The Polynomial Method Strikes Back: Tight Quantum Query Bounds via Dual Polynomials*

Mark Bun
Princeton University
Princeton, NJ, USA
mbun@cs.princeton.edu

Robin Kothari
Microsoft Research
Redmond, WA, USA
robin.kothari@microsoft.com

Justin Thaler
Georgetown University
Washington, DC, USA
justin.thaler@georgetown.edu

ABSTRACT

The approximate degree of a Boolean function f is the least degree of a real polynomial that approximates f pointwise to error at most $1/3$. The approximate degree of f is known to be a lower bound on the quantum query complexity of f (Beals et al., FOCS 1998 and J. ACM 2001).

We resolve or nearly resolve the approximate degree and quantum query complexities of several basic functions. Specifically, we show the following:

k-distinctness: For any constant k , the approximate degree and quantum query complexity of *k*-distinctness is $\Omega(n^{3/4-1/(2k)})$. This is nearly tight for large k , as Belovs (FOCS 2012) has shown that for any constant k , the approximate degree and quantum query complexity of *k*-distinctness is $O(n^{3/4-1/(2^{k+2}-4)})$.

Image Size Testing: The approximate degree and quantum query complexity of testing the size of the image of a function $[n] \rightarrow [n]$ is $\tilde{\Omega}(n^{1/2})$. This proves a conjecture of Ambainis et al. (SODA 2016), and it implies tight lower bounds on the approximate degree and quantum query complexity of the following natural problems. *k*-junta testing: A tight $\tilde{\Omega}(k^{1/2})$ lower bound for *k*-junta testing, answering the main open question of Ambainis et al. (SODA 2016). Statistical Distance from Uniform: A tight $\tilde{\Omega}(n^{1/2})$ lower bound for approximating the statistical distance of a distribution from the uniform distribution, answering the main question left open by Bravyi et al. (STACS 2010 and IEEE Trans. Inf. Theory 2011). Shannon entropy: A tight $\tilde{\Omega}(n^{1/2})$ lower bound for approximating Shannon entropy up to a certain additive constant, answering a question of Li and Wu (2017).

Surjectivity: The approximate degree of the Surjectivity function is $\tilde{\Omega}(n^{3/4})$. The best prior lower bound was $\Omega(n^{2/3})$. Our result matches an upper bound of $\tilde{O}(n^{3/4})$ due to Sherstov (STOC 2018), which we reprove using different techniques. The quantum query complexity of this function is known to be $\Theta(n)$ (Beame and Mouchi, Quantum Inf. Comput. 2012 and Sherstov, FOCS 2015).

Our upper bound for Surjectivity introduces new techniques for approximating Boolean functions by low-degree polynomials. Our lower bounds are proved by significantly refining techniques recently introduced by Bun and Thaler (FOCS 2017).

CCS CONCEPTS

• Theory of computation $\rightarrow$ Oracles and decision trees;

KEYWORDS

Polynomial method, quantum query complexity, approximate degree, dual polynomials

ACM Reference Format:

Mark Bun, Robin Kothari, and Justin Thaler. 2018. The Polynomial Method Strikes Back: Tight Quantum Query Bounds via Dual Polynomials. In *Proceedings of 50th Annual ACM SIGACT Symposium on the Theory of Computing (STOC'18)*. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3188745.3188784>

1 INTRODUCTION

Approximate degree. The approximate degree of a Boolean function $f: \{-1, 1\}^n \rightarrow \{-1, 1\}$, denoted $\widetilde{\deg}(f)$, is the least degree of a real polynomial p such that $|p(x) - f(x)| \leq 1/3$ for all $x \in \{-1, 1\}^n$. Approximate degree is a basic measure of the complexity of a Boolean function, and has diverse applications throughout theoretical computer science.

Upper bounds on approximate degree are at the heart of the most powerful known learning algorithms in a number of models [9, 40, 41, 43, 44, 56, 62], algorithmic approximations for the inclusion-exclusion principle [39, 64], and algorithms for differentially private data release [30, 81]. A recent line of work [78, 79] has used approximate degree upper bounds to show new lower bounds on the formula and graph complexity of explicit functions.

Lower bounds on approximate degree have enabled progress in several areas of complexity theory, including communication complexity [24, 31–33, 36, 57, 63, 66, 67, 69], circuit complexity [53, 65], oracle separations [14, 20], and secret-sharing [19]. Most importantly for this paper, approximate degree lower bounds have been critical in shaping our understanding of *quantum query complexity* [1, 3, 12].

In spite of the importance of approximate degree, major gaps remain in our understanding. In particular, the approximate degrees of many basic functions are still unknown. Our goal in this paper is to resolve the approximate degrees of many natural functions which had previously withstood characterization.

Quantum query complexity. While resolving the approximate degree of basic functions of interest is a test of our understanding

*The full version of this paper is available at <https://arxiv.org/abs/1710.09079>.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

STOC'18, June 25–29, 2018, Los Angeles, CA, USA

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.

ACM ISBN 978-1-4503-5559-9/18/06...\$15.00

<https://doi.org/10.1145/3188745.3188784>

Table 1: Our lower bounds on quantum query complexity and approximate degree vs. prior work.

Problem	Best Prior Upper Bound	Our Lower Bound	Best Prior Lower Bound
k -distinctness	$O(n^{3/4-1/(2^{k+2}-4)})$ [15]	$\tilde{\Omega}(n^{3/4-1/(2k)})$	$\tilde{\Omega}(n^{2/3})$ [3]
Image Size Testing	$O(\sqrt{n} \log n)$ [8]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [8]
k -junta Testing	$O(\sqrt{k} \log k)$ [8]	$\tilde{\Omega}(\sqrt{k})$	$\tilde{\Omega}(k^{1/3})$ [8]
SDU	$O(\sqrt{n})$ [21]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [3, 21]
Shannon Entropy	$\tilde{O}(\sqrt{n})$ [21, 51]	$\tilde{\Omega}(\sqrt{n})$	$\tilde{\Omega}(n^{1/3})$ [51]

of approximate degree, it is also motivated by the study of quantum algorithms. In the quantum query model, a quantum algorithm is given query access to the bits of an input x , and the goal is to compute some function f of x while minimizing the number of queried bits. Quantum query complexity captures much of the power of quantum computing, and most quantum algorithms were discovered in or can easily be described in the query setting.

Approximate degree was one of the first general lower bound techniques for quantum query complexity. In 1998, Beals et al. [12] observed that the bounded-error quantum query complexity of a function f is lower bounded by (one half times) the approximate degree of f . Since polynomials are sometimes easier to understand than quantum algorithms, this observation led to a number of new lower bounds on quantum query complexity. This method of proving quantum query lower bounds is called the *polynomial method*.

After several significant quantum query lower bounds were proved via the polynomial method (including the work of Aaronson and Shi [3], who proved optimal lower bounds for the Collision and Element Distinctness problems), the polynomial method took a back seat. Since then, the positive-weights adversary method [5, 11, 46, 85] and the newer negative-weights adversary method [38, 48, 59] have become the tools of choice for proving quantum query lower bounds (with some notable exceptions, such as Zhandry's recent tight lower bound for the set equality problem [84]). This leads us to our second goal for this work.

In this work, we seek to resolve several open problems in quantum query complexity using approximate degree as the lower bound technique. A distinct advantage of proving quantum query lower bounds with the polynomial method is that any such bound can be “lifted” via Sherstov’s *pattern matrix method* [66] to a quantum *communication* lower bound (even with unlimited shared entanglement [49]); such a result is not known for any other quantum query lower bound technique. More generally, using approximate degree as a lower bound technique for quantum query complexity has other advantages, such as the ability to show lower bounds for zero-error and small-error quantum algorithms [22], unbounded-error quantum algorithms [12], and time-space tradeoffs [42].

Quantum query complexity and approximate degree. In this work we illustrate the power of the polynomial method by proving optimal or nearly optimal bounds on several functions studied in the quantum computing community. These results are summarized in Table 1, and definitions of the problems considered can be found in Section 1.1. Since the upper bounds for these functions were shown

using quantum algorithms, our results resolve both the quantum query complexity and approximate degree of these functions.

For most of the functions studied in this paper, the positive-weights adversary bound provably cannot show optimal lower bounds due to the certificate complexity barrier [77, 85] and the property testing barrier [38]. While these barriers do not apply to the negative-weights variant (which is actually capable for proving tight quantum query lower bounds for *all* functions [48, 59]), the negative-weights adversary method is often challenging to apply to specific problems, and the problems we consider have withstood characterization for a long time.

For the functions presented in Table 1, the approximate degree and quantum query complexity are essentially the same. This is not the case for the Surjectivity function, which has played an important role in the literature on approximate degree and quantum query complexity. Specifically, Beame and Machmouchi [13] showed that Surjectivity has quantum query complexity $\tilde{\Theta}(n)$. On the other hand, Sherstov recently showed that Surjectivity has approximate degree $\tilde{O}(n^{3/4})$ [74]. Surjectivity is the only known example of a “natural” function separating approximate degree from quantum query complexity; prior examples of such functions [2, 6] were contrived, and (unlike Surjectivity) specifically constructed to separate the two measures.

Our final result gives a full characterization of the approximate degree of Surjectivity. We prove a new lower bound of $\tilde{\Omega}(n^{3/4})$, which matches Sherstov’s upper bound up to logarithmic factors. We also give a new construction of an approximating polynomial of degree $\tilde{O}(n^{3/4})$, using very different techniques than [74]. We believe that our proof of this $\tilde{O}(n^{3/4})$ upper bound is of independent interest. In particular, our lower bound proof for Surjectivity is specifically tailored to showing optimality of our upper bound construction, in a sense that can be made formal via complementary slackness. We are optimistic that our approximation techniques will be useful for showing additional tight approximate degree bounds in the future.

1.1 Our Results

We now describe our results and prior work on these functions in more detail.

1.1.1 Functions Considered. We now informally describe the functions studied in this paper. These functions are formally defined in Section 2.4.

Let R be a power of two and $N \geq R$, and let $n = N \cdot \log_2 R$. Most of the functions that we consider interpret their inputs in $\{-1, 1\}^n$

Table 2: Our bounds on the approximate degree of Surjectivity vs. prior work.

Problem	Best Prior Upper Bound	Our Upper Bound	Our Lower Bound	Best Prior Lower Bound
Surjectivity	$\tilde{O}(n^{3/4})$ [74]	$\tilde{O}(n^{3/4})$	$\tilde{\Omega}(n^{3/4})$	$\tilde{\Omega}(n^{2/3})$ [3]

as a list of N numbers from a range $[R]$, and determine whether this list satisfies various natural properties. We let the *frequency* f_i of range item $i \in R$ denote the number of times i appears in the input list.

In this paper we study the following functions in which the input is N numbers from a range $[R]$:

- **Surjectivity (SURJ):** Do all range items appear at least once?
- **k -distinctness:** Is there a range item that appears k or more times?
- **Image Size Testing:** Decide if all range items appear at least once or if at most $\gamma \cdot R$ range items appear at least once, under the promise that one of these is true.
- **Statistical distance from uniform (SDU):** Interpret the input as a probability distribution p , where $p_i = f_i/N$. Compute the statistical distance of p from the uniform distribution over R up to some small additive error ϵ .
- **Shannon entropy:** Interpret the input as a probability distribution p , where $p_i = f_i/N$. Compute the Shannon entropy $\sum_{i \in R} p_i \cdot \log(1/p_i)$ of p up to additive error ϵ .

An additional function we consider that does not fit neatly into the framework above is k -junta testing.

- **k -junta testing:** Given an input in $\{-1, 1\}^n$ representing the truth table of a function $\{-1, 1\}^{\log n} \rightarrow \{-1, 1\}$, determine whether this function depends on at most k of its input bits, or is at least ϵ -far from any such function.

We resolve or nearly resolve the quantum query complexity and/or approximate degree of all of the functions above. Our lower bounds for SURJ, k -distinctness, Image Size Testing, SDU, and entropy approximation all require N to be “sufficiently larger” than R , by a certain constant factor. For simplicity, throughout this introduction we do not make this requirement explicit, and for this reason we label the theorems in this introduction informal.

1.1.2 Results in Detail.

Surjectivity. In the Surjectivity problem we are given N numbers from $[R]$ and must decide if every range item appears at least once in the input.

