Quantum Query-To-Communication Simulation Needs a Logarithmic Overhead

Sourav Chakraborty
Indian Statistical Institute, Kolkata, India
sourav@isical.ac.in

Arkadev Chattopadhyay
Tata Institute of Fundamental Research, Mumbai, India
arkadev.c@tifr.res.in

Nikhil S. Mande
Georgetown University, Washington, D.C., USA
nikhil.mande@georgetown.edu

Manaswi Paraashar
Indian Statistical Institute, Kolkata, India
manaswi.isi@gmail.com

Abstract

Buhrman, Cleve and Wigderson (STOC'98) observed that for every Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ and $\bullet : \{-1,1\}^2 \to \{-1,1\}$ the two-party bounded-error quantum communication complexity of $(f \circ \bullet)$ is $O(Q(f) \log n)$, where $Q(f)$ is the bounded-error quantum query complexity of f. Note that the bounded-error randomized communication complexity of $(f \circ \bullet)$ is bounded by $O(R(f))$, where $R(f)$ denotes the bounded-error randomized query complexity of f. Thus, the BCW simulation has an extra $O(\log n)$ factor appearing that is absent in classical simulation. A natural question is if this factor can be avoided. Razborov (IZV MATH'03) showed that the bounded-error quantum communication complexity of Set-Disjointness is $\Omega(\sqrt{n})$. The BCW simulation yields an upper bound of $O(\sqrt{n}\log n)$. Høyer and de Wolf (STACS'02) showed that this can be reduced to $c^{\log^* n}$ for some constant c, and subsequently Aaronson and Ambainis (FOCS'03) showed that this factor can be made a constant. That is, the quantum communication complexity of the Set-Disjointness function (which is $\mathsf{NOR}_n \circ \wedge$) is $O(Q(\mathsf{NOR}_n))$.

Perhaps somewhat surprisingly, we show that when $\bullet = \oplus$, then the extra $\log n$ factor in the BCW simulation is unavoidable. In other words, we exhibit a total function $F : \{-1,1\}^n \to \{-1,1\}$ such that $Q^{cc}(F \circ \oplus) = \Theta(Q(F) \log n)$.

To the best of our knowledge, it was not even known prior to this work whether there existed a total function F and 2-bit function $\bullet$, such that $Q^{cc}(F \circ \bullet) = \omega(Q(F))$.

2012 ACM Subject Classification Theory of computation → Quantum query complexity; Theory of computation → Quantum communication complexity

Keywords and phrases Quantum query complexity, quantum communication complexity, approximate degree, approximate spectral norm

Digital Object Identifier 10.4230/LIPIcs.CCC.2020.32

Related Version A full version of the paper is available at `https://arxiv.org/pdf/1909.10428.pdf`.

Acknowledgements We thank Ronald de Wolf and Srinivasan Arunachalam for various discussions on this problem. They provided us with a number of important pointers that were essential for our understanding of the problem and its importance. We thank Justin Thaler for referring us to Claim 31. We thank the anonymous CCC 2020 reviewers for invaluable comments. In particular, we thank one of the reviewers for pointing out the contrasts between the classical query model and the quantum query model that our results imply (Remark 5).

COMPUTATIONAL COMPLEXITY CONFERENCE
35th Computational Complexity Conference (CCC 2020).
Editor: Shubhangi Saraf; Article No. 32; pp. 32:1–32:15
LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Classical communication complexity, introduced by Yao [24], is aptly called the "swiss-army-knife" for understanding, especially the limitations of, classical computing. Quantum communication complexity holds the same promise with regards to quantum computing. Yet, there are many problems that remain open. One broad theme is to understand the fundamental differences between classical randomized and quantum protocols, especially for computing total functions.

Recall a standard way to derive a communication problem from a function $f : \{-1,1\}^n \to \{-1,1\}$. Each input bit of f is *encoded* between the two players Alice and Bob, using an instance of a binary primitive, denoted by $\bullet : \{-1,1\} \times \{-1,1\} \to \{-1,1\}$, giving rise to the communication problem of evaluating $f \circ \bullet$. Each input bit to f is obtained by evaluating $\bullet$ on the relevant bit of Alice and that of Bob, i.e. $\big(f \circ \bullet\big)\big(x,y\big) = f\big(\bullet(x_1,y_1),\ldots,\bullet(x_n,y_n)\big)$ and x, y are each n-bit strings given to Alice and Bob respectively. Many well known functions in communication complexity are derived in this way: Set-Disjointness is $\mathsf{NOR} \circ \wedge$, Inner-Product being $\mathsf{PARITY} \circ \wedge$, Equality being $\mathsf{NOR} \circ \oplus$.[1] Set-Disjointness is also a standard total function where quantum protocols provably yield a significant cost saving over their classical counterpart.

A natural and well studied question in this regard is what is the relationship between the query complexity of f and the communication problem of $f \circ \bullet$, when the $\bullet$ is $\wedge$ or $\oplus$. This question has been studied for particular interesting functions or special classes of functions. Classically, it is folklore that

$$R^{cc}(f \circ \bullet) = O(R(f)),$$

where $R(f)$ denotes the bounded-error randomized query complexity of f and $R^{cc}(f \circ \bullet)$ denotes the bounded-error randomized communication complexity for computing $f \circ \bullet$. In an influential work, Buhrman, Cleve and Wigderson [9] observed that a general and natural recipe exists for constructing a quantum communication protocol for $f \circ \bullet$, using a quantum query algorithm for f as a black-box.

▶ **Theorem 1** ([9]). *For any Boolean function $f : \{-1,1\}^n \to \{-1,1\}$, we have*

$$Q^{cc}\big(f \circ \bullet\big) = O\big(Q(f) \cdot \log n\big),$$

where $\bullet$ is either $\wedge$ or $\oplus$.

Here $Q(f)$ denotes the bounded-error quantum query complexity of f, and $Q^{cc}(f \circ \bullet)$ denotes the bounded-error quantum communication complexity for computing $f \circ \bullet$. We remark here that the BCW simulation works for any constant-sized primitive $\bullet : \{-1,1\}^k \times \{-1,1\}^k \to \{-1,1\}$, but the focus of this paper is on the case where $k = 1$. Thus in the quantum world one incurs a logarithmic factor in the natural BCW simulation while no such factor is needed in the randomized setting. The basic question that arises naturally and which we completely answer in this work, is the following: analogous to the classical model, can this multiplicative $\log n$ blow-up in the communication cost be always avoided by designing quantum communication protocols that more cleverly simulate quantum query algorithms?

[1] Here $\wedge$ and $\oplus$ are the AND function and the XOR functions on 2 bits, respectively.

A priori, it is not clear what the answer to this question ought to be. For certain special functions and some classes of functions, quantum protocols exist where the $\log n$ factor can be saved. Theorem 1 implies a communication upper bound of $O(\sqrt{n}\log n)$ for the Set-Disjointness function. Høyer and de Wolf [15] designed a quantum protocol for Set-Disjointness of cost $O(\sqrt{n}c^{\log^* n})$, speeding up the BCW simulation significantly. Later, Aaronson and Ambainis [1] gave a more clever protocol that only incurred a constant factor overhead from Grover's search using more involved ideas, matching an $\Omega(\sqrt{n})$ lower bound due to Razborov [20].

