Electronic Colloquium on Computational Complexity, Report No. 107 (2022)

Fractional certificates for bounded functions

Shachar Lovett*
Computer Science Department
University of California, San Diego
slovett@ucsd.edu

Jiapeng Zhang†
Computer Science Department
University of Southern California
jiapengz@usc.edu

July 20, 2022

Abstract

A folklore conjecture in quantum computing is that the acceptance probability of a quantum query algorithm can be approximated by a classical decision tree, with only a polynomial increase in the number of queries. Motivated by this conjecture, Aaronson and Ambainis (Theory of Computing, 2014) conjectured that this should hold more generally for any bounded function computed by a low degree polynomial.

In this work we prove two new results towards establishing this conjecture: first, that any such polynomial has a small fractional certificate complexity; and second, that many inputs have a small sensitive block. We also give two new conjectures that, if true, would imply the Aaronson and Ambainis conjecture given our results.

On the technical side, many classical techniques used in the analysis of Boolean functions seem to fail when applied to bounded functions. Here, we develop a new technique, based on a mix of combinatorics, analysis and geometry, and which in part extends a recent technique of Knop et al. (STOC 2021) to bounded functions.

1 Introduction

Aaronson and Ambainis [2] popularized the conjecture that quantum query algorithms can be approximated by classical query algorithms, on most inputs, with only a polynomial increase in the number of queries. This captures the informal belief that quantum algorithms can only achieve exponential speedup on highly structured inputs. Moreover, since the acceptance probability of quantum query algorithms can be computed by low degree polynomials, they conjectured that this holds more generally for any bounded function computed by a low degree polynomial.

*Research supported by NSF awards 1953928 and 2006443.
†Research supported by NSF CAREER award 2141536.

ISSN 1433-8092

A bit more formally, let $f : \{0,1\}^n \to [0,1]$ be a function which computes for each input x the acceptance probability of a quantum query algorithm. If the quantum algorithm makes at most q queries, then Beals et al. [7] showed that f is computed by a real polynomial of degree at most $d = 2q$. Aaronson and Ambainis conjectured that any such f can be approximated by a shallow decision tree.

Conjecture 1.1 (Aaronson-Ambainis (AA) conjecture [2])**.** *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial, and let $\varepsilon > 0$. Then there exists a decision tree T of depth poly$(d, 1/\varepsilon)$, such that*

$$\mathbb{E}_{x\in\{0,1\}^n}[|f(x) - T(x)|] \le \varepsilon.$$

The AA conjecture is known to be true for Boolean functions. Specificially, the seminal work of Nisan and Szegedy [17] showed that for every Boolean function $f : \{0,1\}^n \to \{0,1\}$, its decision tree complexity and its polynomial degree are equivalent, up to polynomial factors. However, their proof technique does not extend to bounded functions. In fact, many techniques used to study Boolean functions seem to fail when attempting to extend them to bounded functions.

We prove two new results in this paper, which we view as stepping stones towards a better understanding of bounded low degree polynomials:

1. In a bounded low degree polynomial, all inputs have a small fractional certificate complexity.
2. In a bounded low degree polynomial of large variance, many inputs have a small sensitive block.

We note that the first result holds for *all inputs*, whereas the AA conjecture only claims that $f(x) \approx T(x)$ for *most inputs*, and as such the two are incomparable; and that the second result is a direct corollary of the AA conjecture, as it trivially holds for decision trees. We show that it also follows from bounding the fractional certificate complexity.

1.1 Our results

We start with defining the above notions more precisely. Let $f : \{0,1\}^n \to [0,1]$ be a bounded function, let $x \in \{0,1\}^n$ be an input, and $\varepsilon > 0$ be a tolerance parameter. The *ε-certificate complexity* of f at x is the minimal size of a set $I \subset [n]$, such that any input y which agrees with x on I satisfies $|f(y) - f(x)| \le \varepsilon$. The *$\varepsilon$-fractional certificate complexity*[1] is its linear relaxation, where we replace a set I with a distribution π over $[n]$, and require that any y that is close to x under π satisfies $|f(y) - f(x)| \le \varepsilon$ (see Section 2 for formal definitions).

It is known that for Boolean functions, certificate complexity and fractional certificate complexity are equivalent, up to polynomial factors [1,3,19]. However, for bounded functions

[1] A similar notion called randomized certificate complexity was introduced by Aaronson [1].

they are not. Consider for example the linear function $f(x) = (x_1 + \ldots + x_n)/n$. For any constant ε, its ε-certificate complexity is $\Omega(n)$. In contrast, its ε-fractional certificate complexity is $O(1)$.

Motivated by this example, we explore the connections between fractional certificate complexity (which as we will see, is equivalent to fractional block sensitivity) and polynomial degree for bounded functions. Our first result is a bound on the ε-fractional certificate complexity that is polynomial in the degree d, tolerance parameter ε and logarithmic in the number of variables n.

Theorem 1.2 (Informal version of Theorem 2.9)**.** *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial, and let $\varepsilon > 0$. The ε-fractional certificate complexity of f is at most poly$(d, 1/\varepsilon, \log n)$.*

Comment 1.3. *The $\log n$ term appearing in Theorem 1.2 but not in Conjecture 1.1 might seem out of place. However, as Conjecture 1.1 allows for approximation of the function f, it is well-known that any degree d bounded function can be approximated up to error ε by its restriction to poly$(2^d, 1/\varepsilon)$ inputs [8]. So, if we allow approximation, then we may assume that $\log n = O(d + \log(1/\varepsilon))$.*

Next, we show that bounded functions with a small fractional certificate complexity and large variance have an interesting property - many inputs have small sensitive blocks.

Theorem 1.4 (Informal version of Theorem 4.2)**.** *Let $f : \{0,1\}^n \to [0,1]$, $\varepsilon > 0$ and assume Var$[f] = \Omega(\varepsilon)$. Then for at least an ε-fraction of inputs $x \in \{0,1\}^n$, there is a block $B \subset [n]$ of size $|B| \le r$ such that*

$$|f(x) - f(x \oplus B)| \ge \varepsilon,$$

where r is polynomial in the ε-fractional certificate complexity of f and in $\log(1/\varepsilon)$.

