Sensitivity, Affine Transforms and Quantum Communication Complexity

Krishnamoorthy Dinesh[*]   Jayalal Sarma[*]

Abstract

In this paper, we study the Boolean function parameters sensitivity (s), block sensitivity (bs), and alternation (alt) under specially designed affine transforms and show several applications. For a function $f : \mathbb{F}_2^n \to \{0,1\}$, and $A = Mx + b$ for $M \in \mathbb{F}_2^{n \times n}$ and $b \in \mathbb{F}_2^n$, the result of the transformation g is defined as $\forall x \in \mathbb{F}_2^n, g(x) = f(Mx + b)$.

As a warm up, we study alternation under linear shifts (when M is restricted to be the identity matrix) called the *shift invariant alternation* (the smallest alternation that can be achieved for the Boolean function f by shifts, denoted by $\mathsf{salt}(f)$). By a result of Lin and Zhang [LZ17], it follows that $\mathsf{bs}(f) \leq O(\mathsf{salt}(f)^2 \mathsf{s}(f))$. Thus, to settle the Sensitivity Conjecture ($\forall f, \mathsf{bs}(f) \leq \mathsf{poly}(\mathsf{s}(f))$), it suffices to argue that $\forall f, \mathsf{salt}(f) \leq \mathsf{poly}(\mathsf{s}(f))$. However, we exhibit an explicit family of Boolean functions for which $\mathsf{salt}(f)$ is $2^{\Omega(\mathsf{s}(f))}$.

Going further, we use an affine transform A, such that the corresponding function g satisfies $\mathsf{bs}(f, 0^n) \leq \mathsf{s}(g)$, to prove that for $F(x,y) \overset{\text{def}}{=} f(x \wedge y)$, the bounded error quantum communication complexity of F with prior entanglement, $Q_{1/3}^*(F)$ is $\Omega(\sqrt{\mathsf{bs}(f, 0^n)})$. Our proof builds on ideas from Sherstov [She10] where we use specific properties of the above affine transformation. Using this, we show the following.

(a) For a fixed prime p and an ϵ, $0 < \epsilon < 1$, any Boolean function f that depends on all its inputs with $\deg_p(f) \leq (1-\epsilon)\log n$ must satisfy $Q_{1/3}^*(F) = \Omega\left(\frac{n^{\epsilon/2}}{\log n}\right)$. Here, $\deg_p(f)$ denotes the degree of the multilinear polynomial over $\mathbb{F}_p$ which agrees with f on Boolean inputs.

(b) For Boolean function f such that there exists primes p and q with $\deg_q(f) \geq \Omega(\deg_p(f)^\delta)$ for $\delta > 2$, the deterministic communication complexity - $\mathsf{D}(F)$ and $Q_{1/3}^*(F)$ are polynomially related. In particular, this holds when $\deg_p(f) = O(1)$. Thus, for this class of functions, this answers an open question (see [BdW01]) about the relation between the two measures.

Restricting back to the linear setting, we construct linear transformation A, such that the corresponding function g satisfies, $\mathsf{alt}(f) \leq 2\mathsf{s}(g) + 1$. Using this new relation, we exhibit Boolean functions f (other than the parity function) such that $\mathsf{s}(f)$ is $\Omega(\sqrt{\mathsf{sparsity}(f)})$ where $\mathsf{sparsity}(f)$ is the number of non-zero coefficients in the Fourier representation of f.

1 Introduction

For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, *sensitivity* of f on $x \in \{0,1\}^n$, is the maximum number of indices $i \in [n]$, such that $f(x \oplus e_i) \neq f(x)$ where $e_i \in \{0,1\}^n$ with exactly the ith bit

---

[*]Indian Institute of Technology Madras, Chennai, India. {kdinesh,jayalal}@cse.iitm.ac.in

as 1. The *sensitivity* of f (denoted by $\mathsf{s}(f)$) is the maximum sensitivity of f over all inputs. A related parameter is the *block sensitivity* of f (denoted by $\mathsf{bs}(f)$), where we allow disjoint blocks of indices to be flipped instead of a single bit. Another parameter is the deterministic *decision tree complexity* (denoted by $\mathsf{DT}(f)$) which is the depth of an optimal decision tree computing the function f. The *certificate complexity* of f (denoted by $\mathsf{C}(f)$) is the non-deterministic variant of the decision tree complexity. The parameter $\mathsf{s}(f)$ was originally studied by Cook *et al.* [CDR86] in connection with the CREW-PRAM model of computation. Subsequently, Nisan and Szegedy [NS92] (see also [Nis91]) introduced the parameters $\mathsf{bs}(f)$ and $\mathsf{C}(f)$ and conjectured that for any function $f : \{0,1\}^n \to \{0,1\}$, $\mathsf{bs}(f) \leq \mathsf{poly}(\mathsf{s}(f))$ - known as the SENSITIVITY CONJECTURE. Later developments, which revealed several connections between sensitivity, block sensitivity and the other Boolean function parameters, demonstrated the fundamental nature of the conjecture (see [HKP11] for a survey and several equivalent formulations of the conjecture). The best known upper bound for $\mathsf{bs}(f)$ in terms of $\mathsf{s}(f)$ is $\mathsf{bs}(f) \leq \frac{8}{9}(1 + o(1))\mathsf{s}(f)2^{\mathsf{s}(f)-1}$ due to He *et al.* [HLS17] improving a result of Ambainis *et al.* [And16].

Shi and Zhang [ZS10] studied the parity complexity variants of $\mathsf{bs}(f), \mathsf{C}(f)$ and $\mathsf{DT}(f)$ and observed that such variants have the property that they are invariant under arbitrary invertible linear transforms (over $\mathbb{F}_2^n$). They also showed existence of Boolean functions where under *all* invertible linear transforms of the function, the decision tree depth is linear while their parity variant of decision tree complexity is at most logarithmic in the input length.

Our Results : While the existing studies focus on understanding the Boolean function parameters under the effect of arbitrary invertible affine transforms, in this work, we study the relationship between the above parameters of Boolean functions $f : \mathbb{F}_2^n \to \{0,1\}$, under specific affine transformations over $\mathbb{F}_2^n$. More precisely, we explore the relationship of the above parameters for the function $g : \mathbb{F}_2^n \to \{0,1\}$ and f, where g is defined as $g(x) = f(Mx + b)$ for specific $M \in \mathbb{F}_2^{n \times n}$ and $b \in \mathbb{F}_2^n$. We show the following results, and their corresponding applications, which we explain along with the context in which they are relevant.

Alternation under shifts : We study the parameters when the transformation is very structured - namely the matrix M is the identity matrix and $b \in \mathbb{F}_2^n$ is a linear shift. More precisely, we study $f_b(x) \stackrel{\text{def}}{=} f(x+b)$ where b is the shift. Observe that all the parameters mentioned above are invariant under shifts. A Boolean function parameter which is neither shift invariant nor invariant under invertible linear transforms is the *alternation*, a measure of non-monotonicity of Boolean function (see Section 2 for a formal definition). To see this for the case of shifts, if we take f as the majority function on n bits, then there exists shifts $b \in \{0,1\}^n$ where $\mathsf{alt}(f_b) = \Omega(n)$ while $\mathsf{alt}(f) = 1$.

A recent result related to SENSITIVITY CONJECTURE by Lin and Zhang [LZ17] shows that $\mathsf{bs}(f) \leq O(\mathsf{s}(f)\mathsf{alt}(f)^2)$. This bound for $\mathsf{bs}(f)$, implies that to settle the SENSITIVITY CONJECTURE, it suffices to show that $\mathsf{alt}(f)$ is upper bounded by $\mathsf{poly}(\mathsf{s}(f))$ for all Boolean functions f. However, the authors [DS18] ruled this out, by exhibiting a family of functions where $\mathsf{alt}(f)$ is at least $2^{\Omega(\mathsf{s}(f))}$.