For partial functions, tightness of the BCW simulation is known in *some* settings. For example, consider the Deutsch-Jozsa (DJ) problem, where the input is an n-bit string with the promise that its Hamming weight is either 0 or $n/2$, and DJ outputs -1 if the Hamming weight is $n/2$, and 1 otherwise. DJ has quantum query complexity 1 whereas the *exact* quantum communication complexity of $(\mathsf{DJ} \circ \oplus)$ is $\log n$. Note that it is unclear whether the $\log n$ factor loss here is additive or multiplicative.[2] Montanaro, Nishimura and Raymond [19] exhibited a partial function for which the BCW simulation is tight (up to constants) in the *exact* and *non-deterministic* quantum settings. They also observed the existence of a total function for which the BCW simulation is tight (up to constants) in the *unbounded-error* setting.

As far as we know, there was no (partial or total) Boolean-valued function f known prior to our work for which the *bounded-error* quantum communication complexity of $f \circ \bullet$ (i.e. $Q^{cc}(f \circ \bullet)$) is even $\omega(Q(f))$, where $\bullet$ is either $\wedge$ or $\oplus$.

In this paper, we exhibit the first *total function* witnessing the tightness of the BCW simulation in arguably the most well-known quantum model, which is the bounded-error model.

▶ **Theorem 2.** *There exists a total function $F : \{-1,1\}^n \to \{-1,1\}$ for which,*

$$Q^{cc}(F \circ \oplus) = \Theta(Q(F)\log n). \tag{1}$$

The statement above does not necessarily guarantee that a function exists that both satisfies Equation 1 and has bounded-error quantum query complexity (as a function of n) arbitrarily close to n. We answer this question by proving a more general result, from which Theorem 2 follows.

▶ **Theorem 3** (Main Theorem). *For any constant $0 < \delta < 1$, there exists a total function $F : \{-1,1\}^n \to \{-1,1\}$ for which $Q(F) = \Theta(n^\delta)$ and*

$$Q^{cc}(F \circ \oplus) = \Theta(Q(F)\log n).$$

1.1 Overview of our approach and techniques

To demonstrate the tightness of the BCW simulation for a total function in the quantum bounded-error setting we have to find a function F such that $Q^{cc}(F \circ \bullet) = \Theta(Q(F)\log n)$ for some choice of $\bullet$ (that is, either $\bullet$ is $\wedge$ or $\oplus$). This requires us to prove an upper bound of $Q(F)$ and a lower bound on $Q^{cc}(F \circ \bullet)$. We consider the case when $\bullet$ is the $\oplus$ function.

[2] Indeed, there are well-known situations where complexity of 1 vs. $\log n$ can be deceptive. The classical private-coin randomized communication complexity of Equality is $\Theta(\log n)$, whereas the public-coin cost is well known to be $O(1)$. Newman's Theorem shows that this difference in costs, in general, is *not multiplicative* but merely *additive*.

For the inner function the $\oplus$ function is preferred over the $\wedge$ function for one crucial reason: we have an analytical technique for proving lower bounds on $Q^{cc}(F \circ \oplus)$, due to Lee and Shraibman [18]. They reduced the problem of lower bounding the bounded-error quantum communication complexity of $(F \circ \oplus)$ to proving lower bounds on an analytic property of F, called its *approximate spectral norm*. The ϵ-approximate spectral norm of F, denoted by $\|\hat{F}\|_{1,\epsilon}$, is defined to be the minimum ℓ_1-norm of the coefficients of a polynomial that approximates F uniformly to error ϵ (see Definition 18). Lee and Shraibman [18] showed that $Q^{cc}(F \circ \oplus) = \Omega(\log \|\widehat{F}\|_{1,1/3})$. Thus, the lower bound on Theorem 3 follows immediately from our result below.

▶ **Theorem 4.** *For any constant $0 < \delta < 1$, there exists a total function $F : \{-1,1\}^n \to \{-1,1\}$ for which $Q(F) = \Theta(n^\delta)$ and*

$$\log\left(\|\widehat{F}\|_{1,1/3}\right) = \Theta(Q(F) \log n).$$

▶ Remark 5. It is interesting to note that, in contrast, it is well known that the extra $\log n$ factor does not appear for classical randomized query complexity. That is, if $R(F)$ denotes the bounded-error randomized query complexity of F, then $\log(\|\widehat{F}\|_{1,1/3}) = O(R(F))$. This indicates a significant structural difference between the classical and quantum query models.

There are not many techniques known to bound the approximate spectral norm of a function. This sentiment was expressed both in [18] and in the work of Ada, Fawzi and Hatami [2]. On the other hand, classical approximation theory offers tools to prove bounds on a simpler and better known concept called *approximate degree* which has been invaluable, particularly for quantum query complexity. The ϵ-approximate degree of f, denoted by $\widetilde{\deg}_\epsilon(f)$, is the minimum degree required by a real polynomial to uniformly approximate f to error ϵ (see Definition 17). Recently, two of the authors [11] devised a way of lifting approximate degree bounds to approximate spectral norm bounds. We first show here that technique works a bit more generally, to yield the following: let $\mathsf{ADDR}_{m,t} : \{-1,1\}^m \to [t]$ be a (possibly partial) addressing function (see Definition 13). For any function $f : \{-1,1\}^n \to \{-1,1\}$, define the (partial) function $f^{\mathsf{ADDR}_{m,t}} : \{-1,1\}^{n\times t} \times \{-1,1\}^{n\times m} \to \{-1,1\}$ as follows (formally defined in Definition 15):

$$f^{\mathsf{ADDR}_{m,t}}(x,y) = f\left(x_{1,\mathsf{ADDR}_{m,t}(y_1)}, x_{2,\mathsf{ADDR}_{m,t}(y_2)}, \ldots, x_{n,\mathsf{ADDR}_{m,t}(y_n)}\right).$$

Our main result on lower bounding the spectral norm is stated below.

▶ **Lemma 6** (extending [11]). *Let $t > 1$ be any integer, $\mathsf{ADDR}_{m,t}$ be any (partial) addressing function and $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Then,*

$$\log\left(\|\widehat{f^{\mathsf{ADDR}_{m,t}}}\|_{1,1/3}\right) = \Omega\left(\widetilde{\deg}(f)\log t\right).$$

The functions F constructed for the proof of Theorem 4 are completions of instances of $\mathsf{PARITY}^{\mathsf{ADDR}_{\ell,\ell}}$, and hence Lemma 6 yields lower bounds on the approximate spectral norm of F in terms of the approximate degree of PARITY (which is known to be maximal).

