CHICAGO JOURNAL OF THEORETICAL COMPUTER SCIENCE
http://cjtcs.cs.uchicago.edu/

Vanishing-Error Approximate Degree and QMA Complexity

Alexander A. Sherstov[*] Justin Thaler[†]

April 15, 2022

Abstract: The ε-*approximate degree* of a function $f\colon X \to \{0,1\}$ is the least degree of a multivariate real polynomial p such that $|p(x) - f(x)| \le \varepsilon$ for all $x \in X$. We determine the ε-approximate degree of the element distinctness function, the surjectivity function, and the permutation testing problem, showing they are $\Theta(n^{2/3}\log^{1/3}(1/\varepsilon))$, $\tilde{\Theta}(n^{3/4}\log^{1/4}(1/\varepsilon))$, and $\Theta(n^{1/3}\log^{2/3}(1/\varepsilon))$, respectively. Previously, these bounds were known only for constant ε.

We also derive a connection between vanishing-error approximate degree and quantum Merlin–Arthur (QMA) query complexity. We use this connection to show that the QMA complexity of permutation testing is $\Omega(n^{1/4})$. This improves on the previous best lower bound of $\Omega(n^{1/6})$ due to Aaronson (*Quantum Information & Computation*, 2012), and comes somewhat close to matching a known upper bound of $O(n^{1/3})$.

1 Introduction

The ε-*approximate degree* of a function $f\colon X \to \{0,1\}$, denoted $\deg_\varepsilon(f)$, is the least degree of a multivariate real-valued polynomial p such that $|p(x) - f(x)| \le \varepsilon$ for all inputs $x \in X$. Lower bounds on approximate degree have many applications in theoretical computer science, ranging from quantum query and communication lower bounds, to oracle separations and cryptographic secret sharing schemes. Upper bounds on approximate degree have important algorithmic implications in learning theory and differential

[*] Supported by NSF grant CCF-1814947.
[†] Supported by NSF CAREER award CCF-1845125.

Key words and phrases: approximate degree, QMA, Merlin-Arthur, permutation testing

privacy, and underlie state-of-the-art circuit and formula size lower bounds. The interested reader can find a bibliographic overview of these applications in [12, 21].

This paper focuses on three well-studied functions whose approximation by polynomials has applications to quantum computing and beyond. The first function is *element distinctness* ED_n, where the input is a list of n numbers from $\{1,2,\ldots,n\}$ and the objective is to determine if the numbers are pairwise distinct. The second function is *surjectivity* $\mathrm{SURJ}_{n,r}$, where the input is a list of n numbers from the range $\{1,2,\ldots,r\}$ and the goal is to check whether every range element appears on the list. The canonical setting is $r=\lfloor cn\rfloor$ for some constant $0<c<1$. The third problem that we study is *permutation testing* $\mathrm{PTP}_{n,\alpha}$, parameterized by a constant $0<\alpha<1$. Here, the input is a list of n numbers from $\{1,2,\ldots,n\}$, and the objective is to distinguish the case when the list contains every range element from the case when the list contains at most αn range elements. In the context of polynomial approximation, it is customary to represent the input to these functions as a Boolean matrix $x=[x_{i,j}]$, where $x_{i,j}=1$ if and only if the ith element on the list equals j.[1]

Vanishing-error approximate degree

Much work in the area has focused on *bounded-error* approximate degree, defined for a Boolean function f as the quantity $\deg_{1/3}(f)$. The choice of constant $1/3$ here is arbitrary, as $\deg_\varepsilon(f)=\Theta(\deg_{1/3}(f))$ for all constants $0<\varepsilon<1/2$. In particular, the bounded-error approximate degrees of element distinctness, surjectivity, and permutation testing are known to be $\Theta(n^{2/3})$, $\tilde{\Theta}(n^{3/4})$, and $\Theta(n^{1/3})$, respectively [3, 4, 17, 1, 21, 12].[2] Our understanding of approximate degree with vanishing error, $\varepsilon=o(1)$, is far less complete. Among the very few functions whose vanishing-error approximate degree has been determined is the n-bit AND function, with the asymptotic bound $\deg_\varepsilon(\mathrm{AND}_n)=\Theta(n^{1/2}\log^{1/2}(1/\varepsilon))$ due to Buhrman et al. [11]. We give a new and entirely different proof of their result. Our technique further allows us to settle the vanishing-error approximate degrees of the much more complicated functions of element distinctness, surjectivity, and permutation testing:

Theorem 1.1. *Let $0<c<1$ and $0<\alpha<1$ be arbitrary constants. Then*

$$\deg_\varepsilon(\mathrm{ED}_n)=\Omega\left(n^{2/3}\left(\log\frac{1}{\varepsilon}\right)^{1/3}\right),$$

$$\deg_\varepsilon(\mathrm{SURJ}_{n,\lfloor cn\rfloor})=\tilde{\Omega}\left(n^{3/4}\left(\log\frac{1}{\varepsilon}\right)^{1/4}\right),$$

$$\deg_\varepsilon(\mathrm{PTP}_{n,\alpha})=\Omega\left(n^{1/3}\left(\log\frac{1}{\varepsilon}\right)^{2/3}\right)$$

for all $1/3^n\leq\varepsilon\leq1/3$.

This theorem is optimal with respect to all parameters. The lower bounds for element distinctness and surjectivity match the vanishing-error constructions in [21], whereas the lower bound for permutation

[1] See [22, Section 3] for a detailed explanation of this convention and how it relates to applications of approximate degree bounds.

[2] Throughout this manuscript, $\tilde{O}$, $\tilde{\Omega}$, and $\tilde{\Theta}$ notation hides factors polylogarithmic in n.

testing is tight by a quantum query argument which we include as Theorem 3.8. A comment is in order on ε-approximate degree in the complementary range, $\varepsilon < 1/3^n$. Routine interpolation gives an exact representation for each of the functions in Theorem 1.1 as a polynomial of degree at most n. Theorem 1.1 shows that this upper bound is asymptotically tight, settling the ε-approximate degree for $\varepsilon < 1/3^n$ as well.

We prove a result analogous to Theorem 1.1 for *k-element distinctness* ED_n^k, a well-studied generalization of ED_n. Specifically, we prove that if ED_n^k has bounded-error approximate degree $\Omega(n^\ell)$, then it has ε-approximate degree $\Omega(n^\ell \log^{1-\ell}(1/\varepsilon))$. The state-of-the-art lower bound on the bounded-error approximate degree of ED_n^k is $\tilde{\Omega}(n^{3/4-1/(2k)})$ [12], so this yields

$$\deg_\varepsilon(\mathrm{ED}_n^k) = \tilde{\Omega}\left(n^{\frac{3}{4}-\frac{1}{2k}}\left(\log\frac{1}{\varepsilon}\right)^{\frac{1}{4}+\frac{1}{2k}}\right).$$

For large k, this comes close to the best known upper bound [21]:

$$\deg_\varepsilon(\mathrm{ED}_n^k) = O\left(n^{\frac{3}{4}-\frac{1}{4(2^k-1)}}\left(\log\frac{1}{\varepsilon}\right)^{\frac{1}{4}+\frac{1}{4(2^k-1)}}\right).$$

Our techniques are quite general, and we are confident that they will find other applications in the area. The technical core of our results establishes that for any function f that contains $\mathrm{AND}_k \circ f_{\lfloor n/k \rfloor}$ as a subfunction for each $k \leq n$, any bounded-error approximate degree lower bound for f automatically implies a strong lower bound for the ε-approximate degree of f.[3] This allows us to prove tight lower bounds on the vanishing-error approximate degrees of AND_n, ED_n, ED_n^k, and $\mathrm{SURJ}_{n,r}$. To handle $\mathrm{PTP}_{n,\alpha}$, we generalize our technique to other outer functions. Our analysis is based on the so-called method of dual polynomials, whereby one proves approximate degree lower bounds by constructing explicit dual solutions to a certain linear program capturing the approximate degree of the given function.