The quantum query complexity of this problem was studied by Beame and Machmouchi [13], who proved a lower bound of $\tilde{\Omega}(n)$, which was later improved by Sherstov to the optimal $\Theta(n)$ [73]. Beame and Machmouchi [13] explicitly leave open the question of characterizing the approximate degree of Surjectivity. Recently, Sherstov [74] showed an upper bound of $\tilde{O}(n^{3/4})$ on the approximate degree of this function. The best prior lower bound was $\tilde{\Omega}(n^{2/3})$ [3, 29].

We give a completely different construction of an approximating polynomial for Surjectivity with degree $\tilde{O}(n^{3/4})$. We also prove a matching lower bound, which shows that the approximate degree of the Surjectivity function is $\tilde{\Theta}(n^{3/4})$.

THEOREM 1 (INFORMAL). *The approximate degree of SURJ is $\tilde{\Theta}(n^{3/4})$.*

k -distinctness. In this problem, we are given N numbers in $[R]$ and must decide if any range item appears at least k times in the list (i.e., is there an $i \in [R]$ with $f_i \geq k$?). This generalizes the well-studied Element Distinctness problem, which is the same as 2-distinctness.

Ambainis [7] first used quantum walks to give an $O(n^{k/(k+1)})$ upper bound on the quantum query complexity of any problem with certificates of size k , including k -distinctness and k -sum.¹ Later, Belovs introduced a beautiful new framework for designing quantum algorithms [16] and used it to improve the upper bound for k -distinctness to $O(n^{3/4-1/(2^{k+2}-4)})$ [15]. Several subsequent works have used Belovs’ k -distinctness algorithm as a black-box subroutine for solving more complicated problems (e.g., [51, 54]).

As for lower bounds, Aaronson and Shi [3] established an $\tilde{\Omega}(n^{2/3})$ lower bound on the approximate degree of k -distinctness for any $k \geq 2$. Belovs and Špalek used the adversary method to prove a lower bound of $\Omega(n^{k/(k+1)})$ on the quantum query complexity of k -sum, showing that Ambainis’ algorithm is tight for k -sum. They asked whether their techniques can prove an $\omega(n^{2/3})$ quantum query lower bound for k -distinctness. We achieve this goal, but using the polynomial method instead of the adversary method. Our main result is the following:

THEOREM 2 (INFORMAL). *For any $k \geq 2$, the approximate degree and quantum query complexity of k -distinctness is $\tilde{\Omega}(n^{3/4-1/(2^k)})$.*

This is nearly tight for large k , as it approaches Belovs’ upper bound of $O(n^{3/4-1/(2^{k+2}-4)})$. Note that both bounds approach $\Theta(n^{3/4})$ as $k \rightarrow \infty$. It remains an intriguing open question to close the gap between $n^{3/4-1/(2^{k+2}-4)}$ and $n^{3/4-1/(2^k)}$, especially for small values of $k \geq 3$.

Our k -distinctness lower bound also implies an $\tilde{\Omega}(n^{3/4-1/(2k)})$ lower bound on the quantum query complexity of approximating the maximum frequency, F_∞ , of any element up to relative error less than $1/k$ [54], improving over the previous best bound of $\tilde{\Omega}(n^{2/3})$.

Image Size Testing. In this problem, we are given N numbers in $[R]$ and $\gamma > 0$, and must decide if every range item appears at least once or if at most $\gamma \cdot R$ range items appear at least once. We show for any $\gamma > 0$, the problem has approximate degree and quantum query complexity $\tilde{\Omega}(\sqrt{n})$. This holds as long as $N = c \cdot R$ for a certain constant $c > 0$.

THEOREM 3 (INFORMAL). *The approximate degree and quantum query complexity of Image Size Testing is $\tilde{\Omega}(\sqrt{n})$.*

This lower bound is tight, matching a quantum algorithm of Ambainis, Belovs, Regev, and de Wolf [8], and resolves a conjecture from their work. The previous best lower bound was $\Omega(n^{1/3})$.

¹In the k -sum problem, we are given N numbers in $[R]$ and asked to decide if any k of them sum to 0 (mod R).

[8] obtained via reduction to the Collision lower bound [3]. The classical query complexity of this problem is $\Theta(n/\log n)$ [83].

The version of image size testing we define is actually a special case of the one studied in [8]. The version we define is solvable via the following simple algorithm making $O(\sqrt{n})$ queries: pick a random range item, and Grover search for an instance of that range item. The fact that our lower bound holds even for this special case of the problem considered in prior works obviously only makes our lower bound stronger.

This lower bound also serves as a starting point to establish the next three lower bounds.

k-junta Testing. In this problem, we are given the truth table of a Boolean function and have to determine if the function depends on at most k variables or if it is ϵ -far from any such function.

The best classical algorithm for this problem uses $O(k \log k + k/\epsilon)$ queries [18]. The problem was first studied in the quantum setting by Atıcı and Servedio [10], who gave a quantum algorithm making $O(k/\epsilon)$ queries. This was later improved by Ambainis et al. [8] to $\tilde{O}(\sqrt{k}/\epsilon)$. They also proved a lower bound of $\Omega(k^{1/3})$. Via a connection established by Ambainis et al., our image size testing lower bound implies a $\tilde{\Omega}(\sqrt{k})$ lower bound on the approximate degree and quantum query complexity of k -junta testing (for some $\epsilon = \Omega(1)$).

THEOREM 4 (INFORMAL). *The approximate degree and quantum query complexity of k -junta testing is $\tilde{\Omega}(\sqrt{k})$.*

This matches the upper bound of [8], resolving the main open question from their work.

Statistical Distance From Uniform (SDU). In this problem, we are given N numbers in $[R]$, which we interpret as a probability distribution p , where $p_i = f_i/N$, the fraction of times i appears. The goal is to compute the statistical distance between p and the uniform distribution to error ϵ .

This problem was studied by Bravyi, Harrow, and Hassidim [21], who gave an $O(\sqrt{n})$ -query quantum algorithm approximating the statistical distance between two input distributions to additive error $\epsilon = \Omega(1)$. We show that the approximate degree and quantum query complexity of this task are $\tilde{\Omega}(\sqrt{n})$, even when one of the distributions is known to be the uniform distribution.

THEOREM 5 (INFORMAL). *There is a constant $c > 0$ such that the approximate degree and quantum query complexity of approximating the statistical distribution of a distribution over a range of size n from the uniform distribution over the same range to additive error c is $\tilde{\Omega}(\sqrt{n})$.*

This matches the upper bound of Bravyi et al. [21] and answers the main question left open from that work. Note that the classical query complexity of this problem is $\Theta(n/\log n)$ [83].

Entropy Approximation. As in the previous problem, we interpret the input as a probability distribution, and the goal is to compute its Shannon entropy to additive error ϵ . The classical query complexity of this problem is $\Theta(n/\log n)$ [83]. We show that, for some $\epsilon = \Omega(1)$, the approximate degree and quantum query complexity are $\tilde{\Omega}(\sqrt{n})$.

THEOREM 6 (INFORMAL). *There is a constant $c > 0$ such that the approximate degree and quantum query complexity of approximating*

the Shannon entropy of a distribution over a range of size n to additive error c is $\tilde{\Omega}(\sqrt{n})$.

This too is tight, answering a question of Li and Wu [51].

1.2 Prior Work on Lower Bounding Approximate Degree

A relatively new lower-bound technique for approximate degree called the *method of dual polynomials* plays an essential role in our paper. This method of dual polynomials dates back to work of Sherstov [70] and Špalek [76], though dual polynomials had been used earlier to resolve longstanding questions in communication complexity [31, 50, 65, 66, 75]. To prove a lower bound for a function f via this method, one exhibits an explicit *dual polynomial* for f , which is a dual solution to a certain linear program capturing the approximate degree of f .

A notable feature of the method of dual polynomials is that it is lossless, in the sense that it can exhibit a tight lower bound on the approximate degree of any function f (though actually applying the method to specific functions may be highly challenging). Prior to the method of dual polynomials, the primary tool available for proving approximate degree lower bounds was symmetrization, introduced by Minsky and Papert [53] in the 1960s. Although powerful, symmetrization is not a lossless technique.

Most prior work on the method of dual polynomials can be understood as establishing hardness amplification results. Such results show how to take a function f that is “somewhat hard” to approximate by low-degree polynomials, and turn f into a related function g that is much harder to approximate. Here, harder means either that g requires larger degree to approximate to the same error as f , or that approximations to g of a given degree incur much larger error than do approximations to f of the same degree.

Results for Block-Composed Functions. Until very recently, the method of dual polynomials had been used exclusively to prove hardness amplification results for *block-composed* functions. That is, the harder function g would be obtained by block-composing f with another function h , i.e., $g = h \circ f$. Here, a function $g: \{-1, 1\}^{n \cdot m} \rightarrow \{-1, 1\}$ is the block-composition of $h: \{-1, 1\}^n \rightarrow \{-1, 1\}$ and $f: \{-1, 1\}^m \rightarrow \{-1, 1\}$ if g interprets its input as a sequence of n blocks, applies f to each block, and then feeds the n outputs into h .

The method of dual polynomials turns out to be particularly suited to analyzing block-composed functions, as there are sophisticated ways of “combining” dual witnesses for h and f individually to give an effective dual witness for $h \circ f$ [20, 25, 26, 68, 70, 72, 73, 75, 80]. Prior work on analyzing block-composed functions has, for example, resolved the approximate degree of the function $f(x) = \bigvee_{i=1}^n \bigwedge_{j=1}^m x_{ij}$, known as the AND-OR tree, which had been open for 19 years [25, 68], established new lower bounds for AC^0 under basic complexity measures including discrepancy [26, 72, 73, 80], sign-rank [28], and threshold degree [72, 73], and resolved a number of open questions about the power of statistical zero knowledge proofs [20].

Beyond Block-Composed Functions. While the aforementioned results led to considerable progress in complexity theory, many basic questions require understanding the approximate degree of *non-block-composed* functions. One prominent example with many

applications is to exhibit an AC^0 circuit over n variables with approximate degree $\Omega(n)$. Until very recently, the best result in this direction was Aaronson and Shi's well-known $\tilde{\Omega}(n^{2/3})$ lower bound on the approximate degree of the Element Distinctness function (which is equivalent to k -distinctness for $k = 2$) [3]. However, Bun and Thaler [29] recently achieved a near-resolution of this problem by proving the following theorem.

THEOREM 7 (BUN AND THALER [29]). *For any constant $\delta > 0$, there is an AC^0 circuit with approximate degree $\Omega(n^{1-\delta})$.*

The reason that Theorem 7 required moving beyond block-composed functions is the following result of Sherstov [71].

THEOREM 8 (SHERSTOV). *For any Boolean functions f and h , $\widetilde{\deg}(h \circ f) = O(\widetilde{\deg}(h) \cdot \widetilde{\deg}(f))$.*

Theorem 8 implies that the approximate degree of $h \circ f$ (viewed as a function of its input size) is never higher than the approximate degree of f or h individually (viewed as a function of their input sizes). For example, if f and h are both functions on n inputs, and both have approximate degree $O(n^{1/2})$, then $h \circ f$ has $N := n^2$ inputs, and by Theorem 8, $\widetilde{\deg}(h \circ f) = O(n^{1/2} \cdot n^{1/2}) = O(N^{1/2})$.

This means that block-composing multiple AC^0 functions does not result in a function of higher approximate degree (as a function of its input size) than that of the individual functions. Bun and Thaler [29] overcome this hurdle by introducing a way of analyzing functions that cannot be written as a block-composition of simpler functions.

Bun and Thaler's techniques set the stage to resolve the approximate degree of many basic functions using the method of dual polynomials. However, they were not refined enough to accomplish this on their own. Our lower bounds in this paper are obtained by refining and extending the methods of [29].

1.3 Our Techniques

In order to describe our techniques, it is helpful to explain the process by which we discovered the tight $\tilde{\Theta}(n^{3/4})$ lower and upper bounds for Surjectivity (cf. Theorem 1). It has previously been observed [25, 29, 80] that optimal dual polynomials for a function f tend to be tailored (in a sense that can be made precise via complementary slackness) to showing optimality of some specific approximation technique for f . Hence, constructing a dual polynomial for f can provide a strong hint as to how to construct an optimal approximation for f , and vice versa.

