THEORY OF COMPUTING, Volume 12 (16), 2016, pp. 1–34
www.theoryofcomputing.org

# Dual Polynomials for Collision and Element Distinctness

Mark Bun*   Justin Thaler†

Received August 17, 2015; Revised May 8, 2016; Published October 13, 2016

**Abstract:** The approximate degree of a Boolean function $f: \{-1,1\}^n \to \{-1,1\}$ is the minimum degree of a real polynomial that approximates $f$ to within error $1/3$ in the $\ell_\infty$ norm. In an influential result, Aaronson and Shi (*J. ACM*, 2004) proved tight $\widetilde{\Omega}(n^{1/3})$ and $\widetilde{\Omega}(n^{2/3})$ lower bounds on the approximate degree of the COLLISION and ELEMENT DISTINCTNESS functions, respectively. Their proof was non-constructive, using a sophisticated symmetrization argument and tools from approximation theory.

More recently, several open problems in the study of approximate degree have been resolved via the construction of dual polynomials. These are explicit dual solutions to an appropriate linear program that captures the approximate degree of any function. We reprove Aaronson and Shi's results by constructing explicit dual polynomials for the COLLISION and ELEMENT DISTINCTNESS functions. Our constructions are heavily inspired by Kutin's (*Theory of Computing*, 2005) refinement and simplification of Aaronson and Shi's results.

**ACM Classification:** F.1.2, F.1.3

**AMS Classification:** 68Q12, 68Q17

**Key words and phrases:** polynomials, polynomial approximation, approximate degree, quantum computing, collision problem, element distinctness

*Harvard University, School of Engineering and Applied Sciences. Supported by an NDSEG Fellowship and NSF grant CNS-1237235.

†Yahoo Research. Parts of this work were performed while the author was a Research Fellow at the Simons Institute for the Theory of Computing. Supported in part by a Research Fellowship from the Simons Institute for the Theory of Computing.


DOI: 10.4086/toc.2016.v012a016

# 1 Introduction

The $\varepsilon$-approximate degree of a Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ is the least degree of a real polynomial that approximates $f$ to within error $\varepsilon$ in the $\ell_\infty$ norm. Approximate degree is a fundamental measure of the complexity of a Boolean function, and has wide-ranging applications in theoretical computer science. For example, approximate degree upper bounds underlie several of the best known algorithms for PAC learning [26], agnostic learning [24, 25], learning in the presence of irrelevant information [27, 35], and differentially private data release [49, 21]. Meanwhile, lower bounds on approximate degree imply many optimal lower bounds on quantum query complexity, circuit complexity, and communication complexity (see for example [11, 37, 5, 18, 45, 38, 13, 12, 36]).

In an influential result, Aaronson and Shi proved tight $\widetilde{\Omega}(n^{1/3})$ and $\widetilde{\Omega}(n^{2/3})$ lower bounds on the approximate degree of the COLLISION and ELEMENT DISTINCTNESS functions [5].[1] The COLLISION lower bound matched an earlier $O(n^{1/3})$ upper bound due to Brassard et al. [17], while the lower bound for ELEMENT DISTINCTNESS was later shown to be tight by Ambainis [9].

The COLLISION lower bound subsequently found many applications and extensions in quantum complexity theory; Aaronson recently provided a retrospective overview of these developments [4]. Moreover, the $\widetilde{\Omega}(n^{2/3})$ lower bound for ELEMENT DISTINCTNESS remains the best known approximate degree lower bound for any function in $\mathsf{AC}^0$.

Aaronson and Shi proved their lower bound for COLLISION with a symmetrization argument. This style of argument proceeds in two steps. First, a polynomial $p$ on $n$ variables (which is assumed to approximate the target function $f$) is transformed into a polynomial $q$ on $m < n$ variables in such a way that $\deg(q) \leq \deg(p)$. Second, a lower bound on $\deg(q)$ is proved, typically by applying Markov-Bernstein type inequalities from approximation theory. Aaronson and Shi's proof of the COLLISION lower bound is a particularly sophisticated application of this style of argument.

The lower bound for ELEMENT DISTINCTNESS follows from a reduction to the lower bound for COLLISION. This reduction is discussed in Section 5.

## 1.1 The method of dual polynomials

Despite the many applications of approximate degree in theoretical computer science, significant gaps remain in our understanding of this complexity measure, and there are many simple functions whose approximate degree remains unknown. The slow nature of progress can be attributed in part to the limitations of symmetrization arguments. At an intuitive level, the process of symmetrization is inherently lossy: by turning a polynomial $p$ on $n$ variables into a polynomial $q$ on $m < n$ variables, information about $p$ is necessarily thrown away. Hence, several works have identified that an important research direction is to develop techniques beyond symmetrization for lower bounding the approximate degree of Boolean functions [2, 41, 19].

The last few years have seen significant progress toward this goal. In particular, a series of works has proved new approximate degree lower bounds for important classes of functions by constructing

[1] Aaronson established a lower bound of $\widetilde{\Omega}(n^{1/5})$ for the COLLISION function in a paper that appeared in STOC 2002 [1], and Shi improved it to the tight $\widetilde{\Omega}(n^{1/3})$ in a FOCS paper that same year [44]. A joint journal paper appeared in 2004 [5]. The proof was simplified and extended to the "small range" case by Kutin [28]. Ambainis [7] independently extended Aaronson and Shi's lower bound to the small range case, using different techniques than Kutin.

explicit *dual polynomials*, which are dual solutions to a certain linear program capturing the approximate degree of any function. These polynomials act as certificates of the high approximate degree of a function. Moreover, strong LP duality implies that the technique is lossless, in contrast to symmetrization. That is, for any function $f$ and any $\varepsilon$, there is always some dual polynomial $\phi$ that witnesses a tight approximate degree lower bound for $f$; the challenge is to construct $\phi$.

This "method of dual polynomials" was recently used to resolve the approximate degree of the AND-OR tree [40, 19], closing a long line of incrementally larger lower bounds [44, 7, 23, 41, 30]. It has also been used to establish several "hardness amplification" results for approximate degree [48, 20, 42], and to prove new *threshold degree* lower bounds for several important classes of functions, including the intersection of two majorities [31, 41] and $\mathsf{AC}^0$ [42, 43]. The latter result represented the first superlogarithmic improvement over Minsky and Papert's seminal $\Omega(n^{1/3})$ lower bound from 1969 on the threshold degree of an $\mathsf{AC}^0$ function. We also note that dual polynomials have been used via the *pattern matrix method* [38] to resolve several longstanding open problems in communication complexity (see the survey of Sherstov [36]). The pattern matrix method uses dual polynomials to construct distributions under which various communication problems are hard (see the next section for more details).

## 1.2 Contribution and motivation

We reprove Aaronson and Shi's results by constructing explicit dual polynomials for the COLLISION and ELEMENT DISTINCTNESS functions.[2] First, we give a direct construction of a dual polynomial for COLLISION. The construction of this dual polynomial is heavily inspired by Kutin's refined proof of the COLLISION lower bound [28]. In Section 2.5, we give an overview of the ideas that go into this construction. In Section 5, we give a generic reduction which shows show how to turn any dual polynomial $\psi$ for COLLISION into a dual polynomial $\varphi$ for ELEMENT DISTINCTNESS. We construct $\varphi(x)$ by averaging $\psi(y)$ over a carefully constructed set of extensions from each $x$ to a longer input $y$.

We have four main motivations for reproving Aaronson and Shi's lower bound in this manner.

1. First, only a handful of techniques are currently known for the construction of dual polynomials, especially for the case where $\varepsilon = \Theta(1)$. To date, dual polynomials have been constructed only for symmetric functions [46, 19] and a handful of highly structured block-composed functions [20, 19, 40, 41, 42, 43, 39]. (A *block-composed* function $F\colon \{-1,1\}^{M\cdot N} \to \{-1,1\}$ is a function of the form of the form $F = g(f(x_1),\ldots,f(x_M))$ for some $g\colon \{-1,1\}^M \to \{-1,1\}$ and $f\colon \{-1,1\}^N \to \{-1,1\}$.) The COLLISION and ELEMENT DISTINCTNESS functions fall into neither category; our constructions of dual polynomials for these problems introduce several new techniques that we are optimistic will prove useful in future applications.

   In particular, we hope that our techniques will prove useful for establishing new, stronger approximate degree lower bounds for $\mathsf{AC}^0$. Up to polylogarithmic factors, the best lower bound on the approximate degree of any function in $\mathsf{AC}^0$ is the $\widetilde{\Omega}(n^{2/3})$ lower bound for the ELEMENT DISTINCTNESS function. Meanwhile, no $o(n)$ upper bound is known. Resolving the approximate degree of $\mathsf{AC}^0$ remains a significant open problem with many complexity-theoretic applications, and the new techniques that we use to construct dual polynomials may help close the gap.

[2] Like Kutin's simplification and refinement of Aaronson and Shi's original proof of the COLLISION lower bound, our construction yields a dual polynomial for the COLLISION function even in the "small-range" case.

2. A second motivation is to shed new light on the COLLISION lower bound itself. The earlier symmetrization-based proof [5, 28], while shorter than ours, is non-constructive and relies on Markov-Bernstein inequalities from approximation theory. In contrast, our proof is constructive and entirely elementary. We also believe that our analysis illuminates some of the more miraculous aspects of the earlier symmetrization-based proof—see Section 2.6 for further discussion of this point.

3. Our third motivation is to make explicit the proofs of several of the best known bounds in the complexity-theoretic study of constant-depth circuits. In particular, very recent work of Sherstov [43] exhibits a function $F$ in $\mathsf{AC}^0$ with threshold degree $\Omega(n^{1/2})$. The function $F$ is obtained by block-composing the ELEMENT DISTINCTNESS function with a certain constant-depth Boolean formula, and Sherstov's proof is via the method of dual polynomials. There is only one non-explicit element in Sherstov's construction of a dual polynomial $\phi$ witnessing the fact that the threshold degree of $F$ is in $\Omega(n^{1/2})$; he uses, in a black-box manner, the *existence* of a dual polynomial for ELEMENT DISTINCTNESS. In giving the first explicit construction of such a dual polynomial, we render Sherstov's construction of $\phi$ fully explicit. This has the following implication.

   By applying the pattern matrix method to Sherstov's function $F$, one obtains a function $G$ in $\mathsf{AC}^0$ with *discrepancy* $\exp(-\Omega(n^{1/2}))$.[3] This is the strongest known bound on the discrepancy of a function in $\mathsf{AC}^0$, improving over a lengthy line of work (see Table 1). However, because the construction of the dual polynomial $\phi$ in [43] is not entirely explicit, combining Sherstov's result [43] with the pattern matrix method does not give an explicit distribution under which $G$ has low discrepancy. Our construction of a dual polynomial for ELEMENT DISTINCTNESS remedies this situation, yielding the first explicit distribution under which an $\mathsf{AC}^0$ function has discrepancy $\exp(-\Omega(n^{1/2}))$.

   We remark that several other applications of the method of dual polynomials have utilized the existence of a dual polynomial for ELEMENT DISTINCTNESS in a black-box fashion [42, 20]. Our results render explicit these dual polynomial constructions as well.

4. Finally, following the initial dissemination of this work, Bogdanov et al. [16] used our dual polynomial for ELEMENT DISTINCTNESS to design an explicit secret sharing scheme with a reconstruction algorithm computable by an $\mathsf{AC}^0$ circuit. An $(n,d)$-secret sharing scheme with reconstruction advantage $\varepsilon$ is a procedure that enables a dealer to share a secret bit between $n$ parties in such a way that any coalition of at most $d$ parties learns nothing about the secret, but the $n$ parties can combine all of their shares to recover a (randomly chosen) secret with probability at least $1/2+\varepsilon$. Bogdanov et al. [16] showed that if the $\varepsilon$-approximate degree of a function $f : \{-1,1\}^n \to \{-1,1\}$ is at least $d$, then there exists a $(n,d)$-secret sharing scheme with reconstruction advantage at least $\varepsilon/2$, in which the $n$ parties apply $f$ to their shares in order to reconstruct the secret bit. Moreover, the distributions from which the dealer samples shares are exactly determined by a dual polynomial for $f$. Thus, for any constant $\varepsilon < 1$, our dual

[3] Low discrepancy implies high communication cost in nearly every communication model. We refer the reader to [38, Section 10] and [42, Section 8] for the definition of discrepancy and applications of discrepancy upper bounds to communication complexity, computational learning theory, and circuit complexity.

| Discrepancy | Reference | Explicit Distribution Given? |
|---|---|---|
| $\exp(-\Omega(n^{1/3}))$ | [18] | No |
| $\exp(-\Omega(n^{1/3}))$ | [38] | Yes |
| $\exp(-\Omega(n^{2/5}))$ | [20] | No |
| $\exp(-\Omega(n^{1/2-\delta}))$ for any constant $\delta > 0$ | [42] | Yes |
| $\exp(-\Omega(n^{1/2}))$ | [43] | No |
| $\exp(-\Omega(n^{1/2}))$ | This work | Yes |

Table 1: Upper bounds on the discrepancy of circuits of constant depth and polynomial size.

polynomial for ELEMENT DISTINCTNESS yields an explicit $(n, \widetilde{\Omega}(n^{2/3}))$-secret sharing scheme with reconstruction advantage at least $\varepsilon/2$. Bogdanov et al. additionally showed that the structure of our dual polynomial enables these shares to be sampled by $\mathsf{AC}^0$ circuits.

### 1.3 Related work on quantum query complexity

Aaronson and Shi's original motivation for studying the approximate degree of the COLLISION function was to understand its quantum query complexity. (Recall that approximate degree provides a *lower bound* on quantum query complexity [11]. However, it is known that the lower bound is not always tight [8].) Subsequent to Aaronson and Shi's work, a body of lower bound techniques collectively known as *adversary methods* were developed for quantum query complexity [22, 6, 47, 8, 52, 10]. It is now known that one of the most general forms of the adversary method, called the *negative-weights adversary method* [22], always gives a tight characterization of quantum query complexity [33, 34]. (Moreover, the polynomial method can be viewed as a special case of the multiplicative adversary method [29].)