In the remainder of the introduction, we focus on an application of Theorem 1.1 to quantum Merlin–Arthur complexity.

The Merlin–Arthur model

The Merlin–Arthur (MA) model of query complexity features a function f and two asymmetric players, Merlin and Arthur. Arthur's goal is to compute f on some unknown input x while querying as few bits of x as possible. Merlin, who knows x, can help Arthur compute $f(x)$ by sending him a single witness, i.e., an arbitrary message of some bit length m. However, Merlin is untrusted. The model requires that, for any $x \in f^{-1}(1)$, there is some Merlin message causing Arthur to output 1 with probability at least $2/3$, and for any $x \in f^{-1}(0)$, no Merlin message can cause Arthur to output 1 with probability more than $1/3$. The cost of the protocol is the sum of the witness length m and the number of bits of x queried by Arthur. In *quantum* Merlin-Arthur (QMA) query complexity, the witness sent by Merlin is allowed to be an arbitrary m-qubit quantum message, and Arthur is permitted to query bits of the input x in superposition. The MA and QMA query models have important analogues in communication complexity and Turing machine

[3] When we say that f contains g as a subfunction, we mean that there is a restriction f' of f such that the domain of g is a subset of the domain of f', and $f'(x) = g(x)$ for all x in the domain of g.

complexity. In the former setting, Arthur is replaced by two parties Alice and Bob, and the input x is split between them.

The complexity class QMA is a quantum analog of NP and accordingly has received considerable attention. It is well known that any QMA protocol can be simulated by an $\mathsf{SBQP} \subseteq \mathsf{PP}$ protocol with at most a quadratic blowup in cost, i.e., $\mathsf{QMA}(f) \geq \Omega(\mathsf{SBQP}(f)^{1/2})$ [23].[4] In turn, the existence of an SBQP query protocol that makes at most c queries implies that the *one-sided* (1/3)-approximate degree of f is at most $O(c)$. Here, the one-sided ε-approximate degree of f is the least degree of a real polynomial p such that $|p(x)| \leq \varepsilon$ for all $x \in f^{-1}(0)$, and $p(x) \geq 1-\varepsilon$ for all $x \in f^{-1}(1)$ [13] (observe that $p(x)$ is permitted to take very large values on inputs in $f^{-1}(1)$). As a consequence, one can prove QMA query lower bounds for f by lower bounding the one-sided approximate degree of f.

Only a handful of additional results are known about QMA query and communication complexity. Raz and Shpilka [20] showed that AND_n has QMA query complexity $\Theta(\sqrt{n})$. Klauck [15] showed that the QMA communication complexity of the disjointness problem is $\Omega(n^{1/3})$. Neither of these results follows from a naïve application of the bound $\mathsf{QMA}(f) \geq \Omega(\sqrt{\mathsf{SBQP}(f)})$.

QMA complexity of permutation testing

The permutation testing problem $\text{PTP}_{n,\alpha}$ has played an important role in the study of interactive proof systems because it possesses a simple non-interactive perfect zero knowledge (NIPZK) protocol of logarithmic cost, yet is a hard problem in many other models. Hence, it has been used to prove a variety of complexity class separations. In particular, Aaronson [1] showed that the QMA query complexity of $\text{PTP}_{n,\alpha}$ is $\Omega(n^{1/6})$, and thereby gave an oracle separating NIPZK from QMA. Bouland et al. [8] built on Aaronson's result to give an oracle separating non-interactive *statistical* zero knowledge (NISZK) from the complexity class UPP, answering a question of Watrous from 2002. Gur, Liu, and Rothblum [14] showed that the MA query complexity of $\text{PTP}_{n,\alpha}$ is $\Omega(n^{1/4})$. Despite this progress, the precise QMA complexity of $\text{PTP}_{n,\alpha}$ has remained open, with the best upper bound being $O(n^{1/3})$ [10, 1] and the best lower bound being Aaronson's $\Omega(n^{1/6})$. We obtain a polynomially stronger lower bound.

Theorem 1.2. *Let $0 < \alpha < 1$ be an arbitrary constant. Then any QMA query protocol for $\text{PTP}_{n,\alpha}$ with witness length m has query cost $\Omega(n/m)^{1/3}$. In particular, $\text{PTP}_{n,\alpha}$ has QMA complexity $\Omega(n^{1/4})$.*

This result quantitatively matches the MA lower bound of Gur et al. [14] but holds in the more powerful quantum setting. Theorem 1.2 comes reasonably close to matching the known QMA query upper bound of $O(n^{1/3})$, which holds even if Merlin does not send any message to Arthur; see Theorem 3.8.

To prove Theorem 1.2, we derive a connection between QMA query complexity and vanishing-error approximate degree for a class of functions that includes AND_n, ED_n, and $\text{PTP}_{n,\alpha}$. This connection amounts to the observation that, for these particular functions, the one-sided ε-approximate degree is *equal* to the ε-approximate degree. Prior work on QMA complexity (e.g., [15]) has implicitly exploited a similar observation in the special case of AND_n. Our analysis substantially generalizes the insights of prior work, and makes explicit the key phenomenon at play, namely the equivalence of one-sided vs. standard approximate degree for these functions. Combining this connection with our new vanishing-error approximate degree lower bounds in Theorem 1.1 establishes Theorem 1.2.

[4] An SBQP protocol $\mathcal{A}$ is a quantum protocol for which there is some α such that $\mathcal{A}$ accepts every input in $f^{-1}(1)$ with probability at least α, and every input in $f^{-1}(0)$ with probability at most $\alpha/2$ [16].

2 Preliminaries

For a function f, we let $\operatorname{dom} f$ and $\operatorname{im} f$ stand for the domain and image of f, respectively. We view Boolean functions as mappings $f\colon X \to \{0,1\}$ for a finite set X. For functions $g\colon X \to Y$ and $f\colon Y^n \to Z$, we let $f \circ g$ denote the block-composition of f and g. In more detail, $f \circ g\colon X^n \to Z$ is the function that maps $(x_1,\dots,x_n) \in X^n$ to $f(g(x_1),\dots,g(x_n))$. We generalize block-composition to the case when the domain of f is properly contained in Y^n by defining the domain of $f \circ g$ as the set of $(x_1,\dots,x_n) \in X^n$ such that $(g(x_1),\dots,g(x_n)) \in \operatorname{dom} f$.

2.1 Polynomial approximation

For a multivariate real polynomial $p\colon \mathbb{R}^n \to \mathbb{R}$, we let $\deg p$ denote the total degree of p, i.e., the largest degree of any monomial of p. It will be convenient to define the degree of the zero polynomial by $\deg 0 = -\infty$. For two functions $f,\psi\colon X \to \mathbb{R}$, let $\langle f,\psi\rangle = \sum_{x\in X} f(x)\psi(x)$ denote the correlation of f and ψ, and let $\|\psi\|_1 = \sum_{x\in X}|\psi(x)|$. For a real-valued function ϕ supported on a finite subset of $\mathbb{R}^n$, we define the *orthogonal content of* ϕ, denoted $\operatorname{orth}\phi$, to be the minimum degree of a real polynomial p for which $\langle \phi, p\rangle \neq 0$. We adopt the convention that $\operatorname{orth}\phi = \infty$ if no such polynomial exists. For any real-valued function $\psi\colon X \to \mathbb{R}$, its k-th tensor power $\psi^{\otimes k}\colon X^k \to \mathbb{R}$ is given by $\psi^{\otimes k}(x_1,\dots,x_k) = \psi(x_1)\cdots\psi(x_k)$.

