which comes quite close to our conjecture for the case of the **Permanent** function and uniformly distributed $\{\pm1\}^{n\times n} = \mathbb{T}_2^{n\times n}$ matrix. More specifically, for the above purpose of relating the hardness of additive solutions to the hardness of multiplicative solutions, we would need an upper bound of any inverse polynomial δ, instead of a fixed $n^{-0.1}$.

6 Sampling from Distributions with Probabilities Proportional to $[-k,k]$ Evaluations of Efficiently Specifiable Polynomials

In the prior sections we discussed quantum sampling from distributions in which the probabilities are proportional to evaluations of Efficiently Specifiable polynomials evaluated

at points in $\mathbb{T}_\ell^n$. In this section we show how to generalize this to quantumly sampling from distributions in which the probabilities are proportional to evaluations of Efficiently Specifiable polynomials evaluated at polynomially bounded integer values. In particular, we show a simple way to take an Efficiently Specifiable polynomial with n variables and create another Efficiently Specifiable polynomial with kn variables, in which evaluating this new polynomial at $\{-1,+1\}^{kn}$ is equivalent to evaluation of the old polynomial at $[-k,k]^n$.

▶ **Definition 14** (k-valued equivalent polynomial)**.** For every Efficiently Specifiable polynomial Q with m monomials and every fixed $k > 0$ consider the polynomial $Q'_k : \mathbb{T}_2^{kn} \to \mathbb{R}$ defined by replacing each variable x_i in Q with the sum of k new variables $x_i^{(1)} + x_i^{(2)} + ... + x_i^{(k)}$. We will call Q'_k the k-valued equivalent polynomial with respect to Q.

Note that a uniformly chosen $\{\pm 1\}$ assignment to the variables in Q'_k induces an assignment to the variables in Q, distributed from a distribution we call $\mathcal{B}(0,k)$:

▶ **Definition 15** ($\mathcal{B}(0,k)$)**.** For k a positive integer, we define the distribution $\mathcal{B}(0,k)$ supported over the odd integers in the range $[-k,k]$ (if k is odd), or even integers in the range $[-k,k]$ (if k is even), so that:

$$\Pr_{\mathcal{B}(0,k)}[y] = \begin{cases} \frac{\binom{k}{\frac{k+y}{2}}}{2^k} & \text{if y and k are both odd or both even} \\ 0 & \text{otherwise} \end{cases}$$

▶ **Theorem 16.** *Given an Efficiently Specifiable polynomial Q with n variables and m monomials, let Q'_k be its k-valued equivalent polynomial. For all $\ell < exp(n)$, we can quantumly sample from the distribution $\mathcal{D}_{Q'_k,\ell}$ in time* $\mathrm{poly}(n,k)$.

Proof. Our proof follows from the following lemma, which proves that Q'_k is Efficiently Specifiable.

▶ **Lemma 17.** *Suppose Q is an n-variate, homogeneous degree d Efficiently Specifiable polynomial with m monomials relative to a function $h : [m] \to \{0,1\}^n$. Let $k \leqslant poly(n)$ and let Q'_k be the k-valued equivalent polynomial with respect to Q. Then Q'_k is Efficiently Specifiable with respect to an efficiently computable function $h' : [m] \times [k]^d \to \{0,1\}^{kn}$.*

Proof. We first define and prove that h' is efficiently computable. We note that if there are m monomials in Q, there are mk^d monomials in Q'_k. As above, we'll think of the new variables in Q'_k as indexed by a pair of indices, a "top index" in $[k]$ and a "bottom index" in $[n]$. Equivalently we are labeling each variable in Q'_k as $x_i^{(j)}$, the j-th copy of the i-th variable in Q.

We can think of each monomial in Q'_k (and hence the input to h') as being indexed by a value $r \in [m]$ and $y_1, y_2, ..., y_d \in [k]^d$. We can obtain the variables in any particular monomial of Q'_k by simply using the output of $h(r)$ to obtain the "bottom" indices of the variables, and use the values of $y_1, y_2, ..., y_d$ to obtain the "top" indices for each of the d variables.

We will now show that h'^{-1} is efficiently computable. As before we will think of $z \in \{0,1\}^{kn}$ as being indexed by a pair, a "top index" in $[k]$ and a "bottom index" in $[n]$. Then we compute $h'^{-1}(z)$ by first obtaining from z the bottom indices $j_1, j_2, ..., j_d$ and the corresponding top indices, $i_1, i_2, ..., i_d$. Then obtain from the bottom indices the string $x \in \{0,1\}^n$ corresponding to the variables used in Q and output the concatenation of $h^{-1}(x)$ and $j_1, j_2, ..., j_d$. ◀

Theorem 16 now follows from Lemma 17, where we established that Q'_k is Efficiently Specifiable, and Theorem 5, where we established that we can sample from $\mathcal{D}_{Q'_k,\ell}$ quantumly. ◀

