A simplified proof of the Aaronson–Ambainis conjecture

Shiva Kintali*

November 6, 2026

Abstract

We give a shorter exposition of Turturean's proof of the Aaronson–Ambainis influence conjecture. For a function $f : \{-1,1\}^n \to [0,1]$ of degree at most $d \geq 1$ and variance $V > 0$, we show that some coordinate i satisfies $\mathbb{E}(f(X) - f(X^i))^2 \geq 4V^4/(9 \cdot 36^4 d^8)$, where X is uniform and X^i is obtained by flipping its ith coordinate. The proof combines Turturean's determinant construction with degree bounds on fractional certificates, then uses threshold integration and direct averaging to obtain the influence estimate. This is a simplified proof with a weaker quantitative bound.

1 Introduction

How much can a quantum algorithm gain by querying its input in superposition? To study this question on uniformly random Boolean inputs, Beals et al. [1] represent the acceptance probability of any T-query quantum algorithm as a polynomial of degree at most $2T$. Its values lie in $[0,1]$ on the whole Boolean cube. Aaronson and Ambainis [2] proposed the following conjecture about such bounded polynomials; it is now a theorem by Turturean [18].

Aaronson–Ambainis influence conjecture. There are universal constants $c, C > 0$ such that, for every $n \geq 1$ and every $f : Q_n \to [0,1]$ of degree at most $d \geq 1$ and variance $V > 0$, some coordinate i satisfies

$$\mathbb{E}\big(f(X) - f(X^i)\big)^2 \geq c \left(\frac{V}{d}\right)^C.$$

Here $Q_n = \{-1,1\}^n$ and $[n] = \{1,\ldots,n\}$; X is uniform on Q_n, X^i is X with its ith coordinate flipped, and $V = \mathbb{E}\big(f(X) - \mathbb{E}f(X)\big)^2$. The degree is that of the multilinear polynomial representing f on the cube. We denote the left side by $J_i(f)$; it is the mean squared change caused by flipping coordinate i, called its *influence*. Thus a bounded low-degree polynomial with substantial variance has an influential coordinate, with a bound independent of the number of variables. The precise bound proved in this note appears in Theorem 1.

Theorem 1. *Let $n \geq 1$ and let $f : Q_n \to [0,1]$ have degree at most $d \geq 1$ and variance $V > 0$. Then some $i \in [n]$ satisfies*

$$J_i(f) = \mathbb{E}\big(f(X) - f(X^i)\big)^2 \geq \frac{4V^4}{9 \cdot 36^4 d^8},$$

where X is uniformly random in Q_n.

*Email: shiva.kintali@gmail.com Homepage: shivakintali.github.io

The proof appears at the end of Section 4. The hypothesis $V > 0$ excludes constant functions, and hence forces the degree to be at least one. Since $f^2 \le f$, writing $\mu = \mathbb{E}f$ gives $0 < V \le \mu(1-\mu) \le 1/4$. The conclusion is independent of n: non-negligible variance and low degree force at least one coordinate to have non-negligible influence. The particular powers and constant are not optimized; their polynomial dependence on V and $1/d$ is what establishes the conjecture. The affine bijection $z \mapsto 2z - \mathbf{1}$ from $\{0,1\}^n$ to Q_n preserves uniform measure and coordinate flips. Substitution by this map and its inverse does not increase degree, so it preserves multilinear degree as well, and the same bound holds on $\{0,1\}^n$.

Comparison of bounds. In the same range and influence normalization, Turturean [18] proves the stronger quantitative estimate

$$\begin{aligned} &\text{This exposition:} && \max_i J_i(f) \ge c_0 \frac{V^4}{d^8}, && c_0 = \frac{4}{9 \cdot 36^4}, \\ &\text{Turturean:} && \max_i J_i(f) \ge c_1 \frac{V^{5/2}}{d^6}, && c_1 > 0. \end{aligned}$$

Here c_1 is an absolute constant. The present bound is weaker; the aim is to make the final passage from certificates to influence shorter.

Aaronson and Ambainis showed that the conjecture implies a classical simulation of quantum acceptance probabilities on most inputs: for every T-query quantum algorithm and $0 < \varepsilon, \delta < 1$, a deterministic classical algorithm can make $\mathrm{poly}(T, 1/\varepsilon, 1/\delta)$ queries and estimate the acceptance probability to within additive error ε on at least a $1 - \delta$ fraction of uniformly random inputs. The guarantee concerns the number of queries; it does not by itself give a polynomial running-time bound or accuracy on every input. It constrains quantum query advantages under the uniform distribution while allowing large advantages on inputs satisfying special promises.