Observing that the parameters $\mathsf{s}(f), \mathsf{bs}(f)$ are invariant under shifts, we define a new quantity *shift-invariant alternation*, $\mathsf{salt}(f)$, which is the minimum alternation of any function g obtained from f by shifting by a vector $b \in \{0,1\}^n$ (see Definition 3.1). By the aforementioned bound on $\mathsf{bs}(f)$ of Lin and Zhang [LZ17], it is easy to observe that $\mathsf{bs}(f) \leq O(\mathsf{s}(f)\mathsf{salt}(f)^2)$. We also show that there exists a family of Boolean functions f with $\mathsf{bs}(f) = \Omega(\mathsf{s}(f)\mathsf{salt}(f))$ (See Proposition 3.5).

It is conceivable that $\mathsf{salt}(f)$ is much smaller compared to $\mathsf{alt}(f)$ for a Boolean function f and hence that $\mathsf{salt}(f)$ can potentially be upper bounded by $\mathsf{poly}(\mathsf{s}(f))$ thereby settling the SENSITIVITY CONJECTURE. However, we rule this out by showing the following stronger gap, about the same

family of functions demonstrated in [DS18] (see also [GSW16]).

Proposition 1.1. *There exists an explicit family of Boolean functions for which* $\mathsf{salt}(f)$ *is* $2^{\Omega(\mathsf{s}(f))}$

Block Sensitivity under Affine Transformations : We now generalize our theme of study to the affine transforms over $\mathbb{F}_2^n$. In particular, we explore how to design affine transformations in such a way that block sensitivity of the original function (f) is upper bounded by the sensitivity of the new function (g).

Lemma 1.2. *For any* $f : \mathbb{F}_2^n \to \{-1, 1\}$ *and* $a \in \{0, 1\}^n$, *there exists an affine transform* $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$ *such that for* $g(x) = f(A(x))$,

(a) $\mathsf{bs}(f, a) \le \mathsf{s}(g, 0^n)$, *and*

(b) $g(x) = f((x_{i_1}, x_{i_2}, \ldots, x_{i_n}) \oplus a)$ *where* $i_1, \ldots, i_n \in [n]$ *are not necessarily distinct.*

The above transformation is used in Nisan and Szegedy (see Lemma 7 of [NS92]) to show that $\mathsf{bs}(f) \le 2\mathsf{deg}(f)^2$. Here, $\mathsf{deg}(f)$ is the degree of the multilinear polynomial over reals that agrees with f on Boolean inputs. We show another application of Lemma 1.2 in the context of quantum communication complexity, a model for which was introduced by Yao [Yao93]. In this model, two parties Alice and Bob have to compute a function $F : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}$, where Alice is given an $x \in \{0, 1\}^n$ and Bob is given a $y \in \{0, 1\}^n$. Both the parties have to come up with a *quantum protocol* where they communicate qubits via a quantum channel and compute f while minimizing the number of qubits exchanged (which is the *cost* of the quantum protocol) in the process. In this model, we allow protocols to have prior entanglement. We define $Q_{1/3}^*(F)$ as the minimum cost quantum protocol computing F with prior entanglement. For more details on this model, see [Raz03]. The corresponding analog in the classical setting is the bounded error randomized communication model where the parties communicate with $0, 1$ bits and share an unbiased random source. We define $\mathsf{R}_{1/3}(F)$ as the minimum cost randomized protocol computing F with error at most $1/3$. It can be shown that $Q_{1/3}^*(F) \le \mathsf{R}_{1/3}(F) \le \mathsf{D}(F)$.

One of the fundamental goals in quantum communication complexity is to see if there are functions where their randomized communication complexity is significantly larger than their quantum communication complexity. It has been the conjectured by Shi and Zhu [SZ09] that this is not the case in general (which they called as the Log-Equivalence Conjecture). In this work, we are interested in the case when $F(x, y)$ is of the form $f(x \wedge y)$ where $f : \{0, 1\}^n \to \{0, 1\}$ and $x \wedge y$ is the string obtained by bitwise AND of x and y.

Question 1.3. *For* $f : \{0, 1\}^n \to \{0, 1\}$, *let* $F : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}$ *be defined as* $F(x, y) = f(x \wedge y)$. *Is it true that for any such* F, $\mathsf{D}(F) \le \mathsf{poly}(Q_{1/3}^*(F))$ *?*

Since $\mathsf{R}_{1/3}(F) \le \mathsf{D}(F)$, answering the above question in positive would show that the classical randomized communication model is as powerful as the quantum communication model for the class of functions $F(x, y) = f(x \wedge y)$. This question for such restricted F has also been proposed by Klauck [Kla07] as a first step towards answering the general question (see also [BdW01]). In this direction, Razborov [Raz03] showed that for the special case when f is symmetric, $F(x, y) = f(x \wedge y)$ satisfy $\mathsf{D}(F) \le O(Q_{1/3}^*(F)^2)$. In the process, Razborov developed powerful techniques to obtain lower bounds on $Q_{1/3}^*(F)$ which were subsequently generalized by Sherstov [She08], Shi and Zhu [SZ09]. Subsequently, in a slightly different direction, Sherstov [She10] showed that instead

of computing $F(x, y) = f(x \wedge y)$ alone, if we consider F to be the problem of computing both of $F_1(x, y) = f(x \wedge y)$ and $F_2(x, y) = f(x \vee y)$, then $\mathsf{D}(F) = O(Q^*_{1/3}(F)^{12})$ for all Boolean functions f where $Q^*_{1/3}(F) = \max\left\{Q^*_{1/3}(F_1), Q^*_{1/3}(F_2)\right\}$ and $\mathsf{D}(F) = \max\{\mathsf{D}(F_1), \mathsf{D}(F_2)\}$. Using Lemma 1.2, we build on the ideas of Sherstov [She10] and obtain a lower bound for $Q^*_{1/3}(F)$ where $F(x, y) = F_1(x, y) = f(x \wedge y)$.

Theorem 1.4. *Let* $f : \{0, 1\}^n \to \{-1, 1\}$ *and* $F(x, y) = f(x \wedge y)$, *then,*

$$Q^*_{1/3}(F) = \Omega\left(\sqrt{\mathsf{bs}(f, 0^n)}\right)$$

In this context, we make an important comparison with a result of Sherstov [She10]. He proved that for $F'(x, y) = f_b(x \wedge y)$, where $b \in \{0, 1\}^n$ is the input on which $\mathsf{bs}(f, x)$ is maximum, $Q^*_{1/3}(F') = \Omega(\sqrt{\mathsf{bs}(f)}) \geq \Omega(\sqrt{\mathsf{bs}(f, 0^n)})$ (Corollary 4.5 of [She10]). Notice that F and F' differ by a linear shift of f with b.[1] Moreover, $Q^*_{1/3}(F)$ can change drastically even under such (special) linear shifts of f. For example, consider $f = \wedge_n$. Since $\mathsf{bs}(f)$ is maximized at 1^n, $b = 1^n$. Hence, the function F' is the disjointness function for which $Q^*_{1/3}(F') = \Omega(\sqrt{n})$ [Raz03] whereas, $Q^*_{1/3}(F) = O(1)$. The same counterexample also shows that $Q^*_{1/3}(F) = \Omega(\sqrt{\mathsf{bs}(f)})$ cannot hold for all f (see Remark 4.2). Since the lower bounds shown on quantum communication complexity are on different functions, Theorem 1.4 is incomparable with the result of Sherstov (Corollary 4.5 of [She10]).

Using the above result, for a prime p, we show that if f has small degree when expressed as a polynomial over $\mathbb{F}_p$ (denoted by $\deg_p(f)$), the quantum communication complexity of F is large.

Theorem 1.5. *Fix a prime* p. *Let* $f : \{0, 1\}^n \to \{-1, 1\}$ *where* f *depends on all the variables. Let* $F(x, y) = f(x \wedge y)$. *For any* $0 < \epsilon < 1$ *such that* $\deg_p(f) \leq (1 - \epsilon) \log n$, *we have*

$$Q^*_{1/3}(F) = \Omega\left(\frac{n^{\epsilon/2}}{\log n}\right)$$

Observe that, though Theorem 1.4 does not answer Question 1.3 in positive for all functions, we could show a class of Boolean function for which $\mathsf{D}(F)$ and $Q^*_{1/3}(F)$ are polynomially related. More specifically, we show this for the set of all Boolean functions f such that there exists two distinct primes p, q with $\deg_p(f)$ and $\deg_q(f)$ are sufficiently far apart (Theorem 1.6).