The negative-weights adversary method for lower bounding quantum query complexity is closely analogous to the method of dual polynomials for approximate degree; the former is characterized by a semidefinite program, and a solution to this semidefinite program is known as an *adversary matrix*. A recent line of work, similar in spirit to our own, has proved or reproved optimal quantum query complexity lower bounds for several functions by constructing explicit adversary matrices. In particular, Belovs and Rosmanis [14] constructed an optimal adversary matrix for the COLLISION function in the "large range" case (note that the dual polynomial that we construct applies even in the "small range" case), and Belovs and Špalek constructed an optimal adversary matrix for the ELEMENT DISTINCTNESS function [15].

Very recently, Zhandry [51] (improving on work of Yuen [50]) proved a tight lower bound of $\Omega(N^{1/3})$ on the quantum query complexity of finding a collision in a randomly chosen function.

## 2 Preliminaries

### 2.1 Notation

For any positive integer $n$, we denote the set $\{1,\dots,n\}$ by $[n]$, and the set $\{0,1,\dots,n\}$ by $[[n]]$. For a function $f : D \to \mathbb{R}$, define the $L_1$ norm $\|f\|_1 = \sum_{x \in D} |f(x)|$. For any subset $S \subseteq [n]$, we let $\chi_S \colon \{-1,1\}^n \to \{-1,1\}$ denote the parity function on $S$, i. e., $\chi_S(x) = \prod_{i \in S} x_i$.

## 2.2 Approximate degree and its dual characterization

Let $D \subseteq \{-1,1\}^n$, and let $f\colon D \to \{-1,1\}$ be a partial Boolean function defined on $D$. A real polynomial $p\colon \{-1,1\}^n \to \mathbb{R}$ is said to $\varepsilon$-approximate $f$ if

1. $|p(x) - f(x)| \le \varepsilon$ for all $x \in D$, and
2. $|p(x)| \le 1+\varepsilon$ for all $x \in \{-1,1\}^n$.

The $\varepsilon$-approximate degree of $f$, denoted $\widetilde{\deg}_\varepsilon(f)$, is the minimum degree of an $\varepsilon$-approximation for $f$. We use $\widetilde{\deg}(f)$ to denote $\widetilde{\deg}_{1/3}(f)$, and refer to this quantity without qualification as the *approximate degree* of $f$. The choice of $1/3$ is arbitrary, as $\widetilde{\deg}(f)$ is related to $\widetilde{\deg}_\varepsilon(f)$ by a constant multiplicative factor for any constant $\varepsilon \in (0,1)$.

Given a partial Boolean function $f$, let $p$ be a real polynomial that attains the smallest $\varepsilon$ subject to the constraints above, over all polynomials of degree at most $d$. Since we work over $x \in \{-1,1\}^n$, we may assume without loss of generality that $p$ is multilinear with the representation

$$p(x) = \sum_{|S| \le d} c_S \chi_S(x)\,,$$

where the coefficients $c_S$ are real numbers. Then $p$ is an optimum of the following linear program.

$$\begin{array}{lll}
\text{min} & \varepsilon & \\
\text{such that} & \left| f(x) - \sum_{|S|\le d} c_S \chi_S(x) \right| \le \varepsilon & \text{for each } x \in D \\
 & \left| \sum_{|S|\le d} c_S \chi_S(x) \right| \le 1+\varepsilon & \text{for each } x \in \{-1,1\}^n \setminus D \\
 & c_S \in \mathbb{R} & \text{for each } |S| \le d \\
 & \varepsilon \ge 0 &
\end{array}$$

The dual linear program is as follows.

$$\begin{array}{lll}
\text{max} & \sum_{x\in D} \phi(x) f(x) - \sum_{x \in \{-1,1\}^n \setminus D} |\phi(x)| & \\
\text{such that} & \sum_{x \in \{-1,1\}^n} |\phi(x)| = 1 & \\
 & \sum_{x \in \{-1,1\}^n} \phi(x)\chi_S(x) = 0 & \text{for each } |S| \le d \\
 & \phi(x) \in \mathbb{R} & \text{for each } x \in \{-1,1\}^n
\end{array}$$

Strong LP-duality thus implies the following dual characterization of approximate degree.

**Theorem 2.1.** *Let $f : D \to \{-1,1\}$ be a partial Boolean function. Then $\widetilde{\deg}_\varepsilon(f) > d$ if and only if there is a polynomial $\phi\colon \{-1,1\}^n \to \mathbb{R}$ such that*

$$\sum_{x\in D} f(x)\phi(x) - \sum_{x\in\{-1,1\}^n \setminus D} |\phi(x)| > \varepsilon \cdot \sum_{x \in \{-1,1\}^n} |\phi(x)|\,, \tag{2.1}$$

*and*

$$\sum_{x\in\{-1,1\}^n} \phi(x)\chi_S(x) = 0 \text{ for each } |S| \le d\,. \tag{2.2}$$

If $\phi$ satisfies (2.1), we say that $\phi$ has *correlation greater than* $\varepsilon$ with $f$. If $\phi$ satisfies (2.2), i. e., all Fourier coefficients of $\phi$ below degree $d$ are zero, we say $\phi$ has *pure high degree* $d$. We refer to any feasible solution $\phi$ to the dual linear program as an $(\varepsilon,d)$-*dual polynomial* for $f$.

### 2.3 The COLLISION and ELEMENT DISTINCTNESS functions

Let $[N]=\{1,\dots,N\}$, and fix a triple of positive integers $n,N,R$ such that $R \geq N$, and $n = N\cdot\log_2 R$. For simplicity throughout, we assume that $R$ is a power of 2. The COLLISION and ELEMENT DISTINCTNESS functions are typically thought of as *properties* of functions mapping $[N]$ to $[R]$. However, it will be convenient for us to think of them instead as functions on the Boolean hypercube $\{-1,1\}^n$. To this end, given an input $x \in \{-1,1\}^n$, we interpret $x$ as the evaluations of a function $g_x$ mapping $[N]\to[R]$. That is, we break $x$ up into $N$ blocks, each of length $\log_2 R$, and regard each block $x_i$ as the binary representation of $g_x(i)$.

**Definition 2.2** (COLLISION Function)**.** A function $g_x\colon [N]\to[R]$ is said to be $k$-to-1 if for every $i\in[N]$, there exist exactly $k-1$ values $j\neq i$ such that $g_x(i)=g_x(j)$. Let $T_k := \{x\in\{-1,1\}^n : g_x \text{ is } k\text{-to-1}\}$ (clearly, $T_k$ is non-empty only if $k \mid N$). The COLLISION function, which we denote by $\mathrm{Col}_{N,R}$, is the partial Boolean function defined on $T_1\cup T_2 \subseteq \{-1,1\}^n$ such that $\mathrm{Col}_{N,R}(x)=1$ if and only if $x\in T_1$. That is, $\mathrm{Col}_{N,R}$ is the partial Boolean function corresponding to the property that $g_x$ is a 1-to-1 function, with the promise that $g_x$ is either 1-to-1 or 2-to-1.

**Definition 2.3** (ELEMENT DISTINCTNESS Function)**.** The ELEMENT DISTINCTNESS function, denoted $\mathrm{ED}_{N,R}$, is the total Boolean function defined such that $\mathrm{ED}_{N,R}(x)=1$ if and only if $g_x$ is 1-to-1. That is, $\mathrm{ED}_{N,R}$ is the total Boolean function corresponding to the property that $g_x$ is 1-to-1.

Let $B\subset\{-1,1\}^n$ denote the set of inputs $x$ such that $g_x$ is neither 1-to-1 nor 2-to-1. Then an $(\varepsilon,d)$-dual polynomial $\phi$ for $\mathrm{Col}_{N,R}$ has the following properties (cf. Section 2.2):

1. $\sum_{x\in T_1}\phi(x)-\sum_{x\in T_2}\phi(x)-\sum_{x\in B}|\phi(x)| > \varepsilon\cdot\sum_{x\in\{-1,1\}^n}|\phi(x)|$.
2. $\sum_{x\in\{-1,1\}^n}\phi(x)\chi_S(x)=0$ for all $|S|\leq d$.

Similarly, an $(\varepsilon,d)$-dual polynomial for $\mathrm{ED}_{N,R}$ satisfies the following conditions.

1. $\sum_{x\in T_1}\phi(x)-\sum_{x\notin T_1}\phi(x) > \varepsilon\cdot\sum_{x\in\{-1,1\}^n}|\phi(x)|$.
2. $\sum_{x\in\{-1,1\}^n}\phi(x)\chi_S(x)=0$ for all $|S|\leq d$.

### 2.4 Overview of the symmetrization-based proof of the COLLISION lower bound

Kutin's simplified proof of the COLLISION lower bound [28] proceeds in two steps. The first step is a symmetrization step, which establishes the following remarkable result. (We state this result slightly informally in this overview.)

**Lemma 2.4** (Informal version of Lemma 2.5)**.** *Call a triple $(m,a,b)$* valid *if $a\mid m$ and $b\mid(N-m)$. For any triple $(m,a,b)$, let $R_{m,a,b}$ denote the set of inputs $x\in\{-1,1\}^n$ such that $g_x\colon [N]\to[R]$ maps $m$ of its inputs to $[R]$ in an $a$-to-1 manner, and maps the remaining $N-m$ of its inputs to $[R]$ in a $b$-to-1 manner.*

*Let $p(x)$ be a real polynomial over $\{-1,1\}^n$ of degree $d$. Then there is a trivariate polynomial $P$ of total degree at most $d$ such that for every valid triple $(m,a,b)$, it holds that*

$$P(m,a,b) = \mathbb{E}_{x \in R_{m,a,b}}[p(x)] .$$

Note in the above lemma that the sets $R_{m,a,b}$ are not uniquely determined; for instance $R_{m,1,1} = R_{0,a,1} = R_{N,1,b} = T_1$ for every triple $(m,a,b)$.

The second step of Kutin's proof argues that if $p$ is a $(1/3)$-approximating polynomial for the COLLISION function, then $P$ must have degree $\Omega(N^{1/3})$. Hence by Lemma 2.4, $p$ must have degree $\Omega(N^{1/3})$ as well.

In more detail, the second step of Kutin's proof proceeds via a case analysis. Four cases are considered.

**Case 1:** $P(N/2,1,2) \geq 1/2$, and $|P(N/2,1,b)| \leq 2$ for all $b \in [N^{2/3}]$. In this case, Kutin is able to apply Markov's inequality from approximation theory to conclude that the degree of $P$ in its third variable is $\Omega(N^{1/3})$.

**Case 2:** $P(N/2,1,2) \geq 1/2$, and $|P(N/2,1,b)| > 2$ for some $b \in [N^{2/3}]$. In this case, Kutin is able to apply Bernstein's inequality from approximation theory to conclude that the degree of $P$ in its first variable is $\Omega(N^{1/3})$.

**Case 3:** $P(N/2,1,2) < 1/2$, and $|P(N/2,a,2)| \leq 2$ for all $a \in [N^{2/3}]$. In this case, Kutin is able to apply Markov's inequality to conclude that the degree of $P$ in its second variable is $\Omega(N^{1/3})$.

**Case 4:** $P(N/2,1,2) < 1/2$, and $|P(N/2,a,2)| > 2$ for some $a \in [N^{2/3}]$. In this case, Kutin is able to apply Bernstein's inequality to conclude that the degree of $P$ in its first variable is $\Omega(N^{1/3})$.

A key technical complication that must be dealt with in the argument above is that $|P(m,a,b)|$ may be much larger than 1 for *invalid* triples $(m,a,b)$. This may seem like a minor technicality, but in fact it is a central issue: if $P(m,a,b)$ were bounded for all invalid triples, then it would be possible to argue that the total degree of $P$ is $\Omega(N^{1/2})$, which would imply a (false) lower bound of $\Omega(N^{1/2})$ on the approximate degree of $\mathrm{Col}_{N,R}$.

## 2.5 Overview of our construction for the COLLISION function

As with Kutin's proof, our construction also makes essential use of Lemma 2.4. Whereas Kutin used Lemma 2.4 to reduce to a setting where Markov-Bernstein inequalities could be applied in a non-constructive manner, we instead use Lemma 2.4 to argue that the dual polynomial $\phi$ that we construct has pure high degree $\Omega(N^{1/3})$.

In more detail, we present our construction in two stages, in order to highlight distinct ideas that go into the proof. In the first stage, we construct a simpler dual polynomial $\phi\colon \{-1,1\}^n \to \{-1,1\}$ that exhibits an $\Omega(\sqrt{\log N/\log\log N})$ lower bound on the approximate degree of $\mathrm{Col}_{N,R}$. The second stage constructs a dual polynomial $\psi$ exhibiting the optimal $\Omega(N^{1/3})$ lower bound.

**Overview of the first stage.** Let $H_k \subseteq \{-1,1\}^n$ denote the set of inputs of Hamming weight $k$. The symmetrization-based proof of the COLLISION lower bound from [5, 28] carries the strong intuition that the sets $T_k$ should play the same role that $H_k$ plays in Nisan and Szegedy's seminal symmetrization-based lower bound for the OR function [30]. We direct the interested reader to Aaronson's lecture notes [3] for a detailed explanation of this intuition. The construction of our simpler dual witness $\phi$ instantiates this intuition in the dual setting.

Recall that a dual polynomial $\phi$ witnessing the fact that $\widetilde{\deg}(\mathrm{Col}_{N,R}) \geq d$ must satisfy two properties: (1) it must have correlation greater than $\varepsilon$ with $\mathrm{Col}_{N,R}$, and (2) it must have pure high degree at least $d$. We define $\phi$ in a way that mimics the structure of known dual witnesses for symmetric functions, even though $\phi$ is not itself symmetric. Specifically, our construction ensures that the analysis establishing Properties (1) and (2) becomes similar to the analyses of known dual polynomials for the OR function [46, 19].

In more detail, our prior work [19] built on work of Špalek [46] to give a dual witness $\gamma$ for the fact that $\widetilde{\deg}_\varepsilon(\mathrm{OR}_n) = \Omega(\sqrt{n})$ for any constant $\varepsilon < 1$; moreover, $\gamma$ places non-zero weight only on sets $H_k$, for values of $k$ equal (up to scaling factors) to perfect squares. The pure high degree of $\gamma$ is shown to be equal to (at least) the number of sets $H_k$ upon which $\gamma$ places non-zero weight.