▶ **Theorem 18.** *Let* $\mathrm{Var}\left[Q(X)\right] = \mathrm{Var}\left[Q(X_1, X_2, ..., X_n)\right]$ *denote the variance of the distribution over* $\mathbb{R}$ *induced by* Q *with assignments distributed from* $\mathcal{B}(0,k)^n$*. Given a Sampler* S *with respect to* $\mathcal{D}_{Q'_k,2}$*, we can find a randomized procedure computing an* $\epsilon\mathrm{Var}\left[Q(X)\right]$*-additive approximate* δ*-average case solution to* Q^2 *with respect to* $\mathcal{B}(0,k)^n$ *in time* $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ *with access to an* **NP** *oracle.*

Proof. We begin by noting that Q'_k is a polynomial of degree d that has kn variables and $m' = mk^d$ monomials. By Theorem 10 we get that a Sampler with respect to $\mathcal{D}_{Q'_{k,2}}$ implies there exists A, an $\epsilon m'$-additive approximate δ-average case solution to ${Q'_k}^2$ with respect to $U_{\{\pm 1\}^{kn}}$ that runs in time $poly(n, 1/\epsilon, 1/\delta)$ with access to an **NP** oracle. We need to show the existence of an A', an $\epsilon m'$-additive approximate δ-average case solution to ${Q'_k}^2$ with respect to the $\mathcal{B}(0,k)^n$ distribution.

We think of A' as receiving an input, $z \in [-k,k]^n$ drawn from $\mathcal{B}(0,k)^n$. A' picks y uniformly from the orbit of z over $\{\pm 1\}^{kn}$ and outputs $A(y)$. Now:

$$\Pr_{z\sim\mathcal{B}(0,k)^n}\left[\left|A'(z) - Q^2(z)\right| \leqslant \epsilon m'\right] = \Pr_{z\sim\mathcal{B}(0,k)^n, y\sim_R \mathrm{orbit}(z)}\left[\left|A(y) - Q^2(z)\right| \leqslant \epsilon m'\right] \quad (1)$$

$$= \Pr_{y\sim U_{\{\pm 1\}^{kn}}}\left[|A(y) - Q'_k(y)| \leqslant \epsilon m'\right] \geqslant 1 - \delta \quad (2)$$

$$(3)$$

Thus, because a uniformly chosen $\{\pm 1\}^{kn}$ assignment to the variables in Q'_k induces a $\mathcal{B}(0,k)^n$ distributed assignment to the variables in Q, this amounts to an $\epsilon m'$-additive approximate δ-average case solution to Q^2 with respect to $\mathcal{B}(0,k)^n$. In Appendix A we prove that $\mathrm{Var}\left[Q(X)\right]$ is m' as desired. ◀

7 The "Compressed" QFT

In this section we begin to prove that quantum algorithms can sample efficiently from distributions with probabilities proportional to evaluations of Efficiently Specifiable polynomials at points in $[-k,k]^n$ for $k \in \exp(n)$. Note that in the prior quantum algorithm of Section 4 we would need to invoke the QFT over $\mathbb{Z}_2^{kn}$, of dimension doubly-exponential in n. Thus we need to define a new Polynomial Transform that can be obtained from the standard Quantum Fourier Transform over $\mathbb{Z}_2^n$, which we refer to as the "Compressed QFT". Now we describe the unitary matrix which implements the Compressed QFT.

Consider the $2^k \times 2^k$ matrix D_k, whose columns are indexed by all possible 2^k multilinear monomials of the variables $x_1, x_2, ..., x_k$ and the rows are indexed by the 2^k different $\{-1,+1\}$ assignments to the variables. The (i,j)-th entry is then defined to be the evaluation of the j-th monomial on the i-th assignment. As we noted earlier, defining $\bar{D}_k$ to be the matrix whose entries are the entries in D_k normalized by $1/\sqrt{2^k}$ gives us the Quantum Fourier Transform matrix over $\mathbb{Z}_2^k$. It is clear, by the unitarity of the Quantum Fourier Transform, that the columns (and rows) in D_k are pairwise orthogonal.

Now we define the "Elementary Symmetric Polynomials":

▶ **Definition 19** (Elementary Symmetric Polynomials)**.** We define the j-th Elementary Symmetric Polynomial on k variables for $j \in [0,k]$ to be:

$$p_j(X_1, X_2, ..., X_k) = \sum_{1 \leqslant \ell_1 < \ell_2 < ... < \ell_j \leqslant k} X_{\ell_1} X_{\ell_2} ... X_{\ell_j} .$$

In this work we will care particularly about the first two elementary symmetric polynomials, p_0 and p_1 which are defined as $p_0(X_1, X_2, ..., X_k) = 1$ and $p_1(X_1, X_2, ..., X_k) = \sum_{1 \leqslant \ell \leqslant k} X_\ell$. Consider the $(k+1) \times (k+1)$ matrix, $\tilde{D}_k$, whose columns are indexed by elementary symmetric polynomials on k variables and whose rows are indexed by equivalence classes of assignments in $\mathbb{Z}_2^k$ under S_k symmetry. We obtain $\tilde{D}_k$ from D_k using two steps.