Theorem 1.6. *Let* $f : \{0, 1\}^n \to \{-1, 1\}$ *with* $F(x, y) = f(x \wedge y)$. *Fix* $0 < \epsilon < 1$. *If there exists distinct primes* p, q *such that* $\deg_q(f) = \Omega(\deg_p(f)^{\frac{2}{1-\epsilon}})$, *then* $\mathsf{D}(F) = O(Q^*_{1/3}(F)^{2/\epsilon})$.

By the result of Gopalan *et al.* (Theorem 1.2, [GLS09]), any Boolean function f with $\deg_p(f) = o(\log n)$ must have $\deg_q(f) = \Omega(n^{1-o(1)})$ thereby satisfying the condition of Theorem 1.6. Hence for all such functions, Theorem 1.6 answers Question 1.3 in positive. Observe that the same can also be derived from Theorem 1.5.

Alternation under Linear Transforms : We now restrict our study to linear transforms. Again, the aim is to design special linear transforms which transforms the parameters of interest for us. In

---

[1] More importantly, this b in Corollary 4.5 of [She10] cannot be fixed to 0^n for all Boolean functions to conclude Theorem 1.4. See Appendix A.1 for details.

particular, in this case, we show linear transforms for which we can upper bound the alternation of the original function in terms of the sensitivity of the resulting function. More precisely, we prove the following lemma:

Lemma 1.7. *For any $f : \mathbb{F}_2^n \to \{0, 1\}$, there exists an invertible linear transform $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that for $g(x) = f(L(x))$,*

$$\mathsf{alt}(f) \leq 2\mathsf{s}(g) + 1$$

We show an application of the above result in the context of sensitivity. Nisan and Szegedy [NS92] showed that for any Boolean function f, $\mathsf{s}(f) \leq 2\mathsf{deg}(f)^2$. However, the situation is quite different for $\mathsf{deg}_2(f)$ - noticing that for f being parity on n variables, $\mathsf{deg}_2(f) = 1$ and $\mathsf{s}(f) = n$ - the gap can even be unbounded. Though parity may appear as a corner case, there are other functions like the Boolean inner product function[2] IP_n whose $\mathbb{F}_2$-degree is constant while sensitivity is $\Omega(n)$ thereby ruling out the possibility that $\mathsf{s}(f) \leq \mathsf{deg}_2(f)^2$. It is known that if f is not the parity on n variables (or its negation), $\mathsf{deg}_2(f) \leq \log \mathsf{sparsity}(f)$ [BC99, GOS+09]. Hence, as a structural question about the two parameters, we ask : for f other than the parity function, is it true that $\mathsf{s}(f) \leq \mathsf{poly}(\log \mathsf{sparsity}(f))$. Observe that IP_n has high sparsity and hence does not rule this out. We use Lemma 1.7, which is in the theme of studying alternation and sensitivity in the context of linear transformations, to improve this gap and show that there is a family of functions where this gap is exponential.

Theorem 1.8. *There exists a family of functions $\{g_k \mid k \in \mathbb{N}\}$ such that*

$$\mathsf{s}(g_k) \geq \frac{\sqrt{\mathsf{sparsity}(g_k)}}{2} - 1$$

2   Preliminaries

In this section, we define the notations used. Define $[n] = \{1, 2, \ldots, n\}$. For $S \subseteq [n]$, define $e_S \in \{0, 1\}^n$ to be the indicator vector of the set S. For $x, y \in \{0, 1\}^n$, we denote $x \wedge y$ (resp. $x \oplus y$) $\in \{0, 1\}^n$ as the string obtained by bitwise AND (resp. XOR) of x and y. We use x_i to denote the ith bit of x.

We now define the Boolean function parameters we use. Let $f : \{0, 1\}^n \to \{0, 1\}$ and $a \in \{0, 1\}^n$, we define, 1) the *sensitivity* of f on a as $\mathsf{s}(f, a) = |\{i \mid f(a \oplus e_i) \neq f(a), i \in [n]\}|$, 2) the *block sensitivity* of f on a, $\mathsf{bs}(f, a)$ to be the maximum number of disjoint blocks $\{B_i \mid B_i \subseteq [n]\}$ such that $f(a \oplus e_{B_i}) \neq f(a)$ and 3) the *certificate complexity* of f on a, $\mathsf{C}(f, a)$ to be the size of the smallest set $S \subseteq [n]$ such that fixing f according to a on the location indexed by S causes the function to become constant. For $\phi \in \{\mathsf{s}, \mathsf{bs}, \mathsf{C}\}$, we define $\phi(f) = \max_{a \in \{0,1\}^n} \phi(f, a)$ and are respectively called the sensitivity, the block sensitivity and the certificate complexity of f. By definition, the three parameters are shift invariant, by which we mean $\forall b \in \{0, 1\}^n$, $\phi(f_b) = \phi(f)$ for $\phi \in \{\mathsf{s}, \mathsf{bs}, \mathsf{C}\}$ where $f_b(x) \stackrel{\text{def}}{=} f(x \oplus b)$. Also, it can be shown that $\mathsf{s}(f) \leq \mathsf{bs}(f) \leq \mathsf{C}(f)$.

For $x, y \in \{0, 1\}^n$, define $x \prec y$ if $\forall i \in [n]$, $x_i \leq y_i$. We define a *chain* $\mathcal{C}$ on $\{0, 1\}^n$ as $(0^n = x^{(0)}, x^{(1)}, \ldots, x^{(n-1)}, x^{(n)} = 1^n)$ such that for all $i \in [n]$, $x^{(i)} \in \{0, 1\}^n$ and $x^{(i-1)} \prec x^{(i)}$. We define *alternation* of f for a chain $\mathcal{C}$, denoted $\mathsf{alt}(f, \mathcal{C})$ as the number of times the value of f changes in the chain. We define alternation of a function $\mathsf{alt}(f)$ as $\max_{\text{chain } \mathcal{C}} \mathsf{alt}(f, \mathcal{C})$.

---

[2] $\mathsf{IP}_n(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n) = \sum_i x_i y_i \mod 2$

Every Boolean function f can be expressed uniquely as a multilinear polynomial $p(x)$ in $\mathbb{F}[x_1, \ldots, x_n]$ over any field $\mathbb{F}$ such that $p(x) = f(x) \ \forall x \in \{0,1\}^n$. Fix a prime p. We denote $\mathsf{deg}(f)$ (resp. $\mathsf{deg}_p(f)$) to be the degree of the multilinear polynomial computing f over reals (resp. $\mathbb{F}_p$). We define $\mathsf{DT}(f)$ as the depth of an optimal decision tree computing f. It is known that for all Boolean functions f, $\mathsf{deg}_p(f) \leq \mathsf{deg}(f) \leq \mathsf{DT}(f) \leq \mathsf{bs}(f)^3$.

Sparsity of a Boolean function $f : \{0,1\}^n \to \{-1, 1\}$ (denoted by $\mathsf{sparsity}(f)$) is the number of non-zero Fourier coefficients in the Fourier representation of f. For more details on this parameter, see [O'D14]. For more details on $\mathsf{DT}(f)$ and other related parameters, see the survey by Buhrman, de Wolf [BdW02] and Hatami *et al.* [HKP11].

We consider the two party classical communication model. Given a function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, Alice is given an $x \in \{0,1\}^n$ and Bob is given $y \in \{0,1\}^n$. They can communicate with each other and their aim is to compute $f(x,y)$ while communicating minimum number of bits. We call the procedure employed by Alice and Bob to computing f as the *protocol*. We define $\mathsf{D}(f)$ as the minimum cost of a deterministic protocol computing f. For functions of the form $F(x,y) = f(x \wedge y)$, it is known that $\mathsf{D}(F) \leq 2\mathsf{DT}(f)$ [MO09]. For more details on communication complexity of Boolean functions, refer [KN06].