The ε-*approximate degree* of a function $f\colon X \to \mathbb{R}$, denoted $\deg_\varepsilon(f)$, is the least degree of a polynomial $p\colon X \to \mathbb{R}$ such that $|p(x) - f(x)| \le \varepsilon$ for all $x \in X$. We emphasize that no restriction is placed on the behavior of p at inputs outside f's domain of definition, X. For most functions of interest to us, the domain X is a proper subset of $\{0,1\}^n$ and thus their approximating polynomials may take on arbitrary values on $\{0,1\}^n \setminus X$. The following dual characterization of approximate degree is well known and can be verified using linear programming duality.

Fact 2.1. *Fix $d > 0$ and a function $f\colon X \to \mathbb{R}$. Then $\deg_\varepsilon(f) \ge d$ if and only if there exists a function $\psi\colon X \to \mathbb{R}$ such that*

$$\begin{aligned}&\langle f,\psi\rangle > \varepsilon\|\psi\|_1,\\ &\operatorname{orth}\psi \ge d.\end{aligned}$$

The simplest function of interest to us is $\mathrm{AND}_n\colon \{0,1\}^n \to \{0,1\}$, given as usual by $\mathrm{AND}_n(x) = \bigwedge_{i=1}^n x_i$. Its bounded-error approximate degree was determined by Nisan and Szegedy [19], as follows.

Theorem 2.1. *For all $n \ge 1$,*

$$\deg_{1/3}(\mathrm{AND}_n) = \Theta(\sqrt{n}).$$

2.2 Surjectivity

Let $\mathcal{D}_{n,r}$ stand for the set of Boolean matrices of size $n \times r$ in which every row has exactly one 1. Every matrix $x \in \mathcal{D}_{n,r}$ has a natural interpretation as specifying a mapping $\phi\colon \{1,2,\dots,n\} \to \{1,2,\dots,r\}$, where $\phi(i) = j$ if and only if $x_{i,j} = 1$. Our next three functions are defined on $\mathcal{D}_{n,r}$ and can thus be regarded as “function properties." To start with, the *surjectivity problem* with n elements and range size r is defined as $\mathrm{SURJ}_{n,r}\colon \mathcal{D}_{n,r} \to \{0,1\}$, where

$$\mathrm{SURJ}_{n,r}(x) = \bigwedge_{j=1}^{r} \bigvee_{i=1}^{n} x_{i,j}.$$

Thus, $\mathrm{SURJ}_{n,r}$ takes as input an $n \times r$ Boolean matrix in which every row contains exactly one 1, and evaluates to 1 if and only if every column of the input contains at least one 1. Interpreting the input matrix as a mapping, $\mathrm{SURJ}_{n,r}$ evaluates to 1 if and only if that mapping is surjective. This surjectivity property is trivially false for $r > n$, and the standard setting of parameters is $r = \lfloor cn \rfloor$ for some constant $0 < c < 1$. The choice of constant c is unimportant because it affects $\deg_{1/3}(\mathrm{SURJ}_{n,\lfloor cn \rfloor})$ by at most a multiplicative constant. It was shown in [21] that the surjectivity function has bounded-error approximate degree $O(n^{3/4})$. Bun et al. [12] gave an alternate proof of this upper bound and additionally proved that it is tight up to a polylogarithmic factor. We thus have:

Theorem 2.2. *Let $0 < c < 1$ be an arbitrary constant. Then*

$$\deg_{1/3}(\mathrm{SURJ}_{n,\lfloor cn \rfloor}) = \tilde{\Theta}(n^{3/4}).$$

2.3 Element distinctness

Another well-studied function is *element distinctness* $\mathrm{ED}_{n,r}\colon \mathcal{D}_{n,r} \to \{0,1\}$, defined by $\mathrm{ED}_{n,r}(x) = 1$ if and only if every column of the input matrix x has at most one 1. Switching to the interpretation of x as a mapping, $\mathrm{ED}_{n,r}(x)$ evaluates to true if and only if the mapping is one-to-one. This property is trivially false for $r < n$. In the complementary case, Ambainis [4] proved that for any given ε, the ε-approximate degree of $\mathrm{ED}_{n,r}$ is the same for all $r \geq n$. This means that one may without loss of generality focus on the special case $r = n$, with the shorthand notation $\mathrm{ED}_n = \mathrm{ED}_{n,n}$. Aaronson and Shi [3], Ambainis [4], and Kutin [17] showed that ED_n has bounded-error approximate degree $\Omega(n^{2/3})$, matching an upper bound of Ambainis [5].

Theorem 2.3. *For all $n \geq 1$,*

$$\deg_{1/3}(\mathrm{ED}_n) = \Theta(n^{2/3}).$$

Element distinctness generalizes in a natural way to a function called *k-element distinctness*, denoted $\mathrm{ED}_n^k\colon \mathcal{D}_{n,n} \to \{0,1\}$. This new function evaluates to true if and only if the input matrix has no column with k or more 1s. Viewing the input as a mapping, ED_n^k evaluates to true if and only if no range element occurs k or more times. With these definitions, we have $\mathrm{ED}_n = \mathrm{ED}_n^2$. Bun, Kothari, and Thaler [12] proved the following lower bound for ED_n^k.

Theorem 2.4. *Let $k > 1$ be any positive integer. Then*

$$\deg_{1/3}\left(\mathrm{ED}_n^k\right) = \tilde{\Omega}\left(n^{\frac{3}{4}-\frac{1}{2k}}\right).$$

2.4 Permutation testing

The final problem of interest to us is a restriction of element distinctness ED_n. In more detail, fix an integer $n \geq 1$ and a real number $0 < \alpha < 1$. The domain of the *permutation testing problem* $\mathrm{PTP}_{n,\alpha}$ is the set of all matrices $x \in \mathcal{D}_{n,n}$ in which the number of columns containing a 1 is either exactly n or at most αn. The function evaluates to true in the former case and to false in the latter. Equivalently, $\mathrm{PTP}_{n,\alpha}(x) = 1$ if and only if x is a permutation matrix. In the regime of interest to us, $0 < \alpha < 1$ is a constant independent of n.

The permutation testing problem was introduced by Aaronson [1], who defined it somewhat differently. In his variant of permutation testing, which we denote by $\mathrm{PTP}^*_{n,\delta}$, one is given a matrix $x \in \mathcal{D}_{n,n}$ that is either (i) a permutation matrix, or (ii) disagrees from every permutation matrix in at least δn rows. The function evaluates to true in case (i) and to false in case (ii). As the following proposition shows, Aaronson's $\mathrm{PTP}^*_{n,\delta}$ is precisely the same function as our $\mathrm{PTP}_{n,1-\delta}$.

Proposition 2.5. *Let $0 < \delta < 1$ and $n \geq 1$ be given. Then as functions,*

$$\mathrm{PTP}^*_{n,\delta} = \mathrm{PTP}_{n,1-\delta}.$$

Specifically, the l.h.s. and r.h.s. have the same domain and agree at every point thereof.

Proof. This claim is easiest to verify by interpreting an input $x \in \mathcal{D}_{n,n}$ as a mapping $\phi\colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$. A moment's reflection shows that ϕ disagrees from every permutation $\{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ in at least $n - |\operatorname{im}\phi|$ points, and there is a permutation that achieves this lower bound. Restating this in matrix terminology, a matrix $x \in \mathcal{D}_{n,n}$ disagrees from every permutation matrix in at least δn rows if and only if the number of columns of x containing a 1 is at most $n - \delta n$. □

By adapting earlier analyses of element distinctness, Aaronson [1] obtained the following result.