First obtain a $2^k \times (k+1)$ rectangular matrix $\tilde{D}_k^{(1)}$ whose rows are indexed by assignments to the variables $x_1, x_2, ..., x_k \in \{\pm 1\}^k$ and columns are the entry-wise sum of the entries in each column of D_k whose monomial is in each respective elementary symmetric polynomial. Then obtain the final $(k+1) \times (k+1)$ matrix $\tilde{D}_k$ by taking $\tilde{D}_k^{(1)}$ and keeping only one representative row in each equivalence class of assignments under S_k symmetry. We label the equivalence classes of assignments under S_k symmetry $o_0, o_1, o_2, ..., o_k$ and note that for each $i \in [k]$, $|o_i| = \binom{k}{i}$. Observe that $\tilde{D}_k$ is precisely the matrix whose (i,j)-th entry is the evaluation of the j-th symmetric polynomial evaluated on an assignment in the i-th symmetry class.

▶ **Theorem 20.** *The columns in the matrix $\tilde{D}_k^{(1)}$ are pairwise orthogonal.*

Proof. Note that each column in the matrix $\tilde{D}_k^{(1)}$ is the sum of columns in D_k each of which are orthogonal. We can prove this theorem by observing that if we take any two columns in $D_k^{(1)}$, called c_1, c_2, where c_1 is the sum of columns $\{u_i\}$ of D_k and c_2 is the sum of columns $\{v_i\}$ of D_k. The inner product, $\langle c_1, c_2 \rangle$ can be written:

$$\langle \sum_i u_i, \sum_j v_j \rangle = \sum_{i,j} \langle u_i, v_j \rangle = 0 .$$ ◀

▶ **Theorem 21.** *Let L be the $(k+1) \times (k+1)$ diagonal matrix with i-th entry equal to $\sqrt{o_i}$. Then the columns of $L \cdot \tilde{D}_k$ are orthogonal.*

Proof. Note that the value of the symmetric polynomial at each assignment in an equivalence class is the same. We have already concluded the orthogonality of columns in $\tilde{D}_k^{(1)}$. Therefore if we let a and b be any two columns in the matrix $\tilde{D}_k$, and their respective columns be $\bar{a}, \bar{b}$ in $\tilde{D}_k^{(1)}$, we can see:

$$\sum_{i=0}^{k} (a_i b_i |o_i|) = \sum_{i=0}^{2^k} \bar{a}_i \bar{b}_i = 0 .$$

From this we conclude that the columns of the matrix $L \cdot \tilde{D}_k$, in which the i-th row of $\tilde{D}_k$ is multiplied by $\sqrt{o_i}$, are orthogonal. ◀

▶ **Theorem 22.** *We have just established that the columns in the matrix $L \cdot \tilde{D}_k$ are orthogonal. Let the $k+1 \times k+1$ diagonal matrix R be such that so that the columns in $L \cdot \tilde{D}_k \cdot R$ are orthonormal, and thus $L \cdot \tilde{D}_k \cdot R$ is unitary. Then the first two nonzero entries in R, which we call r_0, r_1, corresponding to the normalization of the column pertaining to the zero-th and first elementary symmetric polynomial, are $1/\sqrt{2^k}$ and $\frac{1}{\sqrt{\sum_{i=0}^{k} [\binom{k}{i}(k-2i)^2]}}$.*

Proof. First we calculate r_0. Since we wish for a unitary matrix, we want the ℓ_2 norm of the first column of $\tilde{D}_k$ to be 1, and so need:

$$r_0^2 \sum_{i=0}^{k} \left(\sqrt{o_i}\right)^2 = r_0^2 \sum_{i=0}^{k} \binom{k}{i} = 1\,.$$

And so r_0 is $1/\sqrt{2^k}$ as desired.

Now we calculate r_1, the normalization in the column of $\tilde{D}_k$ corresponding to the first elementary symmetric polynomial. Note that in i-th equivalence class of assignments we have exactly i negative ones and $k-i$ positive ones. Thus the value of the first symmetric polynomial is the sum of these values, which for the $i-th$ equivalence class is precisely $k-2i$. Then we note the normalization in each row is $\sqrt{\binom{k}{i}}$. Thus we have

$$r_1^2 \sum_{i=0}^{k} \left[\sqrt{\binom{k}{i}}\,(k-2i)\right]^2 = 1\,.$$

Thus $r_1 = \dfrac{1}{\sqrt{\sum_{i=0}^{k}\left[\binom{k}{i}(k-2i)^2\right]}}$ as desired. ◀

8 Using our "Compressed QFT" to Quantumly Sample from Distributions of Efficiently Specifiable Polynomial Evaluations and Hardness Consequences

In this section we use the unitary matrix developed in Section 7 to quantumly sample distributions with probabilities proportional to evaluations of Efficiently Specifiable polynomials at points in $[-k,k]^n$ for $k \in \exp(n)$. Here we assume that we have an efficient quantum circuit decomposition for this unitary. The prospects for this efficient decomposition are discussed in Section 9.