For the upper bound on $Q(F)$ we use two famous query algorithms - Grover's search [14] and the Bernstein-Vazirani algorithm [7]. The use of these algorithms for upper bounding $Q(F)$ is in the same taste as in the work of Ambainis and de Wolf [3] although their motivation was quite different than ours. Interestingly, Ambainis and de Wolf used their function to pin down the minimal *approximate degree* of a total Boolean function, all of whose input variables are influential.

1.2 Intuition behind the function construction

- From Theorem 1 it is known that for all Boolean functions f, $Q^{cc}(f \circ \oplus) \leq O(Q(f) \log n)$.
- In order to prove a matching lower bound, we construct a Boolean function F on n variables such that $\|\widehat{F}\|_{1,1/3} = 2^{\Omega(Q(F)\log n)}$ (Theorem 4). From Theorem 22, this shows that $Q^{cc}(F \circ \oplus) = \Omega(\log \|\widehat{F}\|_{1,1/3}) = \Omega(Q(F) \log n)$. We want to additionally ensure that $Q(F) = \Theta(n^{\delta})$ for a given constant $0 < \delta < 1$. A formal definition of F is given in Figure 1, we attempt to provide an overview on how we arrived at this function below.
- Assume δ is a constant that is least $1/2$, else the argument follows along similar lines by ignoring suitably many input variables when defining the function. A natural first attempt is to try to construct a composed function of the form $F = f^{\mathsf{ADDR}}$, for some addressing function ADDR (see Definition 13) with $\Omega(n^{1-\delta})$ many target bits, for which $Q(f^{\mathsf{ADDR}}) = \Theta(\widetilde{\deg}(f))$. For the lower bound we use Lemma 6 to show that $\log \|\widehat{f^{\mathsf{ADDR}}}\|_{1,1/3} = \Omega(\widetilde{\deg}(f) \log(n^{1-\delta})) = \Omega(\widetilde{\deg}(f) \log n)$.
- Given the upper bound target, we are led to a natural choice of addressing function. We refer the reader to Definition 13 for the definition of an addressing function and the selector function of an addressing function. Let $\mathsf{HADD}_{n^{1-\delta}}$ be the $(n^{1-\delta}, n^{1-\delta})$-addressing function defined as follows. Fix an arbitrary order on the $n^{1-\delta}$-bit Hadamard codewords (see Definition 12) , say $w_1, \ldots, w_{n^{1-\delta}}$. Define the selector function of $\mathsf{HADD}_{n^{1-\delta}}$, which we denote g, by $g(w_i) = i$ for all $i \in [n^{1-\delta}]$, and $g(x) = \star$ for $x \neq w_i$ for any $i \in [n^{1-\delta}]$.
- For any function f on $n^{\delta}/2$ bits, the *partial* function $f^{\mathsf{HADD}_{n^{1-\delta}}}$ on n inputs has quantum query complexity $O(Q(f) + n^{\delta}/2)$, as we sketch in the next step. We select f appropriately such that this is $\Theta(Q(f))$. Finally, we define the *total* function $F = f^{\mathsf{HADD}_{n^{1-\delta}}}$ to be the completion of $f^{\mathsf{HADD}_{n^{1-\delta}}}$ that evaluates to -1 on the non-promise inputs of $f^{\mathsf{HADD}_{n^{1-\delta}}}$.
- We choose the outer function to be $f = \mathsf{PARITY}_{n^{\delta}/2}$ to ensure $Q(F) = \Theta(n^{\delta})$. To prove the upper bound on $Q(F)$, we crucially use the Bernstein-Vazirani and Grover's search algorithms.
  - Run $n^{\delta}/2$ instances of the Bernstein-Vazirani algorithm [7], one on each block. This algorithm guarantees that if the address variables were all Hadamard codewords, then we would receive the correct indices of the target variables with probability 1, and just $n^{\delta}/2$ queries.
  - In the next step, we run Grover's search [14, 8] on two $n/2$-bit strings to test whether the output of the first step was correct. If it was correct, we succeed with probability 1, and proceed to query the $n^{\delta}/2$ selected target variables and output the parity of them. If it was not correct, Grover's search catches a discrepancy with probability at least $2/3$ and we output -1, succeeding with probability at least $2/3$ in this case.
  - The $n^{\delta}/2$ invocations of the Bernstein-Vazirani algorithm use a total of $n^{\delta}/2$ queries, Grover's search uses another $O(\sqrt{n})$ queries, and the final parity (if Grover's search outputs that the strings are equal) uses another $n^{\delta}/2$ queries, for a cumulative total of $O(n^{\delta} + \sqrt{n}) = O(n^{\delta})$ queries (recall that we assume $\delta \geq 1/2$).

1.3 Other implications of our result

Zhang [26] showed that for all Boolean functions f, there must exist gadgets $\bullet_i$, each either $\wedge$ or $\vee$, such that $Q^{cc}(f(\bullet_1, \ldots, \bullet_n)) = \Omega(\mathrm{poly}(Q(f)))$. For monotone f, they showed that either $Q^{cc}(f \circ \wedge) = \Omega(\mathrm{poly}(Q(f)))$ or $Q^{cc}(f \circ \vee) = \Omega(\mathrm{poly}(Q(f)))$. They also state that it is unclear how tight the BCW simulation is. Our result implies that there exists a function for which it is tight up to constants (on composition with $\oplus$).

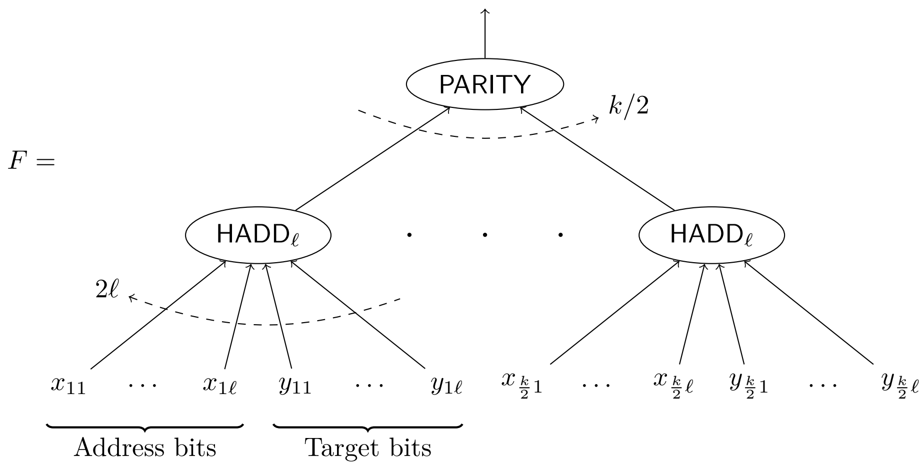

Figure 1 $k = n^{\delta}, \ell = n^{1-\delta}$. If the address bits of an input to the r'th HADD_ℓ is the j'th Hadamard codeword, then y_{rj} is selected. If on an input, there exists at least one HADD_ℓ for which the address bits do not correspond to a Hadamard codeword, F outputs -1. Else it outputs the parity of the $k/2$ selected target bits.