3 Warm up : Alternation under Shifts

In this section, as a warm-up, we study sensitivity and alternation under linear shifts (when the matrix M is the identity matrix). We introduce a parameter, *shift-invariant alternation* (salt). We then show the existence of Boolean functions whose shift-invariant alternation is exponential in its sensitivity (see Proposition 1.1) thereby ruling out the possibility that $\mathsf{salt}(f)$ can be upper bounded by a polynomial in $\mathsf{s}(f)$ for all Boolean functions f.

Recall from the introduction that the parameters s, bs and C are shift invariant while alt is not. We define a variant of alternation which is invariant under shifts.

Definition 3.1 (Shift-invariant Alternation). For $f : \{0,1\}^n \to \{0,1\}$, the *shift-invariant alternation* (denoted by $\mathsf{salt}(f)$) is defined as $\min_{b \in \{0,1\}^n} \mathsf{alt}(f_b)$.

A family of functions with $\mathsf{salt}(f) = \Omega(2^{\mathsf{s}(f)})$: We now exhibit a family of functions $\mathcal{F}$ where for all $f \in \mathcal{F}$, $\mathsf{salt}(f) \geq 2^{\mathsf{s}(f)}$ thereby ruling out the possibility that $\mathsf{salt}(f)$ can be upper bounded by a polynomial in $\mathsf{s}(f)$. The family $\mathcal{F}$ is the same class of Boolean functions for which alternation is at least exponential in sensitivity due to [DS18].

Definition 3.2 (Definition 1 from [DS18]. See also Proof of Lemma A.1 of [GSW16]). Consider the family defined as follows.

$$\mathcal{F} = \left\{ f_k \mid f_k : \{0,1\}^{2^k - 1} \to \{0,1\}, k \in \mathbb{N} \right\}$$

The Boolean function f_k is computed by a decision tree which is a full binary tree of depth k with 2^k leaves. A leaf node is labeled as 0 (resp. 1) if it is the left (resp. right) child of its parent. All of the nodes (except the leaves) are labeled by a distinct variable.

We remark that Gopalan *et al.* [GSW16] demonstrates an exponential lower bound on tree sensitivity (introduced by them as a generalization of the parameter sensitivity) in terms of decision tree depth for the same family of functions in Definition 3.2. We remark that, in general, lower

bound on tree sensitivity need not implies a lower bound on alternation. For instance, if we consider the Majority function[3] Maj_n, the tree sensitivity can be shown to be $\Omega(n)$ while alternation is 1.

The authors [DS18] have shown that for any $f \in \mathcal{F}$, there exists of a chain of large alternation in f. However, this is not sufficient to argue existence of a chain of large alternation under *every* linear shift. We now proceed to prove an exponential lower bound on $\mathsf{salt}(f)$ in terms of $s(f)$ for all $f \in \mathcal{F}$.

Proposition 1.1. For $f_k \in \mathcal{F}$, $\mathsf{salt}(f_k) \geq 2^{\Omega(\mathsf{s}(f_k))}$.

Proof. We show[4] that for $f_k \in \mathcal{F}$ and $n = 2^k - 1$, for all $c \in \{0,1\}^n$, $\mathsf{alt}(f_k(x \oplus c)) \geq 2^{k-2}$. Since $\mathsf{s}(f_k) \leq k$ by construction of f_k, the result follows.

Proof is by induction on k. For $k = 2$, f is a function on 3 variables and it can be verified that for all $c \in \{0,1\}^3$, $\mathsf{alt}(f(x \oplus c)) \geq 1$. Now consider an $f_{k+1} \in \mathcal{F}$ computed by a decision tree T with the variable x_t as its root. Let h_1 and h_2 be the left and right subtrees of x_t in T. Note that $h_1(z')$ and $h_2(z'')$ depends on $n = 2^k - 1$ variables and belongs to $\mathcal{F}$ by construction. Hence by induction, for all $c \in \{0,1\}^n$, $\mathsf{alt}(h_1(z' \oplus c))$ and $\mathsf{alt}(h_2(z'' \oplus c))$ is at least 2^{k-2}. For $m = 2^{k+1} - 1$, consider any $c = (c', b, c'') \in \{0,1\}^m$ where $c', c'' \in \{0,1\}^n$ and $b \in \{0,1\}$. Since h_1 and h_2 are variable disjoint, $\mathsf{alt}(f(x \oplus c)) \geq \mathsf{alt}(h_1(z' \oplus c')) + \mathsf{alt}(h_2(z'' \oplus c'')) \geq 2^{k-2} + 2^{k-2} = 2^{k-1}$ completing the induction. $\square$

A family of functions with $\mathsf{bs}(f) = \Omega(\mathsf{s}(f)\mathsf{salt}(f))$: Lin and Zhang [LZ17] showed that for any Boolean function $f : \{0,1\}^n \to \{0,1\}$,

$$\mathsf{bs}(f) = O(\mathsf{alt}(f)^2 \mathsf{s}(f)) \tag{1}$$

This immediately gives the following proposition.

Proposition 3.3. *For any $f : \{0,1\}^n \to \{0,1\}$, $\mathsf{bs}(f) \leq O(\mathsf{salt}(f)^2 \mathsf{s}(f))$*

We now exhibit a family of functions for which $\mathsf{bs}(f)$ is at least $\frac{\mathsf{s}(f) \cdot \mathsf{salt}(f)}{4}$.

Before proceeding, we show a tight composition result for alternation of Boolean functions when composed with OR_k (which is the k bit Boolean OR function).

For functions $f_1, \ldots, f_k$ where each $f_i : \{0,1\}^n \to \{0,1\}$, define the function $OR_k \circ \overline{f} : \{0,1\}^{nk} \to \{0,1\}$ as $\vee_{i=1}^k f_i(x^{(i)})$ where for each $i \in [k]$, $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)}) \in \{0,1\}^n$ is input to the function f_i.

Lemma 3.4. *Consider k Boolean functions $f_1, \ldots, f_k$ where each $f_i : \{0,1\}^n \to \{0,1\}$ satisfy, $f_i(0^n) = f_i(1^n) = 0$. Then,*

$$\mathsf{alt}(OR_k \circ \overline{f}) = \sum_{i=1}^k \mathsf{alt}(f_i)$$

Proof. Let $f = OR_k \circ \overline{f}$ and $\mathcal{C}$ be a chain in $\{0,1\}^{nk}$ for which $\mathsf{alt}(f, \mathcal{C})$ is maximized. Without loss of generality, let all the functions be non-constant. Let $\mathcal{C}_i$ be the chain in $\{0,1\}^n$ obtained by restricting $\mathcal{C}$ to variables $x_1^{(i)}, \ldots, x_n^{(i)}$ of f_i. Observe that if f changes it value, it must be that

---

[3] $\mathsf{Maj}_n(x) = 1 \iff \sum_i x_i \geq \lceil n/2 \rceil$

[4] In this proof, for simplicity, we abuse the notation $f_k(x \oplus c)$ to denote the function obtained by shifting f_k by c.

at least one of the f_i's have changed their evaluation along the chain $\mathcal{C}$. Since the functions are variable disjoint, such a change must be witnessed in the chain $\mathcal{C}_i$ for some i. Hence

$$\mathsf{alt}(f) = \mathsf{alt}(f,\mathcal{C}) \leq \sum_{i=1}^{k} \mathsf{alt}(f_i,\mathcal{C}_i) \leq \sum_{i=1}^{k} \mathsf{alt}(f_i)$$

To show that $\mathsf{alt}(f) \geq \sum_{i=1}^{k} \mathsf{alt}(f_i)$, we exhibit a chain $\mathcal{C}$ in $\{0,1\}^{nk}$ of alternation $\sum_{i=1}^{k} \mathsf{alt}(f_i)$. Let $\mathcal{C}_i = (0^n = z^{(i0)} \prec z^{(i1)} \prec z^{(i2)} \prec \ldots \prec z^{(in)} = 1^n)$ be a chain in $\{0,1\}^n$ for which f_i achieves maximum alternation. We construct a chain $\mathcal{C}$ by "gluing" together these k chains. More precisely, let $\mathcal{C}$ by the chain such that for all $i \in [k]$, when restricted to the variables $x_1^{(i)}, \ldots, x_n^{(i)}$, we get a chain given by,

$$\overbrace{0^n \prec \ldots \prec 0^n}^{n(i-1) \text{ times}} \prec z^{(i0)} \prec z^{(i1)} \prec z^{(i2)} \prec \ldots \prec z^{(in)} \prec \overbrace{1^n \prec \ldots \prec 1^n}^{n(k-i) \text{ times}}$$