For convenience, we'll define a map $\psi : [-k,k] \to [0,k]$, for k even, with

$$\psi(y) = \begin{cases} \frac{k+y}{2} & \text{if } y \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

▶ **Definition 23.** Suppose Q is an Efficiently Specifiable polynomial Q with n variables and m monomials, and, for $k \leqslant exp(n)$, let Q'_k be its k-valued equivalent polynomial. Let $\mathrm{Var}\left[Q(X)\right]$ be the variance of the distribution over $\mathbb{R}$ induced by Q with assignments to the variables distributed over $\mathcal{B}(0,k)^n$ (or equivalently, this is $\mathrm{Var}\left[Q'_k\right]$ where each variable in Q'_k is independently uniformly chosen from $\{\pm 1\}$), as calculated in Appendix A. Then we define the of distribution $\mathcal{D}_{Q'_k}$ over n tuples of integers in $[-k,k]$ by:

$$\Pr_{\mathcal{D}_{Q'_k}}[y] = \frac{Q(y)^2 \binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}\cdots\binom{k}{\psi(y_n)}}{2^{kn}\mathrm{Var}\left[Q(X)\right]}\,.$$

▶ **Theorem 24.** *By applying $(L \cdot \tilde{D}_k \cdot R)^{\otimes n}$ in place of the Quantum Fourier Transform over $\mathbb{Z}_2^n$ in Section 4 we can quantumly sample from $\mathcal{D}_{Q'_k}$.*

Proof. Since we are assuming Q is Efficiently Specifiable, let $h : [m] \to \{0,1\}^n$ be the invertible function describing the variables in each monomial. We start by producing the

state over $k+1$ dimensional qudits:

$$\frac{1}{\sqrt{m}} \sum_{z \in [m]} |h(z)\rangle$$

which we prepare via the procedure described in Lemma 1.

Instead of thinking of h as mapping an index of a monomial from $[m]$ to the variables in that monomial, we now think of h as taking an index of a monomial in Q to a polynomial expressed in the $\{1, x^{(1)} + x^{(2)} + ... + x^{(k)}\}^n$ basis.

Now take this state and apply the unitary (which we assume can be realized by an efficient quantum circuit) $(L \cdot \tilde{D}_k \cdot R)^{\otimes n}$. Notice each $y \in [-k,k]^n$ has an associated amplitude:

$$\alpha_y = \frac{r_0^{n-d} r_1^d Q(y) \sqrt{\binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}...\binom{k}{\psi(y_n)}}}{\sqrt{m}}.$$

Letting $p_y = \Pr_{\mathcal{D}_{Q'_k}}[y]$, note that, by plugging in r_0, r_1 from Section 7:

$$\begin{aligned}
\alpha_y^2 &= \frac{Q(y)^2 \binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}...\binom{k}{\psi(y_n)} r_0^{2(n-d)} r_1^{2d}}{m} \\
&= \frac{Q(y)^2 \binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}...\binom{k}{\psi(y_n)}}{m 2^{k(n-d)} \left(\sum_{i=0}^{k} \left[\binom{k}{i} (k-2i)^2 \right] \right)^d} \\
&= \frac{Q(y)^2 \binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}...\binom{k}{\psi(y_n)}}{2^{kn-kd} \mathrm{Var}\left[Q(X)\right] 2^{kd}} = \frac{Q(y)^2 \binom{k}{\psi(y_1)}\binom{k}{\psi(y_2)}...\binom{k}{\psi(y_n)}}{2^{kn} \mathrm{Var}\left[Q(X)\right]} = p_y
\end{aligned}$$

◀

Furthermore, using a similar argument to Theorem 10 we can obtain the following theorem, which now gives our hardness result for the existence of Sampler for this class of distributions, whose proof we give in Appendix C:

▶ **Lemma 25.** *Given an Efficiently Specifiable polynomial Q with n variables and m monomials, let Q'_k be its k-valued equivalent polynomial, for some fixed $k \leqslant exp(n)$. Suppose we have a Sampler S with respect to our quantumly sampled distribution class, $\mathcal{D}_{Q'_k}$, and let* $\mathrm{Var}\left[Q(X)\right]$ *denote the variance of the distribution over $\mathbb{R}$ induced by Q with assignments distributed from $\mathcal{B}(0,k)^n$. Then we can find a randomized procedure computing an $\epsilon \mathrm{Var}\left[Q(X)\right]$-additive approximate δ-average case solution to Q^2 with respect to $\mathcal{B}(0,k)^n$ in time* $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ *with access to an* **NP** *oracle.*

9 Putting it All Together

In this section we put our results in perspective and conclude. As mentioned before, our goal is to find a class of distributions $\{\mathcal{D}_n\}_{n>0}$ that can be sampled exactly in $\mathrm{poly}(n)$ time on a Quantum Computer, with the property that there does not exist a (classical) Sampler relative to that class of distributions, $\{\mathcal{D}_n\}_{n>0}$. Using the results in Sections 5 and 6 we can quantumly sample from a class of distributions $\{\mathcal{D}_{Q'_k}\}_{n>0}$, where $k \in \mathrm{poly}(n)$ with the property that, if there exists a classical Sampler relative to this class of distributions, there exists an $\epsilon \mathrm{Var}\left[Q(X)\right]$-additive δ-average case solution to the Q^2 function with respect to the $\mathcal{B}(0,k)^n$ distribution. If we had an efficient decomposition for the "Compressed QFT" unitary matrix, we could use the results from Sections 8 and Appendix C to make k as large as $\exp(n)$. We would like this to be an infeasible proposition, and so we conjecture:

▶ **Conjecture 26.** *There exists some Efficiently Specifiable polynomial Q on n variables, so that $\epsilon \mathrm{Var}\left[Q(X)\right]$-additive δ-average case solutions with respect to $\mathcal{B}(0,k)^n$, for any fixed $k \leqslant \exp(n)$, to Q^2, cannot be computed in (classical) randomized $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ time with a* **PH** *oracle.*

At the moment we don't know of such a decomposition for the "Compressed QFT". However, we do know that we can classically evaluate a related fast (time $n \log^2 n$) polynomial transform by a theorem of Driscoll, Healy, and Rockmore [8]. We wonder if there is some way to use intuition gained by the existence of this fast polynomial transform to show the existence of an efficient decomposition for our "Compressed QFT".

Additionally, if we can prove the Anti-Concentration Conjecture (Conjecture 11) relative to some Efficiently Specifiable polynomial Q and the $\mathcal{B}(0,k)^n$ distribution, we appeal to Theorem 12 to show that it suffices to prove:

▶ **Conjecture 27.** *There exists some Efficiently Specifiable polynomial Q with n variables, so that Q satisfies Conjecture 11 relative to $\mathcal{B}(0,k)^n$, for some fixed $k \leqslant \exp(n)$, and ϵ-multiplicative δ-average case solutions, with respect to $\mathcal{B}(0,k)^n$, to Q^2 cannot be computed in (classical) randomized $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ time with a* **PH** *oracle.*

We would be happy to prove that either of these two solutions (additive or multiplicative) are #**P**-hard. In this case we can simply invoke Toda's Theorem [20] to show that such a randomized classical solution would collapse **PH** to some finite level. We note that at present, both of these conjectures seem out of reach, because we do not have an example of a polynomial that is #**P**-hard to approximate (either multiplicatively or additively) on average, in the sense that we need.

References

1. Scott Aaronson. BQP and the polynomial hierarchy. In Leonard J. Schulman, editor, *STOC*, pages 141–150. ACM, 2010.
2. Scott Aaronson. The equivalence of sampling and searching. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:128, 2010.
3. Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. *Theory of Computing*, 9:143–252, 2013.
4. Andrew C Berry. The accuracy of the gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 49(1):122–136, 1941.
5. G. E. P. Box and M. E. Muller. A note on the generation of random normal deviates. *Annals of Mathematical Statistics*, 29:610–611, 1958.
6. Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2010. doi:10.1098/rspa.2010.0301.
7. Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Average-case complexity versus approximate simulation of commuting quantum computations. *CoRR*, abs/1504.07999, 2015. URL: http://arxiv.org/abs/1504.07999.
8. James R. Driscoll, Dennis M. Healy Jr., and Daniel N. Rockmore. Fast discrete polynomial transforms with applications to data analysis for distance transitive graphs. *SIAM J. Comput.*, 26(4):1066–1099, 1997.
9. Edward Farhi and Aram W Harrow. Quantum supremacy through the quantum approximate optimization algorithm, 2016.

10 Bill Fefferman, Ronen Shaltiel, Christopher Umans, and Emanuele Viola. On beating the hybrid argument. *Theory of Computing*, 9:809–843, 2013. doi:10.4086/toc.2013.v009a026.

11 Bill Fefferman and Chris Umans. On pseudorandom generators and the BQP vs PH problem. *QIP*, 2011.

12 Richard Jozsa and Marrten Van Den Nest. Classical simulation complexity of extended clifford circuits. *Quantum Info. Comput.*, 14(7&8):633–648, May 2014. URL: http://dl.acm.org/citation.cfm?id=2638682.2638689.

13 A.Y Kitaev, A.H Shen, and M.N Vyalyi. *Quantum and Classical Computation*. AMS, 2002.

14 Donald E. Knuth. *The Art of Computer Programming, Volume III: Sorting and Searching*. Addison-Wesley, 1973.

15 Tomoyuki Morimae, Keisuke Fujii, and Joseph F. Fitzsimons. Hardness of classically simulating the one-clean-qubit model. *Phys. Rev. Lett.*, 112:130502, Apr 2014. doi:10.1103/PhysRevLett.112.130502.

16 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quanum Information*. Cambridge U.P., 2000.

17 Larry J. Stockmeyer. On approximation algorithms for #P. *SIAM J. Comput.*, 14(4):849–861, 1985.

18 Terence Tao and Van Vu. On the permanent of random Bernoulli matrices. In *Advances in Mathematics*, page 75, 2008.

19 Barbara M. Terhal and David P. DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. *CoRR*, quant-ph/0205133, 2002. URL: http://arxiv.org/abs/quant-ph/0205133.