Theorem 2.6. *Let $0 < \delta < 1$ be an arbitrary constant. Then*

$$\deg_{1/3}(\mathrm{PTP}^*_{n,\delta}) = \Omega(n^{1/3}).$$

This result is stated in [1] specifically for $\delta = 1/8$, but the proof actually allows any $0 < \delta < 1$. Combining this theorem with Proposition 2.5 gives the following corollary.

Corollary 2.7. *Let $0 < \alpha < 1$ be an arbitrary constant. Then*

$$\deg_{1/3}(\mathrm{PTP}_{n,\alpha}) = \Omega(n^{1/3}).$$

We close this section with a remark on input encoding. In this work, functions like $\mathrm{SURJ}_{n,r}$ take as input a Boolean matrix x in which every row has exactly one 1. Some other works [7, 12] represent the input as a list $y_1,\ldots,y_n \in \{0,1\}^{\lceil \log r \rceil}$, where y_i encodes the location of the unique 1 in the i-th row of the matrix representation x. Switching to this alternate representation affects the approximate degree by at most a logarithmic factor. See [21] for a detailed treatment of the relationship between these representations.

3 Approximate Degree Lower Bounds

In this section, we study the vanishing-error approximate degree of element distinctness, surjectivity, and permutation testing, and in particular settle Theorem 1.1 from the introduction. The core of our technique is the following auxiliary result.

Proposition 3.1. *For any $\varepsilon \geq 0$ and any function $f\colon X \to \mathbb{R}$ on a finite subset X of Euclidean space,*

$$\deg_{\varepsilon^k}(f^{\otimes k}) \geq k \deg_{\varepsilon}(f), \qquad k = 1,2,3,\ldots.$$

In particular, every function $f\colon X \to \{0,1\}$ satisfies

$$\deg_{\varepsilon^k}(\mathrm{AND}_k \circ f) \geq k \deg_{\varepsilon}(f), \qquad k = 1,2,3,\ldots.$$

Proof. We may assume that $\deg_{\varepsilon}(f) \neq 0$ since the proposition is trivial otherwise. Let ψ be an ε-error dual polynomial for f, as guaranteed by Fact 2.1:

$$\langle f, \psi\rangle > \varepsilon \|\psi\|_1,$$
$$\operatorname{orth} \psi = \deg_{\varepsilon}(f).$$

Then

$$\begin{aligned}\langle f^{\otimes k}, \psi^{\otimes k}\rangle &= \langle f, \psi\rangle^k \\ &> (\varepsilon\|\psi\|_1)^k \\ &= \varepsilon^k \|\psi^{\otimes k}\|_1.\end{aligned}$$

Applying Fact 2.1 once again,

$$\begin{aligned}\deg_{\varepsilon^k}(f^{\otimes k}) &\geq \operatorname{orth} \psi^{\otimes k} \\ &\geq k \operatorname{orth} \psi \\ &\geq k \deg_{\varepsilon}(f).\end{aligned}$$

Here, the penultimate inequality holds by the following reasoning. Let

$$p(x_1,\ldots,x_k)\colon X^k \to \mathbb{R}$$

be any polynomial of degree less than $k \operatorname{orth} \psi$, where variable x_i takes values in X. We must show that

$$\langle \psi^{\otimes k}, p\rangle = 0. \tag{3.1}$$

Consider any monomial of p. We may express this monomial as $\prod_{i=1}^k p_i(x_i)$, where $\sum_{i=1}^k \deg(p_i) \leq \deg(p)$. By the pigeonhole principle, there is some $i \in \{1,\ldots,k\}$ such that $\deg(p_i) < \operatorname{orth} \psi$, and hence $\langle \psi, p_i\rangle = 0$. It follows that $\langle \psi^{\otimes k}, \prod_{i=1}^k p_i(x_i)\rangle = \prod_{i=1}^k \langle \psi, p_i\rangle = 0$. Since p is a sum of monomials, Equation (3.1) follows by linearity. □

The proof of Proposition 3.1 applies more generally to the conjunction of k distinct functions, but we will not need this generalization.

3.1 Warmup

To illustrate our technique in the simplest possible setting, we consider the well-studied AND_n function. Buhrman et al. [11] proved that its ε-error approximate degree is $\Theta(\sqrt{n\log(1/\varepsilon)})$. We give a new and simple proof of their lower bound.

Theorem 3.2. *For all* $1/3^n \le \varepsilon \le 1/3$,

$$\deg_\varepsilon(\mathrm{AND}_n) = \Omega\left(\sqrt{n\log\frac{1}{\varepsilon}}\right). \tag{3.2}$$

Proof. For $k = 1,2,\ldots,n$, we have

$$\begin{aligned}\deg_{3^{-k}}(\mathrm{AND}_n) &\ge \deg_{3^{-k}}(\mathrm{AND}_k \circ \mathrm{AND}_{\lfloor n/k\rfloor})\\ &\ge k\deg_{1/3}(\mathrm{AND}_{\lfloor n/k\rfloor})\\ &= k\cdot\Omega\left(\sqrt{\frac{n}{k}}\right)\\ &= \Omega(\sqrt{nk}),\end{aligned}$$

where the first, second, and third steps use the identity $\mathrm{AND}_{n_1n_2} = \mathrm{AND}_{n_1} \circ \mathrm{AND}_{n_2}$, Proposition 3.1, and Theorem 2.1, respectively. This directly implies (3.2). □

3.2 Element distinctness

Our next result is a tight lower bound on the vanishing error approximate degree of element distinctness, matching the upper bound from [21].

Theorem 3.3. *For all* $1/3^n \le \varepsilon \le 1/3$,

$$\deg_\varepsilon(\mathrm{ED}_n) = \Omega\left(n^{2/3}\left(\log\frac{1}{\varepsilon}\right)^{1/3}\right). \tag{3.3}$$

Proof. For any $k = 1,2,3,\ldots,n$, we claim that $\mathrm{AND}_k \circ \mathrm{ED}_{\lfloor n/k\rfloor}$ is a subproblem of ED_n. That is, we identify a restriction f' of f such that the domain of g is a subset of the domain of f', and $f'(x) = g(x)$ for all x in the domain of g.

To see why, recall that the input to ED_n is an $n\times n$ Boolean matrix in which every row i contains exactly one 1, corresponding to the value of the ith element. Now, fix $k \in \{1,2,\ldots,n\}$ and consider the restriction of ED_n to input matrices that are *block-diagonal,* with k blocks of size $\lfloor n/k\rfloor$ each and an additional block of $n-k\lfloor n/k\rfloor$ ones on the diagonal. Each of the first k blocks corresponds to an instance of $\mathrm{ED}_{\lfloor n/k\rfloor}$, and the overall problem amounts to computing the AND of these k instances. Therefore, $\mathrm{AND}_k \circ \mathrm{ED}_{\lfloor n/k\rfloor}$ is a subproblem of ED_n, and

$$\deg_\varepsilon(\mathrm{ED}_n) \ge \deg_\varepsilon(\mathrm{AND}_k \circ \mathrm{ED}_{\lfloor n/k\rfloor}) \tag{3.4}$$

for all ε and all $k = 1,2,3,\ldots,n$.