Another implication of our result is related to the Entropy Influence Conjecture, which is an interesting question in the field of analysis of Boolean functions, posed by Friedgut and Kalai [13]. This conjecture is open for general functions. A much weaker version of this conjecture is called the Min-Entropy Influence Conjecture. For the statement of the conjecture we need to consider the Fourier expansion of Boolean functions $f : \{-1,1\}^n \to \{-1,1\}$ as

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S(x),$$

where $\{\chi_S : S \subseteq [n]\}$ are the *parity* functions ($\chi_S(x) = \Pi_{i \in S} x_i$, when $x = (x_1, \ldots, x_n) \in \{-1,1\}^n$) and $\{\widehat{f}(S) : S \subseteq [n]\}$ are the corresponding Fourier coefficients.

▶ **Conjecture 7** (Min-Entropy Influence Conjecture). *For any Boolean function $f : \{-1,1\}^n \to \{-1,1\}$ there exists a non-zero Fourier coefficient $\widehat{f}(S)$ such that*

$$\log\left(1/|\widehat{f}(S)|\right) = O(I(f)),$$

where $I(f)$ denotes the influence (or average sensitivity) of f ($I(f) = \sum_{S \subseteq [n]} |S| \widehat{f}(S)^2$).

While this conjecture is also open, some attempts have been made to prove it and various implications of it [5, 16]. One interesting implication of the Min-Entropy Influence Conjecture that is still open is that the min-entropy of the Fourier spectrum (that is, $\log\left(1/\max_{S \subseteq [n]} |\widehat{f}(S)|\right)$) is less than $O(Q(f))$. In [5] using a primal-dual technique it was shown that the min-entropy of the Fourier spectrum is less than a constant times $\log(\|\hat{f}\|_{1,\epsilon})$, where the constant depends on ϵ. Thus if it were the case that $\log(\|\hat{f}\|_{1,\epsilon}) = O(Q(f))$, we would have upper bounded the min-entropy of Fourier spectrum by $O(Q(f))$. As $\widetilde{\deg}(f) \leq 2Q(f)$ [6], the following was stated in [5] as a possible approach and was left as an open problem.

▶ **Question 8** ([5, Section 4]). *Is it true that for all Boolean functions $f : \{-1,1\}^n \to \{-1,1\}$,*

$$\log(\|\hat{f}\|_{1,\epsilon}) = O(\widetilde{\deg}(f))?$$

While their conjecture holds true for certain special classes of functions like the symmetric functions (proof given in Appendix A), our result in this paper nullifies this approach for general Boolean functions, since Theorem 4 yields the following along with the fact that $\widetilde{\deg}(f) \leq 2Q(f)$ [6].

▶ **Theorem 9.** *For any constant $0 < \delta < 1$, there exists a total function $F : \{-1,1\}^n \to \{-1,1\}$ for which $\widetilde{\deg}(F) = O(n^\delta)$ and*

$$\log \|\widehat{F}\|_{1,1/3} = \Omega(\widetilde{\deg}(F) \log n).$$

2 Preliminaries

For any positive integer n, we denote the set $\{1, \ldots, n\}$ by $[n]$. For $d \leq n$ we use the notation $\binom{n}{\leq d} := \binom{n}{0} + \cdots + \binom{n}{d}$. Note that $\binom{n}{\leq d} < (n+1)^d$.

In this section we review the necessary preliminaries. We first review some basic notions of Fourier analysis on the Boolean cube. Consider the vector space of functions from $\{-1,1\}^n$ to $\mathbb{R}$, equipped with an inner product defined by

$$\langle f, g \rangle := \mathbb{E}_{x \in \{-1,1\}^n}[f(x)g(x)] = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x)g(x)$$

for every $f, g : \{-1,1\}^n \to \mathbb{R}$. For any set $S \subseteq [n]$, define the associated *parity* function χ_S by $\chi_S(x) = \prod_{i \in S} x_i$. The set of parity functions $\{\chi_S : S \subseteq [n]\}$, forms an orthonormal basis for this vector space. Thus, every function $f : \{-1,1\}^n \to \mathbb{R}$ has a unique multilinear expression as $f = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S$. The coefficients $\{\widehat{f}(S) : S \subseteq [n]\}$ are called the *Fourier coefficients* of f. We also use the notation PARITY_n to denote the function $\chi_{[n]} : \{-1,1\}^n \to \{-1,1\}$.

▶ **Fact 10** (Parseval's Identity). *For any function $f : \{-1,1\}^n \to \mathbb{R}$, we have $\sum_{S \subseteq [n]} \widehat{f}(S)^2 = \frac{\sum_{x \in \{-1,1\}^n} f(x)^2}{2^n}$.*

▶ **Definition 11** (Spectral Norm). *For any function $f : \{-1,1\}^n \to \mathbb{R}$, define its* spectral norm*, which we denote $\|\widehat{f}\|_1$, to be the sum of absolute values of the Fourier coefficients of f. That is, $\|\widehat{f}\|_1 := \sum_{S \subseteq [n]} |\widehat{f}(S)|$.*

▶ **Definition 12** (Hadamard Codeword). *If an ℓ-bit string $(x_1, \ldots, x_\ell) \in \{-1,1\}^\ell$ (alternatively, view the indices of x as subsets of $[\log \ell]$) is of the form $x_S = \prod_{i \in S} z_i$ for all $S \subseteq [\log \ell]$ for some $z \in \{-1,1\}^{\log \ell}$, then define such an $x = x_1 \ldots x_\ell$ to be the ℓ-*bit Hadamard codeword *$h(z)$ of the $(\log \ell)$-bit string z.*

2.1 Addressing functions

▶ **Definition 13** ((m,k)-addressing function). *We define a (partial) function $f : \{-1,1\}^{m+k} \to \{-1,1,\star\}$ to be an (m,k)-*addressing function *if there exists $g : \{-1,1\}^m \to \{[k] \cup \star\}$ such that*

- *$f(x_1, \ldots, x_m, y_1, \ldots, y_k) = y_{g(x_1,\ldots,x_m)}$ if $g(x_1, \ldots, x_m) \in [k]$, $f(x_1, \ldots, x_m, y_1, \ldots, y_k) = \star$ otherwise.*
- *For all $j \in [k]$, there exists $(x_1, \ldots, x_m) \in \{-1,1\}^m$ such that $g(x_1, \ldots, x_m) = j$.*

We call the variables $\{x_1, \ldots, x_m\}$ the address variables *and the variables $\{y_1, \ldots, y_k\}$ the* target variables*. The function g is called the* selector *function of f.*

▶ **Definition 14** (Indexing Function). *The Indexing function, which we denote by* IND_k*, is a* $(k, 2^k)$*-addressing function defined by* $\mathsf{IND}(x_1, \ldots, x_k, y_1, \ldots, y_{2^k}) = y_{\mathrm{bin}(x)}$*, where* $\mathrm{bin}(x)$ *denotes the integer represented by the binary string* $x_1, \ldots, x_k$*.*