Call an input $x \in \{-1,1\}^n$ *valid* if it is in $R_{m,a,b}$ for some valid triple $(m,a,b)$. By analogy with $\gamma$, the dual witness $\phi$ that we construct in Stage 1 places weight only on inputs $x \in T_k$ for divisors $k$ of $N$ that are also (up to scaling factors) perfect squares. In particular, our definition of $\phi$ ensures that

$$\phi(x) = 0 \text{ for all invalid inputs } x. \tag{2.3}$$

We are able to combine (2.3) with Lemma 2.4 and a basic combinatorial identity (cf. Lemma 3.2) to show that the pure high degree of $\phi$ is at least $|S|$, where $S$ denotes the set of $T_k$'s upon which $\phi$ places non-zero weight. In more detail, given any polynomial $p$ of degree at most $|S|$, Lemma 2.4 gives us a measure of control over $p$'s value on valid inputs. Lemma 2.4 provides *no* control over $p$'s value on invalid inputs, but this is rendered irrelevant by (2.3), which guarantees that $\phi$ essentially ignores such inputs.

Moreover, our definition of $\phi$ is carefully chosen to ensure that its correlation with $\mathrm{Col}_{N,R}$ is large; the precise calculation is closely analogous to the analysis from [46, 19] showing that $\gamma$ is well-correlated with the OR function [46, 19].

**Overview of the second stage.** In the second stage, we construct a dual polynomial $\psi$ that exhibits the optimal $\Omega(N^{1/3})$ lower bound. Rather than only weighting inputs in $T_k$ for some some divisors $k$ of $N$, $\psi$ weights inputs in $R_{m,a,b}$ for many valid triples $(m,a,b)$. There are two key ideas that go into the construction of $\psi$.

The first idea is to define $\psi$ as the sum of two simpler dual polynomials $\psi_1$ and $\psi_2$, each with pure high degree $\Omega(N^{1/3})$—then the sum $\psi$ also has pure high degree $\Omega(N^{1/3})$ (see Lemma 4.5). The first polynomial $\psi_1$ places a large constant fraction (close to $1/2$) of its $L_1$ mass on $T_1$, whereas $\psi_2$ places a large constant fraction of its $L_1$ mass on $T_2$. Neither $\psi_1$ nor $\psi_2$ is well-correlated with $\mathrm{Col}_{N,R}$ in the sense of (2.3). However, they each place a constant fraction of their $L_1$ mass on $R_{N/2,2,1}$, and they are designed so that their values exactly cancel out on inputs in $R_{N/2,2,1}$. This allows us to show that $\psi = \psi_1 + \psi_2$ satisfies (2.3), even though $\psi_1$ and $\psi_2$ individually do not.

The second idea goes into the construction of $\psi_1$ and $\psi_2$ themselves. Specifically, we think of $\psi_1$ and $\psi_2$ as each being constructed in a two-step process. We focus on $\psi_1$ in this discussion, since the construction of $\psi_2$ is similar. Very roughly speaking, in the first step, we consider a "polynomial" $\psi'$ of pure high degree $\Omega(N^{1/3})$ that places a large constant fraction of its $L_1$ mass on $T_1$; the construction of $\psi'$ is closely related to our construction of the simpler dual polynomial $\phi$ from Stage 1.

The reason we place the term "polynomial" in quotes above is that there is an important technical caveat to our construction of $\psi'$: we think of $\psi'$ as placing weight on sets $R_{N/2,a,1}$ for many *invalid* triples $(N/2,a,1)$, in addition to some valid ones. Of course, if $(N/2,a,1)$ is invalid, then $R_{N/2,a,1} = \emptyset$, so $\psi'$ cannot place non-zero weight on the set. To address this issue, in Step 2, we add to $\psi'$ a sequence of polynomials $\psi''_{N/2,a,1}$, each of pure high degree $\Omega(N^{1/3})$. For each invalid triple $(N/2,a,1)$, $\psi''_{N/2,a,1}$ is specifically constructed to cancel out the weight that $\psi'$ "places" on $R_{N/2,a,1}$.

Analogously to how our constructions of $\phi$ and $\psi'$ were closely related to the dual witness for OR constructed in our earlier work [19], our construction of $\psi''_{N/2,a,1}$ is closely related to a dual witness $\eta$ for the Majority function, MAJ, that we constructed in the same work. Each $\psi''_{N/2,a,1}$ places additional non-zero mass on (non-empty) sets of the form $R_{m,a,1}$ for some $a \neq 1$ and $m \in [N]$, but we are able to show that the total mass placed on such sets is small, using an analysis closely related to the analysis of $\eta$ from [19]. Hence we are able to show that

$$\psi_1 = \psi' + \sum_{\text{invalid triples } (N/2,a,1)} \psi''_{N/2,a,1}$$

still places a large constant fraction of its $L_1$ mass on $T_1$.

## 2.6 Discussion

**On Kutin's second step.** Our construction of the optimal dual witness $\psi$ for the COLLISION function mimics the second step of Kutin's symmetrization argument in three important ways described below. We find this mimicry to be somewhat surprising—in our earlier work [19], we constructed an optimal dual polynomial for symmetric Boolean functions that bore little relation to Paturi's well-known symmetrization-based proof of the same result [32]. We believe that this mimicry sheds new light, or at least gives a new perspective, on why Kutin's proof takes the structure that it does.

Recall that the second step of Kutin's proof (cf. Section 2.4) proceeds via a case analysis. The first "branch" in the case analysis depends on whether the expected value of the assumed $n$-variate approximation $p$ to $\mathrm{Col}_{N,R}$ on the set $R_{N/2,2,1}$ is large or small. This is mimicked in our construction of $\psi$ as a sum of two dual polynomials $\psi_1$ and $\psi_2$, both of which individually place a lot of weight on $R_{N/2,2,1}$, but whose sum places *zero* weight on $R_{N/2,2,1}$.

The second "branch" in Kutin's case analysis depends on whether $|P(N/2,a,1)|$ or $|P(N/2,2,b)|$ is small for all $a,b \leq N^{2/3}$. He needs to consider this second branch because $P(m,a,b)$ is not guaranteed to be bounded for invalid triples $(m,a,b)$.

This branch is mimicked in our construction of $\psi_1$ (respectively, $\psi_2$) as the sum of a single "polynomial" $\psi'$ that tries to place weight on sets $R_{N/2,a,1}$ for invalid triples $(N/2,a,1)$ (respectively, $(N/2,2,b)$), and many other polynomials $\psi''_{N/2,a,1}$ (respectively, $\psi''_{N/2,2,b}$), one for each invalid triple $(N/2,a,1)$ (respectively, $(N/2,2,b)$). In our dual setting, the reason we need to incorporate the polynomials $\psi''_{N/2,a,1}$ is to cancel out the weight that $\psi'$ tries to place on invalid sets $R_{N/2,a,1}$.

Finally, recall that Kutin applied Markov's inequality from approximation theory in two of the four cases considered in his analysis, and Bernstein's inequality in the other two cases. Markov's inequality underlies Nisan and Szegedy's standard symmetrization-based proof that the approximate degree of OR is $\Omega(\sqrt{n})$ [30], while Berstein's inequality underlies Paturi's proof that the approximate degree of MAJ is $\Omega(n)$ [32]. This is mimicked in our construction of $\psi_1$ and $\psi_2$ as the sum of $\psi'$ and the $\psi''_{N/2,a,1}$ and $\psi''_{N/2,2,b}$ polynomials: the construction of $\psi'$ is closely analogous to the dual witness for OR from [19], while the construction of the $\psi''_{N/2,a,1}$ and $\psi''_{N/2,2,b}$ polynomials is based on the dual witness for MAJ from [19].

**On the first step, or why $k$-to-1 inputs matter.** As noted by several authors (e. g., [2, Slide 36]), the most miraculous element of the symmetrization-based proof of the COLLISION lower bound is the first step (cf. Lemma 2.4). The crux of this step is to establish, roughly speaking, that for any $n$-variate polynomial $p$ of total degree $d$, the function

$$P(k) := \mathbb{E}_{x \in T_k}[p(x)]$$

is a polynomial in $k$ of degree at most $d$. Why should this hold? More basically, why should inputs that are $k$-to-1 even play a prominent role in the proof?

We provide some partial intuition for this in Section 6. Specifically, we explain that there is an (asymptotically) optimal approximation $p$ for $\mathrm{Col}_{N,R}$ such that $k$-to-1 inputs correspond to constraints that are made tight by the solution corresponding to $p$ in the primal linear program of Section 2.2. Hence, complementary slackness suggests that there should be a corresponding dual witness $\psi$ that places weight only on inputs that are $k$-to-1, or nearly so, justifying the prominent role that $k$-to-1 inputs play in both the symmetrization-based proof and our new dual proof.

### 2.7 Formal statement of Lemma 2.4

Following Kutin [28], we define a special collection of functions which are $a$-to-1 on one part of the domain and $b$-to-1 on the other part. For $N > 0$, recall that a triple of numbers $(m,a,b)$ is *valid* if $a \mid m$ and $b \mid (N-m)$. For each valid triple $(m,a,b)$, we define

$$g_{m,a,b}(i) = \begin{cases} \lceil i/a \rceil & \text{if } 1 \le i \le m, \\ R - \lfloor (N-i)/b \rfloor & \text{if } m < i \le n. \end{cases}$$

Moreover, for each valid triple $(m,a,b)$, we define a set $R_{m,a,b}$ that is the orbit of $g_{m,a,b}$ under the automorphism group $S_N \times S_R$. Namely,

$$x \in R_{m,a,b} \iff \exists \sigma \in S_N, \tau \in S_R : \quad \tau \circ g_x \circ \sigma = g_{m,a,b}\,.$$

Note that the sets $R_{m,a,b}$ are not uniquely determined; for instance $R_{m,1,1} = R_{0,a,1} = R_{N,1,b} = T_1$ for every $m,a,b$.

**Lemma 2.5.** *Let $p(x)$ be a real polynomial over $\{-1,1\}^n$ of degree $d$. There is a trivariate polynomial $P$ of degree at most $d$ with the property that for all valid triples $(m,a,b)$,*

$$P(m,a,b) = \mathbb{E}_{x \in R_{m,a,b}}[p(x)]\,.$$

The statement of Lemma 2.5 differs slightly from the corresponding lemma in Kutin's work [28] (Lemma 2.7 below). Lemma 2.5 follows by combining Kutin's formulation with the following simple lemma from [20].

**Lemma 2.6** ([20])**.** *Let $p$ be a polynomial over $\{-1,1\}^n$. Consider the map $T\colon \{-1,1\}^n \to \{0,1\}^{N\cdot R}$ defined by $T_{ij}(x) = 1$ if $g_x(i) = j$, and $T_{ij}(x) = 0$ otherwise. Then there is a polynomial $q\colon \{0,1\}^{N\cdot R} \to \mathbb{R}$ with $\deg q \le \deg p$, such that $q(T(x)) = p(x)$ for all $x \in \{-1,1\}^n$.*

**Lemma 2.7** ([28])**.** *Let $q(t)$ be any degree $d$ polynomial in the variables $t_{ij}$. For a valid triple $(m,a,b)$, define $Q(m,a,b)$ by*

$$Q(m,a,b) = \mathbb{E}_{x \in R_{m,a,b}}[q(T(x))]\,.$$

*Then $Q$ is a degree $d$ polynomial in $m,a,b$.*

## 3 An $\Omega(\sqrt{\log N/\log\log N})$ lower bound for the COLLISION function

The following lemma is a refinement of [19, Proposition 14], which was used there to construct a dual polynomial for OR. The function $\omega$ we construct here forms the core of the "first stage" of our construction of a dual polynomial for the COLLISION function.

**Lemma 3.1.** *There exists a constant $\zeta > 0$ such that for all $\delta \in (0,1)$ and $L \ge 1$, there is an explicit $\omega : \{1,\ldots,L\} \to \mathbb{R}$ with the following properties.*

1. $\omega(1) \ge \dfrac{1-\delta}{2}$.
2. $-\omega(2) \ge \dfrac{1-\delta}{2}$.
3. $\displaystyle\sum_{k=1}^{L} |\omega(k)| = 1$.
4. *For every polynomial $p : \{1,\ldots,L\} \to \mathbb{R}$ of degree $d \le \zeta\sqrt{\delta L}$, we have $\sum_{k=1}^{L} p(k)\omega(k) = 0$.*

The proof will make use of the following simple combinatorial identity, a simple proof of which can be found in [31, Appendix A].

**Lemma 3.2.** *For any $L > 0$, let $q\colon \mathbb{R} \to \mathbb{R}$ be a univariate polynomial of degree strictly less than $L$. Then*

$$\sum_{k=0}^{L} (-1)^k \binom{L}{k} q(k) = 0\,.$$

*Proof of Lemma 3.1.* Let $c = \lceil 16/\delta \rceil$. Let $m = \lfloor \sqrt{(L-1)/c} \rfloor$ and define the set

$$T = \{1\} \cup \{ci^2 : 0 \le i \le m\}\,.$$

Note that $|T| = \Omega(\sqrt{L/c})$. Define the function $\hat{\omega} : \{0,1,\dots,L\} \to \mathbb{R}$ by

$$\hat{\omega}(k) = \binom{L}{k} \frac{c^m (m!)^2}{L!} \prod_{j \in [[L]] \setminus T} (j-k) = \begin{cases} \dfrac{(-1)^k \cdot c^m (m!)^2}{\prod_{j \in T \setminus \{k\}} (j-k)} & \text{if } k \in T, \\ \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that $\hat{\omega}(0) = 1$.