The proof of Theorem 1.4 follows from a new connection between fractional certificate complexity, convex geometry and concentration of measure (specifically, Talagrand's concentration inequality [20]).

Combining Theorem 1.2 and Theorem 1.4 gives the following corollary, that shows that for low degree bounded polynomials with large variance, many points have a small sensitive block.

Corollary 1.5. *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial, let $\varepsilon > 0$, and assume Var$[f] = \Omega(\varepsilon)$. Then for at least an ε-fraction of inputs $x \in \{0,1\}^n$, there is a block $B \subset [n]$ of size $|B| \le r$ such that*

$$|f(x) - f(x \oplus B)| \ge \varepsilon,$$

where $r =$ poly$(d, 1/\varepsilon, \log n)$.

1.2 New conjectures

Finally, we propose two conjectures that, if true, would in particular resolve the AA conjecture based on the results we show in this work. However, we believe that both are of independent interest beyond the connection to the AA conjecture.

Stable interpolating sets for polynomials. The first approach is based on *stable interpolating sets* for polynomials. A set $S \subset \{0,1\}^n$ is an interpolating set for degree d polynomials if two degree d polynomials which agree on S must agree on all points in the hypercube. It is well known that Hamming balls of radius d are such an interpolating set. We conjecture that, by mildly increasing their radius, they are *stable* - if two polynomial are close to each other on all points in S, then they are close on most points in the hypercube. This conjecture is equivalent to the conjecture that Theorem 1.4 can be extended to hold for all inputs, and not just an ε-fraction of inputs. We show that this conjecture, if true, when combined with Corollary 1.5 implies the AA conjecture. For details see Section 5.1.

Talagrand's concentration inequality for L_∞ norm. The second approach is inspired by Talagrand's concentration inequality [20]. Talagrand's concentration inequality states that if $X, Y \subset \{0,1\}^n$ are large sets, then most points in X are close in L_2 to the convex hull of Y. We examine what happens if we replace the L_2 norm with the L_∞ norm, and make the following conjecture: if X, Y do not have influential variables, then most points in X are close to the convex hull of Y also in the L_∞ norm. We show that this conjecture, if true, when combined with Theorem 1.2 also implies the AA conjecture. For details see Section 5.2.

1.3 Related works

The notions of block sensitivity and certificate complexity are extensively used in Boolean function analysis, namely, when the output of the function f takes Boolean values. Motivated by the study of quantum query algorithms, which is naturally captured by a bounded function, Aaronson [1] introduced the notion of randomized certificate complexity, which is very close to fractional certificate complexity. Tal [19] introduced the notions of fractional block-sensitivity and fractional certificate complexity (which are LP duals). Fractional block-sensitivity has been used to show a tight connection between the zero-error randomized decision tree complexity and two-sided bounded error randomized decision tree complexity [13]. Subsequent works [3, 4, 10, 12] studied fractional certificate complexity and fractional block-sensitivity, motivated by various applications in Boolean function analysis and communication complexity.

However, to the best of our knowledge, all these works focused only on Boolean functions. In particular, they showed that fractional certificate complexity and certificate complexity are polynomially related for Boolean functions. As we already discussed, this property is false for bounded functions. Motivated by studying quantum query algorithms, e.g., the AA conjecture, we study the fractional certificate complexity for bounded functions.

Previous works on Aaronson-Ambainis conjecture. Besides its importance in quantum computing, the AA conjecture is also a very intriguing problem in the area of Boolean function analysis. This conjecture is known to be true for Boolean functions [15, 18]. For bounded functions, a weaker bound with an exponential dependence on the degree instead of polynomial, can be proved using hyper-contractive inequalities [8]. Montanaro [16] proved a special case of the conjecture for block-multilinear forms where all the coefficients have the same magnitude. Recently, Bansal, Sinha and de Wolf [6] confirmed this conjecture in the case of functions with completely bounded degree-d block-multilinear form.

Paper organization. We define complexity measures for bounded functions in Section 2. We prove Theorem 1.2 in Section 3 and Theorem 1.4 in Section 4. Section 5 is devoted to two approaches towards the proof of the AA conjecture: the first based on stable interpolating sets for polynomials, and the second on extending Talagrand's inequality to the L_∞ norm.

2 Complexity measures of bounded functions

We introduce classic complexity measures of Boolean functions, as well as their linear relaxations, generalized to bounded functions.

Let $f : \{0,1\}^n \to [0,1]$ be a bounded function, $x \in \{0,1\}^n$ an input, and let $\varepsilon > 0$ be a tolerance parameter. Given a block B we denote by $x \oplus B$ the input obtained by flipping the bits in x corresponding to B. A block B is called *ε-sensitive* for x if $|f(x \oplus B) - f(x)| \geq \varepsilon$. The family of ε-sensitive blocks for x is defined as

$$\mathcal{S}_\varepsilon(f,x) = \{B \subset [n] : |f(x \oplus B) - f(x)| \geq \varepsilon\}.$$

Definition 2.1 (ε-block sensitivity)**.** *The ε-block sensitivity of f at x, denoted $BS_\varepsilon(f,x)$, is the maximal number of pairwise disjoint blocks $B_1, \ldots, B_k \in \mathcal{S}_\varepsilon(f,x)$.*

Definition 2.2 (ε-certificate complexity)**.** *The ε-certificate complexity of f at x, denoted $C_\varepsilon(f,x)$, is the minimal size of a set $I \subset [n]$ that intersects all blocks $B \in \mathcal{S}_\varepsilon(f,x)$. Equivalently:*

$$\forall y \in \{0,1\}^n : \quad y_I = x_I \Rightarrow |f(y) - f(x)| < \varepsilon.$$

We next define the linear relaxations of block sensitivity and certificate complexity, called fractional block sensitivity and fractional certificate complexity. These notion were introduced by Tal [19] in the context of Boolean functions. A similar notion to fractional certificate complexity, called randomized certificate, was introduced earlier by Aaronson [1].

