Sensitivity Conjecture and Log-Rank Conjecture for Functions with Small Alternating Numbers

Chengyu Lin and Shengyu Zhang

1 Columbia University, New York, NY, USA
chengyu@cs.columbia.edu
2 The Chinese University of Hong Kong, Hong Kong, China
syzhang@cse.cuhk.edu.hk

Abstract

The Sensitivity Conjecture and the Log-rank Conjecture are among the most important and challenging problems in concrete complexity. Incidentally, the Sensitivity Conjecture is known to hold for monotone functions, and so is the Log-rank Conjecture for \( f(x \wedge y) \) and \( f(x \oplus y) \) with monotone functions \( f \), where \( \wedge \) and \( \oplus \) are bit-wise AND and XOR, respectively. In this paper, we extend these results to functions \( f \) which alternate values for a relatively small number of times on any monotone path from \( 0^n \) to \( 1^n \). These deepen our understandings of the two conjectures, and contribute to the recent line of research on functions with small alternating numbers.

1998 ACM Subject Classification F.1.3 [Complexity Measures and Classes] Relations among Complexity Measures

Keywords and phrases Analysis of Boolean functions, Sensitivity Conjecture, Log-rank Conjecture, Alternating Number

Digital Object Identifier 10.4230/LIPIcs.ICALP.2017.51

1 Introduction

A central topic in Boolean function complexity theory is relations among different combinatorial and computational measures [29]. For Boolean functions, there is a large family of complexity measures such as block sensitivity, certificate complexity, decision tree complexity (including its randomized and quantum versions), degree (including its approximate version), etc, that are all polynomially related [13]. One outlier\(^1\) is sensitivity, which a priori could be exponentially smaller than the ones in that family. The famous Sensitivity Conjecture raised by Nisan and Szegedy [48] says that sensitivity is also polynomially related to the block sensitivity and others in the family. Despite a lot of efforts, the best upper bound we know is still exponential: \( \text{bs}(f) \leq \text{C}(f) \leq \left( \frac{8}{7} + o(1) \right) \text{s}(f)^{2\text{bs}(f)^{-1}} \) from [25], improving upon previous work [54, 2, 3]. See a recent survey [24] about this conjecture and how it has resisted many serious attacks.

Communication complexity quantifies the minimum amount of communication required for computing functions whose inputs are distributed among two or more parties [31].

\(^*\) This work was supported by Research Grants Council of the Hong Kong S.A.R. (Project no. CUHK14239416). The first author was also supported by NSF grants #CNS-1445424 and #CCF-1423306.

\(^1\) Most of this work was done when Chengyu Lin was in Chinese University of Hong Kong.

\(^2\) There are complexity measures, such as \( F_2 \)-degree, polynomial threshold degree, total influence, Boolean circuit depth, CNF/DNF size, that are known not to belong to the polynomially equivalent class. But the position of sensitivity is elusive.
the standard bipartite setting, the function $F$ has two inputs $x$ and $y$, with $x$ given to Alice and $y$ to Bob. The minimum number of bits needed to be exchanged to compute $F(x, y)$ for all inputs $(x, y)$ is the communication complexity $CC(F)$. It has long been known [41] that the logarithm of the rank of communication matrix $M_F \overset{\text{def}}{=} [F(x, y)]_{x,y}$ is a lower bound of $CC(F)$. Perhaps the most prominent and long-standing open question about communication complexity is the Log-rank Conjecture proposed by Lovász and Saks [37], which asserts that $CC(F)$ of any Boolean function $F$ is also upper bounded by a polynomial of $\log \text{rank}(M_F)$. The conjecture has equivalent forms related to chromatic number conjecture in graph theory [37], nonnegative rank [36], Boolean roots of polynomials over real numbers [61], quantum sampling complexities [4, 65], etc. Despite a lot of efforts devoted to the conjecture in the past decades, the best upper bound is $CC(F) = O(\sqrt{\text{rank}(M_F)} \log (\text{rank}(M_F)))$ by Lovett [39], which is still exponentially far from the target.

While these two conjectures are notoriously challenging in their full generality, special classes of functions have been investigated. In particular, the Sensitivity Conjecture is confirmed to hold for monotone functions, as the sensitivity coincides with block sensitivity and certificate complexity for those functions [47]. The Log-rank Conjecture is not known confirmed to hold for monotone functions, as the sensitivity coincides with block sensitivity classes of functions have been investigated. In particular, the Sensitivity Conjecture is being monotone. The first class contains AND functions $f$ by $\text{AND}(x, y) = f(x \land y)$, where $\land$ is the bit-wise AND of $x, y \in \{0, 1\}^n$. Taking the outer function $f$ to be the $n$-bit OR, we get Disjointness, the function that has had a significant impact on both communication complexity theory itself [53] and applications to many other areas such as streaming, data structures, circuit complexity, proof complexity, game theory and quantum computation [15]. The AND functions also contain other well known functions such as Inner Product, AND-OR trees [28, 33, 27, 19], and functions exhibiting gaps between communication complexity and log-rank [49]. The second class is XOR functions $F = f \circ \oplus$, defined by $F(x, y) = f(x \oplus y)$, where $\oplus$ is the bit-wise XOR function. This class includes Equality [62, 46, 1, 7, 11] and Hamming Distance [63, 16, 26, 34, 35] as special cases.

Both AND and XOR functions have recently drawn much attention [38, 12, 67, 32, 42, 55, 35, 59, 66, 50, 64], partly because their communication matrix rank has intimate connections to the polynomial representations of the outer function $f$. Specifically, the rank of $M_{f \circ \land}$ is exactly the Möbius sparsity $\overset{\text{def}}{=} \text{mono}(f)$, the number of nonzero coefficients $\alpha(S)$ in the multilinear polynomial representation $f(x) = \sum_{S \subseteq [n]} \alpha(S) \prod_{i \in S} x_i$ for $f : \{0, 1\}^n \rightarrow \{0, 1\}$ [12]. And the rank of $M_{f \circ \oplus}$ is exactly the Fourier sparsity $\|\hat{f}\|_0$, the number of nonzero Fourier coefficients $\hat{f}(S)$ in the multilinear polynomial representation $f(x) = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} x_i$ for $f : \{+1, -1\}^n \rightarrow \{0, 1\}$.