For $k = 1$, we have

$$|\hat{\omega}(1)| = \frac{c^m (m!)^2}{\prod_{i=1}^m (ci^2 - 1)}\,.$$

Notice that the magnitude $|\hat{\omega}(1)|$ is tightly controlled:

$$\begin{aligned} 1 \le \frac{c^m (m!)^2}{\prod_{i=1}^m (ci^2 - 1)} &= \prod_{i=1}^m \frac{i^2}{i^2 - 1/c} \\ &= \prod_{i=1}^m \left(1 + \frac{1}{ci^2 - 1}\right) \\ &\le \exp\left(\sum_{i=1}^m \frac{2}{ci^2}\right) \\ &\le \mathrm{e}^{8/c} \le 1 + \frac{3}{2}\delta\,. \end{aligned}$$

Here, we have used the fact that $\prod_{i=1}^m (1 + a_i) \le \exp(\sum_{i=1}^m a_i)$ for nonnegative $a_i$. On the other hand, for $k = c\ell^2$ with $\ell > 0$, we may write $|\hat{\omega}(k)|$ as

$$\begin{aligned} \frac{c^m (m!)^2}{(c\ell^2 - 1) \prod_{i \in [[m]] \setminus \{\ell\}} |ci^2 - c\ell^2|} &= \frac{(m!)^2}{(c\ell^2 - 1) \prod_{i \in [[m]] \setminus \{\ell\}} (i+\ell)|i-\ell|} \\ &= \frac{2(m!)^2}{(c\ell^2 - 1)(m+\ell)!(m-\ell)!} \\ &\le \frac{2}{c\ell^2 - 1}\,, \end{aligned}$$

where the last inequality follows because

$$\frac{(m!)^2}{(m+\ell)!(m-\ell)!} = \frac{m}{m+\ell} \cdot \frac{m-1}{m+\ell-1} \cdot \dots \cdot \frac{m-\ell+1}{m+1}$$

is a product of factors that are each smaller than 1. Thus, the total contribution of terms excluding 0 and 1 to the $L_1$ norm of $\hat{\omega}$ is at most

$$\sum_{i=1}^m \frac{2}{ci^2 - 1} < \sum_{i=1}^\infty \frac{4}{ci^2} < \frac{8}{c} \le \frac{\delta}{2}.$$

Now define $\omega \colon \{1,\dots,L\} \to \mathbb{R}$ via

$$\omega(k) = (-1)^{k-1} \hat{\omega}(k-1) / \|\hat{\omega}\|_1\,.$$

Then

$$-\omega(2) \geq \omega(1) \geq \frac{1}{1+|\hat{\omega}(1)|+\delta/2} \geq \frac{1}{2+2\delta} \geq \frac{1-\delta}{2}\,.$$

This yields the first two claims about $\omega$. The third claim follows immediately from the definition. Finally, let $p$ be a polynomial of degree strictly less than $|T|-1$. Then

$$\sum_{k=1}^{L} p(k)\omega(k) = \sum_{k=0}^{L-1}(-1)^k \cdot \binom{L}{k} \cdot \frac{c^m (m!)^2}{L!\|\hat{\omega}\|_1} \cdot p(k+1) \cdot \prod_{j\in[[L]]\setminus T}(j-k) = \sum_{k=0}^{L-1}(-1)^k \binom{L}{k} q(k)\,, \tag{3.1}$$

where

$$q(k) = \frac{c^m (m!)^2}{L!\|\hat{\omega}\|_1} \cdot p(k+1) \cdot \prod_{j\in[[L]]\setminus T}(j-k)$$

is a polynomial of degree less than $L$. Since $q(L)=0$, the right hand side of (3.1) is zero by Lemma 3.2. This gives the last claim. □

Our prior work [19], building on work of Špalek [46], obtained a dual polynomial $\gamma$ for $\mathrm{OR}_L$ by setting the total weight of $\gamma$ on inputs in $H_k$ (the set of inputs of Hamming weight $k$) to be $\omega(k+1)$. In that work, the first three properties of $\omega$ ensured that $\gamma$ had high correlation with OR, while the fourth ensured that it had pure high degree $\Omega(\sqrt{L})$.

Analogously, our dual polynomial $\phi$ for $\mathrm{Col}_{N,R}$ below sets the total weight of $\phi$ on $T_k$ to be $\omega(k)$. Then again, the first three properties of $\omega$ ensure that $\phi$ is well-correlated with $\mathrm{Col}_{N,R}$, and the fourth ensures that it has pure high degree $\Omega(\sqrt{L})$. However, we face the complication that $T_k$ must be non-empty, i. e., $k$ must divide $N$, for every $k$ in the support of $\omega$. To handle this complication, we take $N$ large enough so that all $k=1,2,\ldots,L$ divide $N$, yielding an $\Omega(\sqrt{\log N/\log\log N})$ lower bound.

**Theorem 3.3.** *Let $N=L!$ for some $L$. For $\delta>0$, there exists an explicit $(1-2\delta,d)$ dual polynomial $\phi$ for $\mathrm{Col}_{N,R}$ with $d=\Omega(\sqrt{\delta L})=\Omega(\sqrt{\delta\log N/\log\log N})$.*

*Proof.* First, notice that $k \mid N$ for all $k\in[L]$, so $T_k\neq\emptyset$ for every such $k$. Define $\phi(x)=\omega(k)/|T_k|$ if $x$ is in $T_k$ for some $k\in[L]$, and $\phi(x)=0$ otherwise, where $\omega$ is obtained by applying Lemma 3.1. Note that $\phi(x)$ is well-defined since $|T_k|\neq 0$ for all $k\in[L]$, and each $x\in\{-1,1\}^n$ is in $T_k$ for at most one value of $k$.

We check

$$\sum_{x\in T_1}\phi(x) - \sum_{x\in T_2}\phi(x) = \omega(1)-\omega(2) \geq 1-\delta\,,$$

where the inequality holds by Parts 1 and 2 of Lemma 3.1. Moreover,

$$\sum_{x\in B}|\phi(x)| = \sum_{k=3}^{L}|\omega(k)| \leq \delta,$$

where the inequality holds by combining Parts 1-3 of Lemma 3.1. Thus,

$$\sum_{x\in T_1}\phi(x) - \sum_{x\in T_2}\phi(x) - \sum_{x\in B}|\phi(x)| \geq 1-2\delta\,.$$

Second,

$$\sum_{x\in T_1\cup T_2\cup B}|\phi(x)|=\sum_{k=1}^{L}|\omega(k)|=1\,,$$

where the final equality holds by Part 3 of Lemma 3.1.

Finally, let $d=\zeta\sqrt{\delta L}$ where $\zeta$ is as in the statement of Lemma 3.1, and let $S\subseteq[n]$ with $|S|\le d$. We must show that $\sum_{x\in T_1\cup T_2\cup B}\phi(x)\chi_S(x)=0$. Note that

$$\sum_{x\in T_1\cup T_2\cup B}\phi(x)\chi_S(x)=\sum_{k=1}^{L}\sum_{x\in T_k}\phi(x)\cdot\chi_S(x)=\sum_{k=1}^{L}\sum_{x\in T_k}(\omega(k)/|T_k|)\cdot\chi_S(x)=\sum_{k=1}^{L}\omega(k)\cdot\mathbb{E}_{x\in T_k}[\chi_S(x)]\,,$$

where the first equality holds because $\phi(x)=0$ if $x$ is not in $T_k$ for some $k\in[L]$.

By Lemma 2.5, there is a trivariate polynomial $P$ of total degree at most $d$ such that

$$P(m,a,b)=\mathbb{E}_{x\in R_{m,a,b}}[\chi_S(x)]$$

for all valid triples $(m,a,b)$. In particular, since $k\mid N$ for all $k\in[L]$, $q(k):=P(N,k,1)$ is a *univariate* polynomial in $k$ such that $q(k)=E_{x\in T_k}[\chi_S(x)]$ for all $k\in[L]$. Hence, Part 4 of Lemma 3.1 implies that

$$\sum_{k=1}^{L}\omega(k)\cdot\mathbb{E}_{x\in T_k}[\chi_S(x)]=0\,.$$

□

## 4 An $\Omega(N^{1/3})$ lower bound for the COLLISION function

We now complete the "second stage" of our construction. The following lemma is a refinement of [19, Proposition 10], which constructed an explicit dual polynomial for MAJ.

**Lemma 4.1.** *There exists a constant $\rho>0$ for which the following holds. Let $\delta\in(0,1)$, $N>0$ an even integer, and $k\in[N]$. Then there is an explicit $\eta_k:[[N]]\to\mathbb{R}$ such that*

1. *$\eta_k$ is supported on $\{2k,4k,\dots,2\lfloor N/2k\rfloor k\}\cup\{N/2\}$,*
2. *$\eta_k(N/2)>(1-\delta)/2$,*
3. *$\sum_{r=0}^{N}|\eta_k(r)|=1$,*
4. *for every polynomial $p:\{0,\dots,N\}\to\mathbb{R}$ of degree $d\le\rho\sqrt{\delta}N/k$, we have $\sum_{r=0}^{N}p(r)\eta_k(r)=0$.*

*Proof.* Throughout the proof, we assume for simplicity that $N/2$ is not a multiple of $2k$. The analysis when $N/2$ is a multiple of $2k$ is similar.

Let $c=\lceil 10/\sqrt{\delta}\rceil$ and $t=2\lfloor N/(4k)\rfloor k$ and define the set

$$S=\{t\pm 2c\ell k\colon 0\le\ell\le\lfloor t/(2ck)\rfloor\}\,.$$

Note that $|S| = \Omega(N/ck)$. We claim that

$$\pi_S(i) := \prod_{j \in S, j \neq i} |j - i|$$

is minimized at $i = t$. Notice that translating all points in $S$ by a constant $t$ does not affect $\pi_{S-t}(i-t)$, and scaling all points in $S$ by a constant $a$ does not affect $\operatorname{argmin}_i \pi_{aS}(i)$. Thus, it is enough to show that $\pi_{S^*}(i)$ is minimized at $i = 0$ for the set $S^* = \{\pm\ell \colon \ell \leq q\}$. In this case, $\pi_{S^*}(i)$ takes the simple form $(q-i)!(q+i)!$, and we see that for all $i \in S^*$,

$$\frac{\pi_{S^*}(0)}{\pi_{S^*}(i)} = \frac{(q!)^2}{(q-i)!(q+i)!} = \frac{q}{q+|i|} \cdot \frac{q-1}{q+|i|-1} \cdot \dots \cdot \frac{q-|i|+1}{q+1}$$

is a product of terms smaller than 1, so $\pi_{S^*}(i)$ is indeed minimized at $i = 0$.

Now let $T = S \cup \{t - 2k, N/2\}$ and define the function

$$\hat{\eta}(r) = \binom{N}{r} \frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{N!} \prod_{j \in [[N]] \setminus T} (j - r) = \begin{cases} \frac{(-1)^r (2ck)^{2h}(h!)^2(2k)(N/2-t)}{\prod_{j \in T \setminus \{r\}}(j-r)} & \text{if } r \in T, \\ 0 & \text{otherwise,} \end{cases}$$

where $h = \lfloor t/2ck \rfloor$. The normalization is chosen so that $|\hat{\eta}(t)| = 1$.

The reason that we include *both* $(r - (t-2k))$ and $(r - (N/2))$ in the denominator of $\hat{\eta}$ is to ensure that the rate of decay of $\hat{\eta}(r)$ is at least quadratic as $r$ moves away from $t$. This will ultimately allow us to show that a large fraction of the $\ell_1$ mass of $\hat{\eta}$ comes from the point $r = N/2$.

For $r = t - 2k$, the mass $|\hat{\eta}(r)|$ is

$$\begin{aligned} \frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{2k(N/2-t+2k)\prod_{\ell=1}^{h}(2ck\ell-2k)(2ck\ell+2k)} &= \frac{(N/2-t)}{N/2-t+2k} \prod_{\ell=1}^{h} \left(1 + \frac{1}{(c\ell)^2 - 1}\right) \\ &\leq \frac{1}{2} \exp\left(\sum_{\ell=1}^{h} \frac{2}{c^2 \ell^2}\right) \\ &\leq \frac{1}{2} \exp\left(\frac{\pi^2}{3c^2}\right) \\ &< \frac{1+\delta}{2}, \end{aligned}$$

where the first inequality holds because $N/2 - t \leq 2k$, combined with the fact that $\prod_{\ell=1}^{h}(1 + a_\ell) \leq \exp(\sum_{\ell=1}^{h} a_\ell)$ for nonnegative $a_\ell$.

For $r = N/2$, we may calculate

$$\begin{aligned} |\hat{\eta}(r)| &= \frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{(N/2-t)(N/2-t+2k)\prod_{\ell=1}^{h}(2ck\ell + (N/2-t))(2ck\ell - (N/2-t))} \\ &= \frac{2k}{N/2-t+2k} \prod_{\ell=1}^{h} \left(\frac{(2ck\ell)^2}{(2ck\ell)^2 - (N/2-t)^2}\right) \geq \frac{1}{2}. \end{aligned}$$

Now we analyze the remaining summands, and show that their total contribution is much smaller than 1. Recall that the choice $i = t$ minimizes $\pi_S(i)$, and that $\pi_S(t) = (2ck)^{2h}(h!)^2$. Therefore,

$$|\hat{\eta}(t+2ck\ell)| = \frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{\prod_{j\in T\setminus\{t+2ck\ell\}}|j-(t+2ck\ell)|} \le \frac{2k(N/2-t)}{|2ck\ell+2k|\cdot|2ck\ell-(N/2-t)|} \le \frac{1}{c^2\ell^2-1},$$

where the final inequality exploits the fact that $N/2-t < 2k$. Similarly,

$$|\hat{\eta}(t-2ck\ell)| = \frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{\prod_{j\in T\setminus\{t-2ck\ell\}}|j-(t-2ck\ell)|} \le \frac{2k(N/2-t)}{|2ck\ell-2k|\cdot|2ck\ell+(N/2-t)|} \le \frac{1}{c^2\ell^2-1}.$$

We can use this quadratic decay to bound the total $L_1$ mass of the points outside of $\{t-k, t, N/2\}$.

$$\sum_{j\in S\setminus\{t\}} |\hat{\eta}(j)| \le \sum_{\ell\neq 0} \frac{1}{(c^2\ell^2-1)} \le \frac{2}{c^2-1}\cdot\frac{\pi^2}{6} < \frac{\delta}{2}.$$

Now let $\eta_k(r) = (-1)^{r+h+N/2}\hat{\eta}(r)/\|\hat{\eta}\|_1$. Since $\hat{\eta}$ is supported on $T \subseteq \{2k, 4k, \dots, 2\lfloor N/2k\rfloor k\} \cup \{N/2\}$, the function $\eta_k$ is as well, giving the first claim. Moreover,

$$\eta_k(N/2) \ge \frac{1/2}{(1/2+\delta/2)+1/2+\delta/2} \ge \frac{1-\delta}{2}.$$

This yields the second claim about $\eta_k$. The third claim follows immediately from the definition. Finally, let $p$ be a polynomial of degree strictly less than $|T|$, where $|T| \ge \rho N/k$ for a constant $\rho$. Then

$$\begin{aligned}\sum_{r=0}^{N} p(r)\eta_k(r) &= \sum_{r=0}^{N} p(r)\frac{(-1)^{r+h+N/2}}{\|\hat{\eta}\|_1}\binom{N}{r}\frac{(2ck)^{2h}(h!)^2(2k)(N/2-t)}{N!}\prod_{j\in[[N]]\setminus T}(j-r)\\ &= \sum_{r=0}^{N}(-1)^r\binom{N}{r}q(r)\end{aligned}$$

for a polynomial $q$ of degree strictly less than $N$. This is equal to zero by Lemma 3.2, giving the final claim. □

We obtain our dual polynomial $\psi$ for the $\mathrm{Col}_{N,R}$ as a linear combination of two simpler functions $\psi_1$ and $\psi_2$. The properties of these functions that we will need are captured in the following lemma.