▶ **Definition 15** (Composition with addressing functions). *For any function* $f : \{-1,1\}^n \to \{-1,1\}$ *and an* (m,k)*-addressing function* ADDR*, define the (partial) function* $f^{\mathsf{ADDR}} : \{-1,1\}^{n(m+k)} \to \{-1,1,\star\}$ *by* $f^{\mathsf{ADDR}}(x_1, y_1, \ldots, x_n, y_n) =$

$$\begin{cases} f(\mathsf{ADDR}(x_1,y_1),\ldots,\mathsf{ADDR}(x_n,y_n)) & \text{if } \forall i \in [n], \mathsf{ADDR}(x_i,y_i) \in \{-1,1\} \\ \star & \text{otherwise.} \end{cases}$$

where $x_i \in \{-1,1\}^m$ *and* $y_i \in \{-1,1\}^k$ *for all* $i \in [n]$*.*

▶ **Definition 16** (Hadamard Addressing Function). *We define the* Hadamard addressing function*, which we denote* $\mathsf{HADD}_\ell : \{-1,1\}^{2\ell} \to \{-1,1,\star\}$*, as follows. Fix an arbitrary order on the* ℓ*-many Hadamard codewords of* $(\log \ell)$*-bit strings, say* $w_1, \ldots, w_\ell$*. Define the selector function of* HADD_ℓ *by*

$$g(x) = \begin{cases} i & \text{if } x = w_i \text{ for some } i \in [\ell] \\ \star & \text{otherwise.} \end{cases}$$

Note that HADD_ℓ *is an* (ℓ,ℓ)*-addressing function.*

2.2 Polynomial approximation

▶ **Definition 17** (Approximate Degree). *The* ϵ-approximate degree *of* $f : \{-1,1\}^n \to \{-1,1,\star\}$*, denoted by* $\widetilde{\deg}_\epsilon(f)$ *is defined to be the minimum degree of a real polynomial* $p : \{-1,1\}^n \to \mathbb{R}$ *that satisfies* $|p(x) - f(x)| \le \epsilon$ *for all* $x \in \{-1,1\}^n$ *for which* $f(x) \in \{-1,1\}$.[3] *That is,*

$$\widetilde{\deg}_\epsilon(f) := \min\{d : \deg(p) \le d, |p(x) - f(x)| \le \epsilon\ \forall x \in \{-1,1\}^n \text{ for which } f(x) \in \{-1,1\}\}.$$

Henceforth, we will use the notation $\widetilde{\deg}(f)$ to denote $\widetilde{\deg}_{1/3}(f)$.

▶ **Definition 18** (Approximate Spectral Norm). *The* approximate spectral norm *of a function* $f : \{-1,1\}^n \to \{-1,1,\star\}$*, denoted by* $\|\widehat{f}\|_{1,\epsilon}$ *is defined to be the minimum spectral norm of a real polynomial* $p : \{-1,1\}^n \to \mathbb{R}$ *that satisfies* $|p(x) - f(x)| \le \epsilon$ *for all* $x \in \{-1,1\}^n$ *for which* $f(x) \in \{-1,1\}$*.*

$$\|\widehat{f}\|_{1,\epsilon} := \min\{\|\widehat{p}\|_1 : |p(x) - f(x)| \le \epsilon \text{ for all } x \in \{-1,1\}^n \text{ for which } f(x) \in \{-1,1\}\}.$$

▶ **Lemma 19** ([10]). *Let* $f : \{-1,1\}^n \to \{-1,1\}$ *be a total function. Then for all constants* $0 < \delta, \epsilon < 1$ *we have* $\widetilde{\deg}_\epsilon(f) = \Theta(\widetilde{\deg}_\delta(f))$*. The constant hidden in the* Θ *notation depends on* δ *and* ϵ*.*

The following is a standard upper bound on the approximate spectral norm of a Boolean function in terms of its approximate degree.

---

[3] When dealing with partial functions, another notion of approximation is sometimes considered, where the approximating polynomial p is required to have bounded values even on the non-promise inputs of f. For the purpose of this paper, we do not require this constraint.

▷ Claim 20. For all total functions $f : \{-1,1\}^n \to \{-1,1\}$, we have $\log \|\widehat{f}\|_{1,1/3} = O(\widetilde{\deg}(f) \log n)$.

Proof. Let d denote the approximate degree of f. Take any $1/3$-approximating polynomial of degree d, say p, to f. Then, $\sum_{S \subseteq [n]} \left|\widehat{p}(S)\right| \leq \sqrt{\binom{n}{\leq d}} \cdot \sqrt{\sum_{S:|S| \leq d} \widehat{p}(S)^2} \leq 4/3 \cdot (n+1)^{d/2} = 2^{O(d \log n)}$, where the first inequality follows by the Cauchy-Schwarz inequality, the second inequality follows by Parseval's identity (Fact 10) and the fact that the absolute value of p is at most $4/3$ for any input $x \in \{-1,1\}^n$. ◁

It is easy to exhibit functions $f : \{-1,1\}^n \to \{-1,1\}$ such that $\log \|\widehat{f}\|_{1,1/3} = \Omega(\widetilde{\deg}(f))$. Bent functions satisfy this bound, for example.

Building upon ideas in [17], the approximate spectral norm of $f \circ \mathsf{IND}_1$ was shown to be bounded below by $2^{\Omega(\widetilde{\deg}(f))}$ in [11].

▶ **Theorem 21** ([11]). *Let $f : \{-1,1\}^n \to \{-1,1\}$ be any function. Then $\|\widehat{f \circ \mathsf{IND}_1}\|_{1,1/3} \geq 2^{c \cdot \widetilde{\deg}_{2/3}(f)}$ for any constant $c < 1 - 3/\widetilde{\deg}_{2/3}(f)$.*

2.3 Communication complexity

The classical model of communication complexity was introduced by Yao in [24]. In this model two parties, say Alice and Bob, wish to compute a function whose output depends on both their inputs. Alice is given an input $x \in \mathcal{X}$, Bob is given $y \in \mathcal{Y}$, and they want to jointly compute the value of a given function $F(x, y)$ by communicating with each other. Alice and Bob individually have unbounded computational power and the number of bits communicated is the resource we wish to minimize. Alice and Bob communicate using a *protocol* that is agreed upon in advance. In the randomized model, Alice and Bob have access to unlimited public random bits and the goal is to compute the correct value of $F(x, y)$ with probability at least $2/3$ for all inputs $(x, y) \in \mathcal{X} \times \mathcal{Y}$. The *bounded-error randomized communication complexity* of a function F, denoted $R^{cc}(F)$, is the number of bits that must be communicated in the worst case by any randomized protocol to compute the correct value of the function $F(x, y)$, with probability at least $2/3$, for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

The quantum model of communication complexity was introduced by Yao in [25]. We refer the reader to the survey [23] for details. The *bounded-error quantum communication complexity* of a function F, denoted $Q^{cc}(F)$ is the number of bits that must be communicated by any quantum communication protocol in the worst case to compute the correct value of the function $F(x, y)$, with probability at least $2/3$, for every (x, y) in domain of F. Buhrman, Cleve and Wigderson [9] observed a quantum simulation theorem, which gives an upper bound on the bounded-error quantum communication complexity of a composed function of the form $f \circ \wedge$ or $f \circ \oplus$ in terms of the bounded-error quantum query complexity of f (see Theorem 1).