It is known that the Log-rank Conjecture holds for these two classes of functions when the outer function $f$ is monotone [38, 42], and this work aims to extend these as well as the sensitivity result on monotone functions, to functions that are close to being monotone. One needs to be careful about the distance measure here, since the widely-used (e.g. in property testing and computational learning) normalized Hamming distance $\text{dist}(f, g) = \mathbb{P}_{x \in \{0, 1\}^n} [f(x) \neq g(x)]$ does not meet our requirement. Indeed, if we flip the value $f(x)$ at just one input $x$, then this changes $f$ by an exponentially small amount measured by $\text{dist}$, but the sensitivity would change from a small $s(f)$ to a large $n - s(f)$. Similarly, the Fourier sparsity is also very sensitive to local changes $(\|\hat{f}\|_0$ to $2^n - \|\hat{f}\|_0$, and so is Moubius sparsity

\footnote{Named after the Möbius transform from $f$ to $\alpha$.}
if we flip the value at $0^n$.

One robust distance measure to monotone functions, which has recently drawn an increasingly amount of attention, is the alternating number, defined as follows. View the Boolean hypercube $\{0, 1\}^n$ as a lattice with the partial order $x \leq y$ if $x_i \leq y_i$ for all $i$. A path $x^{(1)} \rightarrow \cdots \rightarrow x^{(k)}$ on $\{0, 1\}^n$ is monotone if $x^{(i)} \prec x^{(i+1)}$ for all $i$. The alternating number of a function $f$ on $\{0, 1\}^n$ is the maximum number of $i$'s with $f(x^{(i)}) \neq f(x^{(i+1)})$, on any monotone path $x^{(0)} \rightarrow \cdots \rightarrow x^{(n)}$ from $0^n$ to $1^n$. It is clear that constant functions have alternating number 0, and monotone functions have alternating number 1. For general functions $f$, we have $\text{alt}(f) \leq n$, thus $\text{alt}(f)$ is a sub-linear complexity measure. The smaller $\text{alt}(f)$ is, the closer it is to monotone functions. Studies of the alternating number dated back to [40], in which Markov showed that the inversion complexity, the minimum number of negation gates needed in any Boolean circuit computing $f$, is exactly $\lceil \log_2(\text{alt}(f) + 1) \rceil$. Later work investigated the inversion complexity/alternating number over computational models such as constant-depth circuit [52], bounded-depth circuit [56], Boolean formula [43], and non-deterministic circuit [44]. It has been recently shown that small alternating number can be exploited in learning Boolean circuits [10]. Also there are some studies in cryptography considering the effect of negation gates [23].

In this paper, we study the Sensitivity and Log-rank Conjectures for functions whose alternating numbers are small, compared to sensitivity, Möbius sparsity and Fourier sparsity. First, the following theorem shows that the Sensitivity Conjecture holds for non-degenerate functions

\begin{equation}
\text{bs}(f) = O(\text{alt}(f)^2 \cdot s(f)).
\end{equation}

Note that if a function is non-degenerate in the sense that it depends on all $n$ variables, then the sensitivity is at least $\Omega(\log n)$ [54], therefore the above theorem also confirms the Sensitivity Conjecture for non-degenerate functions $f$ with $\text{alt}(f) = \text{poly} \log n$.

The next two theorems confirmed the Log-rank Conjecture for $f \circ \oplus$ with $\text{alt}(f) = \text{poly} \log \|f\|_0$, and for $f \circ \land$ with $\text{alt}(f) = O(1)$.

\begin{equation}
\text{Theorem 1. For any function } f : \{0, 1\}^n \rightarrow \{0, 1\}, \text{ it holds that}
\end{equation}

\begin{equation}
\text{bs}(f) = O(\text{alt}(f)^2 \cdot s(f)).
\end{equation}

In the last theorem, the dependence on $\text{alt}(f)$ can be slightly improved (by a factor of 2) if a factor of $\log n$ is tolerated in the communication cost.

\begin{equation}
\text{Theorem 2. For any function } f : \{0, 1\}^n \rightarrow \{0, 1\}, \text{ it holds that}
\end{equation}

\begin{equation}
\text{CC}(f \circ \oplus) \leq 2 \cdot \text{alt}(f) \cdot \log^2 \text{rank}(M_{f \circ \oplus}).
\end{equation}

\begin{equation}
\text{Theorem 3. For any function } f : \{0, 1\}^n \rightarrow \{0, 1\}, \text{ it holds that}
\end{equation}

\begin{equation}
\text{CC}(f \circ \land) \leq O(\log^{\text{alt}(f)+1} \text{rank}(M_{f \circ \land})).
\end{equation}

\begin{equation}
\text{Related work}
\end{equation}

The Sensitivity Conjecture has many equivalent forms, summarized in the survey [24]. Also see the recent paper [18] which tries to solve this conjecture using a communication game approach. At the other end of the spectrum, [51, 5] seek the largest possible separation between sensitivity and block sensitivity, and [9] got a super-quadratic separation between sensitivity and query complexity. Apart from monotone functions [47], the Sensitivity Conjecture has also been confirmed on graph properties [60], cyclically-invariant function [14],

\begin{equation}
\text{ICALP 2017}
\end{equation}
read-once functions [45], functions admitting the Normalized Block property [57] and several cases of read-k formulas [8]. Other than the conjecture itself, some recent work [6, 21, 22] discussed combinatorial and computational structures of low-sensitivity functions.

For the Log-rank Conjecture, apart from the equivalent forms mentioned earlier, some seemingly weaker formulations in terms of the largest monochromatic rectangle size [49], randomized communication complexity and information cost [17] are actually equivalent to the original conjecture. For lower bounds, the best one had been \( \text{CC}(F) = \Omega\left( (\log \text{rank}(M_F))^{\log_2 n} \right) \) (attributed to Kushilevitz in [49]), achieved by an AND function, until the recent result of \( \text{CC}(F) = \tilde{\Omega}\left( \log^2 \text{rank}(M_F) \right) \) [20]. For XOR functions \( f \circ \oplus \), the Log-rank Conjecture is confirmed when \( f \) is symmetric [67], monotone [42], linear threshold functions (LTFs) [42], \( AC^0 \) functions [30], has low \( \mathbb{F}_2 \)-degree [59] or small spectral norm [59]. For AND functions \( f \circ \wedge \), it seems that the conjecture is only confirmed on monotone functions [38].

## 2 Preliminaries

### 2.1 n-bit (Boolean) functions

We use \([n]\) to denote the set \(\{1, 2, \ldots, n\}\). The all-0 \(n\)-bit string is denoted by \(0^n\) and the all-1 \(n\)-bit string is denoted by \(1^n\).