Definition 2.3 (ε-fractional block sensitivity)**.** *The ε-fractional block sensitivity of f at x, denoted $FBS_\varepsilon(f,x)$, is the maximal k such that there exists a distribution ν over $\mathcal{S}_\varepsilon(f,x)$ that satisfies*

$$\forall i \in [n] : \quad \Pr_{B \sim \nu}[i \in B] \leq 1/k.$$

Definition 2.4 (ε-fractional certificate complexity)**.** *The ε-fractional certificate complexity of f at x, denoted $FC_\varepsilon(f,x)$, is the minimal k such that there exists a distribution π over $[n]$ that satisfies:*

$$\forall B \in \mathcal{S}_\varepsilon(f,x) : \quad \Pr_{i \sim \pi}[i \in B] \geq 1/k.$$

In other words:

$$\forall y \in \{0,1\}^n : \quad \left(\Pr_{i\sim\pi}[y_i \neq x_i] < 1/k\right) \Rightarrow |f(y) - f(x)| \leq \varepsilon.$$

Example 2.5. *Let $f(x) = (x_1 + \cdots + x_n)/n$. Fix an input $x \in \{0,1\}^n$ and $\varepsilon > 0$. Let $y \in \{0,1\}^n$ such that $|f(x) - f(y)| \geq \varepsilon$. This implies that the Hamming distance between x, y is at least εn, and hence $\Pr_{i\sim\pi}[x_i \neq y_i] \geq \varepsilon$, where π is the uniform distribution over $[n]$. This implies that $FC_\varepsilon(f, x) \leq 1/\varepsilon$, which in particular is independent of n.*

The following lemma is the classic connection between matchings, covers and their linear relaxations, when specialized to our setting. See [19] for a proof in the special case of block sensitivity, certificate complexity and their fractional relaxations (the proof in [19] is for Boolean functions, but it works equally well in our context).

Lemma 2.6. *$BS_\varepsilon(f, x) \leq FBS_\varepsilon(f, x) = FC_\varepsilon(f, x) \leq C_\varepsilon(f, x)$.*

We need one more definition of block sensitivity where we do not specify the tolerance ε.

Definition 2.7 (Block sensitivity)**.** *The block sensitivity of f at x, denoted $BS(f, x)$, is defined as*

$$BS(f, x) = \max_{B_1,\ldots,B_k} \sum_{i=1}^{k} |f(x) - f(x \oplus B_i)|,$$

where the maximum is over all collections of pairwise disjoint blocks.

Claim 2.8. *$BS(f, x) \geq \varepsilon \cdot BS_\varepsilon(f, x)$ for any $\varepsilon > 0$.*

For any complexity measure $\mathcal{C}$ (such as C_ε,BS_ε, etc), we define $\mathcal{C}(f) = \max_x \mathcal{C}(f, x)$.

2.1 Our results

With the definitions out of the way, we can now formally state our first theorem.

Theorem 2.9. *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial. Then for any $\varepsilon > 0$,*

$$FBS_\varepsilon(f) = FC_\varepsilon(f) \leq O\left(\frac{d^8 \log^{16} n}{\varepsilon^4}\right).$$

It is known that bounded low degree polynomials have bounded block sensitivity. This was first shown by Backurs and Bavarian [5] and then sharpened by Filmus, Hatami, Keller, and Lifshitz [9].

Lemma 2.10 ([9])**.** *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial. Then $BS(f) = O(d^2)$.*

Theorem 2.9 follows by combining Lemma 2.10 with the following theorem, which is our main technical contribution in this context. It upper bounds the integrality gap of block sensitivity for any bounded function (not necessarily computed by a low degree polynomial). A similar result for total Boolean functions is known [1, 3, 19], but their techniques do not seem to migrate well to the setting of bounded functions. Instead, we take a different approach, adapting ideas from [12] to the setting of bounded functions.

Theorem 2.11 (Upper bounding the integrality gap for block sensitivity)**.** *Let* $f : \{0,1\}^n \to [0,1]$ *and set* $B = \max(BS(f), 1)$*. Then for every* $\varepsilon > 0$*,*

$$FBS_{\varepsilon}(f) \leq O\left(\frac{B^4 \log^{16} n}{\varepsilon^4}\right).$$

We note that we did not attempt to optimize the exponents appearing in Theorem 2.9 and Theorem 2.11.

3 Upper bounding the integrality gap of block sensitivity

We prove Theorem 2.11 in this section. Before doing so, it would be convenient to recast the definitions of fractional block sensitivity and fractional certificates in a more systematic way.

3.1 Smoothness and fractional cover

Definition 3.1 (Smooth distribution)**.** *Let* $p \in (0,1)$*. A distribution* $\mathcal{D}$ *over* $\{0,1\}^n$ *is* p-smooth *if it satisfies* $\Pr_{x\sim\mathcal{D}}[x_i = 1] \leq p$ *for all* $i \in [n]$*.*

Definition 3.2 (Smooth probability)**.** *Let* $S \subset \{0,1\}^n$*. We denote by* $p_{smooth}(S)$ *the minimal* p*, such that there exists a* p*-smooth distribution* $\mathcal{D}$ *supported on* S*.*

Definition 3.3 (Cover probability)**.** *Let* $S \subset \{0,1\}^n$*. We denote by* $p_{cover}(S)$ *the maximal* p*, such that there exists a distribution* π *over* $[n]$ *satisfying* $\Pr_{i\sim\pi}[x_i = 1] \geq p$ *for all* $x \in S$*.*

Recall the definition of $\mathcal{S}_{\varepsilon}(f, x) = \{B \subset [n] : |f(x \oplus B) - f(x)| \geq \varepsilon\}$. We can recast the definitions of fractional block sensitivity and fractional certificates as

$$\mathrm{FBS}_{\varepsilon}(f, x) = p_{\mathrm{smooth}}(\mathcal{S}_{\varepsilon}(f, x)), \quad \mathrm{FC}_{\varepsilon}(f, x) = p_{\mathrm{cover}}(\mathcal{S}_{\varepsilon}(f, x)).$$

We next prove a number of useful claims about p_{smooth} and p_{cover}.

Claim 3.4. $p_{smooth}(S) = p_{cover}(S)$ *for any* $S \subset \{0,1\}^n$*.*

Proof. This is the classic LP duality between fractional matching and fractional covers in hypergraphs (see for example [14]). $\square$

Let $p(S) := p_{\text{smooth}}(S) = p_{\text{cover}}(S)$. Note that if $0^n \in S$ then $p(S) = 0$.

