One-Way Communication Complexity of Partial XOR Functions

Vladimir V. Podolskii
Tufts University, Medford, MA, USA

Dmitrii Sluch
Nebius, Tel Aviv, Israel

Abstract

Boolean function $F(x, y)$ for $x, y \in \{0,1\}^n$ is an XOR function if $F(x, y) = f(x \oplus y)$ for some function f on n input bits, where $\oplus$ is a bit-wise XOR. XOR functions are relevant in communication complexity, partially for allowing the Fourier analytic technique. For total XOR functions, it is known that deterministic communication complexity of F is closely related to parity decision tree complexity of f. Montanaro and Osbourne (2009) observed that one-way communication complexity $\mathrm{D}^{\to}_{\mathrm{cc}}(F)$ of F is exactly equal to non-adaptive parity decision tree complexity $\mathrm{NADT}^{\oplus}(f)$ of f. Hatami et al. (2018) showed that unrestricted communication complexity of F is polynomially related to parity decision tree complexity of f.

We initiate the study of a similar connection for partial functions. We show that in the case of one-way communication complexity whether these measures are equal, depends on the number of undefined inputs of f. More precisely, if $\mathrm{D}^{\to}_{\mathrm{cc}}(F) = t$ and f is undefined on at most $O\left(\frac{2^{n-t}}{\sqrt{n-t}}\right)$ inputs, then $\mathrm{NADT}^{\oplus}(f) = t$. We also provide stronger bounds in extreme cases of small and large complexity.

We show that the restriction on the number of undefined inputs in these results is unavoidable. That is, for a wide range of values of $\mathrm{D}^{\to}_{\mathrm{cc}}(F)$ and $\mathrm{NADT}^{\oplus}(f)$ (from constant to $n-2$) we provide partial functions (with more than $\Omega\left(\frac{2^{n-t}}{\sqrt{n-t}}\right)$ undefined inputs, where $t = \mathrm{D}^{\to}_{\mathrm{cc}}$) for which $\mathrm{D}^{\to}_{\mathrm{cc}}(F) < \mathrm{NADT}^{\oplus}(f)$. In particular, we provide a function with an exponential gap between the two measures. Our separation results translate to the case of two-way communication complexity as well, in particular showing that the result of Hatami et al. (2018) cannot be generalized to partial functions.

Previous results for total functions heavily rely on the Boolean Fourier analysis and thus, the technique does not translate to partial functions. For the proofs of our results we build a linear algebraic framework instead. Separation results are proved through the reduction to covering codes.

2012 ACM Subject Classification Theory of computation → Communication complexity; Theory of computation → Oracles and decision trees; Theory of computation → Error-correcting codes

Keywords and phrases Partial functions, XOR functions, communication complexity, decision trees, covering codes

Digital Object Identifier 10.4230/LIPIcs.ICALP.2024.116

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: https://eccc.weizmann.ac.il/report/2023/157/

Acknowledgements We would like to thank the anonymous reviewers for useful comments that helped us improve the presentation.

1 Introduction

In communication complexity model two players, Alice and Bob, are computing some fixed function $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ on a given input (x, y). However, Alice knows only x and Bob knows only y. The main object of studies in communication complexity is the amount of communication $\mathrm{D}_{\mathrm{cc}}(F)$ needed between Alice and Bob to compute the function.

© Vladimir V. Podolskii and Dmitrii Sluch;
licensed under Creative Commons License CC-BY 4.0
51st International Colloquium on Automata, Languages, and Programming (ICALP 2024).
Editors: Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson;
Article No. 116; pp. 116:1–116:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Function F is a XOR-function if for all $x, y \in \{0,1\}^n$ we have $F(x,y) = f(x \oplus y)$ for some $f \colon \{0,1\}^n \to \{0,1\}$, where $x \oplus y$ is a bit-wise XOR of Boolean vectors x and y. XOR-functions are important in communication complexity [27, 19, 25, 26, 3, 13, 15, 1, 23, 21, 5, 2, 8, 11, 9], on one hand, since there are important XOR-functions defined based on Hamming distance between x and y, and on the other hand, since the structure of XOR-functions allows for the Fourier analytic techniques. In particular, this connection suggests an approach for resolving Log-rank Conjecture for XOR-functions [27, 13].

In recent years there was considerable progress in the characterization of communication complexity of a XOR-function F in terms of the complexity of f in parity decision tree model. In this model the goal is to compute a fixed function f on an unknown input $x \in \{0,1\}^n$ and in one step we are allowed to query XOR of any subset of input bits. We want to minimize the number of queries that is enough to compute f on any input x. The complexity of f in this model is denoted by $\mathrm{DT}^{\oplus}(f)$. It was shown by Hatami et al. [13] that for any total f we have $\mathrm{D_{cc}}(F) = \mathrm{poly}(\mathrm{DT}^{\oplus}(f))$.

Even stronger connection holds for one-way communication complexity case. In this setting only very restricted form of communication is allowed: Alice sends Bob a message based on x and Bob has to compute the output based on this message and y. We denote the complexity of F in this model by $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$. The relevant model of decision trees is the model of non-adaptive parity decision trees. In this model we still want to compute some function f on an unknown input and we still can query XORs of any subsets of input bits, but now all queries should be provided at once (in other words, each query cannot depend on the answers to the previous queries). The complexity of f in this model is denoted by $\mathrm{NADT}^{\oplus}(f)$. It follows from the results of Montanaro, Osbourne [19] and Gopalan et al. [10] that for any total XOR-function $F(x,y) = f(x \oplus y)$ we have $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = \mathrm{NADT}^{\oplus}(f)$.

These results on the connection between communication complexity and parity decision trees can be viewed as lifting results. This type of results have seen substantial progress in recent years (see [20]). The usual structure of a lifting result is that we start with a function f that is hard in some weak computational model (for example, a decision tree type model), compose it with some gadget function g to obtain $f \circ g$ (each variable of f is substituted by a copy of g defined on fresh variables) and show that $f \circ g$ is hard in a stronger computational model (for example, a communication complexity type model). The results on XOR-functions can be viewed as lifting results for $g = \mathrm{XOR}$.

The results on the connection between communication complexity of XOR-functions and parity decision trees discussed above are proved only for total functions f for the reason that the proofs heavily rely on the Fourier techniques. However, in communication complexity and decision tree complexity it is often relevant to consider a more general case of partial functions, and many lifting theorems apply to this type of functions as well, see e.g. [7, 17, 4, 22]. In particular, there are some lifting results for partial functions for gadgets that are stronger than XOR: Mande et al. [18] proved such a result for one-way case for inner product gadget (inner product is XOR applied to ANDs of pairs of variables) and Loff, Mukhopadhyay [17] proved a result on lifting with equality gadget for general case (note that equality for inputs of length 1 is practically XOR function). In [17] a conjecture is mentioned that for partial XOR-functions $\mathrm{D_{cc}}(F)$ is approximately equal to $\mathrm{DT}^{\oplus}(f)$ as well.

Our results

In this paper we initiate the studies of the connection between communication complexity for the case of partial XOR functions and parity decision trees. It turns out that for one-way case whether they are equal depends on the number of inputs on which the function is undefined: if the number of undefined inputs is small, then the complexity measures are equal and if it is too large, they are not equal.