For a Boolean function \(f : \{0, 1\}^n \rightarrow \{0, 1\}\), its \(\mathbb{F}_2\)-degree is the degree of \(f\) viewed as a polynomial over \(\mathbb{F}_2\). Such functions \(f\) can be also viewed as polynomials over \(\mathbb{R}\): \(f(x) = \sum_{S \subseteq [n]} \alpha(S)x^S\), where \(x^S = \prod_{i \in S} x_i\). If we represent the domain by \(+1, -1\)^n, then the polynomial (still over \(\mathbb{R}\)) changes to \(f(x) = \sum_{S \subseteq [n]} \hat{f}(S)x^S\), usually called Fourier expansion of \(f\). The coefficients \(\alpha(S)\) and \(\hat{f}(S)\) capture many important combinatorial properties of \(f\). We denote by \(\text{mono}(f)\) the Möbius sparsity, the number of non-zero coefficients \(\alpha(S)\), and by \(\|\hat{f}\|_0\) the Fourier sparsity, the number of non-zero coefficients \(\hat{f}(S)\). Some basic facts used in this paper are listed as follows.

> **Fact 4.** For any \(f : \{0, 1\}^n \rightarrow \{0, 1\}\), \(\deg_2(f) = n\) if and only if \(|f^{-1}(1)|\) is odd.

> **Fact 5.** For any \(f : \{0, 1\}^n \rightarrow \{0, 1\}\), \(\deg_2(f) \leq \log \|\hat{f}\|_0\).

For any input \(x \in \{0, 1\}^n\) and \(i \in [n]\), let \(x^i\) be the input obtained from \(x\) by flipping the value of \(x_i\). For a Boolean function \(f : \{0, 1\}^n \rightarrow \{0, 1\}\) and an input \(x\), if \(f(x) \neq f(x^i)\), then we say that \(x\) is sensitive to coordinate \(i\), and \(i\) is a sensitive coordinate of \(x\). We can also define these for blocks. For a set \(B \subseteq [n]\), let \(x^B\) be the input obtained from \(x\) by flipping \(x_i\) for all \(i \in B\). Similarly, if \(f(x) \neq f(x^B)\), then we say that \(x\) is sensitive to block \(B\), and \(B\) is a sensitive block of \(x\). The sensitivity \(s(f, x)\) of function \(f\) on input \(x\) is the number of sensitive coordinates \(i\) of \(x\): \(s(f, x) = |\{i \in [n] : f(x) \neq f(x^i)\}|\), and the sensitivity of function \(f\) is \(s(f) = \max_x s(f, x)\). It is easily seen that the \(n\)-bit AND and OR functions both have sensitivity \(n\). The block sensitivity \(bs(f, x)\) of function \(f\) on input \(x\) is the maximum of disjoint sensitive blocks of \(x\), and the block sensitivity of function \(f\) is \(bs(f) = \max_x bs(f, x)\). Note that there are always \(bs(f, x)\) many disjoint minimal sensitive blocks, in the sense that any \(B \subseteq B_x\) is not a sensitive block of \(x\).

For a Boolean function \(f : \{0, 1\}^n \rightarrow \{0, 1\}\) and an input \(x \in \{0, 1\}^n\), the certificate complexity \(C(f, x)\) of function \(f\) on input \(x\) is the minimal number of variables restricting the value of which fixes the function to a constant. The certificate complexity of \(f\) is \(C(f) = \max_{x} C(f, x)\), and the minimal certificate complexity of \(f\) is \(C_{\text{min}}(f) = \min_{x} C(f, x)\). The decision tree complexity \(\text{DT}(f)\) of function \(f\) is the minimum depth of any decision tree that computes \(f\).
A subfunction or a restriction of a function \( f \) on \( \{0, 1\}^n \) is obtained from \( f \) by restricting the values of some variables. Sometimes we say to restrict \( f \) to above an input \( d \), or to take the subfunction \( f' \) over \( \{x : x \geq d\} \), then we mean to restrict variables \( x_i \) to be 1 whenever \( d_i = 1 \). Similarly, we say to restrict \( f \) to under an input \( u \), or to take the subfunction \( f' \) over \( \{x : x \leq u\} \), meaning to restrict \( x_i \) to be 0 whenever \( u_i = 0 \).

Let \( \mathcal{F}_n \) be the set of all the real-valued functions on \( \{0, 1\}^n \). A complexity measure 

\[
M : \bigcup_{n=1}^\infty \mathcal{F}_n \to \mathbb{R}
\]

is downward non-increasing if \( M(f') \leq M(f) \) for all subfunction \( f' \) of \( f \). That is, restricting variables does not increase the measure \( M \). It is easily seen that the \( \mathbb{F}_2 \)-degree, the alternating number, the decision tree complexity, the sensitivity, the block sensitivity, the certificate complexity, the Fourier sparsity, are all downward non-increasing. When \( M \) is not downward non-increasing, it makes sense to define the closure 

\[
M^{\text{clo}}(f) = \max_{f'} M(f')
\]

where the maximum is taken over all subfunctions \( f' \) of \( f \). In particular, \( C_{\text{min}}^{\text{clo}}(f) = \max_{f'} C_{\text{min}}(f') \). The next theorem relates decision tree complexity to \( C_{\text{min}}^{\text{clo}} \).

Theorem 6 ([59]). For any \( f : \{0, 1\}^n \to \{0, 1\} \), it holds that 

\[
\text{DT}(f) \leq C_{\text{min}}^{\text{clo}}(f) \deg_2(f).
\]

(The original theorem proved was actually \( \text{PDT}(f) \leq C_{\text{min}}^{\text{clo}}(f) \deg_2(f) \), where \( \text{PDT}(f) \) is the parity decision tree complexity and \( C_{\text{min}}^{\text{clo}}(f) \) is the parity minimum certificate complexity. But as observed by [58], the same argument applies to standard decision tree as well.)

For general Boolean functions \( f \), we have \( s(f) \leq b_s(f) \leq C(f) \). But when \( f \) is monotone, equalities are achieved.

Fact 7. If \( f : \{0, 1\}^n \to \{0, 1\} \) is monotone, then \( s(f) = b_s(f) = C(f) \).

Fact 8 ([42]). If \( f : \{0, 1\}^n \to \{0, 1\} \) is monotone, then \( s(f) \leq \deg_2(f) \).