Claim 3.5. *Let $S \subset \{0,1\}^n \setminus \{0\}^n$. Then $p(S) \geq 1/n$.*

Proof. Let π be the uniform distribution over $[n]$. As $0^n \notin S$ we have $\Pr[x_i = 1] \geq 1/n$ for all $x \in S$. Thus $p(S) = p_{\text{cover}}(S) \geq 1/n$. $\square$

Claim 3.6. *Let $S \subset \{0,1\}^n$ with $p(S) = p$. Let $\mathcal{D}$ be a q-smooth distribution over $\{0,1\}^n$ where $q < p$. Then*

$$\Pr_{x \sim \mathcal{D}}[x \in S] \leq q/p.$$

Proof. Let $\alpha = \Pr_{x \sim D}[x \in S]$. Let $\mathcal{D}'$ be the distribution of $x \sim \mathcal{D}$ conditioned on $x \in S$, namely $\mathcal{D}'(x) = 0$ if $x \notin S$, and $\mathcal{D}'(x) = \mathcal{D}(x)/\alpha$ if $x \in S$. Note that $\mathcal{D}'$ is (q/α)-smooth and supported on S, and hence $q/\alpha \geq p$. $\square$

We identify $\{0,1\}^n$ with subsets of $[n]$. In particular, given $x, y \in \{0,1\}^n$ we identify $x \cup y$, $x \cap y$ and $x \setminus y$ with the usual definition for sets (union, intersection, set difference).

Claim 3.7. *Let $\mathcal{D}$ be a p-smooth distribution over $\{0,1\}^n$. For $k \geq 1$, define a distribution $\mathcal{D}'$ by the following sampling process: sample $y_1, \ldots, y_k \sim \mathcal{D}$ independently and output*

$$z = \bigcup_{i \neq j} y_i \cap y_j.$$

Then $\mathcal{D}'$ is $(pk)^2$-smooth

Proof. This follows from the definition of smoothness. For any coordinate $\ell \in [n]$ we have

$$\Pr_{z \sim D'}[z_\ell = 1] \leq \sum_{i \neq j} \Pr_{y_i \sim \mathcal{D}}[(y_i)_\ell = 1] \Pr_{y_j \sim \mathcal{D}}[(y_j)_\ell = 1] \leq (pk)^2.$$

$\square$

3.2 Bounding the integrality gap

We now turn to prove Theorem 2.11. It will be convenient to allow to mildly change ε. The following is our main technical lemma in this section. To simplify notations, we set $\mathrm{B}_\varepsilon(f) = \max(\mathrm{BS}_\varepsilon(f), 1)$ throughout the section.

Lemma 3.8. *Let $f : \{0,1\}^n \to [0,1]$ and $\varepsilon \in (0,1)$. Then there exists $1 \leq t \leq \log^4 n$ such that*

$$FBS_\varepsilon(f) \leq FBS_{\varepsilon/t}(f) \leq O\left(B_{\varepsilon/t}(f)^4\right).$$

Combining Lemma 3.8 with the bound $\mathrm{BS}_\delta(f) \leq \mathrm{BS}(f)/\delta$ given by Claim 2.8 implies Theorem 2.11. We prove Lemma 3.8 in the remainder of this subsection. The following lemma is an adaptation of [12, Lemma 3.2] to bounded functions.

Lemma 3.9. *Let $f : \{0,1\}^n \to [0,1]$ and $\varepsilon \in (0, 1/3)$. Then*

$$\frac{FBS_{3\varepsilon}(f)}{\sqrt{FBS_{\varepsilon}(f)}} \leq O\left(B_{\varepsilon}(f)\right).$$

Proof. Let $\mathrm{FBS}_{3\varepsilon}(f) = 1/p$ and $\mathrm{FBS}_{\varepsilon}(f) = 1/q$. Note that $0 \leq q \leq p \leq 1$. We may assume that $q \geq 4p^2$, otherwise the claim is trivial. Let $x \in \{0,1\}^n$ so that $\mathrm{FBS}_{3\varepsilon}(f) = \mathrm{FBS}_{3\varepsilon}(f, x)$. Let $S = \mathcal{S}_{3\varepsilon}(f, x)$, and let $\mathcal{D}$ be a p-smooth distribution supported on S. Let k to be determined later, and sample $y_1, \ldots, y_k \sim \mathcal{D}$ independently. Define

$$e = \bigcup_{i \neq j} (y_i \cap y_j).$$

Finally, let $z_i = y_i \setminus e$. Observe that $z_1, \ldots, z_k$ are pairwise disjoint.

Observe that Claim 3.7 implies that e is $(pk)^2$-smooth and set $\delta = (pk)^2/q$. Let $S_0 = \mathcal{S}_{\varepsilon}(f, x)$ and note that by assumption $p(S_0) \geq q$. Claim 3.6 implies that $\Pr[e \in S_0] \leq \delta$, or in other words

$$\Pr[|f(x) - f(x \oplus e)| \geq \varepsilon] \leq \delta. \tag{1}$$

Next, fix $i \in [k]$ and also fix y_i for a moment. Define

$$e_i = \bigcup_{j \neq j', j, j' \neq i} (y_j \cap y_{j'}) \setminus y_i.$$

Applying Claim 3.7 again we get that e_i is also $(pk)^2$-smooth. Let $S_i = \mathcal{S}_{\varepsilon}(f, x \oplus y_i)$, which again satisfies $p(S_i) \geq q$. Applying Claim 3.6 again gives

$$\Pr_{\{y_j\}_{j \neq i}} [|f(x \oplus y_i) - f(x \oplus y_i \oplus e_i)| \geq \varepsilon] \leq \delta.$$

Note that $y_i \oplus e_i = y_i \vee e_i = y_i \vee e = z_i \oplus e$. Averaging also over y_i gives

$$\Pr[|f(x \oplus y_i) - f(x \oplus z_i \oplus e)| \geq \varepsilon] \leq \delta. \tag{2}$$