More specifically, we show that for $t = \mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$ the equality $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = \mathrm{NADT}^{\oplus}(f)$ holds if f is undefined on at most $O\left(\frac{2^{n-t}}{\sqrt{n-t}}\right)$ inputs.

On the other hand, we provide a family of partial functions for which $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) < \mathrm{NADT}^{\oplus}(f)$[1]. More specifically, we show that for any constant $0 < c < 1$ there is a function f with $\mathrm{NADT}^{\oplus}(f) = cn$ and $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) \leq c'n$ for some $c' < c$.

The number of undefined inputs for the function is $O\left(\frac{2^{dn}}{\sqrt{n}}\right)$ if $c > 1/2$, is equal to 2^{n-1} if $c = 1/2$, and is $2^n - O\left(\frac{2^{dn}}{\sqrt{n}}\right)$ if $c < 1/2$, where $0 < d < 1$ is some constant (depending of c).

We provide a function f for which $\mathrm{NADT}^{\oplus}(f) = \sqrt{n \log n}$ and $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) \leq O(\log n)$, the number of undefined inputs for f is $2^n - 2^{\Theta(\sqrt{n}\log^{3/2} n)}$. Thus, we provide an exponential gap between the two measures.

We provide stronger bounds for small and large values of complexity. For $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = 1$ we show that the equality $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = \mathrm{NADT}^{\oplus}(f)$ is true for all partial f. For $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = 2$ the equality is true for at most $2^{n-3} - 1$ undefined inputs. The smallest values of measures for which we provide a separation are $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = 7$ and $\mathrm{NADT}^{\oplus}(f) = 8$. On the other end of the spectrum we show that for any partial function if $\mathrm{NADT}^{\oplus}(f) \geq n - 1$, then $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = \mathrm{NADT}^{\oplus}(f)$. The largest value of $\mathrm{NADT}^{\oplus}$ for which we provide a separation is $n - 2$, this complements the result that starting with $\mathrm{NADT}^{\oplus}(f) = n - 1$ the measures are equal.

All our separation results translate to the setting of two-way communication complexity vs. parity decision trees. In particular, we provide a partial function f with exponential gap between $\mathrm{D}_{\mathrm{cc}}(F)$ and $\mathrm{DT}^{\oplus}(f)$, which refutes the conjecture mentioned in [17]. It is an interesting open problem whether the polynomial relation between these measures discovered by Hatanami et al. for total functions holds for partial functions with some restriction on the number of undefined points.

The techniques behind the results on the connections between communication complexity of XOR-functions and parity decision tree complexity for total functions heavily rely on the Fourier analysis. However, it is not clear how to translate this technique to partial functions. To prove our results, we instead translate the Fourier-based approach of [19, 10] into the language of linear algebra. We design a framework to capture the notion of one-way communication complexity of partial XOR-functions and use this framework to establish equality of $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$ and $\mathrm{NADT}^{\oplus}(f)$ for the small number of undefined points. The separation results can be proved using our framework, but in these version of the paper we provide self-contained proof. The separation results are proved by a reduction to the covering codes.

The rest of the paper is organized as follows. In Section 2 we provide necessary preliminary information and introduce the notations. In Section 3 we introduce our linear-algebraic framework. In Section 4 we prove main results on the equality of complexity measures. In Section 5 we prove separation results. In Section 6 we provide results for extreme cases. Some of the technical proofs are edited out of this version and can be found in the full paper `https://eccc.weizmann.ac.il/report/2023/157/`.

[1] Note that the gap in the other direction is impossible: it is easy to see that $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) \leq \mathrm{NADT}^{\oplus}(f)$ for all f (see Lemma 4 below). Similar inequality (with an extra factor of 2) holds for general communication complexity and parity decision tree complexity.

2 Preliminaries

2.1 Boolean cube

A Boolean cube is a graph on the set $\{0,1\}^n$ of Boolean strings of length n. We connect two vertices with an edge if they differ in a single bit only. The set $\{0,1\}^n$ can also be thought of as the vector space $\mathbb{F}_2^n$, with the bitwise XOR as the group operation. An inner product over this space can be defined as

$$\langle x, y\rangle = \bigoplus_i x_i \wedge y_i.$$

Hamming weight of x denoted $|x|$ is defined as the number of coordinates of x equal to 1. Hamming distance $\mathrm{dist}(x, y)$ between $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$ is the number of coordinates at which x and y differ. The Hamming ball of radius r is a set of vertices of Boolean cube $\{0,1\}^n$ with Hamming weight not exceeding r. We denote by $V(n, r)$ the volume of a Hamming ball in $\{0,1\}^n$ of radius r.

2.2 Isoperimetric inequalities

▶ **Definition 1.** *For a set A we denote the set of neighbors of elements of A as ΓA. We denote $\Gamma' A := \Gamma A \setminus A$.*

We will need the vertex isoperimetric inequality for a Boolean cube known as Harper's theorem. To state it we first define Hales order.

▶ **Definition 2** (Hales order [12, Page 56])**.** *Consider two subsets $x, y \subseteq [m]$ for some natural m. We define $x \prec y$ if $|x| < |y|$ or $|x| = |y|$ and the smallest element of the symmetric difference of x and y belongs to x. In other words, there exists an i such that $i \in x, i \notin y$, and i is the smallest element in which x and y differ. Here is an example of Hales order for $m = 4$:*

$$\varnothing, \{1\}, \{2\}, \{3\}, \{4\}, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\},$$
$$\{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\}.$$

We can induce Hales order on the set $\{0,1\}^m$ by identifying subsets of $[m]$ with their characteristic vectors. We define I_a^m to be the set of the first a elements of $\{0,1\}^m$ in Hales order.

▶ **Theorem 3** (Harper's theorem [12, Theorem 4.2])**.** *Let $A \subseteq \{0,1\}^m$ be a subset of vertices of m-dimensional Boolean cube and denote $a = |A|$. Then $|\Gamma A| \geq |\Gamma I_a^m|$.*

2.3 Communication Complexity and Decision Trees

Throughout this paper, f denotes a partial function $\{0,1\}^n \to \{0,1,\perp\}$, we let $\mathrm{Dom}(f) = f^{-1}(\{0,1\})$. We define an XOR-function $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1,\perp\}$ as

$$F(x, y) = f(x \oplus y).$$

In communication complexity model two players, Alice and Bob, are computing some fixed function $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ on a given input (x, y). However, Alice knows only x and Bob knows only y. The main subject of studies in communication complexity is the amount of communication $\mathrm{D_{cc}}(F)$ needed between Alice and Bob to compute the function. Formal definition of the model can be found in [16].

We will be mostly interested the in one-way communication model. This is a substantially restricted setting, in which only Alice is permitted to send bits to Bob. Formally, the one-way communication complexity $D_{cc}^{\rightarrow}(F)$ is defined to be the smallest integer t, allowing for a protocol where Alice knowing her input x sends t bits to Bob, which together with Bob's input y enable Bob to calculate the value of F.