The rest of the proof is closely analogous to that for AND_n. For $k = 1, 2, \ldots, n$,

$$\begin{aligned}\deg_{3^{-k}}(\mathrm{ED}_n) &\geq \deg_{3^{-k}}(\mathrm{AND}_k \circ \mathrm{ED}_{\lfloor n/k \rfloor}) \\ &\geq k \deg_{1/3}(\mathrm{ED}_{\lfloor n/k \rfloor}) \\ &\geq k \cdot \Omega\left(\frac{n}{k}\right)^{2/3} \\ &= \Omega(n^{2/3} k^{1/3})\end{aligned}$$

where the first three steps use (3.4), Proposition 3.1, and Theorem 2.3, respectively. This directly implies (3.3). □

The previous proof shows more generally that $\mathrm{AND}_k \circ \mathrm{ED}^r_{\lfloor n/k \rfloor}$ is a subfunction of ED^r_n for any $k = 1, 2, \ldots, n$. As a result, our analysis of element distinctness proves the following statement.

Theorem 3.4. *Fix constants $r \geq 2$ and $\ell \in [0,1]$ such that*

$$\deg_{1/3}(\mathrm{ED}^r_n) = \Omega(n^\ell).$$

Then

$$\deg_\varepsilon(\mathrm{ED}^r_n) = \Omega\left(n^\ell \left(\log \frac{1}{\varepsilon}\right)^{1-\ell}\right), \qquad \frac{1}{3^n} \leq \varepsilon \leq \frac{1}{3}.$$

Combining Theorem 3.4 with Theorem 2.4, we conclude that

$$\deg_\varepsilon(\mathrm{ED}^r_n) = \tilde{\Omega}\left(n^{\frac{3}{4} - \frac{1}{2r}} \left(\log \frac{1}{\varepsilon}\right)^{\frac{1}{4} + \frac{1}{2r}}\right)$$

for $1/3^n \leq \varepsilon \leq 1/3$. Moreover, Theorem 3.4 will, in a black-box manner, translate any future improvement in the bounded-error lower bound for ED^r_n into an improved vanishing-error lower bound.

3.3 Surjectivity

An instance x of the surjectivity problem $\mathrm{SURJ}_{n,r}$ can be embedded inside a larger instance of surjectivity in many ways, e.g., by duplicating a row of x or by forming a block-diagonal matrix with blocks x and 1. These two transformations yield

$$\deg_\varepsilon(\mathrm{SURJ}_{n,r}) \leq \deg_\varepsilon(\mathrm{SURJ}_{n+1,r}), \tag{3.5}$$

$$\deg_\varepsilon(\mathrm{SURJ}_{n,r}) \leq \deg_\varepsilon(\mathrm{SURJ}_{n+1,r+1}), \tag{3.6}$$

respectively. We will now prove an essentially tight lower bound on the vanishing-error approximate degree of surjectivity, matching the upper bound from [21] up to a logarithmic factor.

Theorem 3.5. *Let $0 < c < 1$ be an arbitrary constant. Then*

$$\deg_\varepsilon(\mathrm{SURJ}_{n,\lfloor cn \rfloor}) = \tilde{\Omega}\left(n^{3/4} \left(\log \frac{1}{\varepsilon}\right)^{1/4}\right), \qquad \frac{1}{3^n} \leq \varepsilon \leq \frac{1}{3}.$$

Proof. The proof is a cosmetic adaptation of the analysis of element distinctness. To start with, we claim that for any positive integers n,r,k such that $k \mid n$ and $k \mid r$, the composition $\mathrm{AND}_k \circ \mathrm{SURJ}_{n/k,r/k}$ is a subproblem of $\mathrm{SURJ}_{n,r}$. Indeed, the input to $\mathrm{SURJ}_{n,r}$ is an $n \times r$ Boolean matrix in which every row i contains exactly one 1. Consider the restriction of $\mathrm{SURJ}_{n,r}$ to input matrices that are block-diagonal, with k blocks of size $n/k \times r/k$ each. Each of these blocks corresponds to an instance of $\mathrm{SURJ}_{n/k,r/k}$, and the overall problem amounts to computing the AND of these k instances. This settles the claim.

Now let n be arbitrary. Then for all positive integers $k \leq \min\{cn, (1-c)n\}$,

$$\begin{aligned}
\deg_{3^{-k}}(\mathrm{SURJ}_{n,\lfloor cn\rfloor}) &\geq \deg_{3^{-k}}(\mathrm{SURJ}_{n-(\lfloor cn\rfloor - k\lfloor cn/k\rfloor),k\lfloor cn/k\rfloor}) \\
&\geq \deg_{3^{-k}}(\mathrm{SURJ}_{n-k,k\lfloor cn/k\rfloor}) \\
&\geq \deg_{3^{-k}}(\mathrm{SURJ}_{k(\lfloor n/k\rfloor -1),k\lfloor cn/k\rfloor}) \\
&\geq \deg_{3^{-k}}(\mathrm{AND}_k \circ \mathrm{SURJ}_{\lfloor n/k\rfloor -1,\lfloor cn/k\rfloor}) \\
&\geq k\deg_{1/3}(\mathrm{SURJ}_{\lfloor n/k\rfloor -1,\lfloor cn/k\rfloor}) \\
&\geq k\cdot\tilde{\Omega}\left(\frac{n}{k}\right)^{3/4} \\
&= \tilde{\Omega}(n^{3/4}k^{1/4}),
\end{aligned}$$

where the first step uses (3.6); the second and third steps use (3.5); the fourth step applies the claim from the opening paragraph of the proof; the fifth step is valid by Proposition 3.1; and the sixth step invokes Theorem 2.2. This settles the theorem. □

3.4 Permutation testing

We now turn to the permutation testing problem, which requires a more subtle analysis than the functions that we have examined so far. The difficulty is that permutation testing does not admit a self-reduction with AND as an outer function. To address this, we will need to generalize Proposition 3.1 appropriately. For a real $0 \leq \alpha < 1$ and an integer $k \geq 1$, we define $\mathrm{AND}_{k,\alpha}$ to be the restriction of AND_k to inputs whose Hamming weight is either k or at most αk. The following result subsumes Proposition 3.1 as the special case $\alpha = (k-1)/k$.

Proposition 3.6. *Fix a real number $0 \leq \alpha < 1$ and an integer $k \geq 1$. Then for any $\varepsilon \geq 0$ and any function $f\colon X \to \{0,1\}$ on a finite subset X of Euclidean space,*

$$\deg_{\varepsilon^k/\binom{k-1}{\lfloor\alpha k\rfloor}}(\mathrm{AND}_{k,\alpha} \circ f) \geq (\lfloor \alpha k\rfloor + 1)\deg_\varepsilon(f).$$

In particular,

$$\deg_{(\varepsilon/2)^k}(\mathrm{AND}_{k,\alpha} \circ f) \geq \alpha k \deg_\varepsilon(f).$$

Proof. We may assume that $\deg_\varepsilon(f) \neq 0$ since the proposition is trivial otherwise. Let ψ be an ε-error dual polynomial for f, as guaranteed by Fact 2.1:

$$\begin{aligned}
&\langle f,\psi\rangle > \varepsilon\|\psi\|_1, \\
&\operatorname{orth}\psi = \deg_\varepsilon(f).
\end{aligned}$$

Abbreviate $\ell = \lfloor \alpha k \rfloor$ and define $\Psi\colon X^k \to \mathbb{R}$ by

$$\Psi(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} \psi(x_i) \cdot \prod_{i=\ell+1}^{k-1} (f(x_1) + f(x_2) + \cdots + f(x_k) - i).$$