Next, since each $y_i \sim \mathcal{D}$ is supported on $\mathcal{S}_{3\varepsilon}(f, x)$, we have $|f(x) - f(x \oplus y_i)| \geq 3\varepsilon$ with probability one. Combining this with Equations (1) and (2), and setting $w = x \oplus e$, gives

$$\Pr[|f(w) - f(w \oplus z_i)| \geq \varepsilon] \geq 1 - 2\delta. \tag{3}$$

Recall that $\delta = (pk)^2/q$. We choose $k = \Omega(\sqrt{q}/p)$ so that $\delta \leq 1/4$. Let $I = \{i \in [k] : |f(w) - f(w \oplus z_i)| \geq 2\varepsilon\}$. We have $E[|I|] \geq (1 - 2\delta)k \geq k/2$. By averaging, there exists a choice of $y_1, \ldots, y_k$ so that $|I| \geq k/2$. Fix such a choice, and note that it gives

$$\mathrm{BS}_{\varepsilon}(f) \geq \mathrm{BS}_{\varepsilon}(f, w) \geq k/2.$$

$\square$

Claim 3.10. *Fix $\varepsilon \in (0, 1/3)$ and assume $FBS_\varepsilon(f) \geq 2$. Then there exists $1 \leq t \leq \log^4 n$ so that*

$$FBS_{\varepsilon/t}(f) \leq \left(FBS_{3\varepsilon/t}(f)\right)^{4/3}.$$

Proof. Shorthand $h(i) = \mathrm{FBS}_{\varepsilon/3^i}(f)$ for $i \geq 0$. Let $m \geq 0$ be maximal so that for every $i \in [m]$ it holds that $h(i) \geq (h(i-1))^{4/3}$. This implies that $h(m) \geq 2^{(4/3)^m}$. On the other hand, Claim 3.5 implies $\mathrm{FBS}_\delta(f) \leq n$ for any $\delta > 0$, and hence $h(m) \leq n$. Thus $(4/3)^m \leq \log n$ and hence $3^m \leq (\log n)^{\log_{4/3}(3)} \leq \log^4 n$. The claim holds for $t = 3^m$. $\square$

We now prove Lemma 3.8.

Proof of Lemma 3.8. If $\mathrm{FBS}_\varepsilon(f) \leq 2$ we are done. Otherwise, apply Claim 3.10 to get $1 \leq t \leq \log^4 n$ so that $\mathrm{FBS}_{\varepsilon/t}(f) \leq \left(\mathrm{FBS}_{3\varepsilon/t}(f)\right)^{4/3}$. Set $\varepsilon' = \varepsilon/t$, where rearranging the terms gives

$$\frac{\mathrm{FBS}_{3\varepsilon'}(f)}{\sqrt{\mathrm{FBS}_{\varepsilon'}(f)}} \geq \mathrm{FBS}_{\varepsilon'}(f)^{1/4}.$$

Applying Lemma 3.9 for ε' gives

$$\frac{\mathrm{FBS}_{3\varepsilon'}(f)}{\sqrt{\mathrm{FBS}_{\varepsilon'}(f)}} \leq O\left(\mathrm{B}_{\varepsilon'}(f)\right).$$

To conclude the proof note that $\mathrm{FBS}_\varepsilon(f) \leq \mathrm{FBS}_{\varepsilon'}(f)$ since $\varepsilon' \leq \varepsilon$. $\square$

4 Small block sensitivity

A corollary of the AA conjecture is that for low degree bounded functions with a large variance, many inputs have a small sensitive block (as this holds for decision trees). We show that this also follows from having small fractional certificate complexity.

Definition 4.1 (Small block sensitivity)**.** *Let $f : \{0,1\}^n \to [0,1]$. A point $x \in \{0,1\}^n$ is called (r, ε)-sensitive if there exists a block B of size $|B| \leq r$ such that*

$$|f(x) - f(x \oplus B)| \geq \varepsilon.$$

If no such block exists, we say that x is (r, ε)-insensitive.

Theorem 4.2. *Let $f : \{0,1\}^n \to [0,1]$ and $\varepsilon > 0$, and assume $Var[f] \geq \Omega(\varepsilon)$. Then at least an ε-fraction of the points $x \in \{0,1\}^n$ are (r, ε)-sensitive for $r = O\left(FC_\varepsilon(f)^2 \cdot \log(1/\varepsilon)\right)$.*

The first step towards the proof of Theorem 4.2 is to connect fractional certificate complexity to convex geometry. Let $x \in \{0,1\}^n$, $Y \subset \{0,1\}^n$. We denote by $\mathrm{conv}(Y)$ the convex hull of Y in $[0,1]^n$. Given a set $K \subset [0,1]^n$ and $x \in \{0,1\}^n$, define their L_p distance as

$$d_p(x, K) = \min_{y \in K} \|x - y\|_p.$$

We will restrict our attention to two norms: L_2 and L_∞. We first connect the L_∞ norm to fractional certificate complexity.

Lemma 4.3. *Let $f : \{0,1\}^n \to [0,1]$, $x \in \{0,1\}^n$, $\varepsilon > 0$ and $Y = \{y \in \{0,1\}^n : |f(x) - f(y)| \geq \varepsilon\}$. Then*

$$d_\infty(x, conv(Y)) \geq \frac{1}{FC_\varepsilon(f,x)}.$$

Proof. Assume $\mathrm{FC}_\varepsilon(f,x) = k$. This means there is a distribution π over $[n]$, such that for all $y \in Y$,

$$\Pr_{i \sim \pi}[x_i \neq y_i] \geq 1/k.$$

Let $s_i = (-1)^{x_i}$. We can rewrite this condition as

$$\mathbb{E}_{i\sim\pi}[s_i(y_i - x_i)] \geq 1/k.$$

Let $y^* \in \mathrm{conv}(Y)$ be the point closest to x in L_∞. Then by linearity of expectation we have that

$$\mathbb{E}_{i\sim\pi}[s_i(y^*_i - x_i)] \geq 1/k.$$

Let $p = \|x - y^*\|_\infty = d_\infty(x, \mathrm{conv}(Y))$, so that $|y^*_i - x_i| \leq p$ for all i. Then we must have $p \geq 1/k$. □

We next connect the L_2 norm and the L_∞ norm via small block sensitivity.