Related work. For Boolean-valued functions, the conjecture follows by combining the polynomial relation between degree and decision-tree depth of Nisan and Szegedy [3] with the influence inequality of O'Donnell et al. [4]. For general bounded functions, Dinur et al. [5] and Defant et al. [6] obtained influence bounds with exponential dependence on degree. Earlier special cases include block-multilinear forms whose full coefficient array has constant absolute value [7], block-multilinear forms of completely bounded norm at most one and constant variance [8], homogeneous polynomials of Fourier completely bounded norm at most one [9], and polynomials whose Fourier coefficients have constant absolute value on each degree level [10]. O'Donnell and Zhao [11] showed that the general conjecture is equivalent to its restriction to one-block decoupled polynomials. Bhattacharya [12] obtained an influence bound for a non-negligible fraction of random restrictions, subject to a lower bound on the original variance.

Blanc et al. [13] and Liu and Mutreja [14] proved classical simulation results for quantum algorithms with a bounded number of adaptive rounds of parallel queries. Girish et al. [15] proved equivalences between the uniform-distribution conjectures and corresponding formulations for fixed-bias product distributions and fixed-density slices of the cube.

The route used here comes from fractional certificates, which assign coordinate weights that control changes in function values. Lovett and Zhang [16] developed this approach for bounded functions. Agarwal and Ben-David [17, Theorem 1.7] proved that degree-d functions with values in $[0,1]$ admit certificate rows of cost at most $\pi^2 d^2/4$. Turturean [18] proved the full influence conjecture using a determinant construction that strengthens the matching conclusion of Reimer's Butterfly Lemma [19], together with the certificate framework and integration over level sets.

This paper. We present the argument with all the combinatorial and analytic estimates proved here, taking finite linear-programming duality as standard background. We retain Turturean's determinant construction and the Chebyshev–packing certificate argument following Agarwal and Ben-David, but simplify the final passage from certificates to influence by averaging the certificate rows. We also use a larger constant in the certificate bound to keep its proof simple.

Notation. Dimensions and degree parameters are nonnegative integers, with $[0] = \varnothing$ and Q_0 consisting of a single point. For $S \subseteq [n]$, the point x^S is obtained from x by flipping the coordinates in S; in particular, $x^i = x^{\{i\}}$ and $x^\varnothing = x$. Set $\Delta(x, y) = \{i : x_i \neq y_i\}$, so that $y = x^{\Delta(x,y)}$. An *edge* of Q_n in direction i is an unordered pair $\{x, x^i\}$. Expectations on the cube use the uniform probability measure. We also write $V = \mathrm{Var}(f)$ for the variance defined above. Thus J_i uses the full squared flip difference. With the discrete derivative $D_i f(x) = (f(x) - f(x^i))/2$, the other common convention for influence is $\mathbb{E}[(D_i f(X))^2] = J_i(f)/4$. This constant factor does not affect the conjecture, but it matters when comparing explicit bounds.

The functions $x \mapsto \prod_{i \in S} x_i$, indexed by $S \subseteq [n]$, form an orthonormal basis for real functions on Q_n under uniform expectation: distinct products have inner product zero, since some coordinate occurs to an odd power, and these 2^n functions span a space of dimension 2^n. Consequently every f has the unique multilinear representation

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_i, \qquad \widehat{f}(S) = \mathbb{E}\left[f(X) \prod_{i \in S} X_i \right],$$

where the empty product is one. We say that f has *degree at most* d if $\widehat{f}(S) = 0$ whenever $|S| > d$.

Proof strategy and organization. A *certificate row* at x is a vector of nonnegative weights $a^x = (a^x_1, \ldots, a^x_n)$ such that

$$|f(x) - f(y)| \leq \sum_{i \in \Delta(x,y)} a^x_i \qquad \text{for every } y \in Q_n.$$

Its *cost* is $\sum_i a^x_i$: the changed coordinates must carry enough weight to cover the change in the function's value. The proof has two inputs:

$$\begin{aligned} \text{determinant construction} &\longrightarrow \text{Boolean all-pairs inequality,} \\ \text{low degree} &\longrightarrow \text{rows of cost at most } K = 36d^2. \end{aligned}$$

The first is proved in Section 2 and the second in Section 3.