Upper Bound for Surjectivity. In [29], Bun and Thaler constructed a dual polynomial witnessing a suboptimal bound of $\tilde{\Omega}(n^{2/3})$ for SURJ. Even though this dual polynomial is suboptimal, it still provided a major clue as to what an optimal approximation for SURJ should look like: it curiously ignored all inputs failing to satisfy the following condition.

Condition 1. *Every range item has frequency at most T , for a specific threshold $T = O(N^{1/3}) \ll N$.*

This suggested that an optimal approximation for SURJ should treat inputs satisfying Condition 1 differently than other inputs, leading us to the following multi-phase construction (for clarity

and brevity, this overview is simplified). The first phase constructs a polynomial p of degree $O(n^{3/4})$ approximating SURJ on all inputs satisfying Condition 1. However, p may be *exponentially large* on other inputs. The second phase constructs a polynomial q of degree $O(n^{3/4})$ that is *exponentially small* on inputs x that do *not* satisfy Condition 1 (in particular, $q(x) \ll 1/p(x)$ for such x), and is close to 1 otherwise. The product $p \cdot q$ still approximates SURJ on inputs satisfying Condition 1, and is *exponentially small* on all other inputs. Notice that $\deg(p \cdot q) \leq \deg(p) + \deg(q) = O(n^{3/4})$. Combining the above with an additional averaging step (the details of which we omit from this introduction) yields an approximation to SURJ that is accurate on *all* inputs.

Lower Bound for Surjectivity. With the $O(n^{3/4})$ upper bound in hand, we were able to identify the fundamental bottleneck preventing further improvement of the upper bound. This suggested a way to refine the techniques of [29] to prove a matching lower bound. Once the tight lower bound for SURJ was established, we were able to identify additional refinements to analyze the other functions that we consider. We now describe this in more detail.

Bun and Thaler's [29] (suboptimal) lower bound analysis for SURJ proceeds in two stages. In the first stage, proving a lower bound for SURJ (on N input list items and R range items) is reduced to the problem of proving a lower bound for the *block-composed* function $AND_R \circ OR_N$,² under the promise that the input has Hamming weight at most N .³ In this paper, we use this stage of their analysis unmodified.

The second stage proves an $\tilde{\Omega}(R^{2/3})$ lower bound for the latter problem by leveraging much of the machinery developed to analyze the approximate degree of block-composed functions [25, 58, 68]. To describe this machinery, we require the following notion. A dual polynomial that witnesses the fact that $\widetilde{\deg}_\varepsilon(f_n) \geq d$ is a function $\psi: \{-1, 1\}^n \rightarrow \{-1, 1\}$ satisfying three properties:

- $\sum_{x \in \{-1, 1\}^n} \psi(x) \cdot f(x) > \varepsilon$. If ψ satisfies this condition, it is said to be *well-correlated* with f .
- $\sum_{x \in \{-1, 1\}^n} |\psi(x)| = 1$. If ψ satisfies this condition, it is said to have ℓ_1 -norm equal to 1.
- For all polynomials $p: \{-1, 1\}^n \rightarrow \mathbb{R}$ of degree less than d , we have $\sum_{x \in \{-1, 1\}^n} p(x) \cdot \psi(x) = 0$. If ψ satisfies this condition, it is said to have *pure high degree* at least d .

In more detail, the second stage of the analysis from [29] itself proceeds in two steps. First, the authors consider a dual witness ψ for the high approximate degree of $AND_R \circ OR_N$ that was constructed in prior work [25]. ψ is constructed by taking dual witnesses ϕ and γ for the high approximate degrees of AND_R and OR_N individually, and "combining" them in a specific way [47, 70, 75] to obtain a dual witness for the high approximate degree of their block-composition $AND_R \circ OR_N$.

²When it is not clear from context, we use subscripts to denote the number of variables on which a function is defined.

³Note that a reduction the other direction is straightforward: to approximate SURJ, it suffices to approximate $AND_R \circ OR_N$ on inputs of Hamming weight exactly N . This is because SURJ can be expressed as an AND_R (over all range items $r \in [R]$) of the OR_N (over all input bits $i \in [N]$) of "Is input x_i equal to r ?" Each predicate of the form in quotes is computed exactly by a polynomial of degree $\log R$, since it depends on only $\log R$ of the inputs, and exactly N of these predicates (one for each $i \in [N]$) evaluate to TRUE.

Unfortunately, ψ only witnesses a lower bound for $\text{AND}_R \circ \text{OR}_N$ *without* the promise that the Hamming weight of the input is at most N . To address this issue, it is enough to “post-process” ψ so that it no longer “exploits” any inputs of Hamming weight larger than N (formally, $\psi(x)$ should equal zero for any inputs in $\{-1, 1\}^{R \cdot N}$ of Hamming weight more than N). The authors accomplish this by observing that ψ “almost ignores” all such inputs (i.e., it places exponentially little total mass on all such inputs), and hence it is possible to perturb ψ to make it *completely* ignore all such inputs.

Key to this step is the fact that the “inner” dual witness γ for the high approximate degree of the OR_N function satisfies a “Hamming weight decay” condition:

$$|\gamma(x)| \leq 1/\text{poly}(|x|), \quad (1)$$

for a suitable polynomial function.

To improve the lower bound for SURJ from $\tilde{\Omega}(n^{2/3})$ to the optimal $\tilde{\Omega}(n^{3/4})$, we observe that γ in fact satisfies a much stronger decay condition: while the inverse-polynomial decay property of Equation (1) is tight for small Hamming weights $|x|$, $|\gamma(x)|$ actually decays *exponentially quickly* once $|x|$ is larger than a certain threshold t . This observation is enough to obtain the tight $\tilde{\Omega}(n^{3/4})$ lower bound for SURJ.

For intuition, it is worth mentioning that a primal formulation of the dual decay condition that we exploit shows that any low-degree polynomial p that is an accurate approximation to OR_N on low Hamming weight inputs requires large degree, *even if* $|p(x)|$ is allowed to be exponentially large for inputs of Hamming weight more than t .⁴ This is precisely the bottleneck that prevents us from improving our upper bound for SURJ to $\tilde{o}(N^{3/4})$. In this sense, our dual witness is intuitively tailored to showing optimality of the techniques used in our upper bound.

Other Lower Bounds. To obtain the lower bound for k -distinctness, the first stage of the analysis of [29] reduces to a question about the approximate degree of the block composed function $\text{OR}_R \circ \text{THR}_N^k$, under the promise that the input has Hamming weight at most N . Here $\text{THR}_N^k: \{-1, 1\}^N \rightarrow \{-1, 1\}$ denotes the function that evaluates to -1 if and only if the Hamming weight of its input is at least k . By constructing a suitable dual witness for THR_N^k , and combining it with a dual witness for OR_N via similar techniques as in our construction for SURJ, we are able to prove our $\Omega(n^{3/4-1/(2k)})$ lower bound for k -distinctness. (This description glosses over several significant technical issues that must be dealt with to ensure that the combined dual witness is well-correlated with $\text{OR}_R \circ \text{THR}_N^k$.⁵)

Recall that our lower bounds for k -junta testing, SDU, and entropy approximation are derived as consequences of our lower bound for image size testing. The connection between image size

⁴We do not formally describe this primal formulation of the dual decay condition, because it is not necessary to prove any of the results in this paper.

⁵Specifically, our analysis requires the dual witness γ for THR_N^k to be very well-correlated with THR_N^k in a certain one-sided sense (roughly, we need the probability distribution $|\gamma|$ to have the property that, conditioned on γ outputting a negative value, the input to γ is in $(\text{THR}_N^k)^{-1}(-1)$ with probability at least $1 - 1/(3R)$). This property was not required in the analysis for SURJ, which is why our lower bound for SURJ is larger by a factor of $n^{1/(2k)}$ than our lower bound for k -distinctness. This seemingly technical issue is at least partially intrinsic: a polynomial loss compared to the $\Omega(n^{3/4})$ lower bound for SURJ is unavoidable, owing to Belovs' $n^{3/4-\Omega(1)}$ upper bound [15] for k -distinctness.

testing and junta testing was established by Ambainis et al. [8]. The reason that the image testing lower bound implies lower bounds for SDU is the following. Consider any distribution p over $[R]$ such that all probabilities p_i are integer multiples of $1/N$ for some $N = O(R)$. Then if p has full support, p is guaranteed to be somewhat close to uniform, while if p has small support, p must be very far from uniform. We obtain our lower bound for entropy approximation using a simple reduction from SDU due to Vadhan [82].

To obtain our lower bound for Image Size Testing, we observe that the first stage of the analysis of [29] reduces to a question about the approximate degree of the block composed function $\text{GapAND}_R \circ \text{OR}_N$, under the promise that the input has Hamming weight at most N . Here, GapAND_R is the promise function that outputs -1 if all of its inputs equal -1 , outputs $+1$ if fewer than $\gamma \cdot R$ of its inputs are -1 , and is undefined otherwise.

Roughly speaking, we obtain the desired $\tilde{\Omega}(n^{1/2})$ lower bound by combining a dual witness for $\text{GapAND}_R \circ \text{OR}_N$ from prior work [26] with the same techniques as in our construction for SURJ. However, additional technical refinements to the analysis of [29] are required to obtain our results. In particular, the analysis of [29] only provides a lower bound for SURJ if $N = \Omega(R \cdot \log^2(R))$. But in order to infer our lower bound for SDU and entropy approximation (as well as k -junta testing for $\varepsilon = \Omega(1)$), it is essential that the lower bound hold for $N = O(R)$. This is because a distribution with full support is guaranteed to be $\Omega(1)$ -close to uniform if all probabilities are integer multiples of $1/N$ with $N = O(R)$, but this is not the case otherwise. (Consider, e.g., a distribution that places mass $1 - 1/\log^2(R)$ on a single range item, and spreads out the remaining mass evenly over all other range items). Refining the methods of [29] to yield lower bounds even when $N = O(R)$ requires a significantly more delicate analysis than in [29].

1.4 Organization of This Extended Abstract

Due to space constraints, this extended abstract only contains an expanded discussion of our lower bound on the approximate degree of Surjectivity. Proofs of other results may be found in the full version of this paper.

2 PRELIMINARIES

2.1 Notation

For a natural number N , we let $[N] = \{1, 2, \dots, N\}$ and $[N]_0 = \{0, 1, 2, \dots, N\}$. All logarithms are taken in base 2 unless otherwise noted.

We will frequently work with Boolean functions under the promise that their inputs have low Hamming weight. To this end, we introduce the following notation for the set of low-Hamming weight inputs.

Definition 9. For $1 \leq T \leq n$, let $H_{\leq T}^n$ denote the subset of $\{-1, 1\}^n$ consisting of all inputs Hamming weight at most T . We use $|x|$ to denote the Hamming weight of an input $x \in \{-1, 1\}^n$, so $H_{\leq T}^n = \{x \in \{-1, 1\}^n : |x| \leq T\}$.

2.2 Two Variants of Approximate Degree and Their Dual Formulations

There are two natural notions of approximate degree for promise problems (i.e., for functions defined on a strict subset $\mathcal{X}$ of $\{-1, 1\}^n$). One notion requires an approximating polynomial p to be bounded in absolute value even on inputs in $\{-1, 1\}^n \setminus \mathcal{X}$. The other places no restrictions on p outside of the promise $\mathcal{X}$. In this work, we make use of *both* notions. Hence, we must introduce some (non-standard) notation to distinguish the two.

Definition 10 (Approximate Degree With Boundedness Outside of the Promise Required). *Let $\varepsilon > 0$ and $\mathcal{X} \subseteq \{-1, 1\}^n$. The ε -approximate degree of a Boolean function $f : \mathcal{X} \rightarrow \{-1, 1\}$, denoted $\deg_\varepsilon(f)$, is the least degree of a real polynomial $p : \mathcal{X} \rightarrow \mathbb{R}$ such that $|p(x) - f(x)| \leq \varepsilon$ for all $x \in \mathcal{X}$ and $|p(x)| \leq 1 + \varepsilon$ for all $x \in \{-1, 1\}^n \setminus \mathcal{X}$. We use the term approximate degree without qualification to refer to $\deg(f) = \deg_{1/3}(f)$.*

The following standard dual formulation of this first variant of approximate degree can be found in, e.g., [27].