20 Seinosuke Toda. PP is as hard as the Polynomial-Time Hierarchy. *SIAM J. Comput.*, 20(5):865–877, 1991.

A Computation of the Variance of Efficiently Specifiable Polynomial

In this section we compute the variance of the distribution induced by an Efficiently Specifiable polynomial Q with assignments to the variables chosen independently from the $\mathcal{B}(0,k)$ distribution. We will denote this throughout the section by $\mathrm{Var}\left[Q(X)\right]$. Recall, by the definition of Efficiently Specifiable, we have that Q is an n variate homogenous multilinear polynomial with $\{0,1\}$ coefficients. Assume Q is of degree d and has m monomials. Let each $[-k,k]$ valued variable X_i be independently distributed from $\mathcal{B}(0,k)$.

We adopt the notation whereby, for $j \in [m], l \in [d]$, x_{j_l} is the l-th variable in the j-th monomial of Q.

Using the notation we can express $Q(X_1,...,X_n) = \sum_{j=1}^{m} \prod_{l=1}^{d} X_{j_l}$. By independence of these random variables and since they are mean 0, it suffices to compute the variance of each monomial and multiply by m:

$$\mathrm{Var}\left[Q(X)\right] = \mathrm{Var}\left[Q(X_1,...,X_n)\right] = \mathrm{E}\left[\sum_{j=1}^{m}\prod_{l=1}^{d} X_{j_l}^2\right] = \sum_{j=1}^{m}\mathrm{E}\left[\prod_{l=1}^{d} X_{j_l}^2\right] \tag{4}$$

$$= m\mathrm{E}\left[\prod_{l=1}^{d} X_{1_l}^2\right] = m\prod_{l=1}^{d}\mathrm{E}\left[X_{1_l}^2\right] \tag{5}$$

$$= m\left(\mathrm{E}\left[X_{1_1}^2\right]\right)^d \tag{6}$$

Now since these random variables are independent and identically distributed, we can calculate the variance of an arbitrary X_{j_l} for any $j \in [m]$ and $l \in [d]$:

$$\mathrm{E}\left[X_{jl}^2\right] = \frac{1}{2^k}\sum_{i=0}^{k}\left[(k-2i)^2\binom{k}{i}\right] \tag{7}$$

$$\tag{8}$$

Thus, the variance of Q is:

$$m\frac{1}{2^{kd}}\left(\sum_{i=0}^{k}\left[(k-2i)^2\binom{k}{i}\right]\right)^d.$$

It will be useful to calculate this variance in a different way, and obtain a simple closed form. In this way we will consider the k-valued equivalent polynomial $Q'_k : \mathbb{T}_2^{nk} \to \mathbb{R}$ which is a sum of $m' = mk^d$ multilinear monomials, each of degree d. As before we can write $Q'_k(X_1, ..., X_{nk}) = \sum_{j=1}^{m'}\prod_{l=1}^{d} X_{j_l}$. Note that the uniform distribution over assignments in $\mathbb{T}_2^{kn}$ to Q'_k induces $\mathcal{B}(0,k)^n$ over $[-k,k]^n$ assignments to Q. By the same argument as above, using symmetry and independence of random variables, we have:

$$\begin{aligned}\mathrm{Var}\left[Q(X)\right] = \mathrm{Var}\left[Q(X_1, X_2, ..., X_n)\right] =& \mathrm{Var}\left[Q'_k(X_1, X_2, ..., X_{nk})\right] &(9)\\ =& m'\prod_{l=1}^{d}\mathrm{E}\left[X_{1_l}^2\right] &(10)\\ =& m'\mathrm{E}\left[X_{1_1}^2\right]^d = 1^d m' = m' = k^d m &(11)\end{aligned}$$

B Examples of Efficiently Specifiable Polynomials

In this section we give two examples of Efficiently Specifiable polynomials.

▶ **Theorem 28. Permanent** $(x_1, ..., x_{n^2}) = \sum_{\sigma\in S_n}\prod_{i=1}^{n} x_{i,\sigma(i)}$ *is Efficiently Specifiable.*

Proof. We note that it will be convenient in this section to index starting from 0. The theorem follows from the existence of an $h_{\mathbf{Permanent}} : [0, n! - 1] \to \{0,1\}^{n^2}$ that efficiently maps the i-th permutation over n elements to a string representing its obvious encoding as an $n \times n$ permutation matrix. We will prove that such an efficiently computable $h_{\mathbf{Permanent}}$ exists and prove that its inverse, $h^{-1}_{\mathbf{Permanent}}$ is also efficiently computable.

The existence of $h_{\mathbf{Permanent}}$ follows from the so-called "factorial number system" [14], which gives an efficient bijection that associates each number in $[0, n! - 1]$ with a permutation in S_n. It is customary to think of the permutation encoded by the factorial number system as a permuted sequence of n numbers, so that each permutation is encoded in $n \log n$ bits. However, it is clear that we can efficiently transform this notation into the $n \times n$ permutation matrix.