One can associate a partial order \( \preceq \) to the Boolean hypercube \( \{0, 1\}^n \): \( x \preceq y \) if \( x_i \leq y_i \) for all \( i \). We also write \( y \succeq x \) when \( x \preceq y \). If \( x \preceq y \) but \( x \neq y \), then we write \( x \prec y \) and \( y \succ x \). A path \( x^{(1)} \to \cdots \to x^{(k)} \) on \( \{0, 1\}^n \) is monotone if \( x^{(i)} \prec x^{(i+1)} \) for all \( i \).

Definition 9. For any function on \( \{0, 1\}^n \), the alternating number of a path \( x^{(1)} \to \cdots \to x^{(k)} \) is the number of \( i \in \{1, 2, \ldots, k-1\} \) with \( f(x^{(i)}) \neq f(x^{(i+1)}) \). The alternating number \( \text{alt}(f, x) \) of input \( x \in \{0, 1\}^n \) is the maximum alternating number of any monotone path from \( 0^n \) to \( x \), and the alternating number of a function \( f \) is \( \text{alt}(f) = \text{alt}(f, 1^n) \). Equivalently, one can also define \( \text{alt}(f) \) to be the largest \( k \) such that there exists a list \( \{x^{(1)}, x^{(2)}, \ldots, x^{(k+1)}\} \) with \( x^{(i)} \preceq x^{(i+1)} \) and \( f(x^{(i)}) \neq f(x^{(i+1)}) \), for all \( i \in [k] \).

A function \( f : \{0, 1\}^n \to \mathbb{R} \) is monotone if \( f(x) \leq f(y) \), \( \forall x \preceq y \). A function \( f : \{0, 1\}^n \to \mathbb{R} \) is anti-monotone if \( f(x) \leq f(y) \), \( \forall x \succeq y \). It is not hard to see that \( \text{alt}(f) = 0 \) iff \( f \) is constant, and \( \text{alt}(f) = 1 \) iff \( f \) is monotone or anti-monotone.

Definition 10. For a function \( f \) on \( \{0, 1\}^n \), an input \( u \in \{0, 1\}^n - \{1^n\} \) is called a max term if \( f(u) \neq f(1^n) \), and \( f(x) = f(1^n) \) for all \( x > u \). An input \( d \in \{0, 1\}^n - \{0^n\} \) is called a min term if \( f(d) \neq f(0^n) \), and \( f(x) = f(0^n) \) for all \( x < d \).

2.2 Communication complexity

Suppose that for a bivariate function \( F(x, y) \), the input \( x \) is given to Alice and \( y \) to Bob. The (deterministic) communication complexity \( \text{CC}(F) \) is the minimum number of bits needed to be exchanged by the best (deterministic) protocol that computes \( F \) (on the worst-case input).
The rank (over $\mathbb{R}$) of the communication matrix for bit-wise composed functions coincides with some natural parameters of the outer function $f$. For XOR functions $f \circ \oplus$, it holds that $\text{rank}(M_{f \circ \oplus}) = |\hat{f}|_0$, and for AND functions $f \circ \wedge$, it holds that $\text{rank}(M_{f \circ \wedge}) = \text{mono}(f)$. When $f$ is OR function of $n$ variables, we have $\text{rank}(M_{f \circ \lor}) = \text{mono}(\text{OR}_n) = 2^n - 1$.

It is well known that communication can simulate queries. More specifically, for XOR functions and AND functions, we have that

\[ \text{CC}(f \circ \wedge) \leq 2\text{DT}(f) \text{ and } \text{CC}(f \circ \oplus) \leq 2\text{DT}(f). \]  

(1)

In a $\{0,1\}$-communication matrix $M$, for $b \in \{0,1\}$, a $b$-rectangle is a submatrix of all entries equal to $b$. The $b$-covering number $\text{Cover}_b(M)$ of matrix $M$ is the minimum number of $b$-rectangles that can cover all $b$ entries in $M$. (These $b$-rectangles need not be disjoint.) For notational convenience, we sometimes write $\text{Cover}_1(F)$ for $\text{Cover}_1(M_F)$. Lovász [36] showed the following bounds.

**Theorem 11** ([36]). For any Boolean function $F(x,y)$, it holds that

\[ \log \text{Cover}_0(M_F) \leq \text{CC}(F) \leq \log \text{Cover}_0(M_F) \cdot \log \text{rank}(M_F). \]

### 3 The Sensitivity Conjecture

This section is devoted to the proof of Theorem 1. We will first show the following lemma, in which the first statement is used in this section and the second statement will be used in Section 4 for proving the Log-rank Conjecture of XOR functions.

**Lemma 12.** For any $f : \{0,1\}^n \rightarrow \{0,1\}$, it holds that

1. $\max \{C(f,0^n), C(f,1^n)\} \leq \text{alt}(f) \cdot s(f)$
2. $\max \{C(f,0^n), C(f,1^n)\} \leq \text{alt}(f) \cdot \deg_2(f)$

**Proof.** First note that it suffices to prove the two upper bounds for $C(f,0^n)$, because then we can take $g(x) = f(\hat{x})$ to get that $C(f,1^n) = C(g,0^n) \leq \text{alt}(g) \cdot s(g) = \text{alt}(f) \cdot s(f)$.

We prove upper bounds on $C(f,0^n)$ by induction on $\text{alt}(f)$. When $\text{alt}(f) = 1$, the function is either monotone or anti-monotone, thus

\[ C(f,0^n) \leq C(f) = s(f) \leq \deg_2(f), \]

where the first inequality is by definition of $C(f,0^n)$, the middle equality is by Fact 7 and the last inequality is because $s(f) \leq \deg_2(f)$ for monotone $f$ (Fact 8). Now we assume that the inequalities in the lemma hold for $\text{alt}(f) < a$ and we will show that they hold for $f$ with $\text{alt}(f) = a$ as well. Let $u$ be a max term of $f$. Define $S_0(u) \triangleq \{i \in [n] : u_i = 0\}$, and consider the subcube above $u$: $\{x : x \supset u\}$. Let $f'$ be the subfunction obtained by restricting $f$ on this subcube. By the definition of max term $f(u) \neq f(u')$ for all $i \in S_0(u)$. Therefore,

\[ |S_0(u)| \leq s(f,u) \leq s(f). \]  

(2)

We know that any point $z \succ u$ has $f(z) = f(1^n) \neq f(u)$. So the number of 1-inputs of $f'$ is odd, implying that $\deg_2(f') = |S_0(u)|$ (Fact 4). Thus we have

\[ |S_0(u)| = \deg_2(f') \leq \deg_2(f). \]  