Proposition 11. *Let $\mathcal{X} \subseteq \{-1, 1\}^n$, and let $f : \mathcal{X} \rightarrow \{-1, 1\}$. Then $\deg_\varepsilon(f) \geq d$ if and only if there exists a function $\psi : \{-1, 1\}^n \rightarrow \mathbb{R}$ satisfying the following properties.*

$$\sum_{x \in \mathcal{X}} \psi(x) \cdot f(x) - \sum_{x \in \{-1, 1\}^n \setminus \mathcal{X}} |\psi(x)| > \varepsilon, \quad (2)$$

$$\sum_{x \in \{-1, 1\}^n} |\psi(x)| = 1, \text{ and} \quad (3)$$

For every polynomial $p : \{-1, 1\}^n \rightarrow \mathbb{R}$ of degree less than d ,

$$\sum_{x \in \{-1, 1\}^n} p(x) \cdot \psi(x) = 0. \quad (4)$$

We will refer to functions $\psi : \{-1, 1\}^n \rightarrow \mathbb{R}$ as *dual polynomials*. We refer to $\sum_{x \in \{-1, 1\}^n} |\psi(x)|$ as the ℓ_1 -norm of ψ , and denote this quantity by $\|\psi\|_1$. If ψ satisfies Equation (4), it is said to have *pure high degree* at least d .

Given a function $\psi : \{-1, 1\}^n \rightarrow \mathbb{R}$, and a (possibly partial) function $f : \mathcal{X} \rightarrow \{-1, 1\}$, where $\mathcal{X} \subseteq \{-1, 1\}^n$, we let $\langle f, \psi \rangle := \sum_{x \in \mathcal{X}} f(x) \cdot \psi(x) - \sum_{x \in \{-1, 1\}^n \setminus \mathcal{X}} |\psi(x)|$, and refer to this as the correlation of f and ψ . So Condition (2) is equivalent to requiring ψ and f to have correlation greater than ε .

Definition 12 (Approximate Degree With Unboundedness Permitted Outside of the Promise). *Let $\varepsilon > 0$ and $\mathcal{X}$ be a finite set. The ε -unbounded approximate degree of a Boolean function $f : \mathcal{X} \rightarrow \{-1, 1\}$, denoted $\text{ubdeg}_\varepsilon(f)$, is the least degree of a real polynomial $p : \mathcal{X} \rightarrow \mathbb{R}$ such that $|p(x) - f(x)| \leq \varepsilon$ for all $x \in \mathcal{X}$ (if $\mathcal{X}$ is a strict subset of a larger domain, then no constraints are placed on $p(x)$ for $x \notin \mathcal{X}$). We use the term unbounded approximate degree without qualification to refer to $\text{ubdeg}(f) = \text{ubdeg}_{1/3}(f)$.*

The following standard dual formulation of this second variant of approximate degree can be found in, e.g., [66]. A dual polynomial $\psi : \{-1, 1\}^n \rightarrow \{-1, 1\}$ witnessing the fact that $\text{ubdeg}_\varepsilon(f) \geq d$ is the same as a dual witness for $\deg_\varepsilon(f) \geq d$, but with the additional requirement that $\psi(x) = 0$ outside of $\mathcal{X}$.

Proposition 13. *Let $\mathcal{X} \subseteq \{-1, 1\}^n$, and let $f : \mathcal{X} \rightarrow \{-1, 1\}$. Then $\text{ubdeg}_\varepsilon(f) \geq d$ if and only if there exists a function $\psi : \{-1, 1\}^n \rightarrow \mathbb{R}$ satisfying the following properties.*

$$\psi(x) = 0 \text{ for all } x \notin \mathcal{X}, \quad (5)$$

$$\sum_{x \in \mathcal{X}} \psi(x) \cdot f(x) > \varepsilon, \quad (6)$$

$$\sum_{x \in \{-1, 1\}^n} |\psi(x)| = 1, \text{ and} \quad (7)$$

For every polynomial $p : \{-1, 1\}^n \rightarrow \mathbb{R}$ of degree less than d ,

$$\sum_{x \in \{-1, 1\}^n} p(x) \cdot \psi(x) = 0. \quad (8)$$

Observe that $\deg(f)$ and $\text{ubdeg}(f)$ coincide for total functions f . To avoid notational clutter, when referring to the approximate degree of total functions, we will use the shorter notation $\deg(f)$.

2.3 Basic Facts about Polynomial Approximations

The seminal work of Nisan and Szegedy [55] gave tight bounds on the approximate degree of the AND_n and OR_n functions.

Lemma 14. *For any constant $\varepsilon \in (0, 1)$, the functions AND and OR on n bits have ε -approximate degree $\Theta(n^{1/2})$, and the same holds for their negations.*

Approximate degree is invariant under negating the inputs or output of a function, and hence the result for AND implies the result for NAND , OR , etc.

The following lemma, which forms the basis of the well-known symmetrization argument, is due to Minsky and Papert [53].

Lemma 15. *Let $p : \{-1, 1\}^n \rightarrow \{-1, 1\}$ be an arbitrary polynomial and let $[n]_0$ denote the set $\{0, 1, \dots, n\}$. Then there is a univariate polynomial $q : \mathbb{R} \rightarrow \mathbb{R}$ of degree at most $\deg(p)$ such that*

$$q(t) = \frac{1}{\binom{n}{t}} \sum_{x \in \{-1, 1\}^n : |x|=t} p(x)$$

for all $t \in [n]_0$.

2.4 Formal definition of Surjectivity

Definition 16. *For $N \geq R$, define the function $\text{SURJ}_{N,R} : [R]^N \rightarrow \{-1, 1\}$ by $\text{SURJ}_{N,R}(s_1, \dots, s_N) = -1$ iff for every $j \in [R]$, there exists an i such that $s_i = j$.*

When N and R are clear from context, we will often refer to the function SURJ without the explicit dependence on these parameters. It will sometimes be convenient to think of the input to $\text{SURJ}_{N,R}$ as a function mapping $\{-1, 1\}^n \rightarrow \{-1, 1\}$ rather than $[R]^N \rightarrow \{-1, 1\}$. When needed, we assume that R is a power of 2 and an element of $[R]$ is encoded in binary using $\log R$ bits. In this case we will view Surjectivity as a function on $n = N \log R$ bits, i.e., $\text{SURJ} : \{-1, 1\}^n \rightarrow \{-1, 1\}$.

For technical reasons, when proving lower bounds, it will be more convenient to work with a variant of SURJ where the range $[R]$ is augmented by a “dummy element” 0 that is simply ignored by the function. That is, while any of the items $s_1, \dots, s_N$ may take

the dummy value 0, the presence of a 0 in the input is not required for the input to be deemed surjective. We denote this variant of Surjectivity by dSURJ . More formally:

Definition 17. For $N \geq R$, define the function $\text{dSURJ}_{N,R} : [R]_0^N \rightarrow \{-1, 1\}$ by $\text{dSURJ}_{N,R}(s_1, \dots, s_N) = -1$ iff for every $j \in [R]$, there exists an i such that $s_i = j$.

The following simple reduction shows that a lower bound on the approximate degree of dSURJ implies a lower bound for SURJ itself.

Proposition 18. Let $\varepsilon > 0$ and $N \geq R$. Then

$$\widetilde{\deg}_\varepsilon(\text{dSURJ}_{N,R}) \leq \widetilde{\deg}_\varepsilon(\text{SURJ}_{N+1,R+1}) \cdot \log(R+1).$$

PROOF. Let $p : \{-1, 1\}^{N \cdot \log(R+1)} \rightarrow \{-1, 1\}$ be a polynomial of degree d that ε -approximates $\text{SURJ}_{N+1,R+1}$. We will use p to construct a polynomial of degree d that ε -approximates $\text{dSURJ}_{N,R}$. Recall that an input to $\text{dSURJ}_{N,R}$ takes the form $(s_1, \dots, s_N)$ where each s_i is the binary representation of a number in $[R]_0$. Define the transformation $T : [R]_0 \rightarrow [R+1]$ by

$$T(s) = \begin{cases} R+1 & \text{if } s = 0 \\ s & \text{otherwise.} \end{cases}$$

Note that as a mapping between binary representations, the function T is exactly computed by a vector of polynomials of degree at most $\log(R+1)$. For every $(s_1, \dots, s_N) \in [R]_0^N$, observe that

$$\text{dSURJ}_{N,R}(s_1, \dots, s_N) = \text{SURJ}_{N+1,R+1}(T(s_1), \dots, T(s_N), R+1).$$

Hence, the polynomial

$$p(T(s_1), \dots, T(s_N), R+1)$$

is a polynomial of degree $d \cdot \log(R+1)$ that ε -approximates $\text{dSURJ}_{N,R}$. $\square$

2.5 Connecting Symmetric Properties and Block Composed Functions

An important ingredient in [29] is the relationship between the approximate degree of a property of a list of numbers (such as SURJ) and the approximate degree of a simpler block composed function, defined as follows.

Definition 19. For functions $f : Y^n \rightarrow Z$ and $g : X \rightarrow Y$, define the block composition $f \circ g : X^n \rightarrow Z$ by $(f \circ g)(x_1, \dots, x_n) = f(g(x_1), \dots, g(x_n))$, for all $x_1, \dots, x_n \in X$.

Fix $R, N \in \mathbb{N}$, let $f : \{-1, 1\}^R \rightarrow \{-1, 1\}$ and let $g : \{-1, 1\}^N \rightarrow \{-1, 1\}$. Suppose g is a symmetric function, in the sense that for any $x \in \{-1, 1\}^N$ and any permutation $\sigma : [N] \rightarrow [N]$, we have

$$g(x_1, \dots, x_N) = g(x_{\sigma(1)}, \dots, x_{\sigma(N)}).$$

Equivalently, the value of g on any input x depends only on its Hamming weight $|x|$.

The functions f and g give rise to two functions. The first, which we denote by $F^{\text{PROP}} : [R]_0^N \rightarrow \{-1, 1\}$, is a certain property of a list of numbers $s_1, \dots, s_N \in [R]_0$. The second, which we denote by $F^{\leq N} : H_{\leq N}^{N \cdot R} \rightarrow \{-1, 1\}$, is the block composition of f and g

restricted to inputs of Hamming weight at most N . Formally, these functions are defined as:

$$\begin{aligned} F^{\text{PROP}}(s_1, \dots, s_N) &= f(g(\mathbb{1}[s_1 = 1], \dots, \mathbb{1}[s_N = 1]), \\ &\quad \dots, g(\mathbb{1}[s_1 = R], \dots, \mathbb{1}[s_N = R])), \\ F^{\leq N}(x_1, \dots, x_R) &= \begin{cases} f(g(x_1), \dots, g(x_R)) & \text{if } x_1, \dots, x_R \in \{-1, 1\}^N, \\ & |x_1| + \dots + |x_R| \leq N. \\ \text{undefined} & \text{otherwise.} \end{cases} \end{aligned}$$

The following proposition relates the approximate degrees of the two functions F^{PROP} and $F^{\leq N}$.

THEOREM 20 (BUN AND THALER [29]). Let $f : \{-1, 1\}^R \rightarrow \{-1, 1\}$ be any function and let $g : \{-1, 1\}^N \rightarrow \{-1, 1\}$ be a symmetric function. Then for F^{PROP} and $F^{\leq N}$ defined above, and for any $\varepsilon > 0$, we have

$$\deg_\varepsilon(F^{\text{PROP}}) \geq \widetilde{\deg}_\varepsilon(F^{\leq N}).$$

In the case where $f = \text{AND}_R$ and $g = \text{OR}_N$, the function $F^{\text{PROP}}(s_1, \dots, s_N)$ is the Surjectivity function augmented with a dummy item, $\text{dSURJ}_{N,R}(s_1, \dots, s_N)$. Hence,

Corollary 21. Let $N, R \in \mathbb{N}$. Then for any $\varepsilon > 0$,

$$\widetilde{\deg}_\varepsilon(\text{dSURJ}_{N,R}) \geq \widetilde{\deg}_\varepsilon(F^{\leq N})$$

where $F^{\leq N} : H_{\leq N}^{N \cdot R} \rightarrow \{-1, 1\}$ is the restriction of $\text{AND}_R \circ \text{OR}_N$ to $H_{\leq N}^{N \cdot R}$.