Lee and Shraibman [18] showed that the bounded-error quantum communication complexity of $f \circ \oplus$ is bounded below by the logarithm of the approximate spectral norm of f. Also see [11] for an alternate proof.

▶ **Theorem 22** ([18]). *For any Boolean function $f : \{-1,1\}^n \to \{-1,1\}$,*

$$Q^{cc}(f \circ \oplus) = \Omega(\log \|\widehat{f}\|_{1,1/3}).$$

We remark that to the best of our knowledge, it is not known whether the same lower bound holds on the bounded-error quantum communication complexity of $f \circ \wedge$.

3 Proof of Theorem 3

In this section, we prove Theorem 3. We first formally define the function we use.

3.1 Definition of the function

If $\delta < 1/2$, then ignore the last $n - 2n^{2\delta}$ bits of the input, and define the following function on the first $2n^{\delta}$ bits of the input. The same argument as in Sections 3.2 and 3.3 give the required bounds for Theorem 3 and Theorem 4. Hence, we may assume without loss of generality that $\delta \geq 1/2$.

Define the partial function $f : \{-1,1\}^n \to \{-1,1,\star\}$ by $f = \mathsf{PARITY}_{n^{\delta}/2}^{\mathsf{HADD}_{n^{1-\delta}}}$. Define F to be the completion of f that evaluates to -1 on the non-promise domain of f (see Figure 1).

3.2 Upper bound

In this section, we prove the following.

▷ Claim 23. For $F : \{-1,1\}^n \to \{-1,1\}$ defined as in Section 3.1, we have $Q(F) = \Theta(n^{\delta})$.

The upper bound follows along the lines of a proof in [3], and the lower bound just uses the fact that F is at least as hard as $\mathsf{PARITY}_{n^{\delta}/2}$.

Proof. For convenience, set $\ell = n^{1-\delta}$ and $k = n^{\delta}$. Recall that an input to F is viewed as $(x_{11}, \ldots, x_{1,\ell}, y_{11}, \ldots, y_{1,\ell}, \ldots, x_{\frac{k}{2}1}, \ldots, x_{\frac{k}{2}\ell}, y_{\frac{k}{2}1}, \ldots, y_{\frac{k}{2}\ell})$. The following is an $O(n^{\delta})$-query quantum algorithm computing F. Note that since $\delta \geq 1/2$, we have $k = \Omega(\ell)$.

1. Run $k/2$ instances of the Bernstein-Vazirani algorithm on the $(k/2)$-many input strings $(x_{11}, \ldots, x_{1\ell}), \ldots, (x_{\frac{k}{2}1}, \ldots, x_{\frac{k}{2}\ell})$ to obtain $k/2$ strings $z_1, \ldots, z_{k/2}$.
2. Run Grover's search [14, 8] to check equality of the two strings: $h(z_1), \ldots, h(z_\ell)$ and $x_{11}, \ldots, x_{1\ell}, \ldots, x_{\frac{k}{2}1}, \ldots, x_{\frac{k}{2}\ell}$, i.e. to check whether the addressing bits of the input are indeed all Hadamard codewords which are output by the first step.
3. If the step above outputs that the strings are equal, then query the $k/2$ selected variables and output their parity. Else, output -1.
   - If the input was indeed of the form as claimed in the first step, then Bernstein-Vazirani outputs the correct $z_1, \ldots, z_\ell$ with probability 1, and Grover's search verifies that the strings are equal with probability 1. Hence the algorithm is correct with probability 1 in this case.
   - If the input was not of the claimed form, then the two strings for which equality is to be checked in the second step are not equal. Grover's search catches a discrepancy with probability at least 2/3. Hence, the algorithm is correct with probability at least 2/3 in this case.

The correctness of the algorithm is argued above, and the cost is $k/2$ queries for the first step, $O(\sqrt{k\ell})$ queries for the second step, and at most $k/2$ for the third step. Thus, we have $Q(F) = O(k + \sqrt{k\ell}) = O(k)$, since $k = \Omega(\ell)$. The upper bound in the lemma follows.

For the lower bound, we argue that F is at least as hard as $\mathsf{PARITY}_{k/2}$. To see this formally, set all the address variables such that the selected target variables are the first target variable in each block. Under this restriction, F equals $\mathsf{PARITY}(y_{11}, \ldots, y_{\frac{k}{2}1})$. Thus any quantum query algorithm computing F must be able to compute $\mathsf{PARITY}_{k/2}$, and thus $Q(F) = \Omega(k)$. ◁

▶ Remark 24. The same argument as above works when the function f is defined to be g^{HADD_ℓ} for any $g : \{-1,1\}^{n^\delta} \to \{-1,1\}$ satisfying $\widetilde{\deg}(g) = \Omega(n^\delta)$, and F is the completion of f that evaluates to -1 on all non-promise inputs. The same proof of Theorem 3 also goes through, but we fix $g = \mathsf{PARITY}_{n^\delta/2}$ for convenience.

3.3 Lower bound

In this section, we first prove Lemma 6. We require the following observation.

▶ **Observation 25.** *For any $S \subseteq [n]$ and any $j \in S$, we have $\mathbb{E}_{x_j \sim \{-1,1\}}[\chi_S(x)] = 0$, where x_j is distributed uniformly over $\{-1,1\}$.*

Proof of Lemma 6. Let $F = f^{\mathsf{ADDR}_{m,t}}$. Recall that our goal is to show that $\log \|\widehat{F}\|_{1,1/3} = \Omega(\widetilde{\deg}(f) \log t)$. We may assume $\widetilde{\deg}(f) \geq 1$, because the lemma is trivially true otherwise.

Towards a contradiction, suppose there exists a polynomial P of spectral norm strictly less than $2^{\frac{1}{10}\widetilde{\deg}_{0.99}(f)\log t}$ uniformly approximating F to error $1/3$ on the promise inputs (recall that from Lemma 19, we have $\widetilde{\deg}(f) = \Theta(\widetilde{\deg}_{0.99}(f))$).

Let ν be a distribution on the address bits of $\mathsf{ADDR}_{m,t}$ such that ν is supported only on assignments to the address variables that do not select $\star$, and is the uniform distribution over these assignments. Let $\mu = \nu^n$ be the product distribution over the address bits of the addressing functions in F.