**Lemma 4.2.** *Let $N > 0$ be an integer multiple of 4. For $R \ge N$, there exist explicit $\psi_1, \psi_2 \colon \{-1,1\}^n \to \mathbb{R}$ and $d = \Omega(\delta^{1/3}N^{1/3})$ such that*

1. $\displaystyle\sum_{x\in T_1}\psi_1(x) > \frac{1-\delta}{2}$,

2. $\displaystyle -\sum_{x\in T_2}\psi_2(x) > \frac{1-\delta}{2}$,

3. $\|\psi_1\|_1 = \|\psi_2\|_1 = 1$,

4. $\sum_{x \in T_2} |\psi_1(x)| = \sum_{x \in T_1} |\psi_2(x)| = 0$,

5. $\psi_1, \psi_2$ *have pure high degree at least* $d$,

6. $\sum_{x \in T_1} \psi_1(x) = \sum_{x \in R_{N/2,2,1}} \psi_2(x)$,

7. $\sum_{x \in R_{N/2,2,1}} \psi_1(x) = \sum_{x \in T_2} \psi_2(x)$,

8. $\psi_1$ *and* $\psi_2$ *are each constant on each set* $R_{m,a,b}$ *when* $(m,a,b)$ *is valid.*

The functions $\psi_1, \psi_2$ are themselves based on an intermediate construction of a function $\Psi : [N] \times [K] \to \mathbb{R}$, for $K = \Theta(N^{2/3})$, depicted pictorially in Figure 1 below. The function $\Psi$, defined in the proof of Lemma 4.2, combines the constructions of $\omega$ from Lemma 3.1 and $\eta_k$ from Lemma 4.1.

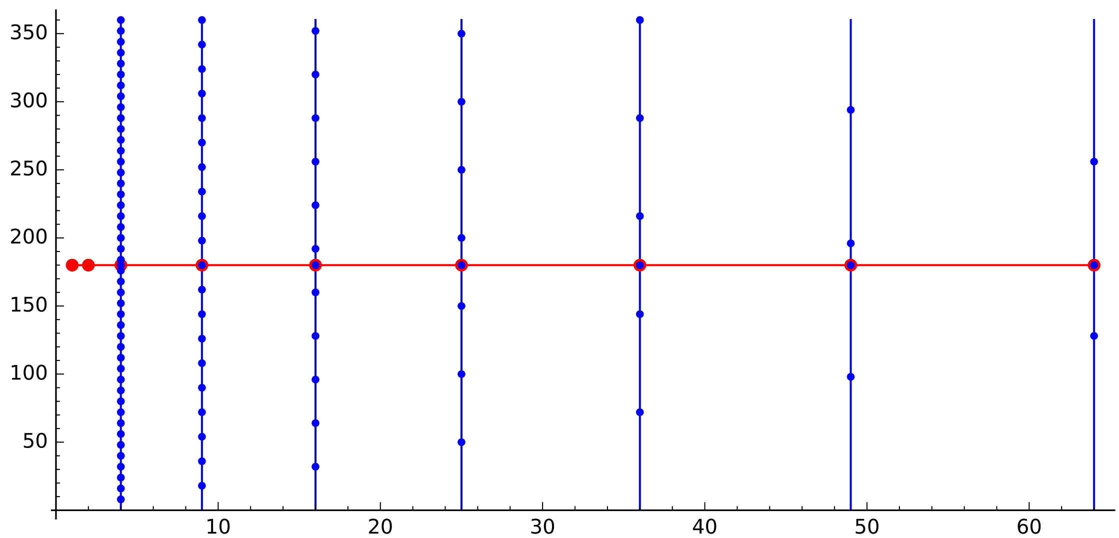


Figure 1: A visualization of the support of the function $\Psi(m,k)$, for $N = 360$ and $K = 64$. The horizontal axis represents values of $k = 1, \dots K$ and the vertical axis represents values of $m = 1, \dots, N$. Red dots indicate values in the support of the function $\omega(k)\mathbb{1}_{m=N/2}$, while blue dots indicate values in the support of the function $\omega(k)\eta_k(m)\mathbb{1}_{k\geq 3}$. When red dots and blue dots overlap, they "cancel out," in the sense that a point with both a red and a blue dot is *not* in the support of $\Psi$.

Together, the functions $\psi_1, \psi_2$ yield the desired dual polynomial for $\mathrm{Col}_{N,R}$.

**Theorem 4.3.** *Let* $N > 0$ *be an integer multiple of 4. For* $R \geq N$*, there exists an explicit* $(1-6\delta, d)$*-dual polynomial* $\psi$ *for* $\mathrm{Col}_{N,R}$ *for* $d = \Omega(\delta^{1/3} N^{1/3})$.

**Remark 4.4.** The dependence of the lower bound Theorem 4.3 on both parameters $\delta$ and $N$ for $1/N \leq \delta \leq 1/10$, is tight up to a logarithmic factor in the size of the range. We show this in Section 7 by constructing an explicit approximating polynomial for $\mathrm{Col}_{N,R}$ of the appropriate degree, by building on the ideas underlying the quantum query algorithm of Brassard et al. [17].

*Proof of Theorem 4.3, assuming Lemma 4.2.* Let

$$a = \sum_{x \in T_1} \psi_1(x)$$

and let

$$b = \sum_{x \in T_2} |\psi_2(x)| = -\sum_{x \in T_2} \psi_2(x),$$

where $\psi_1$ and $\psi_2$ are as in Lemma 4.2. Let $\psi(x) = a\psi_1(x) + b\psi_2(x)$. By Property 5 of Lemma 4.2 and Lemma 4.5 below, $\psi$ also has pure high degree at least $d$. So we need only show that $\psi$ has correlation at least $1-6\delta$ with $\mathrm{Col}_{N,R}$. To this end, note the following.

1. $\sum_{x \in T_1} \psi(x) = a^2 > \dfrac{(1-\delta)^2}{4}$. This inequality uses Properties 1 and 4 of Lemma 4.2.

2. $-\sum_{x \in T_2} \psi(x) = b^2 > \dfrac{(1-\delta)^2}{4}$. This inequality uses Properties 2 and 4 of Lemma 4.2.

3. $\sum_{x \in B} |\psi(x)| \le a \sum_{x \in B \setminus R_{N/2,2,1}} |\psi_1(x)| + b \sum_{x \in B \setminus R_{N/2,2,1}} |\psi_2(x)| \le (a+b)\delta.$

   Here, the first inequality exploits the fact that

$$\sum_{x \in R_{N/2,2,1}} |\psi(x)| = \sum_{x \in R_{N/2,2,1}} |a \cdot \psi_1(x) + b \cdot \psi_2(x)| = 0. \tag{4.1}$$

   The last equality in (4.1) holds because, for all $x \in R_{N/2,2,1}$,

$$\begin{aligned} a \cdot \psi_1(x) + b \cdot \psi_2(x) &= \left( \sum_{x' \in T_1} \psi_1(x') \right) \cdot \psi_1(x) + \left( -\sum_{x' \in T_2} \psi_2(x') \right) \psi_2(x) \\ &= \left( \sum_{x' \in R_{N/2,2,1}} \psi_2(x') \right) \cdot \psi_1(x) + \left( -\sum_{x' \in R_{N/2,2,1}} \psi_1(x') \right) \psi_2(x) \\ &= \left( \sum_{x' \in R_{N/2,2,1}} \psi_2(x') \right) \left( \frac{1}{|R_{N/2,2,1}|} \sum_{x' \in R_{N/2,2,1}} \psi_1(x') \right) + \\ &\qquad \left( -\sum_{x' \in R_{N/2,2,1}} \psi_1(x') \right) \left( \frac{1}{|R_{N/2,2,1}|} \sum_{x' \in R_{N/2,2,1}} \psi_2(x') \right) \\ &= 0, \end{aligned}$$

   where the second equality used Properties 6 and 7 of Lemma 4.2, and the last equality used Property 8.

Thus, the correlation of $\psi$ with $\mathrm{Col}_{N,R}$ is

$$\sum_{x\in T_1}\psi(x)-\sum_{x\in T_2}\psi(x)-\sum_{x\in B}|\psi(x)|\geq a^2+b^2-(a+b)\delta$$
$$\geq\frac{1}{2}-2\delta\geq(1-6\delta)\cdot\|\psi\|_1\,,$$

where the final inequality holds because $\|\psi\|_1\leq a^2+b^2+(a+b)\delta\leq 1/2+\delta$. □

**Lemma 4.5.** *Let $\psi_1,\psi_2:\{-1,1\}^n\to\{-1,1\}$ each have pure high degree at least $d$. Then $\psi=\psi_1+\psi_2$ also has pure high degree at least $d$.*

*Proof.* Let $S\subseteq[n]$ with $|S|\leq d$. Then

$$\sum_{x\in\{-1,1\}^n}\psi(x)\chi_S(x)=\sum_{x\in\{-1,1\}^n}\psi_1(x)\chi_S(x)+\sum_{x\in\{-1,1\}^n}\psi_2(x)\chi_S(x)=0\,.$$ □

What now remains is to prove Lemma 4.2, i. e., to construct the intermediate functions $\psi_1,\psi_2$ that were used to prove Theorem 4.3.

*Proof of Lemma 4.2.* Let $\zeta$ be the constant from Lemma 3.1, let $\rho$ be the constant from Lemma 4.1, and let $\delta'=1/2$. Set

$$K=2\left(\frac{\rho N}{\zeta}\right)^{2/3}\left(\frac{\delta'}{\delta}\right)^{1/3}\quad\text{and let}\quad d=\frac{1}{2}\rho^{1/3}\zeta^{2/3}(\delta')^{1/6}\delta^{1/3}N^{1/3}=\Omega(\delta^{1/3}N^{1/3})\,,$$

noting that $d\leq\zeta(\delta/8)^{1/2}K^{1/2}$ and $d\leq\rho(\delta')^{1/2}N/k$ for every $k\leq K$. Let $\omega\colon\{1,\dots,K\}\to\mathbb{R}$, with correlation constant $\delta/8$, and $\eta_3,\dots,\eta_K\colon\{1,\dots,N\}\to\mathbb{R}$, with correlation constant $\delta'$, be as in the conclusions of those lemmas.

We start by defining a function $\Psi:[N]\times[K]\to\mathbb{R}$ as follows.

$$\Psi(m,k)=\omega(k)\cdot\mathbb{1}_{m=N/2}-\frac{\omega(k)}{\eta_k(N/2)}\mathbb{1}_{k\geq 3}\cdot\eta_k(m)\,.$$

Here,

$$\mathbb{1}_{m=N/2}=\begin{cases}1 & \text{if } m=N/2,\\ 0 & \text{otherwise,}\end{cases}\qquad\textit{and}\qquad\mathbb{1}_{k\geq 3}=\begin{cases}1 & \text{if } k\geq 3,\\ 0 & \text{otherwise.}\end{cases}$$

To help the reader build intuition about the function $\Psi$, we present in Figure 1 a visualization of its support.

We first show how to use $\Psi$ to construct the polynomial $\psi_1$. Analogously to our construction of $\phi$, we want $\psi_1$ to place a total weight of $\Psi(m,a)$ on each set $R_{m,a,1}$. Recall from our overview in Section 2.5 that we think of

$$\psi_1=\psi'+\sum_{\text{invalid triples }(N/2,a,1)}\psi''_{N/2,a,1}\,,$$

where $\psi'$ looks like the simpler "first stage" dual polynomial $\phi$ from our informal overview (which we constructed in Section 3) and each $\psi''_{N/2,a,1}$ cancels out the weight $\phi$ places on values of $k$ the do not divide $N$. This structure underlies our construction of $\Psi$, where we add multiples of the polynomials $\eta_k(m)$ to cancel out the weight $\omega(k)$ places on invalid triples.

Now we construct and analyze the polynomial $\psi_1$. Define

$$\hat{\psi}_1(x) = \begin{cases} \Psi(m,k)/|R_{m,k,1}| & \text{if } x \in R_{m,k,1} \setminus T_1, \\ \Psi(N/2,1)/|T_1| & \text{if } x \in T_1, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\hat{\psi}_1$ is well-defined, because any $x \notin T_1$ is in $R_{m,k,1}$ for at most one triple $(m,k,1)$. We collect several calculations with $\hat{\psi}_1$. First,

$$\sum_{x \in T_1} \hat{\psi}_1(x) = \Psi(N/2,1) = \omega(1) > \frac{1-\delta/8}{2},$$

$$-\sum_{x \in R_{N/2,2,1}} \hat{\psi}_1(x) = -\Psi(N/2,2) = -\omega(2) > \frac{1-\delta/8}{2},$$

and

$$\begin{aligned} \sum_{x \in B \setminus R_{N/2,2,1}} |\hat{\psi}_1(x)| &= \sum_{\{(m,k)\colon k|m\} \setminus \{(N/2,1),(N/2,2)\}} |\Psi(m,k)| \\ &= \sum_{k=3}^{K} \left| \frac{\omega(k)}{\eta_k(N/2)} \right| \sum_{i=1}^{\lfloor N/2k \rfloor} |\eta_k(2ki)| \\ &\leq 4 \sum_{k=3}^{K} |\omega(k)| \\ &\leq \frac{\delta}{2}, \end{aligned}$$

where the penultimate inequality exploits Properties 2 and 3 of Lemma 4.1, and the final inequality exploits Properties 1-3 of Lemma 3.1.