2.6 The Dual Block Method

This section collects definitions and preliminary results on the dual block method [47, 70, 75] for constructing dual witnesses for a block composed function $F \circ f$ by combining dual witnesses for F and f respectively.

Definition 22. Let $\Psi : \{-1, 1\}^M \rightarrow \mathbb{R}$ and $\psi : \{-1, 1\}^m \rightarrow \mathbb{R}$ be functions that are not identically zero. Let $x = (x_1, \dots, x_M) \in (\{-1, 1\}^m)^M$. Define the dual block composition of Ψ and ψ , denoted $\Psi \star \psi : (\{-1, 1\}^m)^M \rightarrow \mathbb{R}$, by

$$(\Psi \star \psi)(x_1, \dots, x_M) = 2^M \cdot \Psi(\dots, \text{sgn}(\psi(x_i)), \dots) \cdot \prod_{i=1}^M |\psi(x_i)|.$$

Proposition 23 [29, 70]). The dual block composition satisfies the following properties:

Preservation of ℓ_1 -norm: If $\|\Psi\|_1 = 1$, $\|\psi\|_1 = 1$, and $\langle \psi, 1 \rangle = 0$, then

$$\|\Psi \star \psi\|_1 = 1. \quad (9)$$

Multiplicativity of pure high degree: If $\langle \Psi, P \rangle = 0$ for every polynomial $P : \{-1, 1\}^M \rightarrow \{-1, 1\}$ of degree less than D , and $\langle \psi, p \rangle = 0$ for every polynomial $p : \{-1, 1\}^m \rightarrow \{-1, 1\}$ of degree less than d , then for every polynomial $q : \{-1, 1\}^{m \cdot M} \rightarrow \{-1, 1\}$,

$$\deg q < D \cdot d \implies \langle \Psi \star \psi, q \rangle = 0. \quad (10)$$

Associativity: For every $\zeta : \{-1, 1\}^{m \cdot \zeta} \rightarrow \mathbb{R}$, $\varphi : \{-1, 1\}^{m \cdot \varphi} \rightarrow \mathbb{R}$, and $\psi : \{-1, 1\}^{m \cdot \psi} \rightarrow \mathbb{R}$, we have

$$(\zeta \star \varphi) \star \psi = \zeta \star (\varphi \star \psi). \quad (11)$$

2.7 A Refinement of a Technical Lemma from Prior Work

The following technical proposition refines techniques of Bun and Thaler [29]. This proposition is useful for “zeroing out” the mass that a dual polynomial ξ places on inputs of high Hamming weight, if ξ is obtained via the dual-block method. The proof of the proposition can be found in the full version.

Proposition 24. *Let $R \in \mathbb{N}$ be sufficiently large, and let $\Phi : \{-1, 1\}^R \rightarrow \mathbb{R}$ with $\|\Phi\|_1 = 1$. For $M \leq R$, let $\omega : [M]_0 \rightarrow \mathbb{R}$ and parameters $1 \leq \alpha \leq R^2, \beta \in (4 \ln^2 R / \sqrt{\alpha} R, 1)$ be such that*

$$\sum_{t=0}^M \omega(t) = 0, \quad (12)$$

$$\sum_{t=0}^M |\omega(t)| = 1, \quad (13)$$

$$|\omega(t)| \leq \alpha \exp(-\beta t) / t^2 \quad \forall t = 1, 2, \dots, M. \quad (14)$$

Let $N = \lceil 20\sqrt{\alpha} \rceil R$, and define $\psi : \{-1, 1\}^N \rightarrow \mathbb{R}$ by $\psi(x) = \omega(|x|) / \binom{N}{|x|}$. If $D < N$ is such that

$$\text{For every polynomial } p \text{ with } \deg p < D, \text{ we have } \langle \Phi \star \psi, p \rangle = 0, \quad (15)$$

then there exist $\Delta \geq \beta \sqrt{\alpha} R / 4 \ln^2 R$ and a function $\zeta : (\{-1, 1\}^N)^R \rightarrow \mathbb{R}$ such that

$$\text{For every polynomial } p \text{ with } \deg p < \min\{D, \Delta\}, \text{ we have } \langle \zeta, p \rangle = 0, \quad (16)$$

$$\|\zeta - \Phi \star \psi\|_1 \leq \frac{2}{9}, \quad (17)$$

$$\|\zeta\|_1 = 1, \quad (18)$$

$$\zeta \text{ is supported on } H_{\leq N}^{N \cdot R}. \quad (19)$$

The key refinement of Proposition 24 relative to the analysis of Bun and Thaler is that Proposition 24 applies when $N = \Theta(R)$ (assuming $\alpha = O(1)$). In contrast, the techniques of Bun and Thaler required $N = \Omega(R \cdot \log^2 R)$. As indicated in Section 1.3, this refinement will be essential in obtaining our lower bounds for SDU, entropy approximation, and junta testing for constant proximity parameter.

3 LOWER BOUND FOR SURJECTIVITY

The goal of this section is to show the following improved lower bound on the approximate degree of the Surjectivity function.

THEOREM 25. *For some $N = O(R)$, the $(1/3)$ -approximate degree of $\text{SURJ}_{N,R}$ is $\tilde{\Omega}(R^{3/4})$.*

To prove Theorem 25, we combine the following theorem with the reductions of Proposition 18 and Corollary 21.

THEOREM 26. *Let $N = c \cdot R$ for a sufficiently large constant $c > 0$. Let $F^{\leq N} : H_{\leq N}^{N \cdot R} \rightarrow \{-1, 1\}$ equal $\text{AND}_R \circ \text{OR}_N$ restricted to inputs in $H_{\leq N}^{N \cdot R} = \{x \in \{-1, 1\}^{N \cdot R} : |x| \leq N\}$. Then $\widetilde{\text{ubdeg}}(F^{\leq N}) \geq \tilde{\Omega}(R^{3/4})$.*

The proof Theorem 26 entails using dual witnesses for the high approximate degree of AND_R and OR_N to construct a dual witness

for the higher approximate degree of $F^{\leq N}$. As indicated in Section 1.3, the construction is essentially the same as in [29], except that we observe that a dual witness for OR constructed and used in prior works satisfies an exponentially stronger decay condition than has been previously realized.

The construction can be thought of as consisting of three steps:

Step 1. We begin by constructing a dual witness ψ for the fact that the unbounded approximate degree of the OR_N function is $\Omega(\sqrt{T})$ even when promised that the input has Hamming weight at most $T = \Theta(\sqrt{R})$. The dual witness ψ is a small variant of the one in [29], but we give a more careful analysis of its tail decay. In particular, we make use of the fact that for all $t \geq 1$, the ℓ_1 weight that ψ places on the t 'th layer of the Hamming cube is upper bounded by $\exp(-\Omega(t/\sqrt{T}))/t^2$.

Step 2. We combine ψ with a dual witness Φ for AND_R to obtain a preliminary dual witness $\Phi \star \psi$ for $F = \text{AND}_R \circ \text{OR}_N$. The dual witness $\Phi \star \psi$ shows that F has approximate degree $\Omega(\sqrt{R} \cdot \sqrt{T}) = \Omega(R^{3/4})$. However, $\Phi \star \psi$ places weight on inputs of Hamming weight larger than N , and hence does not give an *unbounded* approximate degree lower bound for the promise variant $F^{\leq N}$.

Step 3. Using Proposition 24 we zero out the mass that $\Phi \star \psi$ places on inputs of Hamming weight larger than N , while maintaining its pure high degree and correlation with $F^{\leq N}$. This yields the final desired *unbounded* approximate degree dual witness ζ for $F^{\leq N}$, as per Proposition 13.

3.1 Step 1: A Dual Witness for OR_N

We begin by constructing a univariate function which captures the properties we need of our inner dual witness for OR_N . The construction slightly modifies the dual polynomial for OR_N given by Špalek [76]. We provide a careful analysis of its decay as a function of the input's Hamming weight.

Proposition 27. *Let $T \in \mathbb{N}$ and $1/T \leq \delta \leq 1/2$. There exist constants $c_1, c_2 \in (0, 1)$ and a function $\omega : [T]_0 \rightarrow \mathbb{R}$ such that*

$$\omega(0) - \sum_{t=1}^T \omega(t) \geq 1 - \delta \quad (20)$$

$$\sum_{t=0}^T |\omega(t)| = 1 \quad (21)$$

For all univariate polynomials $q : \mathbb{R} \rightarrow \mathbb{R}$,

$$\deg q < c_1 \sqrt{\delta T} \implies \sum_{t=0}^T \omega(t) \cdot q(t) = 0 \quad (22)$$

$$|\omega(t)| \leq \frac{170 \exp(-c_2 t \sqrt{\delta / \sqrt{T}})}{\delta \cdot t^2} \quad \forall t = 1, \dots, T. \quad (23)$$

PROOF OF PROPOSITION 27. By renormalizing, it suffices to construct a function $\omega : [T]_0 \rightarrow \mathbb{R}$ such that

$$\omega(0) - \sum_{t=1}^T \omega(t) \geq (1 - \delta) \|\omega\|_1 \quad (24)$$

For all univariate polynomials $q : \mathbb{R} \rightarrow \mathbb{R}$,

$$\deg q < c_1 \sqrt{\delta T} \implies \sum_{t=0}^T \omega(t) \cdot q(t) = 0 \quad (25)$$

$$|\omega(t)| \leq \frac{170 \exp(-c_2 t \sqrt{\delta}/\sqrt{T}) \|\omega\|_1}{\delta \cdot t^2} \quad \forall t = 1, \dots, T. \quad (26)$$

Let $c = \lceil 8/\delta \rceil$ below. We will freely use the fact that since $\delta \leq 1/2$, we have $c \leq c^2/(c-1) \leq 10/\delta$. Let $m = \lfloor \sqrt{T/2c} \rfloor$ and define the set

$$S = \{1, c\} \cup \{2ci^2 : 0 \leq i \leq m\}.$$

Note that $|S| \geq c_1 \sqrt{\delta T}$ for some absolute constant $c_1 > 0$. Define the function

$$\omega(t) = \frac{(-1)^{t+(T-m)}}{T!} \binom{T}{t} \prod_{r \in [T]_0 \setminus S} (t-r).$$

Property (25) follows from the following combinatorial identity.

Fact 28 (cf. O'Donnell and Servedio [56]). *Let $T \in \mathbb{N}$, and let p be a polynomial of degree less than T . Then*

$$\sum_{t=0}^T (-1)^t \binom{T}{t} p(t) = 0.$$

Expanding out the binomial coefficient in the definition of ω reveals that

$$|\omega(t)| = \begin{cases} \prod_{r \in S \setminus \{t\}} \frac{1}{|t-r|} & \text{for } t \in S, \\ 0 & \text{otherwise.} \end{cases}$$

We now use this characterization to establish the improved tail decay property (26). This clearly holds for $t = 1$ with $c_2 = 1/10$, since $|\omega(1)| \leq \|\omega\|_1$ and

$$\frac{170}{\delta} \exp(-c_2 \sqrt{\delta}/\sqrt{T}) \geq 340 \cdot e^{-\sqrt{2}/10} > 1.$$

For $t = c$, we have

$$\begin{aligned} \frac{|\omega(c)|}{\omega(0)} &= \frac{c \prod_{i=1}^m (2ci^2)}{c(c-1) \prod_{i=1}^m (2ci^2 - c)} \\ &= \frac{1}{c-1} \left(\prod_{i=1}^m \frac{i^2 - 1/2}{i^2} \right)^{-1} \\ &\leq \frac{1}{c-1} \left(1 - \sum_{i=1}^m \frac{1}{2i^2} \right)^{-1} \\ &\leq \frac{1}{c-1} \left(1 - \frac{\pi^2}{12} \right)^{-1} \leq \frac{6}{c-1} \end{aligned} \quad (27)$$

where the first inequality follows from the fact that $\prod_{i=1}^m (1 - a_i) \geq 1 - \sum_{i=1}^m a_i$ for $a_i \in (0, 1)$. Now note that

$$\begin{aligned} \frac{|\omega(c)|}{\|\omega\|_1} &\leq \frac{|\omega(c)|}{\omega(0)} \\ &\leq \frac{6}{c-1} \\ &\leq \frac{60}{\delta \cdot c^2} \quad \text{since } 10/\delta \geq c^2/(c-1) \\ &\leq \frac{170}{\delta \cdot c^2} \cdot e^{-c\sqrt{\delta}/10\sqrt{T}}, \end{aligned}$$

since $\delta \geq 1/T$ and hence $e^{-c\sqrt{\delta}/10\sqrt{T}} \geq e^{-1}$. Thus (26) holds for $t = c$, recalling that $c_2 = 1/10$.