Observe that Ψ is supported on the domain of $\mathrm{AND}_{k,\alpha} \circ f$. Moreover, we have the pointwise inequality

$$\begin{aligned} |\Psi| &\leq |\psi^{\otimes k}| \prod_{i=\ell+1}^{k-1} i \\ &= |\psi^{\otimes k}| \cdot \frac{(k-1)!}{\ell!}. \end{aligned} \tag{3.7}$$

Now

$$\begin{aligned} \langle \Psi, \mathrm{AND}_{k,\alpha} \circ f \rangle &= \langle \Psi, f^{\otimes k} \rangle \\ &= (k-\ell-1)! \, \langle \psi^{\otimes k}, f^{\otimes k} \rangle \\ &> (k-\ell-1)! \, \varepsilon^k \|\psi\|_1{}^k \\ &= (k-\ell-1)! \, \varepsilon^k \|\psi^{\otimes k}\|_1 \\ &\geq \varepsilon^k \cdot \frac{(k-\ell-1)! \, \ell!}{(k-1)!} \|\Psi\|_1 \\ &= \varepsilon^k \binom{k-1}{\ell}^{-1} \|\Psi\|_1, \end{aligned}$$

where the next-to-last step uses (3.7). Applying Fact 2.1 once again,

$$\begin{aligned} \deg_{\varepsilon^k / \binom{k-1}{\ell}} (\mathrm{AND}_{k,\alpha} \circ f) &\geq \operatorname{orth} \Psi \\ &\geq (\ell+1) \operatorname{orth} \psi \\ &= (\ell+1) \deg_{\varepsilon}(f). \end{aligned}$$

□

For $m \leq n$, a permutation testing instance $\phi\colon \{1,2,\ldots,m\} \to \{1,2,\ldots,m\}$ can be extended in a natural way to a larger instance $\Phi\colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ by letting $\Phi(i) = i$ for $i = m+1, m+2, \ldots, n$. This gives

$$\deg_{\varepsilon}(\mathrm{PTP}_{m,\alpha}) \leq \deg_{\varepsilon}\left(\mathrm{PTP}_{n, \frac{m}{n} \cdot \alpha + \frac{n-m}{n}}\right), \qquad m \leq n. \tag{3.8}$$

We are now in a position to prove our lower bound on the ε-approximate degree of permutation testing.

Theorem 3.7. *Let* $0 < \alpha < 1$ *be a given constant. Then*

$$\deg_{\varepsilon}(\mathrm{PTP}_{n,\alpha}) = \Omega\left(n^{1/3} \left(\log \frac{1}{\varepsilon}\right)^{2/3}\right), \qquad \frac{1}{3^n} \leq \varepsilon \leq \frac{1}{3}. \tag{3.9}$$

Proof. Let $0 < \beta < 1$ be arbitrary. We claim that for any positive integers n and k with $k \mid n$, the permutation testing function $\mathrm{PTP}_{n,\beta}$ contains

$$\mathrm{AND}_{k,\beta/2} \circ \mathrm{PTP}_{n/k,\beta/2} \tag{3.10}$$

as a subfunction. The proof is similar to that for element distinctness. Specifically, view instances of (3.10) as block-diagonal matrices with k blocks of size n/k each. Then a positive instance of (3.10) is a permutation matrix and therefore a positive instance of $\mathrm{PTP}_{n,\beta}$. A negative instance of (3.10), on the other hand, features at least $k - \frac{\beta}{2}k$ blocks from $(\mathrm{PTP}_{n/k,\beta/2})^{-1}(0)$ and therefore corresponds to a mapping $\{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ with a range of size at most

$$n - \left(k - \frac{\beta k}{2}\right) \cdot \left(\frac{n}{k} - \frac{\beta n}{2k}\right) \leq \beta n.$$

In particular, any negative instance of (3.10) is also a negative instance of $\mathrm{PTP}_{n,\beta}$. This completes the proof of the claim.

Now for any $\varepsilon \geq 0$ and any $k \in \{1,2,\ldots,\lceil \alpha n/2 \rceil\}$, we have

$$\begin{aligned} \deg_\varepsilon(\mathrm{PTP}_{n,\alpha}) &\geq \deg_\varepsilon(\mathrm{PTP}_{k\lfloor n/k \rfloor,\alpha/2}) \\ &\geq \deg_\varepsilon(\mathrm{AND}_{k,\alpha/4} \circ \mathrm{PTP}_{\lfloor n/k \rfloor,\alpha/4}), \end{aligned} \tag{3.11}$$

where the first inequality uses (3.8), and the second inequality follows from the claim established in the previous paragraph. The rest of the proof is analogous to those for AND_n and ED_n. For $k = 1,2,\ldots,\lceil \alpha n/2 \rceil$,

$$\begin{aligned} \deg_{6^{-k}}(\mathrm{PTP}_{n,\alpha}) &\geq \deg_{6^{-k}}(\mathrm{AND}_{k,\alpha/4} \circ \mathrm{PTP}_{\lfloor n/k \rfloor,\alpha/4}) \\ &\geq \frac{\alpha k}{4} \deg_{1/3}(\mathrm{PTP}_{\lfloor n/k \rfloor,\alpha/4}) \\ &= \frac{\alpha k}{4} \cdot \Omega\left(\frac{n}{k}\right)^{1/3} \\ &= \Omega(n^{1/3}k^{2/3}), \end{aligned}$$

where the first three steps are valid by (3.11), Proposition 3.6, and Corollary 2.7, respectively. This directly implies (3.9). □

We will now show that Theorem 3.7 is optimal with respect to all parameters. In fact, we will prove the stronger result that permutation testing has an ε-error quantum query algorithm with cost $O(n^{1/3}\log^{2/3}(1/\varepsilon))$. Our quantum algorithm is inspired by the well-known algorithm for the collision problem due to Brassard et al. [10].

Theorem 3.8. *Let $0 < \alpha < 1$ be a given constant. Then for all $n \geq 1$ and $1/3^n \leq \varepsilon \leq 1/3$, the permutation testing problem $\mathrm{PTP}_{n,\alpha}$ has an ε-error quantum query algorithm with cost $O(n^{1/3}\log^{2/3}(1/\varepsilon))$. In particular,*

$$\deg_\varepsilon(\mathrm{PTP}_{n,\alpha}) = O\left(n^{1/3}\left(\log\frac{1}{\varepsilon}\right)^{2/3}\right). \tag{3.12}$$

Proof. We give an algorithm whose only quantum component is Grover search. Specifically, we will only use the fact that, given query access to N items of which M are marked, Grover search finds a marked item with probability $2/3$ using $O(\sqrt{N/M})$ queries (see, e.g., [10, 9]). We will follow the convention in the quantum query literature and view the input to $\mathrm{PTP}_{n,\alpha}$ as a function $\phi\colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$, where the algorithm has query access to ϕ.

Let s be an integer parameter to be determined later. Our algorithm starts by choosing a uniformly random subset $S \subseteq \{1,2,\ldots,n\}$ of cardinality $|S| = s$. Next, we query ϕ at every point of S. If ϕ is not one-to-one on S, we output "false." In the complementary case, we execute Grover search $\log(1/\varepsilon)$ times independently, each time looking for a point $i \in \overline{S}$ with the property that $\phi(i) \in \phi(S)$. We output "false" if such a point is found, and "true" otherwise.