Noting that $|\omega(1)| + |\omega(2)| \leq 1$, it follows that $\|\hat{\psi}_1\|_1 \leq 1 + \delta/2$. So setting $\psi_1 = \hat{\psi}_1/\|\hat{\psi}_1\|_1$, it is immediate that $\psi_1$ satisfies the first three properties in the statement of the lemma. $\psi_1$ also satisfies the fourth property, since for any $x \in T_2$, $\psi_1(x) = \Psi(N,2)/|T_2| = 0$.

Now we will show that $\hat{\psi}_1$, and hence $\psi_1$, has pure high degree at least $d$. We require two observations.

a) $\Psi$ is supported on $(m,k)$ for which $k \mid m$. To see this, note first that for any $k \geq 3$,

$$\Psi(N/2,k) = \omega(k) - \frac{\omega(k)}{\eta_k(N/2)} \cdot \eta_k(N/2) = 0.$$

The claim now follows from Property 1 of Lemma 4.1, combined with the fact that $2 \mid N$.

b) $\Psi(m,1)$ is nonzero only for $m = N/2$, and hence

$$\sum_{x \in T_1} \hat{\psi}_1(x) = \Psi(N/2,1) = \sum_{m=1}^{N} \Psi(m,1).$$

Fix any $S \subseteq [n]$ with $|S| \leq d$. Let $Q(m,k)$ be a polynomial of degree at most $d$ in each variable such that, for all pairs $(m,k)$ with $k \mid m$,

$$Q(m,k) = \mathbb{E}_{x \in R_{m,k,1}}[\chi_S(x)] .$$

The existence of such a bivariate polynomial $Q$ is guaranteed by Lemma 2.5. Then the previous two observations together imply that

$$\sum_{x \in \{-1,1\}^n} \hat{\psi}_1(x)\chi_S(x) = \sum_{m=1}^{N} \sum_{k=1}^{N} \Psi(m,k)Q(m,k) . \tag{4.2}$$

We remark that a key point is the derivation of (4.2) is that we have no control over the evaluations $Q(m,k)$ when $k$ does not divide $m$, yet this is rendered irrelevant because $\Psi(m,k) = 0$ for all such pairs.

The right hand side of (4.2) equals

$$\sum_{k=1}^{K} \omega(k)Q(N/2,k) - \sum_{k=3}^{K} \frac{\omega(k)}{\eta_k(N/2)} \left( \eta_k(N/2)Q(N/2,k) + \sum_{i=1}^{\lfloor N/2k \rfloor} \eta_k(2ik)Q(2ik,k) \right) . \tag{4.3}$$

The first sum in (4.3) is zero by Lemma 3.1 since $Q(N/2,k)$ is a polynomial of degree at most $d$ in $k$. The second sum is also zero because for each fixed $k$, $Q(r,k)$ is a polynomial of degree at most $d$ in the variable $r$, and hence the term in parentheses is zero by Lemma 4.1 (Parts 1 and 4). Thus $\hat{\psi}_1$ has pure high degree at least $d$.

The construction of $\psi_2$ is similar. This time, we let

$$\hat{\psi}_2(x) = \begin{cases} \Psi(m,k)/|R_{m,k,2}| & \text{if } x \in R_{m,k,2} \setminus T_2, \\ \Psi(N/2,2)/|T_2| & \text{if } x \in T_2, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\hat{\psi}_2$ is well-defined, because any $x \notin T_2$ is in $R_{m,k,2}$ for at most one triple $(m,k,2)$. We define $\psi_2 = \hat{\psi}_2/\|\hat{\psi}_2\|_1$. Showing that $\psi_2$ satisfies Properties 1-4 of the lemma follows from the same calculations we used for $\psi_1$.

To show that $\psi_2$ has pure high degree at least $d$, we require the following additional observations.

c) $\Psi$ is supported on pairs $(m,k)$ for which $k \mid m$ and $2 \mid (N-m)$. To see the latter property, note that if $\Psi(m,k) \neq 0$, then $m$ is even (this holds because $N/2$ is even, which follows from our requirement that $N$ is a multiple of 4), and hence $N-m$ is as well.

d) $\Psi(m,2)$ is nonzero only for $m = N/2$. It follows that $\sum_{x \in T_2} \hat{\psi}_2(x) = \Psi(N/2,2) = \sum_{m=1}^{N} \Psi(m,2)$.

With these observations in hand, showing that $\psi_2$ has pure high degree $d$ then follows from calculations analogous to the ones we used for $\psi_1$.

Finally, the fact that $\psi_1$ and $\psi_2$ satisfy Properties 6, 7, and 8 of the lemma follows from their definitions, combined with the fact that $R_{N/2,1,2} = R_{N/2,2,1}$. In fact, $\sum_{x \in T_1} \psi_1(x)$ equals $\Psi(N/2,1)$, while

$$\sum_{x \in R_{N/2,2,1}} \psi_2(x)$$

also equals $\Psi(N/2,1)$, giving Property 6. Similarly, $\sum_{x\in T_2}\psi_2(x)$ equals $\Psi(N/2,2)$, while

$$\sum_{x\in R_{N/2,1,2}} \psi_1(x)$$

also equals $\Psi(N/2,1)$. This completes the proof. $\square$

## 5 A dual polynomial for ELEMENT DISTINCTNESS

In this section, we give a generic construction showing how to transform any dual polynomial for COLLISION into a dual polynomial for ELEMENT DISTINCTNESS.

We first recall the reduction from COLLISION to ELEMENT DISTINCTNESS given in [5].[4] The reduction shows how to turn a polynomial $p$ approximating $\mathrm{ED}_{M,R}$ into a polynomial $q$ approximating $\mathrm{Col}_{N,R}$, with $N \approx M^2$ and $\deg q \le \deg p$.

We illustrate the reduction for $N = M^2/12$. Let $p : \{-1,1\}^m \to \{-1,1\}$ be an $(1/6)$-approximation of $\mathrm{ED}_{M,R}$, with $m = M\log R$. Define a polynomial $q : \{-1,1\}^n \to \{-1,1\}$ for $n = N\log R$ by

$$q(y_1,\dots,y_N) = \frac{1}{\binom{N}{M}} \sum_{1\le i_1<i_2<\cdots<i_M\le N} p(y_{i_1},y_{i_2},\dots,y_{i_M})\,.$$

That is, $q(y)$ is the expected value of $p(x)$ where $x$ is the concatenation of a random subset of $M$ of the blocks $y_1,\dots,y_N$. To simplify notation, for a set $S=\{i_1,i_2,\dots,i_M\}$, let $y|_S = (y_{i_1},y_{i_2},\dots,y_{i_M})$. Note that $\deg q \le \deg p$. Moreover, since $q$ is an average of values in $[-7/6,7/6]$, it is always in $[-7/6,7/6]$ itself. To finish arguing that $q$ is a $(1/3)$-approximation to $\mathrm{Col}_{N,R}$, we consider two cases.

1. If $y \in T_1$, i. e., $y$ is a 1-to-1 input, then $y|_S$ is always 1-to-1. Hence $p(y|_S) \in [5/6,7/6]$ for every subset of indices, so $q(y) \in [2/3,4/3]$.

2. If $y \in T_2$, i. e., $y$ is a 2-to-1 input, then with high probability $y|_S$ is not 1-to-1. This follows from the "birthday bound":
$$\Pr_{|S|=M}[\mathrm{ED}(y|_S)=1] \le \exp(-M^2/4N) \le \frac{1}{12}\,.$$
Therefore, $q(y) \le (11/12)(-5/6)+(1/12)(7/6) \le -2/3$.

The construction we give in this section takes a dual view of the reduction above. Namely, we show how to transform a dual polynomial $\psi$ for $\mathrm{Col}_{N,R}$ into a dual polynomial $\varphi$ for $\mathrm{ED}_{M,R}$, with $M^2 \approx N$. In the primal reduction, we constructed $q(y)$ from $p(x)$ by averaging $p$ over all subsets of size $M$. The right analogue in the dual reduction is to construct $\varphi(x)$ by averaging $\psi(y)$ over a carefully constructed set of *extensions* from $x$ to a longer input $y$. In particular, $\varphi(x)$ averages $\psi(y)$ over all $y$ for which $x$ could have been produced by taking a subset of $M$ blocks of $y$.

We give this reduction formally below.

[4] While the reduction given in Aaronson and Shi's paper is stated in terms of quantum query algorithms, it is straightforward to rephrase the reduction in terms of approximating polynomials instead.

**Theorem 5.1.** *Let $\psi : \{-1,1\}^n \to \{-1,1\}$ be a $(1-\delta,d)$-dual polynomial for $\mathrm{Col}_{N,R}$. Then $\psi$ can be used to construct $\varphi : \{-1,1\}^m \to \{-1,1\}$ that is an $(1-2\delta,d)$-dual polynomial for $\mathrm{ED}_{M,R}$ when $M \geq 2\sqrt{N\log(2/\delta)}$.*

**Corollary 5.2.** *For any $\delta > 0$, there is an explicit $(1-\delta,d)$-dual polynomial for $\mathrm{ED}_{M,R}$ with*

$$d = \Omega\left(\left(\frac{\delta}{\log(1/\delta)}\right)^{1/3} M^{2/3}\right).$$

**Remark 5.3.** The dependence of Corollary 5.2 on $\delta$ is essentially tight for $\delta = O(M^{-2})$. See Section 7 for details.

*Proof of Theorem 5.1.* Given a set $S = \{i_1,\dots,i_M\} \subset [N]$ with $i_1 < i_2 < \dots < i_M$ and a bit string $y = (y_1,\dots,y_N) \in \{-1,1\}^n$, define the restriction of $y$ to the set $S$, denoted by $y|_S \in \{-1,1\}^m$, to be the string of length $m = M\log R$ obtained by concatenating the blocks $y_i$ for $i \in S$, i. e., $y|_S = (y_{i_1}, y_{i_2},\dots,y_{i_M})$. Given a bit string $x \in \{-1,1\}^m$, define the multiset of extensions of $x$, denoted by $\mathrm{ext}(x)$, to be the $\binom{N}{M}R^{N-M}$ strings $y \in \{-1,1\}^n$ where $y|_S = x$ for some $|S| = M$. Restrictions and extensions are related by the equality of the multisets

$$\{(x,y) : x \in \{-1,1\}^m, y \in \mathrm{ext}(x)\} = \{(x,y) : y \in \{-1,1\}^n, x = y|_S \text{ for some } |S| = m\}. \tag{5.1}$$

For $x \in \{-1,1\}^m$, define the polynomial

$$\varphi(x) = \frac{1}{\binom{N}{M}} \sum_{y \in \mathrm{ext}(x)} \psi(y).$$

Let $\varphi(x) = 0$ for $x \notin \{-1,1\}^m$. We claim that $\varphi$ is a good dual polynomial for the ELEMENT DISTINCTNESS function ED, which requires us to show

1. $\sum_{x\in\{-1,1\}^m} \varphi(x)\,\mathrm{ED}(x) > (1-2\delta)\cdot \sum_{x\in\{-1,1\}^m} |\varphi(x)|$, and
2. $\sum_{x\in\{-1,1\}^m} \varphi(x)\chi_S(x) = 0$ for all $|S| \leq d$.

To verify the first property, define

$$A(y) = \frac{1}{\binom{N}{M}} \sum_{|S|=M} \mathrm{ED}(y|_S).$$

We collect a few observations about $A$.

1. $|A(y)| \leq 1$ for all $y$.
2. If $y \in T_1$, then $A(y) = 1$.
3. If $y \in T_2$, then
$$\Pr_{|S|=M}[\mathrm{ED}(y|_S) = 1] \leq \exp(-M^2/4N).$$
Hence,
$$A(y) \leq -1 + 2\exp(-M^2/4N) \leq -1+\delta.$$

Therefore we get

$$\begin{aligned}
\sum_{x\in\{-1,1\}^m} \varphi(x)\,\mathrm{ED}(x) &= \frac{1}{\binom{N}{M}} \sum_{x\in\{-1,1\}^m} \sum_{y\in \mathrm{ext}(x)} \psi(y)\,\mathrm{ED}(x) \\
&= \frac{1}{\binom{N}{M}} \sum_{y\in\{-1,1\}^n} \sum_{|S|=M} \psi(y)\,\mathrm{ED}(y|_S) && \text{by (5.1)} \\
&= \sum_{y\in\{-1,1\}^n} A(y)\psi(y) && \text{by definition of } A.
\end{aligned}$$

By observations (1)-(3) about $A$, this expression is

$$\begin{aligned}
&\geq \left(\sum_{y\in T_1} \psi(y) - \sum_{y\in T_2} \psi(y) - \sum_{y\in B} |\psi(y)|\right) - \delta \sum_{y\in T_2} |\psi(y)| \\
&\geq (1-2\delta) \sum_{y\in\{-1,1\}^n} |\psi(y)| && \text{as } \psi \text{ is a dual polynomial for } \mathrm{Col}_{N,R} \\
&= (1-2\delta) \sum_{y\in\{-1,1\}^n} \frac{1}{\binom{N}{M}} \sum_{|S|=M} |\psi(y)| \\
&\geq \frac{1-2\delta}{\binom{N}{M}} \sum_{x\in\{-1,1\}^m} \sum_{y\in \mathrm{ext}(x)} |\psi(y)| && \text{by (5.1)} \\
&\geq (1-2\delta) \sum_{x\in\{-1,1\}^m} |\varphi(x)|\,.
\end{aligned}$$

For the second property, let $T$ be a subset of $[N]$ with $|T| \leq d$. Then

$$\begin{aligned}
\sum_{x\in\{-1,1\}^m} \varphi(x)\chi_T(x) &= \frac{1}{\binom{N}{M}} \sum_{x\in\{-1,1\}^m} \sum_{y\in \mathrm{ext}(x)} \psi(y)\chi_T(x) \\
&= \frac{1}{\binom{N}{M}} \sum_{|S|=M} \sum_{y\in\{-1,1\}^n} \psi(y)\chi_T(y|_S) && \text{by (5.1)} \\
&= \frac{1}{\binom{N}{M}} \sum_{|S|=M} \sum_{y\in\{-1,1\}^n} \psi(y)\chi_{T|_S}(y) \\
&= 0\,,
\end{aligned}$$

where $T|_S$ denotes the subset of $T$ contained in the blocks specified by $S$. □

## 6 On complementary slackness

Recalling that any bounded-error quantum query algorithm can be converted into an approximating polynomial [11], the collision-finding algorithm of Brassard, Høyer, and Tapp [17] yields an explicit, asymptotically optimal approximating polynomial for $\mathrm{Col}_{N,R}$. We describe this polynomial $p$ below.