For $t = 2cj^2$ with $j \geq 1$, we get

$$\begin{aligned} \frac{|\omega(t)|}{\omega(0)} &= \frac{c \prod_{i=1}^m (2ci^2)}{(2cj^2 - 1)(2cj^2 - c) \prod_{i \in [m]_0 \setminus \{j\}} |2ci^2 - 2cj^2|} \\ &= \frac{c(m!)^2}{(4c^2j^4 - (2c^2 + 2c)j^2 + c) \prod_{i \in [m]_0 \setminus \{j\}} (i+j)|i-j|} \\ &= \frac{c}{4c^2j^4 - (2c^2 + 2c)j^2 + c} \cdot \frac{(m!)^2}{(m+j)!(m-j)!}. \end{aligned}$$

For $j \geq 1$, the first factor is bounded by

$$\frac{c}{4c^2j^4 - (2c^2 + 2c)j^2 + c} \leq \frac{3c}{(2cj^2)^2},$$

using the fact that $c \geq 2$. We control the second factor by

$$\begin{aligned} \frac{(m!)^2}{(m+j)!(m-j)!} &= \frac{m}{m+j} \cdot \frac{m-1}{m+j-1} \cdots \frac{m-j+1}{m+1} \\ &\leq \left(\frac{m}{m+j} \right)^j \\ &\leq \left(1 - \frac{j}{2m} \right)^j \\ &\leq e^{-j^2/2m}, \end{aligned}$$

where the last inequality uses the fact that $1 - x \leq e^{-x}$ for all x . Since

$$\frac{|\omega(t)|}{\|\omega\|_1} \leq \frac{|\omega(2cj^2)|}{\omega(0)} \leq \frac{3c}{(2cj^2)^2} \cdot e^{-2cj^2/(4cm)} \leq \frac{170}{\delta \cdot t^2} \cdot e^{-t\sqrt{\delta}/10\sqrt{T}},$$

this establishes (26).

What remains is to perform the correlation calculation to establish (24). For $t = 1$, we observe

$$\frac{|\omega(1)|}{\omega(0)} = \frac{c \prod_{i=1}^m (2ci^2)}{(c-1) \prod_{i=1}^m (2ci^2 - 1)} \geq \prod_{i=1}^m \frac{i^2}{i^2 - 1/(2c)} \geq 1.$$

Next, we observe that the total contribution of $t > c$ to $\|\omega\|_1/\omega(0)$ is at most

$$\sum_{t>c} \frac{|\omega(t)|}{\omega(0)} \leq \sum_{j=1}^m \frac{3c}{(2cj^2)^2} < \sum_{j=1}^{\infty} \frac{3}{4cj^4} = \frac{\pi^4}{120c}. \quad (28)$$

Next, we calculate

$$\begin{aligned}
\omega(0) - \sum_{t=1}^T \omega(t) &\geq \omega(0) - \omega(1) - \left(\sum_{t=c}^T |\omega(t)| \right) \\
&\geq \omega(0) - \omega(1) - \left(\omega(c) + \omega(0) \cdot \frac{\pi^4}{120c} \right) \quad \text{by (28)} \\
&\geq -\omega(1) + \omega(0) \left(1 - \frac{6}{c-1} - \frac{\pi^4}{120c} \right) \quad \text{by (27)} \\
&\geq -\omega(1) + (1-\delta)\omega(0) \text{ by our choice of } c \geq 8/\delta. \quad (29)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\|\omega\|_1 &\leq \omega(0) - \omega(1) + \omega(2) + \omega(0) \cdot \frac{\pi^4}{120c} \quad \text{by (28)} \\
&\leq -\omega(1) + \omega(0) \left(1 + \frac{6}{c-1} + \frac{\pi^4}{120c} \right) \quad \text{by (27)} \\
&\leq -\omega(1) + (1+\delta)\omega(0) \quad \text{since } c \geq 8/\delta. \quad (30)
\end{aligned}$$

Combining (29) and (30), and using the fact that $-\omega(1) \geq \omega(0)$ shows that

$$\frac{\omega(0) - \sum_{t=1}^k \omega(t)}{\|\omega\|_1} \geq \frac{-\omega(1) + (1-\delta)\omega(0)}{-\omega(1) + (1+\delta)\omega(0)} \geq \frac{2-\delta}{2+\delta} \geq 1-\delta.$$

This establishes (24), completing the proof. $\square$

The following construction of a dual polynomial for OR_N , with $N \geq T$, is an immediate consequence of Minsky-Papert symmetrization (Lemma 15), combined with Proposition 27.

Proposition 29. *Let $T, N \in \mathbb{N}$ with $T \leq N$, and let $\delta > 1/T$. Define ω as in Proposition 27. Define the function $\psi : \{-1, 1\}^N \rightarrow \{-1, 1\}$ by $\psi(x) = \omega(|x|)/\binom{N}{|x|}$ for $x \in H_{\leq T}^N$ and $\psi(x) = 0$ otherwise. Then*

$$\langle \psi, \text{OR}_N \rangle \geq 1 - \delta \quad (31)$$

$$\|\psi\|_1 = 1 \quad (32)$$

For any polynomial $p : \{-1, 1\}^N \rightarrow \mathbb{R}$,

$$\deg p < c_1 \sqrt{\delta T} \implies \langle \psi, p \rangle = 0 \quad (33)$$

3.2 Step 2: Constructing a Preliminary Dual Witness for $\text{AND}_R \circ \text{OR}_N$

The following proposition, when combined with Proposition 23, shows that there is a function $\Phi : \{-1, 1\}^R \rightarrow \{-1, 1\}$ such that the dual block composition $\Phi \star \psi$ is a good dual polynomial for $\text{AND}_R \circ \text{OR}_N$. In the next section, we will modify $\Phi \star \psi$ to zero out the weight it places outside $H_{\leq N}^{N \cdot R}$.

Proposition 30. *Let $\psi : \{-1, 1\}^N \rightarrow \{-1, 1\}$ be a function such that $\|\psi\|_1 = 1$ and $\langle \psi, \text{OR}_N \rangle \geq 19/20$. Then there exists a function $\Phi : \{-1, 1\}^R \rightarrow \{-1, 1\}$ with pure high degree $\Omega(\sqrt{R})$ and $\|\Phi\|_1 = 1$ such that*

$$\langle \Phi \star \psi, \text{AND}_R \circ \text{OR}_N \rangle > 2/3.$$

The proof of Proposition 30 is implicit in the results of [25, 68].

3.3 Step 3: Constructing the Final Dual Witness

Proposition 31. *Let R be sufficiently large. There exist $N = O(R)$, $D = \tilde{\Omega}(N^{3/4})$, and $\zeta : (\{-1, 1\}^N)^R \rightarrow \mathbb{R}$ such that*

$$\zeta(x) = 0 \text{ for all } x \notin H_{\leq N}^{N \cdot R}, \quad (34)$$

$$\sum_{x \in H_{\leq N}^{N \cdot R}} \zeta(x) \cdot (\text{AND}_R \circ \text{OR}_N)(x) > 1/3, \quad (35)$$

$$\|\zeta\|_1 = 1, \text{ and} \quad (36)$$

For every polynomial $p : (\{-1, 1\}^N)^R \rightarrow \mathbb{R}$ of degree less than D , we have $\langle p, \zeta \rangle = 0$. (37)

PROOF. We start by fixing choices of several key parameters:

- $d = \Theta(\sqrt{R})$ is the pure high degree of the dual witness Φ for AND_R in Proposition 30,
- $T = \lfloor (R/d)^{1/2} \rfloor^2 = \Theta(\sqrt{R})$,
- $\hat{D} = c_1 \sqrt{T} \cdot d = \Theta(R^{3/4})$, where c_1 is the constant from Proposition 27,
- $\delta = 1/20$
- $\alpha = 170/\delta = 3400$,
- $\beta = c_2 \cdot \sqrt{\delta}/\sqrt{T} = \Theta(1/R^{1/4})$, where c_2 is the constant from Proposition 27,
- $N = \lceil 20\sqrt{\alpha} \rceil R = 693R$.

Let $\psi : \{-1, 1\}^N \rightarrow \{-1, 1\}$ be the function constructed in Proposition 29 with $\delta := 1/20$. Let $\Phi : \{-1, 1\}^R \rightarrow \{-1, 1\}$ be the function constructed in Proposition 30, and define $\xi = \Phi \star \psi$. Then by Proposition 23, ξ satisfies the following properties:

$$\langle \xi, \text{AND}_R \circ \text{OR}_N \rangle > 2/3, \quad (38)$$

$$\|\xi\|_1 = 1, \quad (39)$$

For every polynomial p of degree less than D , we have $\langle \xi, p \rangle = 0$. (40)

Recall that ψ was obtained by symmetrizing the function ω constructed in Proposition 27. Proposition 24 guarantees that for some $\Delta \geq \beta \sqrt{\alpha R}/4 \ln^2 R = \tilde{\Omega}(R^{3/4})$, the function ξ can be modified to produce a function $\zeta : (\{-1, 1\}^N)^R \rightarrow \mathbb{R}$ such that

$$\zeta(x) = 0 \text{ for all } x \notin H_{\leq N}^{N \cdot R},$$

$$\langle \zeta, \text{AND}_R \circ \text{OR}_N \rangle \geq \langle \xi, \text{AND}_R \circ \text{OR}_N \rangle - \|\zeta - \xi\|_1 \geq \frac{2}{3} - \frac{2}{9} > \frac{1}{3},$$

$$\|\zeta\|_1 = 1,$$

For every polynomial p of degree less than $\min\{\hat{D}, \Delta\}$, we have $\langle \zeta, p \rangle = 0$.

Observing that

$$D = \min\{\hat{D}, \Delta\} = \tilde{\Omega}(R^{3/4})$$

shows that the function ζ satisfies the conditions necessary to prove Proposition 31. $\square$

Theorem 26 follows by combining Proposition 31 with the dual characterization of unbounded approximate degree given in Proposition 13. By Corollary 21, we conclude that $\widehat{\deg}(\text{dSURJ}_{N,R}) = \tilde{\Omega}(R^{3/4})$. Theorem 25 follows by Proposition 18.

4 CONCLUSION AND OPEN QUESTIONS

We conclude by briefly describing some additional consequences of our results, as well as a number of open questions and directions for future work.

4.1 Additional Consequences: Approximate Degree Lower Bounds for DNFs and AC^0

For any constant $k > 0$, k -distinctness is computed by a DNF of polynomial size. Our $\tilde{\Omega}(n^{3/4-1/(2k)})$ is the best known lower bound on the approximate degree of polynomial size DNF formulae. The previous best was $\tilde{\Omega}(n^{2/3})$ for Element Distinctness (a.k.a., 2-Distinctness) [3], although Bun and Thaler did establish, for any $\delta > 0$, an $\Omega(n^{1-\delta})$ lower bound on the approximate degree of quasipolynomial size DNFs.

Similarly, for any constant $k \geq 1$, Bun and Thaler exhibited an AC^0 circuit of depth $2k-1$ with approximate degree $\tilde{\Omega}(n^{1-2^{k-2}/3^{k-1}})$. Our techniques can be used to give a polynomial improvement for any fixed $k \geq 2$, to $\tilde{\Omega}(n^{1-2^{-k}})$ (Theorem 1 is the special case of $k = 2$, as SURJ is computed by an AC^0 circuit of depth three). We omit further details of this result for brevity.

4.2 Open Problems

The most obvious direction for future work is to extend our techniques to resolve the approximate degree and quantum query complexity of additional problems of interest in the study of quantum algorithms. These include triangle finding problem [35, 52], graph collision [52], and verifying matrix products [23, 45]. It would also be interesting to close the gap between our $\Omega(n^{3/4-1/(2k)})$ lower bound for k -distinctness and Belovs' $O(n^{3/4-1/(2^{k+2}-4)})$ upper bound, especially for small values of k (e.g., $k = 3$).