(3)

Now consider another restriction of $f$, this time to the subcube under $u$, i.e. $\{x : x \preceq u\}$. This is implemented by restricting all variables in $S_0(u)$ to 0, yielding a subfunction $f''$ with
\[\text{alt}(f') \leq \text{alt}(f) - 1.\] Using induction hypothesis, we have that
\[
C(f'', 0^{[n] - S_0(u)}) \leq \text{alt}(f'') \cdot \min\{s(f''), \deg_2(f'')\} \\
\leq (\text{alt}(f) - 1) \cdot \min\{s(f), \deg_2(f)\}
\]
(4)

Recall that \(f''\) is obtained from \(f\) by restricting \(|S_0(u)|\) variables, thus
\[
C(f, 0^n) \leq |S_0(u)| + C(f'', 0^{[n] - S_0(u)}).
\]
Plugging Eq.(2) and Eq.(4) into the above inequality gives
\[
C(f, 0^n) \leq \text{alt}(f) \cdot \min\{s(f), \deg_2(f)\},
\]
completing the induction.

Now we are ready to prove the following theorem, which gives an explicit constant for Theorem 1.

\textbf{Theorem 13.} For any boolean function \(f\),
\[
bs(f) \leq \begin{cases} 
C_t \cdot s(f) & \text{if } \text{alt}(f) = 2t, \\
(C_t + 1) \cdot s(f) & \text{if } \text{alt}(f) = 2t + 1,
\end{cases}
\]
where \(C_t = \sum_{i=1}^{t} (i + 2) = \frac{1}{2} t(t + 5)\).

\textbf{Proof.} We prove Eq.(5) by induction on \(t = |\text{alt}(f)/2|\). Clearly it holds when \(t = 0\): If \(\text{alt}(f) = 0\) then \(f\) is a constant function and \(bs(f) = s(f) = 0\). When \(\text{alt}(f) = 1\), \(f\) is monotone or anti-monotone, thus \(bs(f) = s(f)\).

Now for any Boolean function \(f\) with \(\text{alt}(f) > 1\), we first consider the case when \(\text{alt}(f) = 2t \geq 2\). We will bound the block sensitivity for each input \(x\). Consider the following possible properties for \(x\):

1. there exists a max term \(u\) of \(f\) such that \(x \preceq u\);
2. there exists a min term \(d\) of \(f\) such that \(x \succeq d\).

\textbf{Case 1:} \(x\) satisfies at least one of the above conditions. Without loss of generality assume it satisfies the first one; the other case can be similarly argued. Fix such a max term \(u \succeq x\).

By definition of max term, we know that \(\text{alt}(f, u) \leq \text{alt}(f) - 1\), and that \(u\) is sensitive to all \(i \in S_0(u) \overset{\text{def}}{=} \{i : u_i = 0\}\). Therefore, \(|S_0(u)| \leq s(f, u) \leq s(f)\).

Let \(f'\) be the subfunction of \(f\) restricted on the subcube \(\{t : t \preceq u\}\), then \(\text{alt}(f') = \text{alt}(f, u) \leq \text{alt}(f) - 1 = 2t - 1 = 2(t - 1) + 1\).

By induction hypothesis and the fact that sensitivity is downward non-increasing, we have
\[
bs(f', x) \leq bs(f') \leq (C_{t-1} + 1) \cdot s(f') \leq (C_{t-1} + 1) \cdot s(f).
\]
(6)

Next it is not hard to see that
\[
bs(f, x) \leq bs(f', x) + |S_0(u)|.
\]
(7)

Indeed, take any disjoint minimal sensitive blocks \(B_1, \ldots, B_\ell \subseteq [n]\) of \(x\) (with respect to \(f\)), where \(\ell = bs(f, x)\). If \(B_i \subseteq [n] - S_0(u)\), then \(x\) is still sensitive to \(B_i\) in \(f'\). As the \(B_i\)'s are disjoint, at most \(|S_0(u)|\) many \(B_i\)'s are not contained in \([n] - S_0(u)\), thus at least

ICALP 2017
bs(f, x) - |S₀(u)| blocks Bᵢ are still sensitive blocks of x in f'. Therefore, bs(f, x) - |S₀(u)| ≤ bs(f', x), as Eq.(7) claimed.

Combining Eq.(6), Eq.(7), and the fact that |S₀(u)| ≤ s(f), we conclude that

\[ bs(f, x) \leq bs(f', x) + \left| S₀(u) \right| \leq (Cₜ₋₁ + 1) \cdot s(f') + s(f) \leq (Cₜ₋₁ + 2) \cdot s(f), \]

which is at most \( Cₜ \cdot s(f) \) by our setting of parameter \( Cₜ = \sum_{i=1}^{t}(i + 2) = Cₜ₋₁ + t + 2 \).

**Case 2:** \( x \) satisfies neither of the conditions 1 and 2. So \( f(x) \) needs to be the same with both \( f(0^n) \) and \( f(1^n) \), and \( f \) is constant on both subcubes \( \{ t : t \geq x \} \) and \( \{ t : t \leq x \} \). Otherwise we can take a minimal \( d \) such that \( d \leq x \) and \( f(d) = f(x) \not= f(0^n) \) and by definition \( d \) is a min term, or take the maximal \( u \) where \( u \geq x \) and \( f(u) = f(x) = f(1^n) \) and by definition \( u \) is a max term.

Fix \( \ell = bs(f, x) \) disjoint minimal sensitive blocks \( \{ B₁, B₂, \ldots, Bₗ \} \) of \( x \). For each block \( Bᵢ \), decompose it into \( Bᵢ = Uᵢ \cup Dᵢ \) where \( Uᵢ = \{ i \in Bᵢ : xᵢ = 1 \} \) and \( Dᵢ = \{ i \in Bᵢ : xᵢ = 0 \} \), as depicted below.

\[ x = \begin{array}{cccc}
D₁ & U₁ & D₂ & U₂ \\
B₁ & (0 \ldots 0 1 \ldots 1) & B₂ & (0 \ldots 1 0 \ldots 1) \\
& & & \\
Dₗ & Uₗ & & \\
Bₗ & (0 \ldots 0 1 \ldots 1) & & (0 \ldots 1 0 \ldots 1)
\end{array} \]