By construction of $\mathcal{C}$, since $f_j(0^n) = f_j(1^n) = 0$ for all $j \in [k]$, at any input of the chain $\mathcal{C}$, there is exactly one f_i that causes f to alternate. Hence,

$$\mathsf{alt}(f,\mathcal{C}) \geq \sum_{i=1}^{k} \mathsf{alt}(f_i,\mathcal{C}_i) = \sum_{i=1}^{k} \mathsf{alt}(f_i)$$

$\square$

Proposition 3.5. *There exists a family of Boolean functions for which* $\mathsf{bs}(f) \geq \frac{\mathsf{s}(f) \cdot \mathsf{salt}(f)}{4}$

Proof. We consider the Rubinstein's function $f_R : \{0,1\}^{n^2} \to \{0,1\}$ [Rub95] where the input is treated as $n \times n$ matrix which evaluates to 1 iff there is a row with two consecutive ones starting at the odd position and rest of the entries being zero. Alternatively, we can view f_R as $OR_n \circ \overline{h}$ with $h : \{0,1\}^n \to \{0,1\}$ where $h(a) = 1$ iff there are two consecutive ones starting at the odd position with rest of the entries as zero in $a \in \{0,1\}^n$. It can be verified that $\mathsf{alt}(h) = 2$. Since $h(0^n) = h(1^n) = 0$, applying Lemma 3.4 with $f_i = h$ for all $i \in [n]$, we get that $\mathsf{alt}(f_R) = \mathsf{alt}(h) \cdot n = 2n$. It is known that $\mathsf{bs}(f_R) \geq \frac{n^2}{2}$ while $\mathsf{s}(f_R) \leq n$ [Rub95], thereby showing that $\mathsf{bs}(f_R) \geq \frac{\mathsf{s}(f_R) \cdot \mathsf{alt}(f_R)}{4} \geq \frac{\mathsf{s}(f_R) \cdot \mathsf{salt}(f_R)}{4}$. $\square$

We remark that the above bound is stronger than what is needed in the context because, $\mathsf{bs}(f_R) \geq \frac{\mathsf{s}(f_R) \cdot \mathsf{alt}(f_R)}{4}$.

4 Affine Transforms : Lower Bounds on Quantum Communication Complexity

In this section, we study the affine transformation in its full generality applied to block sensitivity and sensitivity, and use it to prove Theorem 1.5 and Theorem 1.6 from the introduction. We achieve this using affine transforms as our tool (Section 4.1), by which we derive a new lower bound for $Q_{1/3}^{*}(F)$ in terms of $\mathsf{bs}(f,0^n)$ (Section 4.2). Using this and a lower bound on $\mathsf{bs}(f,0^n)$ (Proposition 4.3), we show that for any Boolean function f, and any prime

p, $Q_{1/3}^*(F) \geq \Omega\left(\frac{\sqrt{\mathsf{DT}(f)}}{\deg_p(f)}\right)$. This immediately implies that if there is a p such that $\deg_p(f)$ is constant, then $\mathsf{D}(F) \leq 2\mathsf{DT}(f) \leq O(Q_{1/3}^*(F)^2)$ thereby answering Question 1.3 in positive for such functions. We relax this requirement and show that if there exists distinct primes p and q for which $\deg_p(f)$ and $\deg_q(f)$ are not very close, then $\mathsf{D}(F) \leq \mathsf{poly}(Q_{1/3}^*(F))$ (Theorem 1.6).

4.1 Upper Bound for Block Sensitivity via Affine Transforms

In this section, we describe our main tool. Given an $f : \{0,1\}^n \to \{0,1\}$ and any $a \in \{0,1\}^n$, we exhibit an affine transform $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that for $g(x) = f(Ax)$, $\mathsf{bs}(f,a) \leq \mathsf{s}(g,0^n)$.

Before describing the affine transform, we note that a linear transform is already known to achieve a weaker bound of $\mathsf{bs}(f) \leq O(\mathsf{s}(g)^2)$ due to Sherstov [She10].

Proposition 4.1 (Lemma 3.3 of [She10]). *For any $f : \mathbb{F}_2^n \to \{0,1\}$, there exists a linear transform $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that for $g(x) = f(Lx)$, $\mathsf{bs}(f) = O(\mathsf{s}(g)^2)$.*

See Observation A.3 in Appendix A.1 for an explicit description of the linear transform achieving the bounds in the above proposition.

Now we describe an affine transform which improves the bound on $\mathsf{bs}(f)$ in the above proposition to linear in $\mathsf{s}(g)$. This affine transform has already been used in Nisan and Szegedy (see Lemma 7 of [NS92]) to show that $\mathsf{bs}(f) \leq 2\deg(f)^2$. Since the exact form of g is relevant in the subsequent arguments, we explicitly prove it here bringing out the structure of the affine transform that we require.

Lemma 1.2. For any $f : \mathbb{F}_2^n \to \{-1,1\}$ and $a \in \{0,1\}^n$, there exists an affine transform $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that for $g(x) = f(A(x))$,

(a) $\mathsf{bs}(f,a) \leq \mathsf{s}(g,0^n)$, and

(b) $g(x) = f((x_{i_1}, x_{i_2}, \ldots, x_{i_n}) \oplus a)$ where $i_1, \ldots, i_n \in [n]$ are not necessarily distinct.

Proof. Let $\mathsf{bs}(f,a) = k$ and $\{B_1, \ldots, B_k\}$ be the sensitive blocks on a. Since the blocks are disjoint, $\{B_i \mid i \in [k]\}$ viewed as vectors over $\mathbb{F}_2^n$ are linearly independent. Hence, there is a linear transform $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that $L(e_i) = B_i$ for $i \in [k]$.[5] Define $A(x) = L(x) \oplus a$. For $g(x) = f(A(x))$,

$$
\begin{aligned}
\mathsf{s}(g,0^n) &= |\{i \mid g(0^n) \neq g(0^n \oplus e_i), i \in [n]\}| \\
&= |\{i \mid f(a) \neq f(a \oplus L(e_i)), i \in [n]\}| = \mathsf{bs}(f,a)
\end{aligned}
$$

which completes the proof of main statement and Item a. Item b holds as the sensitive blocks are disjoint. $\square$

4.2 From Block Sensitivity Lower Bound at 0^n to Quantum Communication Lower Bounds

We now prove a lower bound for $Q_{1/3}^*(F)$ in terms of $\mathsf{bs}(f,0^n)$.

Theorem 1.4. Let $f : \{0,1\}^n \to \{-1,1\}$ and $F(x,y) = f(x \wedge y)$, then,

$$Q_{1/3}^*(F) = \Omega\left(\sqrt{\mathsf{bs}(f,0^n)}\right)$$

---

[5] For completeness of definition of L, for $i \notin [k]$, we define $L(e_i) = 0^n$.

Proof. We first state a weaker version of this result which follows from Theorem 4.2 of Sherstov [She10]. The result, which is based on a powerful method of proving quantum communication lower bounds due to Razborov [Raz03] and Klauck [Kla07], says that for a Boolean function $g : \{0,1\}^n \to \{-1,1\}$ with $G(x,y) = g(x \wedge y)$, if there exists an $z \in \{0,1\}^n$ such that $z_i = 0$ for $i \in [k]$ and $g(z \oplus e_1) = g(z \oplus e_2) = \ldots = g(z \oplus e_k) \neq g(z)$, then $Q^*_{1/3}(G) = \Omega(\sqrt{k})$. This immediately implies that for any $g : \{0,1\}^n \to \{-1,1\}$,

$$Q^*_{1/3}(G) = \Omega\left(\sqrt{\mathsf{s}(g,0^n)}\right) \tag{2}$$

Given an f, we now describe a $g : \{0,1\}^n \to \{-1,1\}$ such that $Q^*_{1/3}(F) \geq Q^*_{1/3}(G)$ and $Q^*_{1/3}(G) = \Omega(\sqrt{\mathsf{bs}(f,0^n)})$ as follows thereby completing the proof.