If ϕ is a permutation, the described algorithm is always correct. In the complementary case when $|\operatorname{im}\phi| \leq \alpha n$, there are at least $(1-\alpha)n$ points $i \in \{1,2,\ldots,n\}$ such that $|\phi^{-1}(\phi(i))| \geq 2$. Call such points *special.* We will henceforth assume that S contains at least $(1-\alpha)s/2$ special points, which happens with probability at least $1-\exp(-\Theta_\alpha(s))$, where the Θ_α notation hides factors that depend only on α. If ϕ is not one-to-one on S, the algorithm correctly outputs "false." If ϕ is one-to-one on S and S contains at least $(1-\alpha)s/2$ special points, then each of the Grover executions has $\geq (1-\alpha)s/2$ eligible points to output from among a total of $|\overline{S}| = n-s$ possibilities; this means that each Grover execution finds an eligible point with probability at least $2/3$ using $O(\sqrt{n/((1-\alpha)s)})$ queries, thereby forcing the correct output. In summary, the described algorithm has error probability at most $\exp(-\Theta_\alpha(s)) + (1/3)^{\log(1/\varepsilon)}$ and query cost $s + O(\sqrt{n/((1-\alpha)s)} \cdot \log(1/\varepsilon))$. In particular, error ε can be achieved with query cost $O(n^{1/3}\log^{2/3}(1/\varepsilon))$ by setting $s = \Theta(n^{1/3}\log^{2/3}(1/\varepsilon))$. This query bound in turn implies (3.12) using the standard transformation of a quantum query algorithm to a polynomial; see, e.g., Beals et al. [6]. □

4 QMA Lower Bounds

The objective of this section is to "lift" the approximate degree lower bound of Theorem 3.7 to QMA query complexity. As our first step, we generalize our lower bound to one-sided approximation. The *one-sided* ε-approximate degree of a function $f\colon X \to \mathbb{R}$, denoted $\deg^+_\varepsilon(f)$, is the least degree of a polynomial $p\colon X \to \mathbb{R}$ such that $|p(x)| \leq \varepsilon$ for all $x \in f^{-1}(0)$, and $p(x) \geq 1-\varepsilon$ for all $x \in f^{-1}(1)$. Thus, p approximates f uniformly on $f^{-1}(0)$ but may take on arbitrarily large values on $f^{-1}(1)$. It is clear from the definition that $\deg^+_\varepsilon(f) \leq \deg_\varepsilon(f)$. The gap between these quantities can be large in general, such as 1 versus $\Omega(\sqrt{n})$ for the bounded-error approximation of OR_n. However, we will show that these two notions of approximation are equivalent for the permutation testing function.

Proposition 4.1. *For all α, ε, and n,*

$$\deg^+_\varepsilon(\mathrm{PTP}_{n,\alpha}) = \deg_\varepsilon(\mathrm{PTP}_{n,\alpha}). \tag{4.1}$$

This equality of approximate degree and one-sided approximate degree for permutation testing has the important consequence that the lower bound of Theorem 3.7 applies to the one-sided setting as well. The proof of Proposition 4.1 is based on the observation that any one-sided approximant for permutation testing can be symmetrized to be constant on $f^{-1}(1)$, effectively making it a two-sided approximant. This technique was used previously in [13, Theorem 2] to argue that $\deg^+_\varepsilon(\mathrm{ED}_n) = \deg_\varepsilon(\mathrm{ED}_n)$.

Proof of Proposition 4.1. Let p be a one-sided approximant for $\mathrm{PTP}_{n,\alpha}$ with error ε, so that $|p| \leq \varepsilon$ on $\mathrm{PTP}_{n,\alpha}^{-1}(0)$ and $p \geq 1-\varepsilon$ on $\mathrm{PTP}_{n,\alpha}^{-1}(1)$. Define

$$p^*(x) = \mathbf{E}\, p(\sigma x \tau), \tag{4.2}$$

where σ, τ are uniformly random permutations on $\{1,2,\ldots,n\}$, and $\sigma x \tau$ denotes the matrix obtained by permuting the rows of x according to σ and the columns according to τ. Then p^* is also a one-sided approximant for $\mathrm{PTP}_{n,\alpha}$ because $\mathrm{PTP}_{n,\alpha}^{-1}(0)$ and $\mathrm{PTP}_{n,\alpha}^{-1}(1)$ are closed under permutations of rows and columns. Moreover, p^* takes on the same value, call it M, at all $x \in \mathrm{PTP}_{n,\alpha}^{-1}(1)$ because $\sigma x \tau$ in (4.2) is a uniformly random permutation matrix in that case. As a result, the normalized polynomial $p^*/\max\{1,M\}$ approximates $\mathrm{PTP}_{n,\alpha}$ pointwise within ε. Finally, $\deg p^* \leq \deg p$ because p^* is an average of polynomials, each obtained from p by permuting the input variables. □

We will also need the following proposition, implicit in Marriott and Watrous's proof [18] of Vyalyi's result [23] on QMA and SBQP. For completeness, we include its short proof.

Proposition 4.2. *Suppose that $f\colon X \to \{0,1\}$ has a QMA query protocol with witness length m and query cost q. Then there is a polynomial $p\colon X \to \mathbb{R}$ such that*

$$\deg p = O(mq), \tag{4.3}$$

$$|p(x)| \leq 2^{-2m} \qquad \textit{for all } x \in f^{-1}(0), \tag{4.4}$$

$$p(x) \geq 2^{-m-1} \qquad \textit{for all } x \in f^{-1}(1). \tag{4.5}$$

Proof. Marriott and Watrous [18] showed that the soundness and completeness errors of the QMA query protocol for f can be driven down to 2^{-2m} without an increase in witness length, and with only a factor of $O(m)$ increase in query cost. This yields a QMA protocol $\mathcal{Q}$ for f that has witness length m, query cost $O(mq)$, and soundness and completeness errors 2^{-2m}. That is, on any input in $f^{-1}(1)$, there exists a witness that causes Arthur to accept with probability at least $1-2^{-2m}$, and on any input in $f^{-1}(0)$, for every witness that might be sent by Merlin, Arthur accepts with probability at most 2^{-2m}.

Now run $\mathcal{Q}$ with the witness fixed to the totally mixed state. This yields a quantum query algorithm $\mathcal{A}$. On inputs in $f^{-1}(0)$, the acceptance probability of $\mathcal{A}$ is at most the soundness error of $\mathcal{Q}$, which is at most 2^{-2m}. On inputs in $f^{-1}(1)$, the acceptance probability of $\mathcal{A}$ is at least $(1-2^{-2m}) \cdot 2^{-m} \geq 2^{-m-1}$. Now (4.3)–(4.5) follow from the well-known result of Beals et al. [6] that the acceptance probability of any T-query quantum algorithm on input x is a polynomial $p(x)$ of degree at most $2T$. □

We have reached our main result on the QMA complexity of permutation testing, stated as Theorem 1.2 in the introduction. For the reader's convenience, we restate the theorem here.

Theorem. *Let $0 < \alpha < 1$ be an arbitrary constant. Then any QMA query protocol for $\mathrm{PTP}_{n,\alpha}$ with witness length m has query cost $\Omega(n/m)^{1/3}$. In particular, $\mathrm{PTP}_{n,\alpha}$ has QMA complexity $\Omega(n^{1/4})$.*

Proof. Fix a QMA query protocol for $\mathrm{PTP}_{n,\alpha}$ with witness length $m \in [3,n]$ and query cost q. Then Proposition 4.2 gives a polynomial p satisfying (4.3)–(4.5). It follows that $2^{m+1}p$ approximates $\mathrm{PTP}_{n,\alpha}$ in a one-sided manner to error 2^{-m+1}, forcing $\deg^+_{2^{-m+1}}(\mathrm{PTP}_{n,\alpha}) = O(mq)$. On the other hand, taking

$\varepsilon = 2^{-m+1}$ in Theorem 3.7 and Proposition 4.1 shows that $\deg^{+}_{2^{-m+1}}(\mathrm{PTP}_{n,\alpha}) = \Omega(n^{1/3}m^{2/3})$. Comparing these complementary bounds on the one-sided approximate degree of permutation testing gives $q = \Omega(n/m)^{1/3}$ and thus $\max\{m,q\} = \Omega(n^{1/4})$. □

We remind the reader that by virtue of Proposition 2.5, the variant of permutation testing studied in this paper is equivalent to Aaronson's permutation testing problem [1]. As a result, Theorems 1.2, 3.7, and 3.8 and Proposition 4.1 remain valid with $\mathrm{PTP}_{n,\alpha}$ replaced by $\mathrm{PTP}^{*}_{n,1-\alpha}$.