First we will show that for each \( i, x^{Uᵢ} \) satisfies condition 1 and \( x^{Dᵢ} \) satisfies condition 2, i.e. there exist some max term \( uᵢ \geq x^{Uᵢ} \) and some min term \( dᵢ \leq x^{Dᵢ} \). (See Figure 1 for an illustration.) Indeed, for any sensitive block \( Bᵢ \) of \( x \), \( f(x^{Bᵢ}) \not= f(x) = f(0^n) = f(1^n) \). Take a maximal \( uᵢ \) such that \( uᵢ \geq x^{Bᵢ} \) and \( f(uᵢ) = f(x^{Bᵢ}) \). By definition \( uᵢ \) is a max term. Similarly we can take a min term \( dᵢ \) where \( dᵢ \leq x^{Bᵢ} \). Then from the definition of \( Uᵢ \) and \( Dᵢ \) we can conclude that \( x^{Uᵢ} \succeq x^{Bᵢ} \succeq uᵢ \) and \( x^{Dᵢ} \succeq x^{Bᵢ} \succeq dᵢ \). Moreover, both \( Uᵢ \) and \( Dᵢ \) cannot be empty, since otherwise either \( x \succeq x^{Dᵢ} = x^{Bᵢ} \succeq uᵢ \) or \( x \succeq x^{Uᵢ} = x^{Bᵢ} \succeq dᵢ \), contradicting our assumption of case 2. This further indicates that \( f(x) = f(x^{Uᵢ}) = f(x^{Dᵢ}) \) as we have taken each \( Bᵢ \) to be a minimal sensitive block.

Next we are going to find some \( Uᵢ \) or \( Dᵢ \) such that \( x^{Uᵢ} \) or \( x^{Dᵢ} \) is sensitive to most \( Bᵢ \)’s. In this case if there are many sensitive blocks of input \( x \), \( x^{Uᵢ} \) or \( x^{Dᵢ} \) must have high block sensitivity. But we have eliminate this possibility in case 1. To achieve this, we count the following two quantities:

- \( \#U \): the number of pairs \((i, j)\) such that \( i \neq j \) and \( f(x^{Uᵢ}) \neq f(x^{Uᵢ∪Bⱼ}) \),
- \( \#D \): the number of pairs \((i, j)\) such that \( i \neq j \) and \( f(x^{Dᵢ}) \neq f(x^{Dᵢ∪Bⱼ}) \).
Recall that $f(x) = f(x^{U_i}) = f(x^{D_i})$ and $f(x) \neq f(x^{B_j})$, thus it is equivalent to counting

$\#U :$ the number of pairs $(i, j)$ such that $i \neq j$ and $f(x^{B_j}) = f(x^{U_i \cup B_j})$,

$\#D :$ the number of pairs $(i, j)$ such that $i \neq j$ and $f(x^{B_j}) = f(x^{D_i \cup B_j})$.

Now we bound the number of such $i$'s for each $j$. Fix a block $B_j$, and consider the subfunction $f^u$ on the subcube $\{z : z \geq x^{B_j}\}$ and the subfunction $f^d$ on the subcube $\{z : z \leq x^{B_j}\}$. Let us look at $f^u$ first. Because $D_i \cap B_j = \emptyset$ whenever $i \neq j$, $x^{D_i \cup B_j} \succeq x^{B_j}$ which lies in the domain of $f^u$. By the definition of certificate complexity of $f$ on input $x^{B_j}$, there is a subcube $C$ of co-dimension $C(f^u, x^{B_j})$ (with respect to $\{z : z \geq x^{B_j}\}$) containing $x^{B_j}$, s.t. $f$ takes a constant 0/1 value on $C$. Denote by $S$ the set of coordinates in this certificate. Then $S \subseteq \{k \in [n] : (x^{B_j})_k = 0\}$ and $|S| = C(f, x^{B_j})$. Now for each $D_i$, if $D_i \cap S = \emptyset$, then $f(x^{B_j}) = f(x^{D_i \cup B_j})$ as the values of the certificate variables are not flipped. As all $\{D_i\}_{i \neq j}$ are disjoint, at most $C(f^u, x^{B_j})$ many of $D_i$’s may intersect $S$. Thus $f(x^{B_j}) = f(x^{D_i \cup B_j})$ for all but at most $C(f^u, x^{B_j})$ many of $D_i$. Similarly we can say that all but at most $C(f^d, x^{B_j})$ many of $U_i$’s $(i \neq j)$ satisfy that $f(x^{B_j}) = f(x^{U_i \cup B_j})$. Applying Lemma 12 (statement 1), we have

\[
C(f^u, x^{B_j}) \leq alt(f^u) \cdot s(f^u) \leq alt(f^u) \cdot s(f),
\]

\[
C(f^d, x^{B_j}) \leq alt(f^d) \cdot s(f^d) \leq alt(f^d) \cdot s(f).
\]

Because $alt(f^u) + alt(f^d) = 2t$, and there are $t$ sensitive blocks $B_i$, thus from the second definition of $\#U$ and $\#D$ we can see that

\[
\#U + \#D \geq t \cdot ((\ell - 1 - alt(f^u) \cdot s(f)) + (\ell - 1 - alt(f^d) \cdot s(f))) \\
\geq 2t \cdot (\ell - 1 - t \cdot s(f)).
\]

(9)

Since there are $2t$ of $U_i$’s and $D_i$’s in total, by pigeonhole principle there must be a $T_i$ (being $U_i$ or $D_i$) that contributes to at least $(\#U + \#D)/2t \geq \ell - 1 - t \cdot s(f)$ to $(\#U + \#D)$. Fix this $T_i$. By definition of $\#U$ and $\#D$, there exist at least $\ell - 1 - t \cdot s(f)$ blocks $B_j$ with $i \neq j$ and $f(x^{T_i}) \neq f(x^{U_i \cup B_j})$. That is, $x^{T_i}$ is sensitive to at least $\ell - 1 - t \cdot s(f)$ blocks $B_j$ where $j \neq i$. Considering that $x^{T_i}$ is also sensitive to $B_i \setminus T_i$, we conclude that

\[
bs(f, x^{T_i}) \geq 1 + (\ell - 1 - t \cdot s(f)) = bs(f, x) - t \cdot s(f).
\]

Finally, recall that we have showed that $x^{T_i}$ satisfies one of the condition 1 and 2. Therefore $x^{T_i}$ is an input falling into case 1. By Eq.(8), we have $bs(f, x^{T_i}) \leq (C_{t-1} + 2) \cdot s(f)$. Putting everything together, we have

\[
bs(f, x) \leq bs(f, x^{T_i}) + t \cdot s(f) \leq (C_{t-1} + 2 + t) \cdot s(f) = C_t \cdot s(f).
\]

This finishes the proof for $alt(f) = 2t$.