To go from an integer $j \in [0, n! - 1]$ to its permutation we:

1. Take j to its "factorial representation", an n number sequence, where the i-th place value is associated with $(i-1)!$, and the sum of the digits multiplied by the respective place value is the value of the number itself. We achieve this representation by starting from $(n-1)!$, setting the leftmost value of the representation to $j' = \lfloor \frac{j}{(n-1)!} \rfloor$, letting the

next value be $\lfloor \frac{j-j' \cdot (n-1)!}{(n-2)!} \rfloor$ and continuing until 0. Clearly this process can be efficiently achieved and efficiently inverted, and observe that the largest each value in the i-th place value can be is i.

2. In each step we maintain a list ℓ which we think of as originally containing n numbers in ascending order from 0 to $n-1$.
3. Repeat this step n times, once for each number in the factorial representation. Going from left to right, start with the left-most number in the representation and output the value in that position in the list, ℓ. Remove that position from ℓ.
4. The resulting n number sequence is the encoding of the permutation, in the standard $n \log n$ bit encoding. ◀

Now we show that the Hamiltonian Cycle Polynomial is Efficiently Specifiable.

Given a graph G on n vertices, we say a Hamiltonian Cycle is a path in G that starts at a given vertex, visits each vertex in the graph exactly once and returns to the start vertex.

We define an n-cycle to be a Hamiltonian cycle in the complete graph on n vertices. Note that there are exactly $(n-1)!$ n-cycles.

▶ **Theorem 29.** **HamiltonianCycle** $(x_1, ..., x_{n^2}) = \sum_{\sigma:\, n-cycle} \prod_{i=1}^{n} x_{i,\sigma(i)}$ *is Efficiently Specifiable.*

Proof. We can modify the algorithm for the Permanent above to give us an efficiently computable $h_{HC} : [0, (n-1)! - 1] \to \{0,1\}^{n^2}$ with an efficiently computable h_{HC}^{-1}.

To go from a number $j \in [0, (n-1)! - 1]$ to its n-cycle we:

1. Take j to its factorial representation as above. Now this is an $n-1$ number sequence where the i-th place value is associated with $(i-1)!$, and the sum of the digits multiplied by the respective place value is the value of the number itself.
2. In each step we maintain a list ℓ which we think of as originally containing n numbers in ascending order from 0 to $n-1$.
3. Repeat this step $n-1$ times, once for each number in the factorial representation. First remove the smallest element of the list. Then going from left to right, start with the left-most number in the representation and output the value in that position in the list, ℓ. Remove that position from ℓ.
4. We output 0 as the n-th value of our n-cycle.

To take an n-cycle to a factorial representation, we can easily invert the process:

1. In each step we maintain a list ℓ which we think of as originally containing n numbers in order from 0 to $n-1$.
2. Repeat this step $n-1$ times. Remove the smallest element of the list. Going from left to right, start with the left-most number in the n-cycle and output the position of that number in the list ℓ (where we index the list starting with the 0 position). Remove the number at this position from ℓ. ◀

C The Hardness of Classical Sampling from the Compressed Distribution

In this section, we use the same ideas used in the analysis of Section 5, to invoke Stockmeyer's Theorem (Theorem 7), together with the assumed existence of a Sampler for $\mathcal{D}_{Q'_k}$ to obtain hardness consequences for classical sampling with $k \leqslant \exp(n)$.

▶ **Lemma 30.** *Given an Efficiently Specifiable polynomial Q with n variables and m monomials, let Q'_k be its k-valued equivalent polynomial, for some fixed $k \leqslant exp(n)$. Suppose we have a*

Sampler S with respect to our quantumly sampled distribution class, $\mathcal{D}_{Q'_k}$, and let $\mathrm{Var}[Q(X)]$ *denote the variance of the distribution over $\mathbb{R}$ induced by Q with assignments distributed from $\mathcal{B}(0,k)^n$. Then we can find a randomized procedure computing an $\epsilon\mathrm{Var}[Q(X)]$-additive approximate δ-average case solution to Q^2 with respect to $\mathcal{B}(0,k)^n$ in time* $\mathrm{poly}(n, 1/\epsilon, 1/\delta)$ *with access to an* **NP** *oracle.*

Proof. Setting $\nu = \epsilon\delta/16$, suppose S samples from a distribution $\mathcal{D}'$ so that $\|\mathcal{D}_{Q'_k} - \mathcal{D}'\| \leqslant \nu$. Let $p_y = \Pr_{\mathcal{D}_{Q'_k}}[y]$ and $q_y = \Pr_{\mathcal{D}'}[y]$.

We define $\phi : \{\pm 1\}^{kn} \to [-k,k]^n$ to be the map from each $\{\pm 1\}^{kn}$ assignment to its equivalence class of assignments, which is n blocks of even integral values in the interval $[-k,k]$. Note that, given a uniformly random $\{\pm 1\}^{kn}$ assignment, ϕ induces the $\mathcal{B}(0,k)$ distribution over $[-k,k]^n$.