The bits communicated by Alice depend only on x, that is Alice's message to Bob is $h(x)$ for some fixed total function $h\colon \{0,1\}^n \to \{0,1\}^t$. Bob computes the output $F(x,y)$ based on $h(x)$ and his input y. That is, Bob outputs $\varphi(h(x),y)$ for some fixed total function $\varphi\colon \{0,1\}^t \times \{0,1\}^n \to \{0,1\}$. If (x,y) is within the domain of F, then the equality $\varphi(h(x),y) = F(x,y)$ must be true.

The notion of parity decision tree complexity is a generalization of the well-known decision tree complexity model. In this model, to evaluate a function f for a given input x the protocol queries the parities of some subsets of the bits in x. The cost of the protocol on specified input x is the number of queries the protocol makes on that input. The cost of the protocol (sometimes referred to as the worst-case cost) is maximum over all inputs x, costs of protocol on the input x. The complexity of problem f in the model of parity decision trees $\mathrm{DT}^{\oplus}(f)$ is the minimal over all valid protocols, cost of a protocol for f.

We consider the non-adaptive parity decision tree complexity $\mathrm{NADT}^{\oplus}(f)$. This version differs from its adaptive counterpart in that all the queries should be fixed at once. In other words, each next query should not depend on the answers to previous queries. Next, we give a more formal definition of $\mathrm{NADT}^{\oplus}(f)$.

The protocol of complexity p is defined by n-bit strings $s_1,\dots,s_p$ and a total function $l\colon \{0,1\}^p \to \{0,1\}$. On input x the protocol queries the values of

$$\langle s_1, x\rangle, \dots, \langle s_p, x\rangle$$

and outputs

$$l(\langle s_1, x\rangle, \dots, \langle s_p, x\rangle).$$

The protocol computes partial function f, if for any $x \in \mathrm{Dom}(f)$ we have

$$l(\langle s_1, x\rangle, \dots, \langle s_p, x\rangle) = f(x).$$

Throughout the paper $t, h, \varphi, p, s_1, \dots, s_p, l$ have the same meaning as defined above.

It is easy to see that there is a simple relation between $\mathrm{NADT}^{\oplus}(f)$ and $\mathrm{D}_{\mathrm{cc}}^{\rightarrow}(F)$.

▶ **Lemma 4.** *For any f we have $\mathrm{D}_{\mathrm{cc}}^{\rightarrow}(F) \le \mathrm{NADT}^{\oplus}(f)$.*

Proof. Alice and Bob can compute $F(x,y)$ by a simple simulation of $\mathrm{NADT}^{\oplus}$ protocol for f. The idea is that they privately calculate the parities of their respective inputs according to $\mathrm{NADT}^{\oplus}$ protocol, then Alice sends the computed values to Bob, who XORs them with his own parities, and then computes the value of F.

More formally, assume that $\mathrm{NADT}^{\oplus}(f) = p$ and the corresponding protocol is given by $s_1,\dots,s_p \in \{0,1\}^n$ and a function l, that is

$$\forall x \in \mathrm{Dom}(f), f(x) = l(\langle s_1, x\rangle, \dots, \langle s_p, x\rangle).$$

For $i \in [p]$, we let

$$h_i(x) := \langle s_i, x\rangle.$$

For the communication protocol of complexity p we let

$$h(x) = (h_1(x), \dots, h_p(x)),$$
$$\varphi(a, y) := l(a_1 \oplus \langle s_1, y\rangle, \dots, a_p \oplus \langle s_p, y\rangle).$$

Then for any (x, y) such that $x \oplus y \in \mathrm{Dom}(f)$ we have

$$\begin{aligned}\varphi(h(x), y) = l(h_1(x) \oplus \langle s_1, y\rangle, \dots, h_p(x) \oplus \langle s_p, y\rangle) =\\ l(\langle s_1, x\rangle \oplus \langle s_1, y\rangle, \dots, \langle s_p, x\rangle \oplus \langle s_p, y\rangle) =\\ l(\langle s_1, x \oplus y\rangle, \dots, \langle s_p, x \oplus y\rangle) = f(x \oplus y) = F(x, y).\end{aligned}$$

We constructed a p-bit communication protocol for F, and thus

$$\mathrm{D}^{\to}_{\mathrm{cc}}(F) \le p = \mathrm{NADT}^{\oplus}(f).$$ ◀

In this paper, we are mainly interested in whether the inequality in the opposite direction is true.

2.4 Covering Codes

▶ **Definition 5.** *A subset $\mathcal{C} \subseteq \{0,1\}^n$ is a (n, K, R) covering code if $|\mathcal{C}| \le K$ and for any $x \in \{0,1\}^n$ there is $c \in \mathcal{C}$ such that $\mathrm{dist}(x, c) \le R$. In other words, all points in $\{0,1\}^n$ are covered by balls of radius R with centers in $\mathcal{C}$.*

The following general bounds on K are known for covering codes.

▶ **Theorem 6** ([6, Theorem 12.1.2])**.** *For any (n, K, R) covering code we have*

$$\log K \ge n - \log V(n, R).$$

For any n and any $R \le n$ there is a (n, K, R) covering code with

$$\log K \le n - \log V(n, R) + \log n.$$

We will use the following well known fact.

▶ **Theorem 7** ([6, Section 2.6])**.** *If $n = 2^m - 1$ for some m, then Boolean cube $\{0,1\}^n$ can be splitted into disjoint balls of radius 1.*

This construction is known as a Hamming error correcting code. Note that it is a $(n = 2^m - 1, \frac{2^n}{n+1}, 1)$ covering code.

▶ **Definition 8.** *For two covering codes $\mathcal{C}_1$ and $\mathcal{C}_2$ their direct sum is*

$$\mathcal{C}_1 \oplus \mathcal{C}_2 = \{(c_1, c_2) \mid c_1 \in \mathcal{C}_1, c_2 \in \mathcal{C}_2\}.$$

▶ **Lemma 9** ([6, Theorem 12.1.2])**.** *If $\mathcal{C}_1$ is a (n_1, K_1, R_1) covering code and $\mathcal{C}_2$ is a (n_2, K_2, R_2) covering code, then $\mathcal{C}_1 \oplus \mathcal{C}_2$ has parameters $(n_1 + n_2, K_1K_2, R_1 + R_2)$.*

We need the following bounds on the sizes of Hamming balls (see, e.g. [14, Appendix A]).

▶ **Lemma 10.** *For any n and $k \le n$ we have*

$$\left(\frac{n}{k}\right)^k \le V(n, k) \le \left(\frac{en}{k}\right)^k.$$

▶ **Lemma 11.** *For any constant* $0 < c < 1$ *we have*

$$\binom{n}{cn} = O\left(\frac{1}{\sqrt{n}} 2^{H(c)n}\right).$$

For any constant $0 < c < 1/2$ *we have*

$$V(n, cn) = O\left(\frac{1}{\sqrt{n}} 2^{H(c)n}\right),$$

where H *is the binary entropy function.*

▶ **Lemma 12** ([24, Section 5.4])**.**

$$V\left(n, \frac{n}{2} - \Theta(\sqrt{n \log n})\right) = \frac{2^n}{poly(n)}.$$

For the binary entropy function $H(x)$ we will use the following simple fact.