Applying Lemma 1.2 with $a = 0^n$ to f, we obtain $g(x) = f(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$. We note that F and G can be viewed as a $2^n \times 2^n$ matrix with (x,y)th entry being $f(x \wedge y)$ and $g(x \wedge y)$ respectively. By construction of g, using the observation that the matrix G appears as a submatrix of F, $Q^*_{1/3}(F) \geq Q^*_{1/3}(G)$. This observation is used in Sherstov (for instance, see proof of Theorem 5.1 of [She10]) without giving details. For completeness, we give the details here. Let $S = \{i_1, \ldots i_n\} \subseteq [n]$ of size k. For $j \in S$, let $B_j = \{t \mid i_t = j\}$. Hence g depends only on these k input variables of S and all the variables with indices in B_j are assigned the variable x_j. This implies that

$$g(x) = f(\oplus_{j \in S} x_j e_{B_j}) \tag{3}$$

We now exhibit a submatrix of F containing G. Consider the submatrix of F with rows and columns restricted to $W = \left\{ a_1 e_{B_1} \oplus a_2 e_{B_2} \oplus \ldots a_k e_{B_k} \mid (a_1, a_2 \ldots, a_k) \in \{0,1\}^k \right\}$. For $u, y \in W$,

$$\begin{aligned} F(u,y) = f(u \wedge y) &= f((u_1 e_{B_1} \oplus \ldots \oplus u_k e_{B_k}) \wedge (y_1 e_{B_1} \oplus \ldots \oplus y_k e_{B_k})) \\ &= f(u_1 \wedge y_1 e_{B_1} \oplus \ldots \oplus u_k \wedge y_k e_{B_k}) && [B_j\text{s are disjoint}] \\ &= g(u \wedge y) && [\text{By Eq. } (3)] \end{aligned}$$

Applying Eq. (2) to the g obtained, we have $Q^*_{1/3}(G) \geq \Omega(\sqrt{\mathsf{s}(g,0^n)})$. Hence, by Item a of Lemma 1.2, as $a = 0^n$, we have $Q^*_{1/3}(G) \geq \Omega(\sqrt{\mathsf{bs}(f,0^n)})$. $\qquad\square$

Remark 4.2. *Observe that for an arbitrary $a \in \{0,1\}^n$ for $g(x) = f(x \oplus a)$, the statement $Q^*_{1/3}(G) \leq Q^*_{1/3}(F)$ does not hold. Otherwise, we would have $Q^*_{1/3}(F) = \Omega(\sqrt{\mathsf{bs}(f)})$ for all f which is not true (see the discussion after Theorem 1.4 in the Introduction).*

4.3 Putting Them Together

We are now ready to prove Theorem 1.5 and Theorem 1.6. A critical component of our proof is the following stronger connection between $\mathsf{DT}(f)$ and $\mathsf{bs}(f,0^n)$. Buhrman and de Wolf, in their survey [BdW02], showed that $\mathsf{DT}(f) \leq \mathsf{bs}(f) \cdot \mathsf{deg}(f)^2$ where the proof is attributed to Noam Nisan and Roman Smolensky. The same proof can be adapted (see Appendix A.2 for details) to show the following strengthening of their result.

Proposition 4.3. *For any $f : \{0,1\}^n \to \{0,1\}$, and any prime p,*

$$\mathsf{DT}(f) \leq \mathsf{bs}(f,0^n) \cdot \mathsf{deg}_p(f)^2$$

We now give a proof of Theorem 1.5 and Theorem 1.6.

Theorem 1.5. Fix a prime p. Let $f : \{0,1\}^n \to \{-1,1\}$ where f depends on all the inputs. Let $F(x,y) = f(x \wedge y)$. For any $0 < \epsilon < 1$ such that $\deg_p(f) \leq (1 - \epsilon) \log n$, we have

$$Q_{1/3}^*(F) = \Omega\left(\frac{n^{\epsilon/2}}{\log n}\right)$$

Proof. Applying Theorem 1.4 and Proposition 4.3, we have $Q_{1/3}^*(F) \geq \Omega\left(\frac{\sqrt{\mathsf{DT}(f)}}{\deg_p(f)}\right)$. As observed in Gopalan *et al.* [GLS09], by a modification to an argument in the proof of Nisan and Szegedy (Theorem 1 of [NS92]), it can be shown that $\deg(f) \geq \frac{n}{2^{\deg_p(f)}}$. Since, $\mathsf{DT}(f) \geq \deg(f)$, we have $\mathsf{DT}(f) \geq \frac{n}{2^{\deg_p(f)}}$. Hence,

$$Q_{1/3}^*(F) = \Omega\left(\frac{\sqrt{n}}{\deg_p(f)2^{\deg_p(f)/2}}\right) = \Omega\left(\frac{n^{\epsilon/2}}{(1-\epsilon)\log n}\right)$$

where the last lower bound follows upon applying the bound on $\deg_p(f)$.   □

As a demonstrative example, we show a weaker lower bound on quantum communication complexity with prior entanglement for the generalized inner product function $\mathsf{GIP}_{n,k}(x,y) \stackrel{\text{def}}{=} \oplus_{i=1}^n \bigwedge_{j=1}^k (x_{ij} \wedge y_{ij})$ when $k = \frac{1}{2} \log n$. We remark that a lower bound of $\Omega(n)$ is known for the inner product function [CvDNT99].

Note that $\mathsf{GIP}_{n,k}$ can be expressed as $f \circ \wedge$, where $f(z) \stackrel{\text{def}}{=} \oplus_{i=1}^n \bigwedge_{j=1}^k z_{ij}$, with $\deg_2(f) = k$. Applying Theorem 1.5 with $\epsilon = 1/2$ and $p = 2$, we have $Q_{1/3}^*(\mathsf{GIP}_{n,\frac{1}{2}\log n}) = \Omega\left(\frac{n^{1/4}}{\log n}\right)$. Though this bound is arguably weak, Theorem 1.5 gives a non-trivial lower bound for a all those Boolean functions f with small $\deg_p(f)$ for some prime p.

Theorem 1.6. Let $f : \{0,1\}^n \to \{-1,1\}$ with $F(x,y) = f(x \wedge y)$. Fix $0 < \epsilon < 1$. If there exists distinct primes p, q such that $\deg_q(f) = \Omega(\deg_p(f)^{\frac{2}{1-\epsilon}})$, then $\mathsf{D}(F) = O(Q_{1/3}^*(F)^{2/\epsilon})$.

Proof. Applying, Theorem 1.4 and Proposition 4.3, we get that for any prime t, $Q_{1/3}^*(F) \geq \Omega\left(\frac{\sqrt{\mathsf{DT}(f)}}{\deg_t(f)}\right)$. By hypothesis, $\deg_p(f) \leq O(\deg_q(f)^{\frac{1-\epsilon}{2}}) \leq O(\mathsf{DT}(f)^{\frac{1-\epsilon}{2}})$ implying that for $t = p$, $\mathsf{D}(F) \leq 2\mathsf{DT}(f) \leq O(Q_{1/3}^*(F)^{2/\epsilon})$ □

5 Linear Transforms : Sensitivity versus Sparsity

Continuing in the theme of affine transforms, in this section, we first establish an upper bound on alternation of a function in terms of sensitivity of function after application of a suitable linear transform. Using this, we show the existence of a function whose sensitivity is asymptotically as large as square root of sparsity (see introduction for a motivation and discussion).