Our procedure picks a $y \in [-k,k]^n$ distributed[2] via $\mathcal{B}(0,k)^n$, and outputs an estimate $\tilde{q}_y$. Equivalently, we analyze this procedure by considering a uniformly distributed $x \in \{\pm 1\}^{kn}$ and then returning an approximate count, $\tilde{q}_{\phi(x)}$ to $q_{\phi(x)}$. We prove that our procedure runs in time $poly(n, 1/\epsilon, 1/\delta)$ with the guarantee that:

$$\Pr_x\left[\frac{|\tilde{q}_{\phi(x)} - p_{\phi(x)}|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{\epsilon}{2^{kn}}\right] \leqslant \delta\,.$$

And by our above analysis of the quantum sampler:

$$p_{\phi(x)} = \frac{Q(\phi(x))^2\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}{2^{kn}\mathrm{Var}[Q(X)]}\,.$$

Note that: $\frac{1}{2}\sum_{y\in[-k,+k]^n}|p_y - q_y| \leqslant \nu$, which, in terms of x, because we are summing over all strings in the orbit under $(S_k)^n$ symmetry, can be written:

$$\frac{1}{2}\sum_{x\in\{\pm 1\}^{kn}}\frac{\left|p_{\phi(x)} - q_{\phi(x)}\right|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} \leqslant \nu\,.$$

First we define for each x, $\Delta_x = \frac{|p_{\phi(x)} - q_{\phi(x)}|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}$ and so $\|\mathcal{D}_{Q'_k} - \mathcal{D}'\| = \frac{1}{2}\sum_x \Delta_x$.

Note that:

$$\mathrm{E}_x[\Delta_x] = \frac{\sum_x \Delta_x}{2^{kn}} = \frac{2\nu}{2^{kn}}\,.$$

And applying Markov, $\forall j > 1$,

$$\Pr_x[\Delta_x > \frac{j2\nu}{2^{kn}}] < \frac{1}{j}\,.$$

Setting $j = \frac{4}{\delta}$ and recalling that $\nu = \frac{\epsilon\delta}{16}$, we have,

$$\Pr_x[\Delta_x > \frac{\epsilon}{2}\cdot\frac{1}{2^{kn}}] < \frac{\delta}{4}\,.$$

[2] We can do this when $k = exp(n)$ by approximately sampling from the Normal distribution, with only $poly(n)$ bits of randomness, and using this to approximate $\mathcal{B}(0,k)$ to within additive error $1/poly(n)$ e.g., [5, 4].

Then use approximate counting (with an **NP** oracle), using Theorem 7 on the randomness of S to obtain an output $\tilde{q}_y$ so that, for all $\gamma > 0$, in time polynomial in n and $\frac{1}{\gamma}$:

$$\Pr[|\tilde{q}_y - q_y| > \gamma \cdot q_y] < \frac{1}{2^n}.$$

Because we can amplify the failure probability of Stockmeyer's algorithm to be inverse exponential.

Equivalently in terms of x:

$$\Pr_x\left[\frac{|\tilde{q}_{\phi(x)} - q_{\phi(x)}|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \gamma \cdot \frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}\right] < \frac{1}{2^n}.$$

And we have:

$$\mathrm{E}_x\left[\frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}\right] \leqslant \frac{\sum_x \frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}}{2^{kn}} = \frac{1}{2^{kn}}.$$

Thus, by Markov,

$$\Pr_x[\frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{j}{2^{kn}}] < \frac{1}{j}.$$

Now, setting $\gamma = \frac{\epsilon\delta}{8}$ and applying the union bound:

$$\begin{aligned}
&\Pr_x\left[\frac{\left|\tilde{q}_{\phi(x)} - p_{\phi(x)}\right|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{\epsilon}{2^{kn}}\right] \\
&\quad\leqslant \Pr_x\left[\frac{\left|\tilde{q}_{\phi(x)} - q_{\phi(x)}\right|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{\epsilon}{2}\cdot\frac{1}{2^{kn}}\right] \\
&\qquad\quad + \Pr_x\left[\frac{\left|q_{\phi(x)} - p_{\phi(x)}\right|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{\epsilon}{2}\cdot\frac{1}{2^{kn}}\right] \\
&\quad\leqslant \Pr_x\left[\frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \frac{j}{2^{kn}}\right] \\
&\qquad\quad + \Pr\left[\frac{|\tilde{q}_{\phi(x)} - q_{\phi(x)}|}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}} > \gamma\cdot\frac{q_{\phi(x)}}{\binom{k}{\psi(\phi(x)_1)}\binom{k}{\psi(\phi(x)_2)}\cdots\binom{k}{\psi(\phi(x)_n)}}\right] \\
&\qquad\quad + \Pr_x\left[\Delta_x > \frac{\epsilon}{2}\cdot\frac{1}{2^{kn}}\right] \\
&\quad\leqslant \frac{1}{j} + \frac{1}{2^n} + \frac{\delta}{4} \\
&\quad\leqslant \frac{\delta}{2} + \frac{1}{2^n} \leqslant \delta.
\end{aligned}$$

◀