▶ **Lemma 13.** *For any constant* $c \in (0, 1)$ *and for any* $\alpha_n \xrightarrow[n\to\infty]{} 0$ *we have*

$$H(c + \alpha_n) = H(c) + O(\alpha_n),$$

where the constant in O*-notation might depend on* c*, but not on* n*.*

This is true since the derivative of H is upper bounded by a constant in any small enough neighborhood of c.

3 Linear-algebraic framework

3.1 Graph-based analysis of one-way communication protocols

Recall that in a one-way communication protocol of complexity t for $F(x, y) = f(x \oplus y)$ Alice on input $x \in \{0, 1\}^n$ first sends to Bob $h(x)$ for some fixed $h: \{0, 1\}^n \to \{0, 1\}^t$. After that Bob computes the output $\varphi(h(x), y)$, where $y \in \{0, 1\}^n$ is Bob's input and $\varphi: \{0, 1\}^t \times \{0, 1\}^n \to \{0, 1\}$.

Let's consider the partition $\mathcal{H} = \{H_a \mid a \in \{0, 1\}^t\}$, where for any $a \in \{0, 1\}^t$

$$H_a = h^{-1}(a).$$

We refer to $\mathcal{H}$ as h*-induced partition.* A class H_a of this partition is the set of inputs for which Alice sends Bob the same message.

Consider two arbitrary inputs $x, y \in \{0, 1\}^n$. We call the vector $\Delta = x \oplus y$ the *shift* between x and y. The intuition is that y is equal to the shift $x \oplus \Delta$ of x by y (and vise versa).

We say that $\Delta \in \{0, 1\}^n$ is a *good shift* if there is a pair $x, y \in \{0, 1\}^n$ such that $x \oplus y = \Delta$ and $h(x) = h(y)$, or equivalently, if x and y belong to the same class of $\mathcal{H}$. Note that f does not necessarily need to be defined on inputs x and y. However, it turns out that on the domain of f the value of f is invariant under good shifts.

▶ **Lemma 14.** *Assume that* Δ *is a good shift. Consider any* $v, u \in \mathrm{Dom}(f)$ *such that* $v \oplus u = \Delta$*. Then,* $f(v) = f(u)$*.*

Proof. Since Δ is good, there are x and y such that $h(x) = h(y)$ and $x \oplus y = \Delta$. Then

$$f(v) = \varphi(h(x), x \oplus v) = \varphi(h(y), x \oplus v) = f(v \oplus x \oplus y) = f(v \oplus \Delta) = f(u).$$ ◀

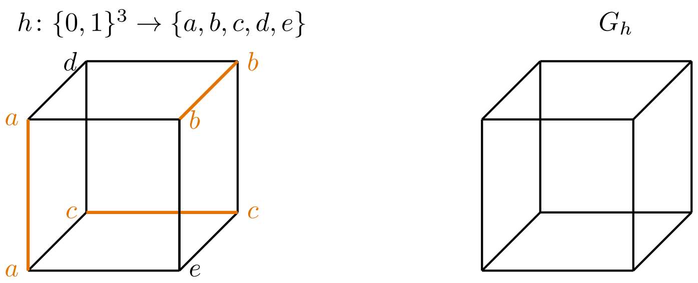

Figure 1 Example of total h-induced graph.

This leads us to the following notion.

▶ **Definition 15.** *For the functions $f : \{0,1\}^n \to \{0,1\}$, $f(x \oplus y) = \varphi(h(x), y)$ let the* total h-induced graph *be the graph with vertices $\{0,1\}^n$ and with an edge between $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$ if $x \oplus y$ is a good shift for h. Now remove vertices where the function f is undefined. The resulting graph is called the* partial h-induced graph*.*

There is an alternative way of thinking about total h-induced graph. Consider a graph with vertices labeled $\{0,1\}^n$ in which we connect two vertices if the value of h on these vertices is the same. Clearly it is a subgraph of the total h-induced graph. Now consider a shift of this graph, that is, a graph in which we XORed labels of all vertices with some fixed vector. This graph is also a subset of the total h-induced graph. By considering all possible shifts and taking the union of all graphs we will get the total h-induced graph. See Figure 1 for an example of total h-induced graph.

By transitivity, if (h, φ) form a valid communication protocol then f assigns identical values to each connected component in partial h-induced graph. The converse is also true.

▶ **Theorem 16.** *Consider a function $f : \{0,1\}^n \to \{0,1\}$. For a function $h : \{0,1\}^n \to \{0,1\}^t$ there is a function $\varphi : \{0,1\}^t \times \{0,1\}^n \to \{0,1\}$ such that (h, φ) form a valid communication protocol for f if and only if f assigns the same value to each connected component in the partial h-induced graph.*

Proof. As discussed above, if (h, φ) forms a valid communication protocol, then f assigns the same value to each connected component of the partial h-induced graph.

It remains to prove the converse statement. We assume that f assigns the same value to each connected component and we need to show that there is φ such that

$$\forall (x, y) \in \mathrm{Dom}(F), \quad F(x, y) = \varphi(h(x), y).$$

The proof idea is the following. Each input $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ to F yields an input $(a, y) \in \{0,1\}^t \times \{0,1\}^n$ to φ where $\alpha = h(x)$. We define φ on (α, y) to be equal to F on a single corresponding F-input (x', y) with $x' \in h^{-1}(\alpha)$. Then we prove that φ defined that way gives a communication protocol computing F correctly on all inputs $(x, y) \in \mathrm{Dom}(f)$

Formally, we define φ as follows. For each $\alpha \in \{0,1\}^t$ and $y \in \{0,1\}^n$, consider $x' \in \{0,1\}^n$ such that $h(x') = \alpha$ and $(x', y) \in \mathrm{Dom}(F)$. If there is no such x' we define $\varphi(\alpha, y)$ arbitrarily. If there is such an x', let

$$\varphi(\alpha, y) := F(x', y).$$

Now we show that the resulting protocol computes $F(x, y)$ correctly for any (x, y).

Consider arbitrary $(x, y) \in \mathrm{Dom}(F)$. Consider x' chosen for $\alpha = h(x)$ and y (it exists, since clearly x itself satisfies all the necessary conditions).

Thus, we have

$$\varphi(h(x), y) = F(x', y).$$

It remains to prove that

$$F(x', y) = F(x, y)$$

or equivalently,

$$f(x' \oplus y) = f(x \oplus y).$$

For XOR of these two inputs of f we have

$$(x' \oplus y) \oplus (x \oplus y) = x' \oplus x.$$

Since $h(x) = h(x')$, we have that $x' \oplus x$ is a good shift. And since

$$(x, y), (x', y) \in \mathrm{Dom}(F),$$

we have

$$x \oplus y, x' \oplus y \in \mathrm{Dom}(f).$$

We have that vertices $x \oplus y$ and $x' \oplus y$ are connected in the partial h-induced graph and by Lemma 14 f assigns the same value to them. Hence, the function φ, together with h, forms a communication protocol for F. ◀

3.2 Using coset structures on a Boolean cube to analyze non-adaptive parity decision trees

We consider the vertices of the Boolean cube as a vector space $\mathbb{F}_2^n$. We show that a $\mathrm{NADT}^{\oplus}$ protocol corresponds to a linear subspace of $\mathbb{F}_2^n$ such that f is constant on each of its cosets (the coset for a linear subspace L and a vector l is defined as the set $\{x + l | l \in L\}$ and denoted $L + l$).