When $alt(f) = 2t + 1$, for any input $x$, $f(x)$ must differ from either $f(0^n)$ or $f(1^n)$ since $f(0^n) \neq f(1^n)$. Without loss of generality, assume that $f(x) \neq f(0^n)$. Take the minimal $d$ such that $d \preceq x$ and $f(d) = f(x) \neq f(0^n)$. By definition $d$ is a min term and $x$ satisfies condition 2. Then using the same analysis above as in case 1, we can show $bs(f, x) \leq (C_t + 1) \cdot s(f)$ and this finishes the proof.

4 The Log-Rank Conjecture

We prove Theorem 2 and 3 in this section. We start with Theorem 2, which is now easy given Lemma 12. Recall that the second statement of Lemma 12 says that $\max\{C(f, 0^n), C(f, 1^n)\} \leq \ell - 1 - \ell t$. Therefore $\ell - 1 - \ell t$ is an upper bound on the certificate complexity of $f$. Since $C(f, 0^n) \leq C(f, 1^n)$, we have $C(f, 0^n) \leq \ell - 1 - \ell t$.
$\text{alt}(f) \cdot \text{deg}_2(f)$, therefore

$$C_{\text{min}}(f) \leq \text{alt}(f) \cdot \text{deg}_2(f). \quad (10)$$

As both $\text{alt}(f)$ and $\text{deg}_2(f)$ are downward non-increasing, applying Eq. (10) to all subfunctions of $f$ yields $C_{\text{min}}(f) \leq \text{alt}(f) \cdot \text{deg}_2(f)$. Since $\text{DT}(f) \leq C_{\text{min}}^{\text{alt}}(f) \cdot \text{deg}_2(f)$ (Theorem 6) we get the following.

**Theorem 14.** For any $f : \{0,1\}^n \rightarrow \{0,1\}$, it holds that $\text{DT}(f) \leq \text{alt}(f) \cdot \text{deg}_2(f)^2$.

Theorem 2 follows from this together with the fact that $\text{CC}(f \circ \oplus) \leq 2\text{DT}(f)$ (Eq. (1)) and that $\text{deg}_2(f) \leq \log \|f\|_0 = \log \text{rank}(M_{f \circ \oplus})$ (Fact 5).

Note that if we use the first statement of Lemma 12, we will get the following corollary, which gives better dependence on $\text{alt}(f)$ for low $F_2$-degree functions.

**Corollary 15.** $\text{DT}(f) \leq \text{alt}(f) \text{sgn}(f) \cdot \text{deg}_2(f)$.

Next we prove Theorem 3 for AND functions. Different than the above approach for XOR functions of going through $\text{DT}(f)$, we directly argue communication complexity of AND functions. Recall that Theorem 3 says that

$$\text{CC}(f \circ \land) \leq \min \{O(\log^{a+1} \text{rank}(M_{f \circ \land})), O(\log^{a-1} \text{rank}(M_{f \circ \land}) \log n)\}.$$  

**Proof of Theorem 3.** Without loss of generality, we can assume that $f(0^n) = 0$ since otherwise we can compute $\neg f$ first and negate the answer (note that $\text{rank}(M_{\neg \circ \land})$ differs from $\text{rank}(M_{f \circ \land})$ by at most 1). For notational convenience let us define $r = \text{mono}(f) = \text{rank}(M_{f \circ \land})$ and $\ell = \log r$. For $b \in \{0,1\}$, further define $C_b^{(a)}$ to be the maximum $\text{Cover}_b(f \circ \land)$ over all functions $f : \{0,1\}^n \rightarrow \{0,1\}$ with alternating number $a$ and $f(0^n) = 0$. We will give three bounds for $C_b^{(a)}$ in terms of $C_b^{(a-1)}$, and combining them gives the claimed result in Theorem 3.

**Bound 1, from max terms.** We apply this bound for $C_b^{(a)}$ when $a$ and $b$ have different parities, that is, when $a$ is even and $b = 1$, and when $a$ is odd and $b = 0$. Consider the first case and the second is similar. Take any Boolean function $f$ with $f(0^n) = 0$ and $\text{alt}(f) = a$ is even, we have $f(1^n) = 0$. Any 1-input is under some max term, so it is enough to cover inputs under max terms when bounding the $\text{Cover}_b(f)$. Take an arbitrary max term $u \in \{0,1\}^n$.

Suppose its Hamming weight is $s$. Considering the subfunction $f' \in \{1 : t \geq u\}$, which is an OR function of $n - s$ variables. In the communication setting, this is the Disjointness function of $n - s$ variables. Thus $\ell = \log \text{rank}(M_{f \circ \land}) \geq n - s$. This implies that all max terms $u$ of $f$ are $\ell$-close to $1^n$ in Hamming distance. Considering that different max terms are incomparable by definition, we know that the number of max terms is at most $\binom{n}{\ell}$.

Next we upper bound the 1-rectangles by giving a partition of set of 1-inputs into 1-rectangles. For each max term $u \in \{0,1\}^n$, let $U = \{i \in [n] : u_i = 1\}$, and $k = n - |U|$, then $k \leq \ell$. The submatrix $\{(x,y) : x,y \in \{0,1\}^n, x \land y \leq u\}$ is partitioned into $3^k$ submatrices as follows. Suppose that the set of 0-coordinates in $u$ is $\{i_1, \ldots, i_k\}$, then for each $i_j$, we can choose $(x_{ij}, y_{ij})$ from the set $\{(0,0), (0,1), (1,0)\}$ to enforce $x_{ij} \land y_{ij} = 0$. Thus there are $3^k$ ways of restricting these $k$ variables in $\bar{U}$, giving $3^k$ submatrices. Let $f_u : \{0,1\}^U \rightarrow \{0,1\}$ be the subfunction of $f$ restricted on the subcube $\{t : t \leq u\}$ where $f_u(z_U) = f(z_U, 0_U)$. (Here the input to $f$ is $x' \land y'$ at $U$ and 0 at $\bar{U}$.) Note that each of the $3^k$ submatrices is still the communication matrix of $f_u \circ \land$ for some max term $u$. Also note that this $f_u$ has $f_u(0_U) = 0$,
but \( f_u(1^u) = 1 \) and \( \alpha(f_u) \leq \alpha(f) - 1 \). Since all the 1-inputs of \( f \) are under some max term \( u \), the 1-covering number \( \text{Cover}_1(f \circ \wedge) \) can be upper bounded by the following:

\[
\text{Cover}_1(f \circ \wedge) \leq \sum_{u \text{max term}} 3^f \cdot \text{Cover}_1(f_u \circ \wedge) \leq \left( \frac{n}{f} \right) \cdot 3^f \cdot \max_{u \text{max term}} \text{Cover}_1(f_u \circ \wedge).
\]