5 Open Problems

A natural next step would be to close the gap between our $\Omega(n^{1/4})$ QMA lower bound for permutation testing and the known upper bound of $O(n^{1/3})$. In addition, we highlight the well-known open question of resolving the QMA communication complexity of set disjointness. The best known lower bound here is $\Omega(n^{1/3})$ [15], while the best upper bound is $O(n^{1/2})$ due to [2]. We believe that both questions highlight significant gaps in our understanding of QMA. Another natural open question is whether the naïve error-reduction method for approximate degree is optimal. Namely, it is well known that $\deg_{\varepsilon}(f) \leq O(\min\{\deg_{1/3}(f)\log(1/\varepsilon), n\})$ for every $f\colon \{0,1\}^n \to \{0,1\}$, yet this bound is not known to be tight for any such f with sublinear approximate degree. It *is* tight for some f whose domain is a proper subset of $\{0,1\}^n$, based for example on approximate counting [11].

References

[1] SCOTT AARONSON: Impossibility of succinct quantum proofs for collision-freeness. *Quantum Information & Computation*, 12(1-2):21–28, 2012. 2, 4, 7, 16

[2] SCOTT AARONSON AND ANDRIS AMBAINIS: Quantum search of spatial regions. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pp. 200–209. IEEE, 2003. 16

[3] SCOTT AARONSON AND YAOYUN SHI: Quantum lower bounds for the collision and the element distinctness problems. *J. ACM*, 51(4):595–605, 2004. [doi:10.1145/1008731.1008735] 2, 6

[4] ANDRIS AMBAINIS: Polynomial degree and lower bounds in quantum complexity: Collision and element distinctness with small range. *Theory of Computing*, 1(1):37–46, 2005. [doi:10.4086/toc.2005.v001a003] 2, 6

[5] ANDRIS AMBAINIS: Quantum walk algorithm for element distinctness. *SIAM Journal on Computing*, 37(1):210–239, 2007. 6

[6] ROBERT BEALS, HARRY BUHRMAN, RICHARD CLEVE, MICHELE MOSCA, AND RONALD DE WOLF: Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001. [doi:10.1145/502090.502097] 14, 15

[7] PAUL BEAME AND WIDAD MACHMOUCHI: The quantum query complexity of AC^0. *Quantum Information & Computation*, 12(7-8):670–676, 2012. 7

[8] ADAM BOULAND, LIJIE CHEN, DHIRAJ HOLDEN, JUSTIN THALER, AND PRASHANT NALINI VASUDEVAN: On the power of statistical zero knowledge. In *Proceedings of the Fifty-Eighth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pp. 708–719, 2017. [doi:10.1109/FOCS.2017.71] 4

[9] GILLES BRASSARD, PETER HØYER, AND ALAIN TAPP: Quantum counting. In *Proceedings of the Twenty-Fifth International Colloquium on Automata, Languages and Programming* (ICALP), pp. 820–831, 1998. [doi:10.1007/BFb0055105] 14

[10] GILLES BRASSARD, PETER HØYER, AND ALAIN TAPP: Quantum algorithm for the collision problem. In *Encyclopedia of Algorithms*, pp. 1662–1664. Springer, 2016. [doi:10.1007/978-1-4939-2864-4_304] 4, 13, 14

[11] HARRY BUHRMAN, RICHARD CLEVE, RONALD DE WOLF, AND CHRISTOF ZALKA: Bounds for small-error and zero-error quantum algorithms. In *Proceedings of the Fortieth Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pp. 358–368, 1999. [doi:10.1109/SFFCS.1999.814607] 2, 9, 16

[12] MARK BUN, ROBIN KOTHARI, AND JUSTIN THALER: The polynomial method strikes back: Tight quantum query bounds via dual polynomials. In *Proceedings of the Fiftieth Annual ACM Symposium on Theory of Computing* (STOC), pp. 297–310, 2018. [doi:10.1145/3188745.3188784] 2, 3, 6, 7

[13] MARK BUN AND JUSTIN THALER: Hardness amplification and the approximate degree of constant-depth circuits. In *Proceedings of the Forty-Second International Colloquium on Automata, Languages and Programming* (ICALP), pp. 268–280, 2015. [doi:10.1007/978-3-662-47672-7_22] 4, 14

[14] TOM GUR, YANG P. LIU, AND RON D. ROTHBLUM: An exponential separation between MA and AM proofs of proximity. In *Proceedings of the Forty-Fifth International Colloquium on Automata, Languages and Programming* (ICALP), pp. 73:1–73:15, 2018. [doi:10.4230/LIPIcs.ICALP.2018.73] 4

[15] HARTMUT KLAUCK: On Arthur Merlin games in communication complexity. In *Proceedings of the Twenty-Sixth Annual IEEE Conference on Computational Complexity* (CCC), 2011. [doi:10.1109/CCC.2011.33] 4, 16

[16] GREG KUPERBERG: How hard is it to approximate the Jones polynomial? *Theory of Computing*, 11:183–219, 2015. [doi:10.4086/toc.2015.v011a006] 4

[17] SAMUEL KUTIN: Quantum lower bound for the collision problem with small range. *Theory of Computing*, 1(1):29–36, 2005. [doi:10.4086/toc.2005.v001a002] 2, 6

[18] CHRIS MARRIOTT AND JOHN WATROUS: Quantum Arthur–Merlin games. *Computational Complexity*, 14(2):122–152, 2005. [doi:10.1007/s00037-005-0194-x] 15

[19] NOAM NISAN AND MARIO SZEGEDY: On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4:301–313, 1994. [doi:10.1007/BF01263419] 5

[20] RAN RAZ AND AMIR SHPILKA: On the power of quantum proofs. In *Proceedings of the Nineteenth Annual IEEE Conference on Computational Complexity* (CCC), pp. 260–274, 2004. [doi:10.1109/CCC.2004.1313849] 4

[21] ALEXANDER A. SHERSTOV: Algorithmic polynomials. In *Proceedings of the Fiftieth Annual ACM Symposium on Theory of Computing* (STOC), pp. 311–324, 2018. [doi:10.1145/3188745.3188958] 2, 3, 6, 7, 9, 10

[22] ALEXANDER A SHERSTOV: The power of asymmetry in constant-depth circuits. *SIAM Journal on Computing*, 47(6):2362–2434, 2018. 2

[23] MIKHAIL N. VYALYI: QMA=PP implies that PP contains PH. *Electronic Colloquium on Computational Complexity*, 10(021), 2003. 4, 15

AUTHORS

Alexander A. Sherstov
Associate Professor
Computer Science Department, UCLA, Los Angeles, CA.
sherstov@cs.ucla.edu
http://web.cs.ucla.edu/~sherstov/

Justin Thaler
Associate Professor
Department of Computer Science, Georgetown University, Washington, D.C.
justin.thaler@georgetown.edu
https://people.cs.georgetown.edu/jthaler/