▶ **Theorem 17.** *Let $f\colon \{0,1\}^n \to \{0,1\}$. There is a p-bit $\mathrm{NADT}^{\oplus}$ protocol for f if and only if there exists an $n - p$ dimensional subspace of $\{0,1\}^n$ such that for each coset of that subspace, f assigns the same value to all inputs of the coset where f is defined.*

Proof. Suppose $s_1, \ldots, s_p, l$ form a valid $\mathrm{NADT}^{\oplus}$ protocol for f. We construct a matrix S with rows $s_1, \ldots, s_p$. If some of the rows are linearly dependent, we add rows arbitrarily to make the rank of S equal to p. When S is multiplied on the right by some vector x, we obtain all inner products of x with vectors $s_1, \ldots, s_p$ (and possibly other bits if we added rows).

Consider the vector subspace $\{x | Sx = 0\}$. This is an $n - p$ dimensional space. For all points in the same coset of this subspace, the tuple consisting of values of the inner products $(\langle s_1, x\rangle, \ldots, \langle s_p, x\rangle)$ is the same, so is the value of $l(\langle s_1, x\rangle, \ldots, \langle s_p, x\rangle)$. For all points where f is defined and lying in the same coset, the value of f must be equal to the value of l and thus the same for all points in the coset.

In the reverse direction, let $\langle e_1, \ldots, e_{n-p}\rangle$ be an $n-p$ dimensional subspace of $\{0,1\}^n$ such that for each of its cosets f is constant on all points of that coset on which it is defined. We can represent this subspace in the form $\{x | Sx = 0\}$ for some matrix S of size $p \times n$.

Vectors x and y are in the same coset of $\langle e_1, \ldots, e_{n-p}\rangle$ iff $Sx = Sy$. Thus, to compute $f(x)$ it is enough to compute the inner product of x with the rows of S. ◀

▶ **Corollary 18.** *Consider a function $f : \{0,1\}^n \to \{0,1\}$ having valid communication protocol $f(x \oplus y) = \varphi(h(x), y)$ where $h : \{0,1\}^n \to \{0,1\}^t, \varphi : \{0,1\}^t \times \{0,1\}^n \to \{0,1\}$. Suppose there is an $n-t$ dimensional subspace L of $\{0,1\}^n$ and consider subgraphs of partial h-induced graph each over vertices belonging to different cosets of L. If all of these subgraphs are connected then* $\mathrm{NADT}^{\oplus}(f) \leq t$.

Proof. By Theorem 16 f is constant on each coset. By Theorem 17 it follows that $\mathrm{NADT}^{\oplus}(f) \leq t$. ◀

4 Equality between $\mathrm{D}^{\to}_{\mathrm{cc}}(F)$ and $\mathrm{NADT}^{\oplus}(f)$

In this section we will show that if $\mathrm{D}^{\to}_{\mathrm{cc}}(F) = t$ and the number of undefined inputs is small, then $\mathrm{NADT}^{\oplus}(f) = t$ as well. More specifically, we prove the following theorem.

▶ **Theorem 19.** *If for the function $f \colon \{0,1\}^n \to \{0,1\}$ we have $\mathrm{D}^{\to}_{\mathrm{cc}}(F) = t$, where $F(x,y) = f(x \oplus y)$, and f is undefined on less than $\binom{n-t+1}{\lfloor \frac{n-t}{2} \rfloor - 1}$ inputs, then* $\mathrm{NADT}^{\oplus}(f) = t$.

By Lemma 11 we have that $\binom{n-t+1}{\lfloor \frac{n-t+1}{2} \rfloor} = O(\frac{2^{n-t}}{\sqrt{n-t}})$ and since $\lfloor \frac{n-t}{2} \rfloor - 1$ differs from $\lfloor \frac{n-t+1}{2} \rfloor$ by only a constant, it is easy to see that the same estimate applies to $\binom{n-t+1}{\lfloor \frac{n-t}{2} \rfloor - 1}$ as well. Thus, the number of undefined inputs is $O(\frac{2^{n-t}}{\sqrt{n-t}})$.

The rest of the section is devoted to the proof of Theorem 19. The idea of the proof is as follows. Consider the h-induced partition $\mathcal{H}$ corresponding to the communication protocol of complexity t. We show that either the partition $\mathcal{H}$ corresponds to the cosets of an $n-t$ dimensional subspace of $\{0,1\}^n$, which allows us to construct an $\mathrm{NADT}^{\oplus}$ protocol, or there exist *many* good shifts. The structure of these good shifts imposes restrictions on f that again allow us to construct an $\mathrm{NADT}^{\oplus}$ protocol.

We start with a simple case.

▶ **Lemma 20.** *If there exists t-bit communication protocol, (h, φ) for a function $f \colon \{0,1\}^n \to \{0,1\}$, and the h-induced partition $\mathcal{H}$ corresponds to cosets of an $n-t$ dimensional subspace L of $\{0,1\}^n$, then* $\mathrm{NADT}^{\oplus}(f) \leq t$.

Proof. Since the partition $\mathcal{H}$ corresponds to the cosets of L, we have that for any inputs x and y, if $h(x) = h(y)$, then $x \oplus y \in L$ and vice versa. In other words, all good shifts are in L and any shift in L is good. Thus, connected components of the total h-induced graph are cosets of L and are fully connected. By Corollary 18 we have that $\mathrm{NADT}^{\oplus}(f) \leq t$. ◀

The structure of the proof for the other case is the following. We show that the total h-induced graph is structured into connected components, each of which is a coset of a k-dimensional subspace of $\{0,1\}^n$ for $k \geq n-t$. We show that there is a bijective graph homomorphism of the k-dimensional Boolean cube onto each component. Furthermore, each vertex in the total h-induced graph has a degree of at least $\frac{2^n}{2^t} - 1$. We show that if we remove fewer than $\binom{n-t+1}{\lfloor \frac{n-t}{2} \rfloor - 1}$ vertices, each coset still contains one connected component. By the way of contradiction, suppose this is not the case and some coset contains more than one connected component. We consider the smallest of these components, denote the set of its nodes by S. We show that the number of neighboring vertices of S in the total h-induced graph (excluding S itself) is not less than $\binom{n-t+1}{\lfloor \frac{n-t}{2} \rfloor - 1}$. This implies that after removing the undefined inputs of f S cannot not be separated from other nodes in the coset. To show this we treat separately cases of large and small $|S|$. For small $|S|$ we use the fact that vertices have high degree. For large $|S|$ we use the vertex-isoperimetric inequality for the Boolean cube.