Although we prove a lower bound of $\tilde{\Omega}(R^{1/2})$ for $SDU_{N,R}^{\gamma_1, \gamma_2}$ for some constants $0 < \gamma_2 < \gamma_1$, we leave open whether or not $SDU_{N,R}^{2/3, 1/3} = \tilde{\Omega}(R^{1/2})$. It may be tempting to suspect that Theorem 5 implies an $\tilde{\Omega}(R^{1/2})$ lower bound on $SDU_{N,R}^{2/3, 1/3}$, by invoking the well-known Polarization Lemma of Sahai and Vadhan [60]. The Polarization Lemma reduces $SDU_{N,R}^{\gamma_1, \gamma_2}$ for any pair of constant γ_1, γ_2 with $\gamma_2 < \gamma_1^2$ to $SDU_{N',R'}^{2/3, 1/3}$ for an appropriate choice of N' and R' . Unfortunately, N' and R' may be polynomially larger than N and R , so this reduction does not give an $\tilde{\Omega}(R^{1/2})$ lower bound for $SDU_{N,R}^{2/3, 1/3}$ itself.

Another important direction is to resolve the approximate degree of specific classes of functions, especially polynomial size DNF formulae, and AC^0 circuits. As mentioned in the previous subsection, our k -distinctness lower bound (Theorem 2) gives the best known lower bound on polynomial size DNFs. A compelling candidate for improving this lower bound is the k -sum function, which may have approximate degree as large as $\Theta(n^{k/(k+1)})$ (it is known that the quantum query complexity of k -sum is $\tilde{\Theta}(n^{k/(k+1)})$ [7, 17]). On the upper bounds side, it may be possible to extend the techniques underlying our $\tilde{O}(n^{3/4})$ upper bound on the approximate degree of SURJ to yield a sublinear upper bound for every DNF formula of polynomial size.

Open Problem 32. For every constant $c > 0$ and every DNF formula $f: \{-1, 1\}^n \rightarrow \{-1, 1\}$ of size at most n^c , is there a $\delta > 0$ (depending only on c) such that $\deg(f) = O(n^{1-\delta})$?

A positive answer to Open Problem 32 would imply a subexponential time algorithm for agnostically learning DNF formulae [40] (and PAC learning depth three circuits [43]) of any fixed polynomial size. These are far more expressive classes than those that are currently known to be non-trivially learnable in these models [37, 61].

For general AC^0 circuits, an $\Omega(n^{1-\delta})$ approximate degree lower bound is already known [29]. It would be very interesting to improve this lower bound to an optimal $\Omega(n)$. Until recently, SURJ was a prime candidate for exhibiting such a lower bound. However, owing to Sherstov's upper bound [74] and Theorem 1, SURJ is no longer a candidate function. However, we are optimistic about the following closely related candidate. An approximate majority function is any total Boolean function that evaluates to -1 (respectively $+1$) whenever at least $2/3$ of its inputs are -1 (respectively $+1$). It is well-known (via the probabilistic method) that there are approximate majorities computable by depth 3 circuits of quadratic size and logarithmic bottom fan-in [4]. It is possible that every approximate majority has approximate degree $\Omega(n)$; proving this would resolve a question of Srinivasan [34].

Finally, an important open question is to determine the *threshold degree* of AC^0 . Here, the threshold degree of a function $f: \{-1, 1\}^n \rightarrow \{-1, 1\}$ is the minimum degree of a polynomial p such that $p(x) \cdot f(x) > 0$ for every $x \in \{-1, 1\}^n$ (equivalently, the threshold degree of f equals the limit of the ε -approximate degree of f as the error approaches 1 from below). The best known lower bound on the threshold degree of AC^0 is $\Omega(n^{1/2})$, for a complicated circuit of depth four [73]. We believe that our analysis of SDU (Theorem 5) can be extended to show that the threshold degree of SURJ is $\Omega(n^{1/2})$, a bound that would be tight for SURJ. Perhaps our techniques will be useful in proving even stronger threshold degree lower bounds for AC^0 functions.

ACKNOWLEDGMENTS

We are grateful to Sasha Sherstov for an inspiring conversation at the BIRS 2017 workshop on Communication Complexity and Applications, II, which helped to spark this work.

Some of this work was performed when R. K. was a postdoctoral associate at MIT and was partly supported by NSF grant CCF-1629809.

REFERENCES

- [1] Scott Aaronson. 2012. Impossibility of succinct quantum proofs for collision-freeness. *Quantum Information & Computation* 12, 1-2 (2012), 21–28. <http://www.rintonpress.com/xxqic12/qic-12-12/0021-0028.pdf>
- [2] Scott Aaronson, Shalev Ben-David, and Robin Kothari. 2016. Separations in query complexity using cheat sheets. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*. 863–876.
- [3] Scott Aaronson and Yaoyun Shi. 2004. Quantum lower bounds for the collision and the element distinctness problems. *J. ACM* 51, 4 (2004), 595–605.
- [4] Miklós Ajtai. 1983. Σ_1^1 -Formulae on finite structures. *Annals of pure and applied logic* 24, 1 (1983), 1–48.
- [5] Andris Ambainis. 2002. Quantum Lower Bounds by Quantum Arguments. *J. Comput. System Sci.* 64, 4 (June 2002), 750–767. <https://doi.org/10.1006/jcss.2002.1826>

- [6] Andris Ambainis. 2003. Polynomial degree vs. quantum query complexity. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*. IEEE, 230–239.
- [7] Andris Ambainis. 2007. Quantum walk algorithm for element distinctness. *SIAM J. Comput.* 37, 1 (2007), 210–239.
- [8] Andris Ambainis, Aleksandrs Belovs, Oded Regev, and Ronald de Wolf. 2016. Efficient quantum algorithms for (gapped) group testing and junta testing. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 903–922.
- [9] Andris Ambainis, Andrew M. Childs, Ben Reichardt, Robert Špalek, and Shengyu Zhang. 2010. Any AND-OR Formula of Size N Can Be Evaluated in Time $N^{1/2+o(1)}$ on a Quantum Computer. *SIAM J. Comput.* 39, 6 (2010), 2513–2530.
- [10] Alp Atilci and Rocco A. Servedio. 2007. Quantum algorithms for learning and testing juntas. *Quantum Information Processing* 6, 5 (2007), 323–348.
- [11] Howard Barnum, Michael Saks, and Mario Szegedy. 2003. Quantum query complexity and semi-definite programming. In *18th Conference on Computational Complexity (CCC 2003)*. 179–193. <https://doi.org/10.1109/CCC.2003.1214419>
- [12] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. 2001. Quantum lower bounds by polynomials. *J. ACM* 48, 4 (2001), 778–797. <https://doi.org/10.1145/502090.502097>
- [13] Paul Beame and Widad Machmouchi. 2012. The quantum query complexity of AC^0 . *Quantum Information & Computation* 12, 7–8 (2012), 670–676. <http://www.rintonpress.com/xxqic12/qic-12-78/0670-0676.pdf>
- [14] Richard Beigel. 1994. Perceptrons, PP, and the Polynomial Hierarchy. *Computational Complexity* 4 (1994), 339–349. <https://doi.org/10.1007/BF01263422>
- [15] Aleksandrs Belovs. 2012. Learning-graph-based quantum algorithm for k -distinctness. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*. IEEE, 207–216.
- [16] Aleksandrs Belovs. 2012. Span programs for functions with constant-sized 1-certificates. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. ACM, 77–84.
- [17] Aleksandrs Belovs and Robert Špalek. 2013. Adversary Lower Bound for the K -sum Problem. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science (ITCS '13)*. ACM, New York, NY, USA, 323–328. <https://doi.org/10.1145/2422436.2422474>
- [18] Eric Blais. 2009. Testing Juntas Nearly Optimally. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing (STOC '09)*. ACM, New York, NY, USA, 151–158. <https://doi.org/10.1145/1536414.1536437>
- [19] Andrej Bogdanov, Yuval Ishai, Emanuele Viola, and Christopher Williamson. 2016. Bounded Indistinguishability and the Complexity of Recovering Secrets. In *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14–18, 2016, Proceedings, Part III (Lecture Notes in Computer Science)*, Matthew Robshaw and Jonathan Katz (Eds.), Vol. 9816. Springer, 593–618. https://doi.org/10.1007/978-3-662-53015-3_21
- [20] Adam Boulund, Lijie Chen, Dhiraj Holden, Justin Thaler, and Prashant Nalini Vasudevan. 2017. On The Power of Statistical Zero Knowledge. In *To Appear In Proceedings of IEEE Symposium on Foundations of Computer Science (FOCS)*. Preliminary version available at <http://eccc.hpi-web.de/report/2016/140>.
- [21] Sergey Bravyi, Aram Wetzroth Harrow, and Avinandan Hassidim. 2011. Quantum Algorithms for Testing Properties of Distributions. *IEEE Trans. Information Theory* 57, 6 (2011), 3971–3981. <https://doi.org/10.1109/TIT.2011.2134250>
- [22] Harry Buhrman, Richard Cleve, Ronald de Wolf, and Christof Zalka. 1999. Bounds for Small-Error and Zero-Error Quantum Algorithms. In *FOCS. IEEE Computer Society*, 358–368.
- [23] Harry Buhrman and Robert Špalek. 2006. Quantum verification of matrix products. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2006, Miami, Florida, USA, January 22–26, 2006*. ACM Press, 880–889. <http://dl.acm.org/citation.cfm?id=1109557.1109654>
- [24] Harry Buhrman, Nikolai K. Vereshchagin, and Ronald de Wolf. 2007. On Computation and Communication with Small Bias. In *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13–16 June 2007, San Diego, California, USA*. IEEE Computer Society, 24–32. <https://doi.org/10.1109/CCC.2007.18>
- [25] Mark Bun and Justin Thaler. 2013. Dual Lower Bounds for Approximate Degree and Markov-Bernstein Inequalities. In *ICALP (1) (Lecture Notes in Computer Science)*, Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg (Eds.), Vol. 7965. Springer, 303–314.
- [26] Mark Bun and Justin Thaler. 2015. Hardness Amplification and the Approximate Degree of Constant-Depth Circuits. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6–10, 2015, Proceedings, Part I (Lecture Notes in Computer Science)*, Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann (Eds.), Vol. 9134. Springer, 268–280. https://doi.org/10.1007/978-3-662-47672-7_22 Full version available at <http://eccc.hpi-web.de/report/2013/151>.
- [27] Mark Bun and Justin Thaler. 2016. Dual Polynomials for Collision and Element Distinctness. *Theory of Computing* 12, 16 (2016), 1–34. <https://doi.org/10.4086/toc.2016.v012a016>
- [28] Mark Bun and Justin Thaler. 2016. Improved Bounds on the Sign-Rank of AC^0 . In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11–15, 2016, Rome, Italy (LIPIcs)*, Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi (Eds.), Vol. 55. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 37:1–37:14. <https://doi.org/10.4230/LIPIcs.ICALP.2016.37>
- [29] Mark Bun and Justin Thaler. 2017. A Nearly Optimal Lower Bound on the Approximate Degree of AC^0 . *arXiv preprint arXiv:1703.05784* (2017). To Appear In *FOCS* 2017.
- [30] Karthekeyan Chandrasekaran, Justin Thaler, Jonathan Ullman, and Andrew Wan. 2014. Faster private release of marginals on small databases. In *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12–14, 2014*. 387–402. <https://doi.org/10.1145/2554797.2554833>
- [31] Arkadev Chattopadhyay and Anil Ada. 2008. Multiparty Communication Complexity of Disjointness. *Electronic Colloquium on Computational Complexity (ECCC)* 15, 002 (2008). <http://eccc.hpi-web.de/eccc-reports/2008/TR08-002/index.html>
- [32] Matei David and Toniann Pitassi. 2008. Separating NOF communication complexity classes RP and NP. *Electronic Colloquium on Computational Complexity (ECCC)* 15, 014 (2008). <http://eccc.hpi-web.de/eccc-reports/2008/TR08-014/index.html>
- [33] Matei David, Toniann Pitassi, and Emanuele Viola. 2009. Improved Separations between Nondeterministic and Randomized Multiparty Communication. *TOCT* 1, 2 (2009). <https://doi.org/10.1145/1595391.1595392>
- [34] Yuval Filmus, Hamed Hatami, Steven Heilman, Elchanan Mossel, Ryan O'Donnell, Sushant Sachdeva, Andrew Wan, and Karl Wimmer. 2014. Real analysis in computer science: A collection of open problems. (2014).
- [35] François Le Gall. 2014. Improved Quantum Algorithm for Triangle Finding via Combinatorial Arguments. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18–21, 2014*. 216–225. <https://doi.org/10.1109/FOCS.2014.31>
- [36] Dmitry Gavinsky and Alexander A. Sherstov. 2010. A Separation of NP and coNP in Multiparty Communication Complexity. *Theory of Computing* 6, 1 (2010), 227–245. <https://doi.org/10.4086/toc.2010.v006a010>
- [37] Lisa Hellerstein and Rocco A. Servedio. 2007. On PAC learning algorithms for rich Boolean function classes. *Theor. Comput. Sci.* 384, 1 (2007), 66–76. <https://doi.org/10.1016/j.tcs.2007.05.018>
- [38] Peter Høyer, Troy Lee, and Robert Špalek. 2007. Negative weights make adversaries stronger. In *Proceedings of the 39th Symposium on Theory of Computing (STOC 2007)*. 526–535. <https://doi.org/10.1145/1250790.1250867>
- [39] Jeff Kahn, Nathan Linial, and Alex Samorodnitsky. 1996. Inclusion-Exclusion: Exact and Approximate. *Combinatorica* 16, 4 (1996), 465–477.
- [40] Adam Tauman Kalai, Adam R. Klivans, Yishay Mansour, and Rocco A. Servedio. 2008. Agnostically Learning Halfspaces. *SIAM J. Comput.* 37, 6 (2008), 1777–1805. <https://doi.org/10.1137/060649057>
- [41] Varun Kanade and Justin Thaler. 2014. Distribution-independent Reliable Learning. In *Proceedings of The 27th Conference on Learning Theory, COLT 2014, Barcelona, Spain, June 13–15, 2014 (JMLR Proceedings)*, Maria-Florina Balcan and Csaba Szepesvári (Eds.), Vol. 35. JMLR.org, 3–24. <http://jmlr.org/proceedings/papers/v35/kanade14.html>
- [42] Hartmut Klauck, Robert Špalek, and Ronald de Wolf. 2007. Quantum and Classical Strong Direct Product Theorems and Optimal Time-Space Tradeoffs. *SIAM J. Comput.* 36, 5 (2007), 1472–1493. <https://doi.org/10.1137/05063235X> arXiv:https://doi.org/10.1137/05063235X
- [43] Adam R. Klivans and Rocco A. Servedio. 2004. Learning DNF in time $2^{\tilde{O}(n^{1/3})}$. *J. Comput. Syst. Sci.* 68, 2 (2004), 303–318. <https://doi.org/10.1016/j.jcss.2003.07.007>
- [44] Adam R. Klivans and Rocco A. Servedio. 2006. Toward Attribute Efficient Learning of Decision Lists and Parities. *Journal of Machine Learning Research* 7 (2006), 587–602. <http://www.jmlr.org/papers/v7/klivans06a.html>
- [45] Robin Kothari and Ashwin Nayak. 2016. Quantum Algorithms for Matrix Multiplication and Product Verification. In *Encyclopedia of Algorithms*. Springer, 1673–1677. https://doi.org/10.1007/978-1-4939-2864-4_303
- [46] Sophie Laplante and Frédéric Magniez. 2004. Lower bounds for randomized and quantum query complexity using Kolmogorov arguments. In *Proceedings of the 19th Conference on Computational Complexity*. 294–304. <https://doi.org/10.1109/CCC.2004.1313852>
- [47] Troy Lee. 2009. A note on the sign degree of formulas. *CoRR* abs/0909.4607 (2009). <http://arxiv.org/abs/0909.4607>
- [48] Troy Lee, Rajat Mittal, Ben W. Reichardt, Robert Špalek, and Mario Szegedy. 2011. Quantum query complexity of state conversion. In *Proceedings of the 52nd Symposium on Foundations of Computer Science (FOCS 2011)*. 344–353. <https://doi.org/10.1109/FOCS.2011.75>
- [49] Troy Lee and Adi Shraibman. 2009. An Approximation Algorithm for Approximation Rank. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15–18 July 2009*. 351–357.
- [50] Troy Lee and Adi Shraibman. 2009. Disjointness is Hard in the Multiparty Number-on-the-Forehead Model. *Computational Complexity* 18, 2 (2009), 309–336. <https://doi.org/10.1007/s00037-009-0276-2> Preliminary version in CCC