Lemma 4.4. *Let $f : \{0,1\}^n \to [0,1]$, $x \in \{0,1\}^n$, $t \geq f(x)$ and $\varepsilon > 0$. Define*

$$Y = \{y \in \{0,1\}^n : f(y) \geq t + \varepsilon\}$$

and

$$Z = \{z \in \{0,1\}^n : f(z) \geq t + 2\varepsilon \text{ and } z \text{ is } (r,\varepsilon)\text{-insensitive}\}.$$

Then

$$d_2(x, conv(Z)) \geq d_\infty(x, conv(Y)) \cdot \sqrt{r}.$$

Proof. Let $p = d_\infty(x, \mathrm{conv}(Y))$. Let $B_r(Z)$ denote the Hamming ball of radius r around Z:

$$B_r(Z) = \{z \oplus B : z \in Z, B \subset [n], |B| \leq r\}.$$

Observe first that $B_r(Z) \subset Y$. To see that, take $z \in Z$ and $|B| \leq r$. We need to show that $z \oplus B \in Y$. Since by assumption z is (r,ε)-insensitive, we have $f(z \oplus B) \geq f(z) - \varepsilon \geq t + \varepsilon$ and hence $z \oplus B \in Y$.

For each $0 \leq \ell \leq r$ let $z^{(\ell)} \in \mathrm{conv}(B_\ell(Z))$ be the closest point to x in L_2. The proof will follow by showing that for all $0 \leq \ell \leq r-1$:

$$\|x - z^{(\ell)}\|_2^2 \geq p^2 + \|x - z^{(\ell+1)}\|_2^2, \tag{4}$$

as this implies

$$d_2(x, \mathrm{conv}(Z))^2 = \|x - z^{(0)}\|_2^2 \geq p^2 r.$$

We next prove Equation (4). Fix ℓ and consider $z^{(\ell)}$. Since $z^{(\ell)} \in \mathrm{conv}(B_\ell(Z)) \subset \mathrm{conv}(Y)$, we must have $\|x - z^{(\ell)}\|_\infty \geq d_\infty(x, \mathrm{conv}(Y)) = p$. Let $i \in [n]$ be a coordinate for which $|x_i - z_i^{(\ell)}| \geq p$. Define $w^{(\ell)} \in [0,1]^n$ as follows: $w_i^{(\ell)} = x_i$ and $w_j^{(\ell)} = z_j^{(\ell)}$ for $j \neq i$. Then

$$\|x - z^{(\ell)}\|_2^2 \geq p^2 + \|x - w^{(\ell)}\|_2^2.$$

To conclude the proof, note that as $z^{(\ell)} \in \mathrm{conv}(B_\ell(Z))$ and $w^{(\ell)}$ differs from $z^{(\ell)}$ in at most one coordinate, then $w^{(\ell)} \in \mathrm{conv}(B_{\ell+1}(Z))$. This implies that $\|x - z^{(\ell+1)}\|_2 \leq \|x - w^{(\ell)}\|_2$ which completes the proof. □

We would need the following simple claim, showing that a bounded random variable which does not deviate much from its expectation, must have a small variance.

Claim 4.5. *Let X be a random variable taking values in $[0,1]$. Assume that for some $a, b > 0$ we have*

$$\Pr[X \geq \mathbb{E}[X] + a] \leq b.$$

Then

$$Var[X] \leq 2(a+b).$$

Proof. Let $Y = X - \mathbb{E}[X]$ so that Y takes values in $[-1,1]$ and $\mathbb{E}[Y] = 0$. We have

$$0 = \mathbb{E}[Y] = \mathbb{E}[\max(Y,0)] - \mathbb{E}[\max(-Y,0)].$$

Therefore, $\mathbb{E}[\max(Y,0)] = \mathbb{E}[\max(-Y,0)]$. By assumption, $\Pr[Y \geq a] \leq b$ and hence

$$\mathbb{E}[\max(Y,0)] \leq a + b$$

which implies

$$\mathbb{E}[\max(-Y,0)] \leq a + b.$$

Thus

$$\mathrm{Var}[X] = \mathbb{E}[Y^2] \leq \mathbb{E}[|Y|] = \mathbb{E}[\max(Y,0)] + \mathbb{E}[\max(-Y,0)] \leq 2(a+b).$$

□

The final piece we need is Talagrand's concentration inequality [20].

Theorem 4.6 (Talagrand [20])**.** *Let $X, Y \subset \{0,1\}^n$. Assume that for all $x \in X$,*

$$d_2(x, conv(Y)) \geq \lambda.$$

Then

$$\frac{|X||Y|}{2^{2n}} \leq \exp(-\lambda^2/4).$$

We now prove Theorem 4.2.

Proof of Theorem 4.2. Let t be the average value of $\{f(x) : x \in \{0,1\}^n\}$. Define

$$\begin{aligned}
X &= \{x \in \{0,1\}^n : f(x) \leq t\}, \\
Y &= \{y \in \{0,1\}^n : f(y) \geq t + \varepsilon\}, \\
Z &= \{z \in \{0,1\}^n : f(z) \geq t + 2\varepsilon\}, \\
W &= \{w \in \{0,1\}^n : f(w) \geq t + 2\varepsilon \text{ and } w \text{ is } (r,\varepsilon)\text{-insensitive}\}.
\end{aligned}$$

The assumption $\mathrm{Var}[f] \geq \Omega(\varepsilon)$ implies by Claim 4.5 that $|X|, |Y|, |Z| \geq 2\varepsilon 2^n$. We will soon show that $|W| \leq \varepsilon 2^n$. This will conclude the proof as all points in $Z \setminus W$ are (r, ε)-sensitive, and there are at least $|Z| - |W| \geq \varepsilon 2^n$ such points.