Recall that any approximating polynomial for $\mathrm{Col}_{N,R}$ represents a feasible solution to the primal linear program considered in Section 2.2. If the polynomial $p$ were an *exactly* optimal $\varepsilon$-approximation for $\mathrm{Col}_{N,R}$, then complementary slackness would imply that the optimal dual polynomial $\psi$ for $\mathrm{Col}_{N,R}$ is supported on the points corresponding to constraints made tight by $p$. That is, $\psi : \{-1,1\}^n \to \{-1,1\}$ is supported on $x \in \{-1,1\}^n$ for which $|p(x) - \mathrm{Col}(x)| = \varepsilon$. We refer to these as the *maximum-error points* of $p$.

While we do not know whether $p$ is an exactly optimal approximating polynomial for $\mathrm{Col}_{N,R}$, we might still expect that an approximate version of complementary slackness holds, in the sense that a "good" dual polynomial should place all or most of its weight on points that are "nearly" maximum-error points of $p$. Indeed, this intuition has proven accurate for all of the dual polynomials constructed in prior work, including for symmetric functions (see [19, Section 4.5]), block-composed functions (see [48, Section 1.2.4]), and the intersection of two majorities [41]. Below, we argue that $k$-to-1 inputs are nearly maximum-error points for $p$, which explains why our dual polynomials for collision are supported on inputs that are roughly $k$-to-1, in addition to why these inputs play a prominent role in the original symmetrization-based proof.

**An asymptotically optimal approximation $p$ for $\mathrm{Col}_{N,R}$.** For a subset $S \subset [N]$, define $\mathsf{cross}_S \colon \{-1,1\}^n \to \mathbb{R}$ via

$$\mathsf{cross}_S(x_1,\ldots,x_N) = |\{i \in S, j \notin S : x_i = x_j\}| = \sum_{i \in S, j \notin S} \mathrm{EQ}(x_i,x_j)\,,$$

where EQ denotes the equality function. That is, $\mathsf{cross}_S(x)$ counts the number of cross-collisions between indices in $S$ and indices outside of $S$. Notice that $\mathrm{EQ}(x_i,x_j)$ is a function of only $2 \cdot \log R$ variables, and hence $\mathsf{cross}_S(x_1,\ldots,x_N)$ is exactly computed by a polynomial of degree $2 \cdot \log R$.

In addition, for a subset $S \subset [N]$, define the function $\mathbb{I}_{\mathrm{ED},S}(x_1,\ldots,x_N)$ to be 1 if $x_i \neq x_j$ for all pairs $i, j \in S$ with $i \neq j$, and 0 otherwise. That is, $\mathbb{I}_{\mathrm{ED},S}$ indicates whether $x$ is 1-to-1 on the indices in $S$. Notice that $\mathbb{I}_{\mathrm{ED},S}$ is a function of only $|S| \cdot \log R$ variables, and hence is exactly computed by a polynomial of degree $|S| \cdot \log R$.

For the remainder of the discussion, let $r = N^{1/3}$—we focus on the quantity $\mathsf{cross}_S(x)$ when $|S| = r$. We will need the following simple observations.

1. If $x \in T_1$, i. e., $x$ is a 1-to-1 input, then $\mathsf{cross}_S(x) = 0$ and $\mathbb{I}_{\mathrm{ED},S}(x) = 1$ for any $S$.
2. If $x \in T_2$, i. e., $x$ is a 2-to-1 input, then $\mathbb{I}_{\mathrm{ED},S}(x) = 1 \Longrightarrow \mathsf{cross}_S(x) = r$.
3. If $x \in T_2$, then, over the random choice of $S$, $\mathbb{I}_{\mathrm{ED},S}(x) = 0$ with probability at most $(N/2) \cdot (r/N)^2 \leq N^{-1/3}$.
4. For all $x \in \{-1,1\}^n$, $\mathbb{I}_{\mathrm{ED},S}(x) = 1 \Longrightarrow \mathsf{cross}_S(x) \leq N - r$.

Let $T_d : \mathbb{R} \to \mathbb{R}$ denote the degree-$d$ Chebyshev polynomial of the first kind. This polynomial has the following properties.

a) $T_d(x) \in [-1,1]$ for $x \in [-1,1]$.

b) $T_d(1 + M/d^2) \geq 10$ for a constant $M$ independent of $d$.

c) The extreme points of $T_d$ in $[-1,1]$ are the degree-$d$ *Chebyshev nodes*, which take the form $\cos(i\pi/d)$ for $0 \le i \le d$.

Truncating the Taylor expansion of $\cos(x) = 1 - x^2/2 + \dots$ after the quadratic term, one sees that the Chebyshev nodes are well-approximated via the expression $\cos(i\pi/d) \approx 1 - (ci^2/d^2)$ for some constant $c$.

Applying an appropriate affine transformation to $T_d$, we obtain a polynomial $A_d$ with the following properties.

a) $A_d(0) = 1$.

b) $A_d(i) \in [-1, -3/4]$ for all real numbers $i \in [1, d^2/M]$.

c) $A_d(i) \in [-1, 1]$ for all real numbers $i \in [0, d^2/M]$.

d) The extreme points of $A_d$ are well approximated by the points $c \cdot i^2$ for $i \in \{0, 1, \dots, \lfloor d \cdot M^{-1/2} \rfloor\}$.

Let $p_S(x) = \mathbb{I}_{\mathrm{ED},S}(x) \cdot A_d(\mathsf{cross}_S(x_1, \dots, x_N)/r)$ for $d = 100 \cdot M \cdot N^{1/3}$, and let

$$p(x) = \mathbb{E}_{|S|=r}[p_S(x)] = \frac{1}{\binom{N}{r}} \sum_{|S|=r} p_S(x)\,.$$

Then $p$ is a polynomial of degree $|S| \log R + 2 \cdot d \cdot \log R = O(N^{1/3} \log R)$. We argue that $p$ approximates $\mathrm{Col}_{N,R}$ to error $\varepsilon$ for some $\varepsilon \le 1/3$. The analysis falls into three cases.

**Case 1:** For $x \in T_1$, $p_S(x) = A_d(0) = -1$ for all $S$, where the first equality follows from Property 1 above. So $p(x) = \mathbb{E}_{|S|=r}[p_S(x)] = 1$.

**Case 2:** For $x \in T_2$, $\mathbb{I}_{\mathrm{ED},S}(x) = 1 \Longrightarrow p_S(x) = A_d(1) \in [-1, -3/4]$, where the equality follows from Property 2 above. Meanwhile, $\mathbb{I}_{\mathrm{ED},S}(x) \ne 1 \Longrightarrow p_S(x) = 0$. Combining these two facts with Property 3 above establishes that $p(x) = \mathbb{E}_{|S|=r}[p_S(x)] \in [-1, -2/3]$.

**Case 3:** For $x \in \{-1, 1\}^n$, $p_S(x) \in [-1, 1]$. This follows from Property 4 above.

**Identifying maximum-error points of** $p$. For any fixed $S$, the maximum-error points of $p_S$ are well-approximated by the $x \in \{1, 1\}^n$ for which the following two equations hold:

$$\mathsf{cross}_S(x) = c \cdot i^2 \cdot r \text{ for some } i \in \{0, 1, \dots, \lfloor d \cdot M^{-1/2} \rfloor\} \tag{6.1}$$

and

$$\mathbb{I}_{\mathrm{ED},S}(x) = 1. \tag{6.2}$$

(This follows from the fact that the extreme points of $A_d$ are roughly of the form $c \cdot i^2$ for $0 \le i \le d \cdot M^{-1/2}$).

However, the maximum-error points for the averaged polynomial $p(x) = \mathbb{E}_{|S|=r}[p_S(x)]$ are the points $x$ that satisfy (6.1) and (6.2) *with high probability* over the choice of $S$. Indeed, for these points $x$, the error of $p(x)$ is at least $\varepsilon \cdot (1 - o(1)) \approx \varepsilon$.

Consider any $k$ of the form $k = c \cdot i^2 + 1$ for some $i \in \{0, 1, \ldots, \lfloor d \cdot M^{-1/2} \rfloor\}$, such that $k = o(N^{1/3})$. Consider any $x \in T_k$; we claim that $x$ satisfies (6.1) and (6.2) with probability $1 - o(1)$ over choice of $S$. To see this, observe that the probability that $\mathbb{I}_{\mathrm{ED},S}(x_S) = 0$ is at most $(N/k) \cdot k^2 \cdot (r/N)^2 = k \cdot r^2/N = o(1)$. And if $\mathbb{I}_{\mathrm{ED},S}(x_S) \neq 0$, then the number of cross-collisions is exactly

$$\mathsf{cross}_S(x_1, \ldots, x_N) = r \cdot (k-1)\,.$$

When $k$ takes the form $k = c \cdot i^2 + 1$, this means that $x$ satisfies (6.1). Hence, $x$ has nearly maximal error even for the averaged polynomial $p$.

## 7 On the tightness of Theorem 4.3 and Corollary 5.2

To complement Theorem 4.3, we construct an approximating polynomial that gives a nearly matching upper bound on the approximate degree of $\mathrm{Col}_{N,R}$. The construction is a refinement of the approximating polynomial given in Section 6.

**Proposition 7.1.** *For $0 \leq \delta \leq 1/N$, there exists a polynomial $p$ of degree $O(\delta^{1/3}N^{1/3}\log R)$ that $(1-\delta)$-approximates $\mathrm{Col}_{N,R}$.*

*Proof sketch.* See Section 6 for the construction of an approximating polynomial of degree $O(N^{1/3}\log R)$ in the case where $\delta$ is constant. In order to obtain an improved upper bound for vanishing $\delta$, we make the following changes to that construction.

1. We instead choose $r = \delta^{1/3}N^{1/3}$. Now if $x$ is a 2-1 input, the probability over the random choice of the set $S$ of obtaining a collision inside $S$, i.e., the probability that $\mathbb{I}_{\mathrm{ED},S} = 0$, is at most $(N/2) \cdot (r/N)^2 \leq \delta/2$.

2. We instead let $A_d$ be an affine transformation of a Chebyshev polynomial with the following properties for some constant $M$:

   - $A_d(0) \geq \delta/2$,
   - $A_d(i) \in [-1, -\delta/2]$ for $i \in [1, d^2/M\delta]$,
   - $A_d(i) \in [-1, 1]$ for $x \in [0, d^2/M\delta]$.

3. Setting $d = 100 \cdot M \cdot r$ ensures that the polynomial $p$ has degree $O(\delta^{1/3}N^{1/3}\log R)$ and is a $(1-\delta)$-approximation of $\mathrm{Col}_{N,R}$. □

We now show that Corollary 5.2 is tight up to a factor of $\log R$, when $\delta \leq 1/M^2$. This gives mild evidence that the lower bound has the right dependence on both parameters $M, \delta$ for vanishing $\delta$.

**Proposition 7.2.** *Let $\delta \leq 1/M^2$. Then there exists a $(1-\delta)$-approximating polynomial for $\mathrm{ED}_{M,R}$ with degree $O(\log R)$.*

*Proof.* We write

$$\mathrm{ED}_{M,R}(x_1,\ldots,x_M) = \bigwedge_{i\neq j} \mathrm{NEQ}(x_i,x_j)\,,$$

where $\mathrm{NEQ}(x_i,x_j) = 1$ if $i$ and inputs $j$ are distinct, and is zero otherwise. The function NEQ can be computed exactly by a polynomial of degree $O(\log R)$. Therefore, the polynomial

$$\frac{1}{\binom{M}{2}}\left(\frac{1}{2} - \sum_{i\neq j} \mathrm{NEQ}(x_i,x_j)\right)$$

has degree $O(\log R)$ and approximates $\mathrm{ED}_{M,R}$ to within error $1-1/M^2$. □

**Acknowledgments.** We are grateful to Guy Kindler, Yaoyun Shi, and Mario Szegedy for several illuminating discussions during the early stages of this work. We thank Scott Aaronson and Emanuele Viola for helpful comments on an earlier version of this manuscript. We also thank Sasha Sherstov and the anonymous reviewers for additional helpful suggestions (including the suggestion to include Figure 1).