- For any assignment z of the address variables from the support of μ, define a relevant (target) variable, with respect to z, to be one that is selected by z. Analogously, define a target variable to be irrelevant if it is not selected by z. Define a monomial to be relevant if it does not contain irrelevant variables, and irrelevant otherwise.
- Note that for any target variable, the probability with which it is selected is exactly $1/t$.
- In the analysis that follows in this proof, we are interested in the *set of target variables* that are present in a monomial χ_S, which we denote by S_{target}. Also define the *target-degree* of S to be $|S_{\text{target}}|$, i.e. the degree of a monomial χ_S when restricted to the target variables.
- Thus under any assignment z drawn from μ, for any monomial of the function P of target-degree $t \geq \widetilde{\deg}_{0.99}(f)$, the probability that it is relevant is at most $1/t^{\widetilde{\deg}_{0.99}(f)}$. Hence

$$\begin{aligned}
&\mathop{\mathbb{E}}_{z\sim\mu}[\ell_1\text{-norm of relevant monomials w.r.t. } z \text{ in } P \text{ of target-degree } \geq \widetilde{\deg}_{0.99}(f)] \\
&= \sum_{|S_{\text{target}}| \geq \widetilde{\deg}_{0.99}(f)} |\widehat{P}(S)| \Pr_{z\sim\mu}[\chi_S \text{ is relevant w.r.t. } z] \\
&\leq \max_{|S_{\text{target}}| \geq \widetilde{\deg}_{0.99}(f)} \{\Pr_{z\sim\mu}[\chi_S \text{ is relevant w.r.t. } z]\} \cdot \|\widehat{P}\|_1 \\
&< \frac{1}{t^{\widetilde{\deg}_{0.99}(f)}} \cdot 2^{\frac{1}{10}\widetilde{\deg}_{0.99}(f)\log t} = 2^{(-\frac{9}{10})\widetilde{\deg}_{0.99}(f)\log t} < \frac{3}{5},
\end{aligned}$$

  where the second last inequality holds because of Claim 20 and the last inequality holds because $t \geq 2$ and $\widetilde{\deg}_{0.99}(f) \geq 1$.
- Fix an assignment to the address variables from the support of μ such that under this assignment, the ℓ_1-norm of relevant monomials in P of degree $\geq \widetilde{\deg}_{0.99}(f)$ is less than $3/5$.
- Note that under this assignment (in fact under any assignment in the support of μ), the restricted F is just the function f on the n variables selected by the addressing functions. Denote by P_1 the polynomial on the target variables obtained from P by fixing address variables as per this assignment.

- Drop the relevant monomials of degree $\geq \widetilde{\deg}_{0.99}(f)$ from P_1 to get a polynomial P_2, which uniformly approximates the restricted F (which is f on n variables) to error $1/3 + 3/5 < 0.99$.
- Take expectation over irrelevant variables (from the distribution where each irrelevant variable independently takes values uniformly from $\{-1, 1\}$). Under this expectation, the value of F does not change (since irrelevant variables do not affect F's output by definition), and all irrelevant monomials of P_2 become 0 (using Observation 25 and linearity of expectation). Hence, under this expectation we have $\mathbb{E}[P_2] = P_3$, where P_3 is a polynomial of degree strictly less than $\widetilde{\deg}_{0.99}(f)$. Furthermore, P_3 uniformly approximates f to error less than 0.99 which is a contradiction. ◀

As a corollary of Lemma 6, we obtain a lower bound on the approximate spectral norm of F, where F is defined as in Section 3.1. This yields a proof of Theorem 4.

Proof of Theorem 4. Construct F as in Section 3.1. Claim 23 implies $Q(F) = \Theta(n^\delta)$.

Let $f = \mathsf{PARITY}_{n^\delta/2}^{\mathsf{HADD}_{n^{1-\delta}}}$. Lemma 6 implies that $\|\widehat{f}\|_{1,1/3} = \Omega(n^\delta \log n)$.

Since F is a completion of f, we have $\|\widehat{F}\|_{1,1/3} = \Omega(n^\delta \log n)$, which proves the lower bound in Theorem 4. The upper bound follows from Theorem 1. ◀

We are now ready to prove our main theorem.

Proof of Theorem 3. It immediately follows from Theorem 4 and Theorem 22. ◀

4 Conclusions

We conclude with the following points: first, we find our main result somewhat surprising that simulating a query algorithm by a communication protocol in the quantum context has a larger overhead than in the classical context. Second, it is remarkable that this relatively fine overhead of $\log n$ can be detected using analytic techniques that are an adaptation of the generalized discrepancy method. Third, the function that we used in this work is an XOR function. Study of this class of functions is proving to be very insightful. A recent example is the refutation of the log-approximate-rank conjecture [12] and even its quantum version [4, 22]. Our work further advocates the study of XOR functions.

An open question that remains is whether there exists a Boolean function $F : \{-1, 1\}^n \to \{-1, 1\}$ such that $Q^{cc}(F \circ \wedge) = \Omega(Q(F) \log n)$. Or does there exist a better quantum communication protocol for $(F \circ \wedge)$ that does not incur the logarithmic factor loss?

It is easy to verify that the constructions of F that yield Theorem 4 for any fixed constant $0 < \delta < 1$, also satisfy $\widetilde{\deg}(F) = \Theta(n^\delta)$. Recall that Theorem 9 states that such functions F satisfy $\log \|\widehat{F}\|_{1,1/3} = \Omega(\widetilde{\deg}(F) \log n)$. This gives a negative answer to Question 8 ([5, Section 4]), where it was asked if any degree-d approximating polynomial to a Boolean function of approximate degree d has spectral norm at most $2^{O(d)}$ (it is interesting to note that their conjecture holds true for symmetric functions, which we prove in Appendix A). Thus to prove min-entropy of the Fourier spectrum of a Boolean function is upper bounded by approximate degree, it cannot follow from their observation that min-entropy is upper bounded by the logarithm of the approximate spectral norm. The following remains an interesting and important open problem: (how) can one prove that the min-entropy of the Fourier spectrum of a Boolean function is upper bounded by a constant multiple of its approximate degree? Such an inequality is implied by the Fourier Entropy Influence (FEI) Conjecture.

References

1 Scott Aaronson and Andris Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1(1):47–79, 2005.

2 Anil Ada, Omar Fawzi, and Hamed Hatami. Spectral norm of symmetric functions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012. Proceedings*, pages 338–349, 2012.

3 Andris Ambainis and Ronald de Wolf. How low can approximate degree and quantum query complexity be for total boolean functions? *Computational Complexity*, 23(2):305–322, 2014.

4 Anurag Anshu, Naresh Goud Boddu, and Dave Touchette. Quantum log-approximate-rank conjecture is also false. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 982–994, 2019.

5 Srinivasan Arunachalam, Sourav Chakraborty, Michal Koucký, Nitin Saurabh, and Ronald de Wolf. Improved bounds on fourier entropy and min-entropy. In *37th International Symposium on Theoretical Aspects of Computer Science, STACS 2020, March 10-13, 2020, Montpellier, France*, pages 45:1–45:19, 2020. `doi:10.4230/LIPIcs.STACS.2020.45`.

6 Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.

7 Ethan Bernstein and Umesh V. Vazirani. Quantum complexity theory. *SIAM J. Comput.*, 26(5):1411–1473, 1997.

8 Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 305:53–74, 2002.

9 Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 63–68, 1998.