▶ **Lemma 21.** *Suppose there exists a t-bit communication protocol (h, φ) for $f\colon \{0,1\}^n \to \{0,1\}$ and the h-induced partition $\mathcal{H}$ classes do not correspond to cosets of an $n-t$-dimensional subspace of $\{0,1\}^n$. Let D be the set of good shifts for h. Then D contains a minimum of $n - t + 1$ linearly independent vectors.*

Proof. Suppose there are at most $n-t$ linearly independent good shifts $e_1, \ldots, e_{n-t}$. Consider a linear subspace of $\{0,1\}^n$ spanned over by these shifts and add some vectors to it to make it exactly $n - t$ dimensional if needed. Denote the resulting subspace L. As classes of $\mathcal{H}$ do not correspond to the cosets of L and there are 2^t of both classes and cosets there exist two elements belonging to the same class and different cosets. Their XOR is a good shift linearly independent with $e_1, \ldots, e_{n-t}$. We got a contradiction implying the lemma. ◀

▶ **Lemma 22.** *Suppose there exists t-bit communication protocol (h, φ) for f. Let D be the set of all good shifts for h and $\{e_1, \ldots, e_k\}$ be the largest linearly independent subset of D. Then the total h-induced graph $\mathcal{H}$ has the following properties.*

- *Cosets of the subspace $\langle e_1, \ldots, e_k \rangle$ are connected components of $\mathcal{H}$.*
- *There is a bijective graph homomorphism of k-dimensional Boolean cube into each coset.*

Proof. It is easy to see that all vertices in any coset are connected to each other. Let's show that no edges exist between vertices of different cosets. Assume by contradiction that there is an edge between vertices v and u from different cosets. Note that $u \oplus v \notin \langle e_1, \ldots, e_k \rangle$. Thus, vectors $e_1, \ldots, e_k, u \oplus v$ form a linearly independent system of size $k + 1$, which is a contradiction.

Now, let's construct a homomorphism q from the Boolean cube $\{0,1\}^k$ into the coset $v + \langle e_1, \ldots, e_k \rangle$ for an arbitrary vertex v. Consider a matrix B that has vectors $e_1, \ldots, e_k$ as its columns and let $q(x) = v \oplus Bx$. The image of q is within the coset $v + \langle e_1, \ldots, e_k \rangle$, as columns of B belong to the subspace $\langle e_1, \ldots, e_k \rangle$. The mapping is bijective on $v + \langle e_1, \ldots, e_k \rangle$, as B's columns are linearly independent. Finally, consider a pair of vertices x, y adjacent in a Boolean cube. Since the vertices are adjacent, they only differ in a single bit i. Thus,

$$q(x) \oplus q(y) = (v \oplus Bx) \oplus (v \oplus By) = B(x \oplus y) = e_i.$$

Since $e_i \in D$, an edge exists between $q(x)$ and $q(y)$, implying that q is a graph homomorphism. ◀

▶ **Lemma 23.** *Suppose there exists t-bit communication protocol (h, φ) for $f\colon \{0,1\}^n \to \{0,1\}$. Then in the total h-induced graph, the degree of any vertex is not less than $\frac{2^n}{2^t} - 1$.*

Proof. Let's consider the largest class in the h-induced partition $\mathcal{H}$. Since the number of classes is at most 2^t, the largest class contains at least $\frac{2^n}{2^t}$ elements. Fix an element of the class and compute its XOR with all elements in the same class $\mathcal{H}$. We have $\frac{2^n}{2^t}$ XORs in total, $\frac{2^n}{2^t} - 1$ of which are non-zero. Since each XOR is computed between elements in the same class, these XORs are good shifts. For all vertices in the h-induced graph for each good shift we draw an edge from the vertex corresponding to this shift. Therefore, the degree of any vertex is at least $\frac{2^n}{2^t} - 1$. ◀

▶ **Lemma 24.** *If A is a subset of k-dimensional Boolean cube satisfying $V\left(m, \lfloor\frac{m-1}{2}\rfloor - 2\right) \leq |A| \leq 2^{k-1}$ for some m, then $|\Gamma' A| \geq \binom{m}{\lfloor\frac{m-1}{2}\rfloor - 1}$.*

The proof of the lemma can be found in the full version of the paper. The proof heavily relies on Theorem 3. Finally, we are ready to prove Theorem 19.

Proof of Theorem 19. We are given t-bit communication protocol (h, φ) for F. By Lemma 21, the h-induced partition $\mathcal{H}$ either corresponds to cosets of an $n-t$ dimensional subspace of $\{0,1\}^n$ (and then by Lemma 20 we have $\mathrm{NADT}^{\oplus}(f) \leq t$), or the set of good shifts D contains at least $n-t+1$ linearly independent vectors. Let $\langle e_1, \ldots, e_k\rangle$, where $k \geq n-t+1$, be the largest subset of linearly independent vectors in D. Consider the cosets of the subspace $\langle e_1, \ldots, e_k\rangle$. We will show that if we remove fewer than $\binom{n-t+1}{\lfloor\frac{n-t}{2}\rfloor - 1}$ vertices from the total h-induced graph, each coset will contain no more than one connected component. Assume by contradiction that after removing the vertices, some coset splits into several connected components. Let A be the smallest of these components. If there are at most $V(n-t+1, \lfloor\frac{n-t}{2}\rfloor - 2) - 1$ vertices in A, consider a vertex a in A. Given the degree of a is at least $2^{n-t} - 1$, a has at least

$$2^{n-t} - V\left(n-t+1, \left\lfloor\frac{n-t}{2}\right\rfloor - 2\right)$$
$$\geq V\left(n-t+1, \left\lfloor\frac{n-t}{2}\right\rfloor\right) - V\left(n-t+1, \left\lfloor\frac{n-t}{2}\right\rfloor - 2\right) \geq \binom{n-t+1}{\lfloor\frac{n-t}{2}\rfloor - 1}$$

neighbors outside A.

On the other hand, suppose A has at least $V(n-t+1, \lfloor\frac{n-t}{2}\rfloor - 2)$ vertices. Since A is the smallest connected component in its coset it also follows that A has no more than 2^{k-1} vertices. By Lemma 24 we have $|\Gamma' A| \geq \binom{n-t+1}{\lfloor\frac{n-t}{2}\rfloor - 1}$, which is more than the number of removed vertices, a contradiction. Thus, cosets cannot be split into several components and by Corollary 18 we have $\mathrm{NADT}^{\oplus}(f) \leq n-k \leq t-1$, which is a contradiction. ◀

5 Separations between $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$ and $\mathrm{NADT}^{\oplus}(f)$

In this section we show that if the number of undefined inputs is large, there is a gap between $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$ and $\mathrm{NADT}^{\oplus}(f)$. That is, we aim to come up with a function f such that $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F)$ is small and $\mathrm{NADT}^{\oplus}(f)$ is large.

The key idea in our construction is that in h-induced graph for the intended communication protocol the edges connect only vertices with small Hamming distance between them. Then, if the function f has 0-inputs and 1-inputs far away from each other, they are not connected and h corresponds to a valid protocol. We will ensure that at the same time f has large $\mathrm{NADT}^{\oplus}$ complexity.

We start with the construction of the functions, then investigate their $\mathrm{NADT}^{\oplus}$ complexity and then prove upper bounds on $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}$ complexity of the corresponding XOR functions. The latter part is through the reduction to covering codes.

▶ **Definition 25.** *For a parameter k define $f_k \colon \{0,1\}^n \to \{0,1,\perp\}$ in the following way.*

$$f_k(x) = \begin{cases} 0 & \text{for } |x| \leq k, \\ \perp & \text{for } k+1 \leq |x| \leq n-1, \\ 1 & \text{for } |x| = n. \end{cases}$$

We denote the corresponding XOR function by F_k.