- 2008.
- [51] Tongyang Li and Xiaodi Wu. 2017. Quantum query complexity of entropy estimation. *arXiv preprint arXiv:1710.06025* (2017). <https://www.cs.umd.edu/~xwu/papers/qentropy.pdf>
 - [52] Frédéric Magniez, Miklos Santha, and Mario Szegedy. 2007. Quantum Algorithms for the Triangle Problem. *SIAM J. Comput.* 37, 2 (2007), 413–424. <https://doi.org/10.1137/050643684>
 - [53] Marvin Minsky and Seymour Papert. 1969. *Perceptrons - an introduction to computational geometry*. MIT Press.
 - [54] Ashley Montanaro. 2016. The quantum complexity of approximating the frequency moments. *Quantum Information & Computation* 16, 13&14 (2016), 1169–1190. <http://www.rintonpress.com/xxqic16/qic-16-1314/1169-1190.pdf>
 - [55] Noam Nisan and Mario Szegedy. 1994. On the Degree of Boolean Functions as Real Polynomials. *Computational Complexity* 4 (1994), 301–313. <https://doi.org/10.1007/BF01263419>
 - [56] Ryan O'Donnell and Rocco A. Servedio. 2010. New degree bounds for polynomial threshold functions. *Combinatorica* 30, 3 (2010), 327–358. <https://doi.org/10.1007/s00493-010-2173-3>
 - [57] Anup Rao and Amir Yehudayoff. 2015. Simplified Lower Bounds on the Multiparty Communication Complexity of Disjointness. In *30th Conference on Computational Complexity, CCC 2015, June 17-19, 2015, Portland, Oregon, USA*. 88–101. <https://doi.org/10.4230/LIPIcs.CCC.2015.88>
 - [58] Alexander A. Razborov and Alexander A. Sherstov. 2010. The Sign-Rank of AC^0 . *SIAM J. Comput.* 39, 5 (2010), 1833–1855. <https://doi.org/10.1137/080744037>
 - [59] Ben W Reichardt. 2011. Reflections for quantum query algorithms. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 560–569.
 - [60] Amit Sahai and Salil Vadhan. 2003. A complete problem for statistical zero knowledge. *Journal of the ACM (JACM)* 50, 2 (2003), 196–249.
 - [61] Rocco Servedio and Li-Yang Tan. 2017. What circuit classes can be learned with non-trivial savings?. In *8th Innovations in Theoretical Computer Science Conference (ITCS)*. 23:1–23:22.
 - [62] Rocco A. Servedio, Li-Yang Tan, and Justin Thaler. 2012. Attribute-Efficient Learning and Weight-Degree Tradeoffs for Polynomial Threshold Functions. In *COLT*. 14:1–14:19.
 - [63] Alexander A. Sherstov. 2008. Communication Lower Bounds Using Dual Polynomials. *Bulletin of the EATCS* 95 (2008), 59–93.
 - [64] Alexander A. Sherstov. 2009. Approximate Inclusion-Exclusion for Arbitrary Symmetric Functions. *Computational Complexity* 18, 2 (2009), 219–247. <https://doi.org/10.1007/s00037-009-0274-4>
 - [65] Alexander A. Sherstov. 2009. Separating AC^0 from Depth-2 Majority Circuits. *SIAM J. Comput.* 38, 6 (2009), 2113–2129. <https://doi.org/10.1137/08071421X>
 - [66] Alexander A. Sherstov. 2011. The Pattern Matrix Method. *SIAM J. Comput.* 40, 6 (2011), 1969–2000. <https://doi.org/10.1137/080733644> Preliminary version in STOC 2008.
 - [67] Alexander A. Sherstov. 2012. The multiparty communication complexity of set disjointness. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*. 525–548. <https://doi.org/10.1145/2213977.2214026>
 - [68] Alexander A. Sherstov. 2013. Approximating the AND-OR Tree. *Theory of Computing* 9, 20 (2013), 653–663.
 - [69] Alexander A. Sherstov. 2013. Communication lower bounds using directional derivatives. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, Dan Boneh, Tim Roughgarden, and Joan Feigenbaum (Eds.). ACM, 921–930. <https://doi.org/10.1145/2488608.2488725>
 - [70] Alexander A. Sherstov. 2013. The Intersection of Two Halfspaces Has High Threshold Degree. *SIAM J. Comput.* 42, 6 (2013), 2329–2374. <https://doi.org/10.1137/100785260>
 - [71] Alexander A. Sherstov. 2013. Making Polynomials Robust to Noise. *Theory of Computing* 9 (2013), 593–615. <https://doi.org/10.4086/toc.2013.v009a018>
 - [72] Alexander A. Sherstov. 2014. Breaking the Minsky-Papert barrier for constant-depth circuits. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, David B. Shmoys (Ed.). ACM, 223–232. <https://doi.org/10.1145/2591796.2591871>
 - [73] Alexander A. Sherstov. 2015. The Power of Asymmetry in Constant-Depth Circuits. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*. 431–450. <https://doi.org/10.1109/FOCS.2015.34>
 - [74] Alexander A. Sherstov. 2018. Algorithmic polynomials. *Electronic Colloquium on Computational Complexity (ECCC) TR18-010* (2018). <https://eccc.weizmann.ac.il/report/2018/010/> To appear in STOC 2018.
 - [75] Yaoyun Shi and Yufan Zhu. 2009. Quantum communication complexity of block-composed functions. *Quantum Information & Computation* 9, 5 (2009), 444–460. <http://www.rintonpress.com/xxqic9/qic-9-5/0444-0460.pdf>
 - [76] Robert Špalek. 2008. A Dual Polynomial for OR. *CoRR* abs/0803.4516 (2008). <http://arxiv.org/abs/0803.4516>
 - [77] Robert Špalek and Mario Szegedy. 2006. All Quantum Adversary Methods are Equivalent. *Theory of Computing* 2, 1 (2006), 1–18. <https://doi.org/10.4086/toc.2006.v002a001>
 - [78] Avishay Tal. 2014. Shrinkage of De Morgan Formulae by Spectral Techniques. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*. 551–560. <https://doi.org/10.1109/FOCS.2014.65>
 - [79] Avishay Tal. 2017. Formula lower bounds via the quantum method. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*. 1256–1268. <https://doi.org/10.1145/3055399.3055472>
 - [80] Justin Thaler. 2016. Lower Bounds for the Approximate Degree of Block-Composed Functions. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016) (Leibniz International Proceedings in Informatics (LIPIcs))*, Vol. 55. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 17:1–17:15. <https://doi.org/10.4230/LIPIcs.ICALP.2016.17>
 - [81] Justin Thaler, Jonathan Ullman, and Salil P. Vadhan. 2012. Faster Algorithms for Privately Releasing Marginals. In *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part I (Lecture Notes in Computer Science)*, Artur Czumaj, Kurt Mehlhorn, Andrew M. Pitts, and Roger Wattenhofer (Eds.), Vol. 7391. Springer, 810–821. https://doi.org/10.1007/978-3-642-31594-7_68
 - [82] Salil Pravin Vadhan. 1999. *A study of statistical zero-knowledge proofs*. Ph.D. Dissertation. Massachusetts Institute of Technology.
 - [83] Gregory Valiant and Paul Valiant. 2011. Estimating the Unseen: An $n/\log(n)$ -sample Estimator for Entropy and Support Size, Shown Optimal via New CLTs. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing (STOC '11)*. ACM, New York, NY, USA, 685–694. <https://doi.org/10.1145/1993636.1993727>
 - [84] Mark Zhandry. 2015. A note on the quantum collision and set equality problems. *Quantum Information & Computation* 15, 7&8 (2015), 557–567. <http://www.rintonpress.com/xxqic15/qic-15-78/0557-0567.pdf>
 - [85] Shengyu Zhang. 2005. On the power of Ambainis lower bounds. *Theoretical Computer Science* 339, 2 (2005), 241–256. <https://doi.org/10.1016/j.tcs.2005.01.019>