10 Harry Buhrman, Ilan Newman, Hein Röhrig, and Ronald de Wolf. Robust polynomials and quantum algorithms. *Theory Comput. Syst.*, 40(4):379–395, 2007.

11 Arkadev Chattopadhyay and Nikhil S. Mande. A lifting theorem with applications to symmetric functions. In *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2017, December 11-15, 2017, Kanpur, India*, pages 23:1–23:14, 2017.

12 Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019.*, pages 42–53, 2019.

13 Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society*, 124(10):2993–3002, 1996.

14 Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 212–219, 1996.

15 Peter Høyer and Ronald de Wolf. Improved quantum communication complexity bounds for disjointness and equality. In *STACS 2002, 19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes - Juan les Pins, France, March 14-16, 2002, Proceedings*, pages 299–310, 2002.

16 Esty Kelman, Guy Kindler, Noam Lifshitz, Dor Minzer, and Muli Safra. Revisiting bourgain-kalai and fourier entropies. *CoRR*, abs/1911.10579, 2019. `arXiv:1911.10579`.

17 Matthias Krause and Pavel Pudlák. On the computational power of depth-2 circuits with threshold and modulo gates. *Theor. Comput. Sci.*, 174(1-2):137–156, 1997.

18 Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–398, 2009.

19 Ashley Montanaro, Harumichi Nishimura, and Rudy Raymond. Unbounded-error quantum query complexity. *Theor. Comput. Sci.*, 412(35):4619–4628, 2011.

20 Alexander A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

21 Alexander A. Sherstov. Algorithmic polynomials. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 311–324, 2018.

22 Makrand Sinha and Ronald de Wolf. Exponential separation between quantum communication and logarithm of approximate rank. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 966–981, 2019.

23 Ronald de Wolf. Quantum communication and complexity. *Theor. Comput. Sci.*, 287(1):337–353, 2002. doi:10.1016/S0304-3975(02)00377-8.

24 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 209–213, 1979.

25 Andrew Chi-Chih Yao. Quantum circuit complexity. In *34th Annual Symposium on Foundations of Computer Science, Palo Alto, California, USA, 3-5 November 1993*, pages 352–361, 1993.

26 Shengyu Zhang. On the tightness of the Buhrman-Cleve-Wigderson simulation. In *Algorithms and Computation, 20th International Symposium, ISAAC 2009, Honolulu, Hawaii, USA, December 16-18, 2009. Proceedings*, pages 434–440, 2009.

A Upper bound on the approximate spectral norm of symmetric functions

Recall from Section 1.3 that Theorem 9 gives a negative answer to Question 8, where it was asked if for all Boolean functions of approximate degree d, there exists an approximating polynomial with spectral norm $2^{O(d)}$. We show in this section that the upper bound in Question 8 does hold true for symmetric functions.

▶ **Definition 26** (Multilinear Polynomial)**.** *A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is a multilinear polynomial if ϕ is of the form:*

$$\phi(x_1, \ldots, x_n) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$$

where $a_S \in \mathbb{R}$.

▶ **Definition 27** (Spectral Norm of a Multilinear Polynomial)**.** *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a multilinear polynomial of the form $\phi(x_1, \ldots, x_n) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$. The spectral norm of ϕ, denoted by $\|\phi\|_1$, is defined as*

$$\|\phi\|_1 = \sum_{S \subseteq [n]} |a_S|.$$

▶ **Fact 28** (Properties of Spectral Norm of Multilinear Polynomials)**.** *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be any symmetric polynomials and let $\alpha \in \mathbb{R}$ be any real number. Then,*

1. $\|\alpha f\|_1 = |\alpha| \|f\|_1$,
2. $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$,
3. $\|fg\|_1 \leq \|f\|_1 \|g\|_1$.

▶ **Lemma 29.** *Let $S \subseteq [n]$ and $\chi_S : \{-1, 1\}^n \to \mathbb{R}$ be the symmetric multilinear polynomial defined as*

$$\chi_S(x_1, \ldots, x_n) = \prod_{i \in S} \frac{(1 - x_i)}{2}.$$

Then $\|\chi_S\|_1 = 1$.

Proof. Since for all $i \in [n]$, the spectral norm of $\frac{(1-x_i)}{2} = 1$, the proof follows from 28 (3). ◀

▶ **Definition 30** (Symmetric Multilinear Polynomial). *A multilinear polynomial $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be symmetric if $\phi(x_1, \ldots, x_n) = \phi(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for all $(x_1, \ldots, x_n) \in X$ and $\sigma \in S_n$.*

Sherstov [21] showed the following upper bound on the spectral norm of symmetric multilinear polynomials.

▷ Claim 31 ([21]). Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a symmetric multilinear polynomial. Then

$$\|\phi\|_1 \leq 8^{\deg(\phi)} \max_{x \in \{0,1\}^n} |\phi(x)|.$$

▶ **Lemma 32.** *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a symmetric Boolean function. Then*

$$\log(\|f\|_{1,1/3}) = O(\widetilde{\deg}(f)).$$

Proof. Let $f' : \{0,1\}^n \to \{-1,1\}$ be defined as $f'(x_1, \ldots, x_n) = f\left(\frac{1-x_1}{2}, \ldots, \frac{1-x_n}{2}\right)$. It is not hard to show, since we have done a linear transformation on the input domain, that $\widetilde{\deg}(f') = \widetilde{\deg}(f)$. Let p' be a polynomial that 1/3-approximates the symmetric function f'. By symmetrization we can assume that p' is symmetric, and is of the form $p'(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$.

Define the polynomial $p : \{-1,1\}^n \to \mathbb{R}$ as follows:

$$p(x_1, \ldots, x_n) = p'\left(\frac{1-x_1}{2}, \ldots, \frac{1-x_n}{2}\right).$$

Clearly p is a 1/3-approximation to f since p' is a 1/3-approximation to f' and we can write

$$p(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} \frac{(1-x_i)}{2}.$$

Thus we can upper bound the ℓ_1-norm of p as follows:

$$\|p\|_1 = \left\| \sum_{S \subseteq [n]} a_S \prod_{i \in S} \frac{(1-x_i)}{2} \right\|_1 \leq \sum_{S \subseteq [n]} \left\| a_S \prod_{i \in S} \frac{(1-x_i)}{2} \right\|_1 \tag{2}$$

$$\leq \sum_{S \subseteq [n]} |a_S| \left\| \prod_{i \in S} \frac{(1-x_i)}{2} \right\|_1 \tag{3}$$

$$= \sum_{S \subseteq [n]} |a_S| = \|p'\|_1, \tag{4}$$

where Equation (2) follows from Fact 28 (2), Equation (3) follows from Fact 28 (1) and Equation (4) follows from 29.

Hence, $\log(\|f\|_{1,1/3}) \leq \log(\|p\|_1) \leq \log(\|p'\|_1) = O(\deg(p'))$, where the last equality follows by Claim 31 since p' is symmetric. Since p' was assumed to have degree $\deg(p') = \widetilde{\deg}(f)$, the lemma follows. ◀