Note, that the number of undefined inputs in f_k is $V(n, n-k-1) - 1$.

It turns out that f_k has reasonably large $\mathrm{NADT}^{\oplus}$ and $\mathrm{DT}^{\oplus}$ complexities.

▶ **Theorem 26.** $\mathrm{NADT}^{\oplus}(f_k) = \mathrm{DT}^{\oplus}(f_k) = k + 1$.

Proof. Since $\mathrm{DT}^{\oplus}(f) \leq \mathrm{NADT}^{\oplus}(f)$ for any f, it is enough to prove that $\mathrm{NADT}^{\oplus}(f_k) \leq k+1$ and $\mathrm{DT}^{\oplus}(f_k) \geq k + 1$.

For the upper bound, observe that it is enough to query variables $x_1, \ldots, x_{k+1}$. If all of them are equal to 1, we output 1, otherwise we output 0. It is easy to see that this protocol computes f_k correctly.

For the lower bound suppose, for the sake of contradiction, that an adaptive parity decision tree exists that can compute the function f with k or fewer queries. Consider the path corresponding to the input $e = (1, \ldots, 1)$. Let's assume that the decision tree queried the parities $\langle s_i, e\rangle$ for $s_1, \ldots, s_k$. The answers to the queries are equal to $\langle s_1, e\rangle, \ldots, \langle s_k, e\rangle$. Consider a matrix $B \subseteq \mathbb{F}^{k\times n}$ consisting of rows $s_1, \ldots, s_k$.

Denote $a = Be$. In particular, we have that a lies in the subspace generated by columns of B. Since the rank of B is at most k (the matrix has k rows), there is a subset of at most k columns generating this subspace. In particular, there is $x \in \{0,1\}^n$ with $|x| \leq k$, such that $a = Bx$. That is, $Be = Bx$ and the protocol behaves the same way on e and x, which is a contradiction, since $f_k(e) = 1$ and $f_k(x) = 0$. ◀

▶ Remark 27. Since f_k has large (adaptive) parity decision tree complexity and for any $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ we have $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) \geq \mathrm{D}_{\mathrm{cc}}(F)$, all separations provided by functions f_k translate into the same separations between $\mathrm{DT}^{\oplus}$ and D_{cc}.

Next, we proceed to the upper bound on the $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k)$.

▶ **Theorem 28.** *Suppose for some n, k and t there is a $(n, 2^t, R)$ covering code $\mathcal{C}$ for $R = \left\lfloor \frac{n-k-1}{2} \right\rfloor$. Then, $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) \leq t$.*

Proof. Split the points of $\{0,1\}^n$ into balls with radius R with centers in the points of $\mathcal{C}$ (if some point belongs to several balls, attribute it to one of them arbitrarily). This results in a partition of the cube into 2^t subsets with the diameter of each subset at most $n - k - 1$.

The proof can be finished through Theorem 16, but to make it more self-contained we directly describe communication protocol.

On input x Alice sends as $h(x)$ the index of the ball containing x. Bob computes $\neg y$, componentwise negation of y, and outputs 1 if it is in the same ball. If this is not the case, Bob outputs 0.

Clearly, the complexity of this protocol is at most t. For the correctness of the protocol, if $f(x \oplus y) = 1$, then $x = \neg y$ and the protocol clearly outputs 1. However, if $f(x \oplus y) = 0$, then $|x \oplus y| \leq k$ and thus $\mathrm{dist}(x, \neg y) \geq n - k$. In this case x and $\neg y$ are not in the same ball and the protocol outputs 0. ◀

▶ **Theorem 29.** *For any n and k we have*

$$\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) \leq n - \log V(n, R) + \log n$$

for $R = \left\lfloor \frac{n-k-1}{2} \right\rfloor$.

Proof. By Theorem 6 there exists a $(n, 2^t, R)$ covering code for

$$\log 2^t = t \leq n - \log V(n, R) + \log n.$$

The theorem follows from Theorem 28. ◀

From this we can get a separation for a wide range of parameters.

▶ **Corollary 30.** *Suppose $k = cn$ for some constant $0 < c < 1$. Then $\mathrm{NADT}^{\oplus}(f_k) = cn + 1$ and*

$$\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) \leq \left(1 - H\left(\frac{1-c}{2}\right)\right) n + O(\log n).$$

In particular, $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) < \mathrm{NADT}^{\oplus}(f_k)$. The number of undefined inputs for f_k is $2^n - O(\frac{2^{H(c)n}}{\sqrt{n}})$ if $c < 1/2$, is equal to $(1+o(1))2^{n-1}$ if $c = 1/2$, and is $O(\frac{2^{H(1-c)n}}{\sqrt{n}})$ if $c > 1/2$.

Proof. The equality for $\mathrm{NADT}^{\oplus}$ is proved in Theorem 26.

For communication complexity bound we apply Theorem 29. We have $R = \left\lfloor \frac{(1-c)n-1}{2} \right\rfloor = \frac{(1-c)n}{2} + O(1)$ and by Lemmas 11 and 13 we have

$$\log V(n, R) = H\left(\frac{1-c}{2}\right) n - O(\log n).$$

By Theorem 29 we have

$$\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) \leq n - \log V(n, R) + \log n = \left(1 - H\left(\frac{1-c}{2}\right)\right) n + O(\log n).$$

To show that $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) < \mathrm{NADT}^{\oplus}(f_k)$ we need to compare $k = cn$ with the bound on communication complexity. It is easy to see that

$$1 - H\left(\frac{1-c}{2}\right) < c$$

for all $0 < c < 1$ (the left hand-side and the right hand-side are equal for $c = 0$ and $c = 1$ and the left hand-side is concave in c).

The bounds on the number of undefined inputs follow easily from Lemma 11. ◀

The largest gap we can get is the following.

▶ **Corollary 31.** *For $k = \Theta(\sqrt{n \log n})$ we have that $\mathrm{NADT}^{\oplus}(f_k) = \Theta(\sqrt{n \log n})$ and $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) = O(\log n)$. The number of undefined inputs for f_k is $2^n - 2^{\Theta(\sqrt{n}\log^{3/2} n)}$.*

Proof. For $k = \Theta(\sqrt{n \log n})$ we have $R = \frac{n}{2} - \Theta(\sqrt{n \log n})$ in Theorem 29. By Lemma 12 we have $V(n, R) = \frac{2^n}{\mathrm{poly}(n)}$ and as a result $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_k) = O(\log n)$.

For the number of undefined inputs, we apply Lemma 10:

$$\left(\frac{n}{k}\right)^k \leq V(n, k) \leq \left(\frac{en}{k}\right)^k.$$

For $k = \Theta(\sqrt{n \log n})$ it is easy to see that both sides are $2^{\Theta(\sqrt{n}\log^{3/2} n)}$. From this the estimate on the number of undefined inputs follows. ◀

6 Extreme Cases

In this section we discuss extreme cases. All proves can be found in the full version of the paper.

For small values of complexity measures we have the following equality results.