Using the fact \( \alpha(f_u) \leq \alpha(f) - 1 \), and that the above inequality holds for any \( f \), we have the following bound on \( C_1^{(a)} \):

\[
\log C_1^{(a)} \leq 3^f \cdot \log n + \log C_1^{(a-1)}, \quad \text{when } a \text{ is even.}
\]

Similarly, when \( a \) is odd, \( f(1^n) = 1 \), and thus any 0-input is under some max term. A similar argument shows the following bound on \( C_0^{(a)} \):

\[
\log C_0^{(a)} \leq 3^f \cdot \log n + \log C_0^{(a-1)}, \quad \text{when } a \text{ is odd.}
\]

**Bound 2, from min terms.** Take any Boolean function \( f \) with \( f(0^n) = 0 \). Then any 1-input must be above some min term. Take any min term \( d \). Let \( D = \{ i : d_i = 1 \} \). If we restrict variables \( x_i \) and \( y_i \) to 1 for all \( i \in D \), then we go to a rectangle \( \{(x, y) : x_i = y_i = 1, \forall i \in D\} \). The union of these rectangles for all min terms \( d \) contains all 1-inputs. Restrict \( f \) on the subcube \( \{z : z \supset d\} \) to get a subfunction \( f_d \), which has \( f_d(0^{|d|}) = 1 \), and \( \alpha(f_d) \leq \alpha(f) - 1 \). Note that for each min term \( d \), we have \( \alpha(d) = \sum_{x \leq d} (-1)^{|d \oplus x|} f(x) = 1 \neq 0 \). which contributes 1 to \( \text{mono}(f) \), thus the number of min terms is at most \( \text{mono}(f) = r \). Since each 1-input of \( f \) is above some min term \( d \), the 1-covering number \( \text{Cover}_1(f) \) has

\[
\text{Cover}_1(f \circ \wedge) \leq \sum_{d \text{min term}} \text{Cover}_1(f_d \circ \wedge) \leq r \cdot \max_{d \text{min term}} \text{Cover}_1(f_d \circ \wedge).
\]

Note that \( \alpha(f_d) \leq \alpha(f) - 1 \), and \( f_d \) takes value 1 on its all-0 input, thus \( \text{Cover}_1(f_d \circ \wedge) = \text{Cover}_0(-f_d \circ \wedge) \leq C_0^{(a-1)} \) (note that the maximum in the definition of \( C_0 \) is over all \( f \) with \( f(0^n) = 0 \)). This implies

\[
\log C_1^{(a)} \leq \ell + \log C_0^{(a-1)}.
\]

Note that this inequality holds as long as \( f(0^n) = 0 \), regardless of the parity of \( a \).

**Bound 3, from CC.** When \( a \) is odd, we have a bound for \( C_0^{(a)} \) by Eq.(12) and a bound for \( C_1^{(a)} \) by Eq.(13). When \( a \) is even, we have two bounds for \( C_0^{(a)} \), Eq.(11) and Eq.(13), but no bound for \( C_1^{(a)} \). Note that we can always use CC to bound \( C_0^{(a)} \):

\[
\log \text{Cover}_0(f \circ \wedge) \leq \text{CC}(f \circ \wedge)
\]

\[
\leq \log \text{rank}(M_{f_{\circ \wedge}}) \cdot \log \text{Cover}_1(f \circ \wedge)
\]

\[
= \ell \cdot \log \text{Cover}_1(f \circ \wedge),
\]

This implies that

\[
\log C_0^{(a)} \leq \ell \cdot \log C_1^{(a)}.
\]

\[\text{If } f(0^n) = 1, \text{ then for each min term } d, \text{ we have } \alpha(d) = \sum_{x \leq d} (-1)^{|d \oplus x|} f(x) = -1, \text{ which is still non-zero.}\]
Similarly it also holds that \( \log C_1^{(a)}(a) \leq \ell \cdot \log C_0^{(a)} \).

Now we combine the three bounds and prove the theorem by induction on \( a \). In the base case of \( a = 0 \), the function is constant 0 and thus \( C_0^{(0)} = 1 \) and \( C_1^{(0)} = 0 \). For general \( a \), we can repeatedly apply Eq.(13) and Eq.(14) to get

\[
\log C_1^{(a)}(a) \leq \sum_{i=1}^{a} \ell^i = (1 + o(1))\ell^a.
\]

Thus \( \text{CC}(f \circ \wedge) \leq \ell \cdot \log C_1^{(a)}(a) \leq (1 + o(1))\ell^{a+1} \).

If we can tolerate a \( \log n \) factor, then the dependence on \( a \) can be made slightly better. Assume that \( a \) is even, we have

\[
\begin{align*}
\log C_1^{(a)}(a) &\leq \ell + \log C_0^{(a-1)} \\
&\leq \ell + 3\ell \log n + \log C_0^{(a-2)} \quad \text{(by Eq.(13))} \\
&\leq \ell + 3\ell \log n + \ell \log C_1^{(a-2)} . \quad \text{(by Eq.(12))}
\end{align*}
\]

Solving this recursion gives \( \log C_1^{(a)}(a) \leq O(\ell^2 \log n) \), and thus \( \text{CC} = O(\ell^{a+1} \log n) \). When \( a \) is odd, we can use Eq.(13) and Eq.(14) to reduce it to the “even \( a \)” case, resulting a bound \( \text{CC} \leq O(\ell^{a+1} \log n) \). Putting these two cases together, we get the claimed bound. ▶

**Acknowledgement.** The authors would like to thank Xin Huang for valuable discussions.

---

**References**

5. Andris Ambainis and Xiaoming Sun. New separation between \( s(f) \) and \( bs(f) \). *CoRR*, abs/1108.3494, 2011.


Penghui Yao. Parity decision tree complexity and 4-party communication complexity of xor-functions are polynomially equivalent. *arXiv*, 1506.02936, 2015.