Lemma 1.7. For any $f : \{0,1\}^n \to \{0,1\}$, there exists an invertible linear transform $L : F_2^n \to \mathbb{F}_2^n$ such that for $g(x) = f(L(x))$, $\mathsf{alt}(f) \leq 2\mathsf{s}(g) + 1$

Proof. Let $0^n \prec x_1 \prec x_2 \ldots \prec x_n = 1^n$ be a chain $\mathcal{C}$ of maximum alternation in the Boolean hypercube of f. Since chain $\mathcal{C}$ has maximum alternation, there must be at least $(\mathsf{alt}(f) - 1)/2$ many zeros and $(\mathsf{alt}(f) - 1)/2$ many ones when the x_is are evaluated on f. Note that the set of n distinct inputs $x_1, x_2, \ldots, x_n$ seen as vectors in $\mathbb{F}_2^n$ are linearly independent and hence form a basis of $\mathbb{F}_2^n$. Hence there exists a invertible[6] linear transform $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ taking standard basis vectors to the these vectors, i,e. $L(e_i) = x_i$ for $i \in [n]$.

To prove the result, we now show that $\mathsf{s}(g, 0^n) \geq \frac{\mathsf{alt}(f) - 1}{2}$. The neighbors of 0^n in the hypercube of g are $\{e_i \mid i \in [n]\}$ and each of them evaluates to $g(e_i) = f(L(e_i)) = f(x_i)$ for $i \in [n]$. Since there are at least $(\mathsf{alt}(f) - 1)/2$ many zero and at least these ones among x_is when evaluated by f, there must be at least $(\mathsf{alt}(f) - 1)/2$ many neighbors of 0^n which differ in evaluation with $g(0^n)$ (independent of the value of $g(0^n)$). Hence $\mathsf{s}(g) \geq s(g, 0^n) \geq \frac{\mathsf{alt}(f) - 1}{2}$ which completes the proof. $\square$

We now describe the family of functions and argue an exponential gap between sensitivity and logarithm of sparsity, as stated in the following Theorem.

Theorem 1.8. There exists a family of functions $\{g_k \mid k \in \mathbb{N}\}$ such that

$$\mathsf{s}(g_k) \geq \frac{\sqrt{\mathsf{sparsity}(g_k)}}{2} - 1$$

Proof. We remark that for the family of functions $f_k \in \mathcal{F}$ (Definition 3.2), $\mathsf{alt}(f_k) \geq 2^{(\log \mathsf{sparsity}(f_k))/2} - 1$ [DS18].

We now use this family $\mathcal{F}$ to describe the family of functions g_k. For every $f_k \in \mathcal{F}$, let $g_k(x) = f_k(L(x))$ such that $\mathsf{alt}(f_k) \leq 2\mathsf{s}(g_k) + 1$ as guaranteed by Lemma 1.7. Since, we have $\mathsf{alt}(f_k) \geq 2^{(\log \mathsf{sparsity}(f_k))/2} - 1$, it must be that

$$\mathsf{s}(g_k) \geq \frac{1}{2}(\mathsf{alt}(f_k) - 1) \geq \frac{1}{2}(2^{(\log \mathsf{sparsity}(f_k))/2} - 2) \geq \frac{\sqrt{\mathsf{sparsity}(f_k)}}{2} - 1$$

As the parameter $\mathsf{sparsity}$ does not change under invertible linear transforms [O'D14], $\mathsf{s}(g_k) \geq 0.5\sqrt{\mathsf{sparsity}(f_k)} - 1 = 0.5\sqrt{\mathsf{sparsity}(g_k)} - 1$. $\square$

6 Discussion and Open Problems

In this paper, we study the Boolean function parameters, namely sensitivity, block sensitivity, and alternation under affine transforms. We showed design of special transforms which achieves structurally revealing statements about the resulting function. We used their properties to show lower bounds on the bounded error quantum communication complexity of Boolean function whose $\mathbb{F}_p$-degree is small. We showed that classical and quantum communication complexity are polynomially related for certain special class of functions. We also demonstrated Boolean functions where sensitivity of the function is as large as the square root of its sparsity.

There are several questions that are opened up in this line of work. Given an f, Observation A.3 says that Sherstov's result exhibits a linear transform L such that for $g(x) = f(Lx)$, $\mathsf{bs}(f) = O(\mathsf{s}(g)^2)$. We obtain an affine transform (in Lemma 1.2) where the corresponding g satisfy, $\mathsf{bs}(f) = O(\mathsf{s}(g))$. In this context, a prominent direction is to use the structure of our linear transformation

---

[6]L is actually the change of basis transform from standard basis vectors to x_is and hence is bijective.

in Lemma 1.2 to establish better upper bounds for $\mathsf{s}(g)$ in terms of $\mathsf{s}(f)$. Upper bounds better than $2^{s(f)}$ would improve the best known upper bound of $\mathsf{bs}(f)$ in terms of $\mathsf{s}(f)$. It is even more interesting to restrict to the case of $\mathsf{bs}_3(f)$ and upper bound the corresponding $\mathsf{s}(g)$ by $\mathsf{s}(f)^{3-\epsilon}$ for some $\epsilon > 0$. By a result of Tal [Tal16], this suffices to improve the best known upper bound for block sensitivity in terms of sensitivity.

7 Acknowledgment

References

[And16]　Andris Ambainis, Krišjānis Prūsis and Jevgēnijs Vihrovs. Sensitivity versus certificate complexity of boolean functions. In *Computer Science - Theory and Applications - 11th International Computer Science Symposium in Russia, CSR 2016, St. Petersburg, Russia, June 9-13, 2016, Proceedings*, pages 16–28, 2016.

[BC99]　Anna Bernasconi and Bruno Codenotti. Spectral analysis of boolean functions as a graph eigenvalue problem. *IEEE Trans. Computers*, 48(3):345–351, 1999.

[BdW01]　Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, Illinois, USA, June 18-21, 2001*, pages 120–130, 2001.

[BdW02]　Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: a survey. *Theor. Comput. Sci.*, 288(1):21–43, 2002.

[CDR86]　Stephen A. Cook, Cynthia Dwork, and Rüdiger Reischuk. Upper and lower time bounds for parallel random access machines without simultaneous writes. *SIAM J. Comput.*, 15(1):87–97, 1986.

[CvDNT99]　Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. Quantum entanglement and the communication complexity of the inner product function. In Colin P. Williams, editor, *Quantum Computing and Quantum Communications*, pages 61–74, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[DS18]　Krishnamoorthy Dinesh and Jayalal Sarma. Alternation, sparsity and sensitivity: Combinatorial bounds and exponential gaps. In *Algorithms and Discrete Applied Mathematics - 4th International Conference, CALDAM 2018, Guwahati, India, February 15-17, 2018, Proceedings*, pages 260–273, 2018.

[GLS09]　Parikshit Gopalan, Shachar Lovett, and Amir Shpilka. On the complexity of boolean functions in different characteristics. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity, CCC 2009, Paris, France, 15-18 July 2009*, pages 173–183, 2009.

[GOS⁺09] Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. In *Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5-12, 2009, Proceedings, Part I*, pages 500–512, 2009.

[GSW16] Parikshit Gopalan, Rocco A. Servedio, and Avi Wigderson. Degree and sensitivity: Tails of two distributions. In *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, pages 13:1–13:23, 2016.

[HKP11] Pooya Hatami, Raghav Kulkarni, and Denis Pankratov. *Variations on the Sensitivity Conjecture*. Number 4 in Graduate Surveys. Theory of Computing Library, 2011.

[HLS17] Kun He, Qian Li, and Xiaoming Sun. A tighter relation between sensitivity complexity and certificate complexity. In *Computing and Combinatorics - 23rd International Conference, COCOON 2017, Hong Kong, China, August 3-5, 2017, Proceedings*, pages 262–274, 2017.

[Kla07] Hartmut Klauck. Lower bounds for quantum communication complexity. *SIAM J. Comput.*, 37(1):20–46, 2007.

[KN06] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, 2nd edition, 2006.

[LZ17] Chengyu Lin and Shengyu Zhang. Sensitivity conjecture and log-rank conjecture for functions with small alternating numbers. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, pages 51:1–51:13, 2017.

[MO09] Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *CoRR*, abs/0909.3392, 2009.