▶ **Theorem 32.** *Suppose F satisfies $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = 1$. It then follows that $\mathrm{NADT}^{\oplus}(f) = 1$.*

▶ **Theorem 33.** *If function f is undefined on fewer than 2^{n-3} inputs and $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = 2$, then $\mathrm{NADT}^{\oplus}(f) = 2$.*

On the other end of the spectrum, we show that if $\mathrm{NADT}^{\oplus}(f)$ is really large, then it is equal for all partial functions.

▶ **Theorem 34.** *For any partial function $f\colon \{0,1\}^n \to \{0,1,\perp\}$, if $\mathrm{NADT}^{\oplus}(f) \geq n-1$, then $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F) = \mathrm{NADT}^{\oplus}(f)$.*

The largest value of $\mathrm{NADT}^{\oplus}$ for which we get separation is $n-2$.

▶ **Theorem 35.** $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_{n-3}) \leq n - \Theta(\log n)$*, whereas* $\mathrm{NADT}^{\oplus}(f_{n-3}) = n-2$*. The number of undefined inputs for f_{n-3} is $\frac{n(n+1)}{2}$.*

The smallest value of $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}$ for which we get a separation is 7.

▶ **Theorem 36.** *For any $n \geq 32$ we have $\mathrm{D}^{\rightarrow}_{\mathrm{cc}}(F_7) \leq 7$, whereas $\mathrm{NADT}^{\oplus}(f_7) = 8$.*

References

1 Anurag Anshu, Naresh Goud Boddu, and Dave Touchette. Quantum log-approximate-rank conjecture is also false. In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 982–994. IEEE Computer Society, 2019. doi:10.1109/FOCS.2019.00063.

2 Arkadev Chattopadhyay, Ankit Garg, and Suhail Sherif. Towards stronger counterexamples to the log-approximate-rank conjecture. In Mikolaj Bojanczyk and Chandra Chekuri, editors, *41st IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2021, December 15-17, 2021, Virtual Conference*, volume 213 of *LIPIcs*, pages 13:1–13:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.FSTTCS.2021.13.

3 Arkadev Chattopadhyay and Nikhil S. Mande. A lifting theorem with applications to symmetric functions. In Satya V. Lokam and R. Ramanujam, editors, *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2017, December 11-15, 2017, Kanpur, India*, volume 93 of *LIPIcs*, pages 23:1–23:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.FSTTCS.2017.23.

4 Arkadev Chattopadhyay, Nikhil S. Mande, Swagato Sanyal, and Suhail Sherif. Lifting to parity decision trees via stifling. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 33:1–33:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ITCS.2023.33.

5 Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. *J. ACM*, 67(4):23:1–23:28, 2020. doi:10.1145/3396695.

6 Gérard D. Cohen, Iiro S. Honkala, Simon Litsyn, and Antoine Lobstein. *Covering Codes*, volume 54 of *North-Holland mathematical library*. North-Holland, 2005.

7 Susanna F. de Rezende, Or Meir, Jakob Nordström, Toniann Pitassi, Robert Robere, and Marc Vinyals. Lifting with simple gadgets and applications to circuit and proof complexity. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 24–30. IEEE, 2020. doi:10.1109/FOCS46700.2020.00011.

8 Uma Girish, Ran Raz, and Avishay Tal. Quantum versus randomized communication complexity, with efficient players. *Comput. Complex.*, 31(2):17, 2022. doi:10.1007/s00037-022-00232-7.

9 Uma Girish, Makrand Sinha, Avishay Tal, and Kewen Wu. Fourier growth of communication protocols for XOR functions. *CoRR*, abs/2307.13926, 2023. doi:10.48550/arXiv.2307.13926.

10 Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing fourier dimensionality and sparsity. *SIAM J. Comput.*, 40(4):1075–1100, July 2011. doi:10.1137/100785429.

11 Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. Dimension-free bounds and structural results in communication complexity. *Israel Journal of Mathematics*, 253:555–616, 2023. doi:10.1007/s11856-022-2365-8.

12 L. H. Harper. *Global Methods for Combinatorial Isoperimetric Problems.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2004. doi:10.1017/CBO9780511616679.

13 Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Structure of protocols for XOR functions. *SIAM J. Comput.*, 47(1):208–217, 2018. doi:10.1137/17M1136869.

14 Stasys Jukna. *Boolean Function Complexity - Advances and Frontiers*, volume 27 of *Algorithms and combinatorics*. Springer, 2012. doi:10.1007/978-3-642-24508-4.

15 Sampath Kannan, Elchanan Mossel, Swagato Sanyal, and Grigory Yaroslavtsev. Linear sketching over f_2. In Rocco A. Servedio, editor, *33rd Computational Complexity Conference, CCC 2018, June 22-24, 2018, San Diego, CA, USA*, volume 102 of *LIPIcs*, pages 8:1–8:37. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.CCC.2018.8.

16 Eyal Kushilevitz and Noam Nisan. *Communication Complexity.* Cambridge University Press, 1996. doi:10.1017/CBO9780511574948.

17 Bruno Loff and Sagnik Mukhopadhyay. Lifting theorems for equality. In Rolf Niedermeier and Christophe Paul, editors, *36th International Symposium on Theoretical Aspects of Computer Science, STACS 2019, March 13-16, 2019, Berlin, Germany*, volume 126 of *LIPIcs*, pages 50:1–50:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.STACS.2019.50.

18 Nikhil S. Mande, Swagato Sanyal, and Suhail Sherif. One-way communication complexity and non-adaptive decision trees. In Petra Berenbrink and Benjamin Monmege, editors, *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference)*, volume 219 of *LIPIcs*, pages 49:1–49:24. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.STACS.2022.49.

19 Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *CoRR*, abs/0909.3392, 2009. arXiv:0909.3392.

20 A. Rao and A. Yehudayoff. *Communication Complexity: and Applications.* Cambridge University Press, 2020. URL: https://books.google.com/books?id=emw8PgAACAAJ.

21 Swagato Sanyal. Fourier sparsity and dimension. *Theory Comput.*, 15:1–13, 2019. doi:10.4086/toc.2019.v015a011.

22 Alexander A. Sherstov, Andrey A. Storozhenko, and Pei Wu. An optimal separation of randomized and quantum query complexity. *SIAM J. Comput.*, 52(2):525–567, 2023. doi:10.1137/22m1468943.

23 Makrand Sinha and Ronald de Wolf. Exponential separation between quantum communication and logarithm of approximate rank. In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 966–981. IEEE Computer Society, 2019. doi:10.1109/FOCS.2019.00062.

24 J.H. Spencer and L. Florescu. *Asymptopia.* Student mathematical library. American Mathematical Society, 2014. URL: https://books.google.com/books?id=uBMLugEACAAJ.

25 Hing Yin Tsang, Chung Hoi Wong, Ning Xie, and Shengyu Zhang. Fourier sparsity, spectral norm, and the log-rank conjecture. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 658–667. IEEE Computer Society, 2013. doi:10.1109/FOCS.2013.76.

26 Shengyu Zhang. Efficient quantum protocols for XOR functions. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 1878–1885. SIAM, 2014. doi:10.1137/1.9781611973402.136.

27 Zhiqiang Zhang and Yaoyun Shi. On the parity complexity measures of boolean functions. *Theor. Comput. Sci.*, 411(26-28):2612–2618, 2010. doi:10.1016/j.tcs.2010.03.027.