## References

[1] SCOTT AARONSON: Quantum lower bound for the collision problem. In *Proc. 34th STOC*, pp. 635–642. ACM Press, 2002. [doi:10.1145/509907.509999] 2

[2] SCOTT AARONSON: The polynomial method in quantum and classical computing. In *Proc. 49th FOCS*, p. 3. IEEE Comp. Soc. Press, 2008. [doi:10.1109/FOCS.2008.91] 2, 11

[3] SCOTT AARONSON: The collision problem: Notes for lecture 13 of MIT course 6.845: Quantum complexity theory, 2010. Available at MIT OCW. 9

[4] SCOTT AARONSON: The collision lower bound after 12 years, 2013. Available on author's website. 2

[5] 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, 4, 9, 23

[6] ANDRIS AMBAINIS: Quantum lower bounds by quantum arguments. *J. Comput. System Sci.*, 64(4):750–767, 2002. Preliminary version in STOC'00. [doi:10.1006/jcss.2002.1826, arXiv:quant-ph/0002066] 5

[7] ANDRIS AMBAINIS: Polynomial degree and lower bounds in quantum complexity: Collision and element distinctness with small range. *Theory of Computing*, 1(3):37–46, 2005. [doi:10.4086/toc.2005.v001a003, arXiv:quant-ph/0305179] 2, 3

[8] ANDRIS AMBAINIS: Polynomial degree vs. quantum query complexity. *J. Comput. System Sci.*, 72(2):220–238, 2006. Preliminary version in FOCS'03. [doi:10.1016/j.jcss.2005.06.006, arXiv:quant-ph/0305028] 5

[9] ANDRIS AMBAINIS: Quantum walk algorithm for element distinctness. *SIAM J. Comput.*, 37(1):210–239, 2007. Preliminary version in FOCS'04. [doi:10.1137/S0097539705447311, arXiv:quant-ph/0311001] 2

[10] HOWARD BARNUM, MICHAEL E. SAKS, AND MARIO SZEGEDY: Quantum query complexity and semi-definite programming. In *Proc. 18th IEEE Conf. on Computational Complexity (CCC'03)*, pp. 179–193. IEEE Comp. Soc. Press, 2003. [doi:10.1109/CCC.2003.1214419] 5

[11] ROBERT BEALS, HARRY BUHRMAN, RICHARD CLEVE, MICHELE MOSCA, AND RONALD DE WOLF: Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001. Preliminary version in FOCS'98. [doi:10.1145/502090.502097, arXiv:quant-ph/9802049] 2, 5, 25

[12] RICHARD BEIGEL: The polynomial method in circuit complexity. In *Proc. 8th IEEE Conf. on Structure in Complexity Theory (SCT'93)*, pp. 82–95. IEEE Comp. Soc. Press, 1993. [doi:10.1109/SCT.1993.336538] 2

[13] RICHARD BEIGEL: Perceptrons, PP, and the polynomial hierarchy. *Comput. Complexity*, 4(4):339–349, 1994. Preliminary version in SCT'92. [doi:10.1007/BF01263422] 2

[14] ALEKSANDRS BELOVS AND ANSIS ROSMANIS: Adversary lower bounds for the collision and the set equality problems, 2013. [arXiv:1310.5185] 5

[15] ALEKSANDRS BELOVS AND ROBERT ŠPALEK: Adversary lower bound for the $k$-sum problem. In *Proc. 4th Conf. on Innovations in Theoret. Comput. Sci. (ITCS'13)*, pp. 323–328. ACM Press, 2013. [doi:10.1145/2422436.2422474, arXiv:1206.6528] 5

[16] ANDREJ BOGDANOV, YUVAL ISHAI, EMANUELE VIOLA, AND CHRISTOPHER WILLIAMSON: Bounded indistinguishability and the complexity of recovering secrets. In *Proc. of 36th Internat. Cryptology Conf. (CRYPTO'16)*, volume 9816 of *LNCS*, pp. 593–618. Springer, 2016. Available at ECCC. [doi:10.1007/978-3-662-53015-3_21] 4

[17] GILLES BRASSARD, PETER HØYER, AND ALAIN TAPP: Quantum algorithm for the collision problem. In *Encyclopedia of Algorithms*, pp. 1–99. Springer, 2008. Preliminary version in ACM SIGACT News. [doi:10.1007/978-0-387-30162-4_304, arXiv:quant-ph/9705002] 2, 19, 25

[18] HARRY BUHRMAN, NIKOLAI K. VERESHCHAGIN, AND RONALD DE WOLF: On computation and communication with small bias. In *Proc. 22nd IEEE Conf. on Computational Complexity (CCC'07)*, pp. 24–32. IEEE Comp. Soc. Press, 2007. [doi:10.1109/CCC.2007.18] 2, 5

[19] MARK BUN AND JUSTIN THALER: Dual lower bounds for approximate degree and Markov-Bernstein inequalities. *Inform. and Comput.*, 243:2–25, 2015. Preliminary version in ICALP'13. [doi:10.1016/j.ic.2014.12.003, arXiv:1302.6191] 2, 3, 9, 10, 11, 12, 14, 15, 26

[20] MARK BUN AND JUSTIN THALER: Hardness amplification and the approximate degree of constant-depth circuits. In *Proc. 42nd Internat. Colloq. on Automata, Languages and Programming (ICALP'15)*, volume 9134 of *LNCS*, pp. 268–280. Springer, 2015. [doi:10.1007/978-3-662-47672-7_22, arXiv:1311.1616] 3, 4, 5, 12

[21] KARTHEKEYAN CHANDRASEKARAN, JUSTIN THALER, JONATHAN ULLMAN, AND ANDREW WAN: Faster private release of marginals on small databases. In *Proc. 5th Conf. on Innovations in Theoret. Comput. Sci. (ITCS'14)*, pp. 387–402. ACM Press, 2014. [doi:10.1145/2554797.2554833, arXiv:1304.3754] 2

[22] PETER HØYER, TROY LEE, AND ROBERT SPALEK: Negative weights make adversaries stronger. In *Proc. 39th STOC*, pp. 526–535. ACM Press, 2007. [doi:10.1145/1250790.1250867, arXiv:quant-ph/0611054] 5

[23] PETER HØYER, MICHELE MOSCA, AND RONALD DE WOLF: Quantum search on bounded-error inputs. In *Proc. 30th Internat. Colloq. on Automata, Languages and Programming (ICALP'03)*, volume 2719 of *LNCS*, pp. 291–299. Springer, 2003. [doi:10.1007/3-540-45061-0_25, arXiv:quant-ph/0304052] 3

[24] ADAM TAUMAN KALAI, ADAM R. KLIVANS, YISHAY MANSOUR, AND ROCCO A. SERVEDIO: Agnostically learning halfspaces. *SIAM J. Comput.*, 37(6):1777–1805, 2008. Preliminary version in FOCS'05. [doi:10.1137/060649057] 2

[25] VARUN KANADE AND JUSTIN THALER: Distribution-independent reliable learning. In *Proc. 27th Ann. Conf. on Learning Theory (COLT'14)*, pp. 3–24, 2014. JMLR. [arXiv:1402.5164] 2

[26] ADAM R. KLIVANS AND ROCCO A. SERVEDIO: Learning DNF in time $2^{\tilde{O}(n^{1/3})}$. *J. Comput. System Sci.*, 68(2):303–318, 2004. Preliminary version in STOC'01. [doi:10.1016/j.jcss.2003.07.007] 2

[27] ADAM R. KLIVANS AND ROCCO A. SERVEDIO: Toward attribute efficient learning of decision lists and parities. *J. Machine Learning Res.*, 7:587–602, 2006. JMLR. Preliminary version in COLT'04. 2

[28] SAMUEL KUTIN: Quantum lower bound for the collision problem with small range. *Theory of Computing*, 1(2):29–36, 2005. [doi:10.4086/toc.2005.v001a002] 2, 3, 4, 7, 9, 11, 12

[29] LOÏCK MAGNIN AND JÉRÉMIE ROLAND: Explicit relation between all lower bound techniques for quantum query complexity. *Internat. J. Quantum Inform.*, 13(4):1350059, 2015. Preliminary versions in STACS'13 and ECCC. [doi:10.1142/S0219749913500597, arXiv:1209.2713] 5

[30] NOAM NISAN AND MARIO SZEGEDY: On the degree of boolean functions as real polynomials. *Comput. Complexity*, 4(4):301–313, 1994. Preliminary version in STOC'92. [doi:10.1007/BF01263419] 3, 9, 11

[31] RYAN O'DONNELL AND ROCCO A. SERVEDIO: New degree bounds for polynomial threshold functions. *Combinatorica*, 30(3):327–358, 2010. Preliminary version in STOC'03. [doi:10.1007/s00493-010-2173-3] 3, 12

[32] RAMAMOHAN PATURI: On the degree of polynomials that approximate symmetric boolean functions (preliminary version). In *Proc. 24th STOC*, pp. 468–474. ACM Press, 1992. [doi:10.1145/129712.129758] 10, 11

[33] BEN W. REICHARDT: Span programs and quantum query complexity: The general adversary bound is nearly tight for every boolean function. In *Proc. 50th FOCS*, pp. 544–551. IEEE Comp. Soc. Press, 2009. [doi:10.1109/FOCS.2009.55, arXiv:0904.2759] 5

[34] BEN W. REICHARDT: Reflections for quantum query algorithms. In *Proc. 22nd Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'11)*, pp. 560–569. ACM Press, 2011. [doi:10.1137/1.9781611973082.44, arXiv:1005.1601] 5

[35] ROCCO A. SERVEDIO, LI-YANG TAN, AND JUSTIN THALER: Attribute-efficient learning and weight-degree tradeoffs for polynomial threshold functions. In *Proc. 25th Ann. Conf. on Learning Theory (COLT'12)*, volume 23, pp. 14.1–14.19, 2012. JMLR. 2

[36] ALEXANDER A. SHERSTOV: Communication lower bounds using dual polynomials. *Bulletin of the EATCS*, 95:59–93, 2008. Available at EATCS. [arXiv:0805.2135] 2, 3

[37] ALEXANDER A. SHERSTOV: Separating $AC^0$ from depth-2 majority circuits. *SIAM J. Comput.*, 38(6):2113–2129, 2009. Preliminary version in STOC'07. [doi:10.1137/08071421X] 2

[38] ALEXANDER A. SHERSTOV: The pattern matrix method. *SIAM J. Comput.*, 40(6):1969–2000, 2011. Preliminary version in STOC'08. [doi:10.1137/080733644, arXiv:0906.4291] 2, 3, 4, 5

[39] ALEXANDER A. SHERSTOV: Strong direct product theorems for quantum communication and query complexity. *SIAM J. Comput.*, 41(5):1122–1165, 2011. Preliminary version in STOC'11. [doi:10.1137/110842661, arXiv:1011.4935] 3

[40] ALEXANDER A. SHERSTOV: Approximating the AND-OR tree. *Theory of Computing*, 9(20):653–663, 2013. Preliminary version in ECCC. [doi:10.4086/toc.2013.v009a020] 3

[41] ALEXANDER A. SHERSTOV: The intersection of two halfspaces has high threshold degree. *SIAM J. Comput.*, 42(6):2329–2374, 2013. Preliminary versions in FOCS'09 and ECCC. [doi:10.1137/100785260, arXiv:0910.1862] 2, 3, 26

[42] ALEXANDER A. SHERSTOV: Breaking the Minsky-Papert barrier for constant-depth circuits. In *Proc. 46th STOC*, pp. 223–232. ACM Press, 2014. Also available at ECCC. [doi:10.1145/2591796.2591871] 3, 4, 5

[43] ALEXANDER A. SHERSTOV: The power of asymmetry in constant-depth circuits. In *Proc. 56th FOCS*, pp. 431–450. IEEE Comp. Soc. Press, 2015. Also available at ECCC. [doi:10.1109/FOCS.2015.34] 3, 4, 5

[44] YAOYUN SHI: Quantum lower bounds for the collision and the element distinctness problems. In *Proc. 43rd FOCS*, pp. 513–519. IEEE Comp. Soc. Press, 2002. [doi:10.1109/SFCS.2002.1181975, arXiv:quant-ph/0112086] 2, 3

[45] YAOYUN SHI AND YUFAN ZHU: Quantum communication complexity of block-composed functions. *Quantum Inf. Comput.*, 9(5):444–460, 2009. ACM DL. [arXiv:0710.0095] 2

[46] ROBERT ŠPALEK: A dual polynomial for OR, 2008. [arXiv:0803.4516] 3, 9, 14

[47] ROBERT ŠPALEK AND MARIO SZEGEDY: All quantum adversary methods are equivalent. *Theory of Computing*, 2(1):1–18, 2006. Preliminary version in ICALP'05. [doi:10.4086/toc.2006.v002a001] 5

[48] JUSTIN THALER: Lower bounds for the approximate degree of block-composed functions. *Electron. Colloq. on Comput. Complexity (ECCC)*, 21(150), 2014. ECCC. 3, 26

[49] JUSTIN THALER, JONATHAN ULLMAN, AND SALIL P. VADHAN: Faster algorithms for privately releasing marginals. In *Proc. 39th Internat. Colloq. on Automata, Languages and Programming (ICALP'12)*, volume 7391 of *LNCS*, pp. 810–821. Springer, 2012. [doi:10.1007/978-3-642-31594-7_68, arXiv:1205.1758] 2

[50] HENRY YUEN: A quantum lower bound for distinguishing random functions from random permutations. *Quantum Inf. Comput.*, 14(13-14):1089–1097, 2014. ACM DL. [arXiv:1310.2885] 5

[51] MARK ZHANDRY: A note on the quantum collision and set equality problems. *Quantum Inf. Comput.*, 15(7-8):557–567, 2015. ACM DL. [arXiv:1312.1027] 5

[52] SHENGYU ZHANG: On the power of Ambainis lower bounds. *Theoret. Comput. Sci.*, 339(2-3):241–256, 2005. Preliminary version in ICALP'04. [doi:10.1016/j.tcs.2005.01.019, arXiv:quant-ph/0311060] 5

## AUTHORS

Mark Bun
Ph. D. Candidate
Harvard University, Cambridge, MA
mbun@seas.harvard.edu
http://people.seas.harvard.edu/~mbun/

Justin Thaler
Research Scientist
Yahoo Research, New York, NY
jthaler@fas.harvard.edu
http://people.seas.harvard.edu/~jthaler/

## ABOUT THE AUTHORS

MARK BUN is a Ph. D. student in the Theory of Computation group at Harvard University, where he is advised by Salil Vadhan. His research interests include computational complexity, differential privacy, cryptozoology, and learning theory. As a native Seattleite, he enjoys coffee, rain, composting, competitive paper airplane throwing, extreme ironing, and wingsuit base jumping.

JUSTIN THALER is a Research Scientist at Yahoo Research in New York City. Previously, Justin spent a year as a Research Fellow at the Simons Institute for the Theory of Computing. He received his Ph. D. in Computer Science from Harvard University in 2013 under the supervision of Michael Mitzenmacher. His research interests include communication complexity, computational learning theory, algorithms for massive datasets, and verifiable computation. He enjoys running and binge watching Veronica Mars (not necessarily at the same time), and his favorite norm is $L_\infty$.