Let $k = \mathrm{FC}_\varepsilon(f)$. Lemma 4.3 gives that for all $x \in X$,

$$d_\infty(x, \mathrm{conv}(Y)) \geq \frac{1}{k}.$$

Lemma 4.4 then gives that

$$d_2(x, \mathrm{conv}(W)) \geq \frac{\sqrt{r}}{k}.$$

Applying Talagrand's inequality (Theorem 4.6) to X, W then gives

$$\frac{|X||W|}{2^{2n}} \leq \exp(-r/4k^2).$$

Choosing $r = O(k^2 \log(1/\varepsilon))$, and recalling that $|X| \geq \varepsilon 2^n$, gives that $|W| \leq \varepsilon 2^n$. This concludes the proof. □

5 Towards the Aaronson-Ambainis conjecture

In this section we present two potential (but somewhat speculative) directions towards the Aaronson-Ambainis conjecture. The first is based on stable interpolating sets, and the second on a conjectured extension of Talagrand's inequality to the L_∞ norm.

Before doing so, it would be convenient for us to recast the AA conjecture in a more amenable way. Similar to the equivalent formulation of the AA conjecture of f having an influential variable, we consider the version of an influential small coalition.

Conjecture 5.1 (AA conjecture: equivalent formulation)**.** *Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial, and let $\varepsilon > 0$. Then there is a set $B \subset [n]$ of size $|B| \leq poly(d, 1/\varepsilon)$ and an assignment $b \in \{0,1\}^B$ such that $Var[f(x)|x_B = b] \leq \varepsilon$.*

Claim 5.2. *Conjecture 1.1 and conjecture 5.1 are equivalent.*

Proof. It is clear that Conjecture 5.1 follows from Conjecture 1.1, by considering a leaf in the decision tree approximating f. The reverse direction also holds by standard techniques: querying the variables in the block B reduces the average block sensitivity for the function. For more details, see for example [11, Lemma 6.1], where although their full proof is wrong, this specific lemma is correct and gives the details for this procedure. □

5.1 Stable interpolating sets

Combining Theorem 2.9 and Theorem 4.2 gives the following corollary:

Corollary 5.3. *Let $f : \{0,1\}^n \to [0,1]$ be computed by a polynomial of degree d, and assume $Var[f] = \Omega(\varepsilon)$. Then at least an ε-fraction of inputs $x \in \{0,1\}^n$ are (r,ε)-sensitive for $r = poly(d, 1/\varepsilon, \log(n))$.*

We conjecture that in fact this should hold for *all inputs*, not just many inputs. We show that if true, then it implies the AA conjecture. We rephrase this conjecture in the language of *stable* interpolating sets for polynomials.

Definition 5.4. *A set $S \subset \{0,1\}^n$ is an (ε,δ)-stable interpolating set for degree d bounded polynomials if it satisfies the following condition. Let $f : \{0,1\}^n \to [0,1]$ be a degree d polynomial, and assume $|f(x)| \le \delta$ for all $x \in S$. Then $\mathbb{E}_{x\in\{0,1\}^n}[f(x)^2] \le \varepsilon$.*

Note that $(0,0)$-stable interpolating sets are exactly the standard definition of interpolating sets. It is known that Hamming balls of radius d are interpolating sets for degree d polynomials. We conjecture that, by mildly increasing their radius, they are also stable interpolating sets.

Conjecture 5.5. *Let $d \ge 1, \varepsilon > 0$. Then for some $r = poly(d/\varepsilon)$ and $\delta = poly(\varepsilon/d)$, Hamming balls of radius r are (ε,δ)-interpolating sets for degree d bounded polynomials.*

Note that Conjecture 5.5 is equivalent to the conjecture that if f is a bounded degree d polynomial with $\mathrm{Var}[f] \ge \varepsilon$, then all inputs $x \in \{0,1\}^n$ are (r,δ)-sensitive.

Claim 5.6. *Conjecture 5.5 implies Conjecture 5.1.*

Proof. Take any input x. As long as $\mathrm{Var}[f] \ge \varepsilon$, there is a block B of size $|B| \le r = \mathrm{poly}(d/\varepsilon)$ that is δ-sensitive for x, where $\delta = \mathrm{poly}(\varepsilon/d)$. Restrict to the subcube of inputs consistent with $x|_B$ and repeat. This process will end after at most $\mathrm{BS}_\delta(f,x) = O(d^2/\delta)$ steps, at which point we get a block satisfying the assumption of Conjecture 5.1. □

5.2 Talagrand inequality for L_∞

Recall Talagrand's inequality (Theorem 4.6), and consider replacing the distance from L_2 to L_∞. What would change? First, the distance can be at most 1. Second, even if X, Y are dense sets, their structure plays a part. Consider the following two motivating examples.

Example 5.7 (Subcubes)**.** *Let $X = \{x : x_1 = 0\}$, $Y = \{x : x_1 = 1\}$. Then $|X| = |Y| = 2^{n-1}$ and $d_\infty(x, conv(Y)) = 1$ for all $x \in X$.*

Example 5.8 (Hamming balls)**.** *Let $X = \{x : |x| \le n/2 - \sqrt{n}\}$, $Y = \{x : |x| \ge n/2 + \sqrt{n}\}$ where $|x|$ denotes the Hamming weight of x. Then $|X| = |Y| = \Omega(2^n)$ and $d_\infty(x, conv(Y)) = O(1/\sqrt{n})$ for x on the boundary of X (namely, with $|x| = n/2 - \sqrt{n}$).*

We conjecture that the main difference between these two examples is that, in the first example X, Y have a variable with large influence, whereas in the second example all variables have influence $O(1/\sqrt{n})$. We conjecture that this is a general phenomenon.

Definition 5.9. *Let $X \subset \{0,1\}^n$. The i-th influence of X is the probability that a random element in X moves outside X when the i-th bit is flipped:*

$$\mathit{Inf}_i[X] = \Pr_{x \in X}[x \oplus e_i \notin X].$$

The maximal influence of X is $\mathit{Inf}_\infty[x] = \max_i \mathit{Inf}_i[X]$.

Conjecture 5.10 (Talagrand for L_∞)**.** *Let $X, Y \subset \{0,1\}^n$. Assume that $\mathit{Inf}_\infty[X], \mathit{Inf}_\infty[Y] \leq \tau$. Then there exists $x \in X$ such that*

$$d_\infty(x, \mathit{conv}(Y)) \leq \mathit{poly}(\tau).$$

We show that Conjecture 5.10 also implies the AA conjecture.