[Nis91] Noam Nisan. CREW PRAMs and decision trees. *SIAM J. Comput.*, 20(6):999–1007, 1991.

[NS92] Noam Nisan and Mario Szegedy. On the degree of Boolean functions as real polynomials. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, STOC '92, pages 462–467, New York, NY, USA, 1992. ACM.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[Raz03] A A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

[Rub95] David Rubinstein. Sensitivity vs. block sensitivity of Boolean functions. *Combinatorica*, 15(2):297–299, 1995.

[She08] Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 85–94, 2008.

[She10]   Alexander A. Sherstov. On quantum-classical equivalence for composed communication problems. *Quantum Information & Computation*, 10(5&6):435–455, 2010.

[SZ09]   Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information & Computation*, 9(5):444–460, 2009.

[Tal16]   Avishay Tal. On the sensitivity conjecture. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016, July 11-15, 2016, Rome, Italy*, pages 38:1–38:13, 2016.

[Yao93]   Andrew Chi-Chih Yao. Quantum circuit complexity. In *34th Annual Symposium on Foundations of Computer Science, Palo Alto, California, USA, 3-5 November 1993*, pages 352–361, 1993.

[ZS10]   Zhiqiang Zhang and Yaoyun Shi. On the parity complexity measures of Boolean functions. *Theor. Comput. Sci.*, 411(26-28):2612–2618, 2010.

A   Appendix

A.1   Quantum communication lower bound from block sensitivity

Sherstov in [She10] showed the following lower bound on quantum communication cost of an affine shift of a Boolean function in terms of its block sensitivity.

Corollary A.1 (Corollary 4.5 of [She10]). *Let $f : \{0,1\}^n \to \{-1,1\}$ be given. Then for some $z \in \{0,1\}^n$, the matrix $F' = [f_z(x \wedge y)]_{x,y} = [f(\ldots,(x_i \wedge y_i) \oplus z_i, \ldots)]_{x,y}$ obeys*

$$Q^*_{1/3}(F') = \Omega(\sqrt{\mathsf{bs}(f)})$$

In this section, we elaborate on why one cannot set $z = 0^n$ for all Boolean functions and obtain Theorem 1.4. The above corollary crucially uses two results. The first one is Lemma 3.3 of [She10] which shows that there exists a Boolean function $g : \{0,1\}^n \to \{0,1\}$ such that $\mathsf{bs}(f) \le O(\mathsf{s}(g)^2)$ which is similar in spirit to Lemma 1.2. The second one is Theorem 4.2 of [She10] which shows a lower bound for $Q^*_{1/3}(G)$ in terms of sensitivity of g (where $G(x,y) = g(x \wedge y)$). We reproduce the respective statements of both below.

Lemma A.2 (Lemma 3.3 of [She10]). *Let $f : \{0,1\}^n \to \{-1,1\}$. Then there exists a $g : \{0,1\}^n \to \{-1,1\}$ such that $s(g) = \Omega(\sqrt{\mathsf{bs}(f)})$ and $g(x) = f(x_{i_1}, \ldots, x_{i_n})$ for some $i_1, \ldots, i_n \in [n]$*

The function g is defined as follows.

Let z be the input on which $\mathsf{bs}(f, z)$ is maximum and $f(z) = 0$. Let $S_1, \ldots, S_k \subseteq [n]$ be the sensitive blocks on z. Define $A_i = \{j \in S_i \mid z_j = 0\}$ and $B_i = \{j \in S_i \mid z_j = 1\}$. Let I be the indices $i \in [k]$ such that both A_i and B_i are both non-empty.

Then

$$g(x) = f\left(\bigoplus_{i \in I} x_{minA_i} e_{A_i} \oplus \bigoplus_{i \in I} x_{minB_i} e_{B_i} \oplus \bigoplus_{i \in [k] \setminus I} x_{minS_i} e_{S_i} \oplus \bigoplus_{i \notin S_1 \cup \ldots \cup S_k} x_i e_i \right)$$

Observation A.3. *We observe that the above result of Sherstov (Lemma 3.3 of [She10]) can be seen as applying a suitable linear transform to the Boolean function f to bound the block sensitivity of f which is similar in spirit to Lemma 1.2.*

More precisely, the g obtained in Lemma 3.3 of [She10] can be described as $f(L(x))$ where L is defined as, for $j \in [n]$,

$$L(e_j) = \begin{cases} e_j & \text{if } j \notin S_1 \cup \ldots \cup S_k \\ e_{A_i} & \text{if } \exists i \in [k], \text{ such that } j = \min\{A_i\} \\ e_{B_i} & \text{if } \exists i \in [k], \text{ such that } j = \min\{B_i\} \\ 0^n & \text{otherwise} \end{cases}$$

By definition g as above, Sherstov showed that $\mathsf{s}(g, z) = \Omega(\sqrt{\mathsf{bs}(f)})$.

Theorem A.4 (Theorem 4.2 of [She10]). *For a Boolean function $g : \{0,1\}^n \to \{-1,1\}$ with $G(x,y) = g(x \wedge y)$, if there exists an $w \in \{0,1\}^n$ such that $w_i = 0$ for $i \in [k]$ and $g(w \oplus e_1) = g(w \oplus e_2) = \ldots = g(w \oplus e_k) \neq g(w)$, then $Q^*_{1/3}(G) = \Omega(\sqrt{k})$.*

To use the above result, one way is to start with a function g for which sensitivity is large at 0^n. To achieve, consider the shifted function f_z where z is the same input on which block sensitivity is maximized as before. This is because, by the choice of z, f_z will have maximum block sensitivity at 0^n which upon applying Lemma 3.3 of [She10] ensures that the function g obtained has a large k (i.e. sensitivity) at 0^n. This is exactly what is achieved in the proof of Corollary 4.5 of [She10].

Hence the choice is z is tied up with the block sensitivity of function f.

A.2 Proof of Proposition 4.3

Proof. We observe that the arguments of Buhrman and deWolf (more specifically, Lemma 5, Lemma 6 and Theorem 12 of [BdW02]), can give a stronger upper bound than $\mathsf{bs}(f) \cdot \mathsf{deg}(f)^2$, namely $\mathsf{bs}(f, 0^n) \cdot \mathsf{deg}_p(f)^2$. This is important in our context since we are able to bound $Q^*_{1/3}(F)$ only by $\mathsf{bs}(f, 0^n)$.

Let $p_f(x) \in \mathbb{F}_p[x_1, \ldots, x_n]$ be an $\mathbb{F}_p$ polynomial representation of f. Let S_f be the collection of all monomials of maximal size in p_f. We show that for any Boolean function f, there are most $\mathsf{bs}(f, 0^n) \cdot \mathsf{deg}_p(f)$ many variables which has a non-empty intersection with all the monomials in S_f. Hence querying these variables results in a function whose $\mathbb{F}_p$-degree is at most $\mathsf{deg}_p(f) - 1$. We repeat this on the resulting function to obtain the desired decision tree where at most $\mathsf{bs}(f, 0^n) \cdot \mathsf{deg}_p(f)^2$ variables gets queried. We now argue the existence of a "hitting set", which has a non-empty intersection with all the monomials in S_f, of size at most $\mathsf{bs}(f, 0^n) \cdot \mathsf{deg}_p(f)$.

Firstly, observe that every monomial m in S_f must have a non-empty set of indices B such that $f(0^n) \neq f(0^n \oplus e_B)$. To see this, restrict f to indices in the monomial m by setting all variables not in the monomial to 0. Let g be the resulting function. By construction, g is non-constant as the monomial m appears in the $\mathbb{F}_p$ representation of g. Hence there must be some setting of the input to g such that its evaluation differs from that of the all zero input.

We construct a hitting set H as follows : for each monomial m in S_f, if no variable in H appear in m, add all the variables in it to H. Since, each such monomial contains a sensitive block on the input 0^n, the number of monomials that gets added to H is at most $\mathsf{bs}(f, 0^n)$. Since each monomial is of size at most $\mathsf{deg}_p(f)$, total size of the hitting set is at most $\mathsf{bs}(f, 0^n) \cdot \mathsf{deg}_p(f)$.

$\square$