Claim 5.11. *Conjecture 5.10 implies Conjecture 5.1.*

Proof. Let $f : \{0,1\}^n \to [0,1]$ be computed by a degree d polynomial, and $\mathrm{Var}[f] \geq \Omega(\varepsilon)$. Let t be the average value of $\{f(x) : x \in \{0,1\}^n\}$. For $\alpha \in [\varepsilon, 2\varepsilon]$ to be determined soon, define

$$X = \{x : f(x) \leq t - \alpha\}, \quad Y = \{x : f(x) \geq t + \alpha\}.$$

The assumption that $\mathrm{Var}[f] \geq \Omega(\varepsilon)$ implies by Claim 4.5 that $|X|, |Y| \geq \varepsilon 2^n$. Theorem 2.9 gives that for all $x \in X$, $d_\infty(x, \mathrm{conv}(Y)) \geq p$ where $p^{-1} = \mathrm{poly}(d, 1/\varepsilon, \log n)$. Conjecture 5.10 then implies that either $\mathrm{Inf}_\infty[X] > \tau$ or $\mathrm{Inf}_\infty[Y] > \tau$ where $\tau^{-1} = \mathrm{poly}(d, 1/\varepsilon, \log n)$.

Assume without loss of generality that $\mathrm{Inf}_\infty[X] > \tau$. This means that there is an index $i \in [n]$ such that $\mathrm{Inf}_i[X] > \tau$. In other words, the *linear threshold function* $\mathrm{sign}(f(x) - t + \alpha)$ has an influential variable x_i. We will now show that by a careful choice of α, this implies that x_i is also an influential variable for f. This in turn is sufficient to prove the AA conjecture.

Let $\beta > 0$ to be determined later (where we will have $\beta^{-1} = \mathrm{poly}(d, 1/\varepsilon, \log n)$). Say that a value α is *good* if $\Pr_{x \in \{0,1\}^n}[0 \leq f(x) - t + \alpha \leq \beta] \leq \varepsilon\tau/2$. Note that if α is good, then we get

$$\begin{aligned}
\mathbb{E}_{x \in \{0,1\}^n}[|f(x \oplus e_i) - f(x)|] &\geq \varepsilon \cdot \mathbb{E}_{x \in X}[|f(x \oplus e_i) - f(x)|] \\
&\geq \varepsilon\beta \cdot \Pr_{x \in X}[f(x \oplus e_i) > t - \alpha + \beta] \\
&= \varepsilon\beta \left(\Pr_{x \in X}[f(x \oplus e_i) > t - \alpha] - \Pr_{x \in X}[0 \leq f(x) - t + \alpha \leq \beta] \right) \\
&\geq \varepsilon\beta \left(\mathrm{Inf}_i[X] - \tau/2\right) \\
&\geq \varepsilon\beta\tau/2.
\end{aligned}$$

This implies that x_i is an influential variable in f, with influence $\mathrm{poly}(d, 1/\varepsilon, \log(n))^{-1}$, as conjectured.

To conclude, we need to show that a good value of α exists. Assume not; then for every $\alpha \in [\varepsilon, 2\varepsilon]$, we have at least a $\varepsilon\tau/2$ mass of $\{f(x) : x \in \{0,1\}^n\}$ lying in the interval $[t-\alpha, t-\alpha+\beta]$. This of course is impossible if we set β small enough, concretely $\beta = O(\tau\varepsilon^2)$. □

References

[1] S. Aaronson. Quantum certificate complexity. *Journal of Computer and System Sciences*, 74(3):313–322, 2008.

[2] S. Aaronson and A. Ambainis. The need for structure in quantum speedups. *Theory of Computing*, 10(6):133–166, 2014.

[3] A. Ambainis, K. Prūsis, and J. Vihrovs. On block sensitivity and fractional block sensitivity. *Lobachevskii Journal of Mathematics*, 39(7):967–969, 2018.

[4] A. Anshu, S. Ben-David, and S. Kundu. On query-to-communication lifting for adversary bounds. *arXiv preprint arXiv:2012.03415*, 2020.

[5] A. Backurs and M. Bavarian. On the sum of L_1 influences. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 132–143. IEEE, 2014.

[6] N. Bansal, M. Sinha, and R. de Wolf. Influence in completely bounded block-multilinear forms and classical simulation of quantum algorithms. *arXiv preprint arXiv:2203.00212*, 2022.

[7] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM (JACM)*, 48(4):778–797, 2001.

[8] I. Dinur, E. Friedgut, G. Kindler, and R. O'Donnell. On the Fourier tails of bounded functions over the discrete cube. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 437–446, 2006.

[9] Y. Filmus, H. Hatami, N. Keller, and N. Lifshitz. On the sum of the L_1 influences of bounded functions. *Israel Journal of Mathematics*, 214(1):167–192, 2016.

[10] J. Gilmer, M. Saks, and S. Srinivasan. Composition limits and separating examples for some Boolean function complexity measures. *Combinatorica*, 36(3):265–311, 2016.

[11] N. Keller and O. Klein. Quantum speedups need structure. *arXiv preprint arXiv:1911.03748*, 2019. Paper was withdrawn because it contained an error.

[12] A. Knop, S. Lovett, S. McGuire, and W. Yuan. Log-rank and lifting for AND-functions. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 197–208, 2021.

[13] R. Kulkarni and A. Tal. On fractional block sensitivity. *Chicago J. Theor. Comput. Sci*, 8:1–16, 2016.

[14] L. Lovász. 2-matchings and 2-covers of hypergraphs. *Acta Mathematica Hungarica*, 26(3-4):433–444, 1975.

[15] G. Midrijanis. On randomized and quantum query complexities. *arXiv preprint quant-ph/0501142*, 2005.

[16] A. Montanaro. Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):122206, 2012.

[17] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational complexity*, 4(4):301–313, 1994.

[18] R. O'Donnell, M. Saks, O. Schramm, and R. A. Servedio. Every decision tree has an influential variable. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 31–39. IEEE, 2005.

[19] A. Tal. Properties and applications of Boolean function composition. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 441–454, 2013.

[20] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, 81(1):73–205, 1995.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il