Section 4 combines these inputs. It applies the Boolean inequality to the threshold functions $g_t(x) = \mathbf{1}_{\{f(x) > t\}}$ and integrates over $0 \leq t \leq 1$. Weights chosen from the average certificate rows and two applications of Cauchy–Schwarz give an influence of at least $4V^4/(9K^4)$. Substituting $K = 36d^2$ proves Theorem 1.

Background needed. The proof uses determinants, elementary probability, polynomial interpolation, and the duality theorem for finite linear programs. Appendix A proves the required endpoint Markov inequality; linear-programming duality is assumed. The later sections include worked examples of the determinant construction and the certificate linear program.

2 The combinatorial inequality

We first work with a Boolean function $g : Q_n \to \{0,1\}$. A *certificate row* at an input x is a vector $\lambda^x \in [0,\infty)^n$ such that

$$\sum_{i \in \Delta(x,y)} \lambda_i^x \geq 1 \quad \text{whenever } g(x) \neq g(y).$$

In words, every set of flips that changes the output has total weight at least one. The condition is vacuous when g is constant. Otherwise, the all-ones vector is a certificate row, since inputs with different outputs differ in at least one coordinate. The content of the later sections is that low degree gives rows of small total weight. We abbreviate $v(S) = \sum_{i \in S} v_i$ for any vector v and any set $S \subseteq [n]$.

The main result of this section is Lemma 3, which bounds a sum over pairs of inputs with different outputs in terms of the edges crossing the Boolean boundary. Its proof rests on a permutation lemma: a determinant will select a single permutation that meets all the required coordinate bounds at once.

Lemma 1 (Lowest-degree terms)**.** *Let q be a nonzero polynomial in finitely many variables with coefficients in a commutative ring with identity, and let u be a polynomial in the same variables with constant term $u(0) = 1$. Then $uq \neq 0$, and its terms of least total degree are exactly those of q, with the same coefficients.*

Proof. Write $q = q_r + q_{r+1} + \cdots$, where q_j is the sum of the monomials of q of total degree j and $q_r \neq 0$, and write $u = u_0 + u_1 + u_2 + \cdots$ in the same way, with $u_0 = 1$. The component of uq of total degree j is $\sum_{a+b=j} u_a q_b$. For $j < r$ every q_b occurring vanishes, and for $j = r$ the only surviving product is $u_0 q_r = q_r$. □

We call the sum of the terms of least total degree of a nonzero polynomial its *lowest homogeneous component*. For example, multiplying $s_2^2 + s_1 s_2^2$ by $1 - s_1^2$ leaves its lowest homogeneous component s_2^2 unchanged. This simple observation is the reason we can modify a determinant while preserving the information needed from it.

The proof begins with a permutation lemma. Its hypothesis describes a subcube contained in a set O: at $x \in O$, every combination of flips from A_x stays in O. Let b_i be the number of edges in direction i with exactly one endpoint in O. We seek one permutation that flips all the prescribed coordinates and, simultaneously for every coordinate i, agrees with its input at most b_i times. A determinant will encode each permutation as a monomial whose exponent of s_i is precisely its number of agreements in coordinate i. The following small example shows the encoding before the general argument.

Example 1 (Reading the agreement matrix)**.** Take $O = \{(-1,-1),(1,-1)\} \subseteq Q_2$. There is one internal edge in direction 1, joining the two points, and two boundary edges in direction 2, joining each point to its neighbour with second coordinate $+1$; so $(b_1, b_2) = (0, 2)$. Define $M[x,z] = \prod_{i : x_i = z_i} s_i$ for $x, z \in O$. In the displayed order of the two points,

$$M = \begin{pmatrix} s_1 s_2 & s_2 \\ s_2 & s_1 s_2 \end{pmatrix}, \qquad \det M = s_2^2 (s_1^2 - 1).$$

The lowest homogeneous component is $-s_2^2$. Its exponents $(0, 2)$ say that the swap permutation never agrees in coordinate 1 and agrees twice in coordinate 2. The negative term in the determinant

comes from this swap. If $A_x = \{1\}$ at both points (the hypothesis holds, since flipping coordinate 1 exchanges the two points), zeroing entries that do not make these flips gives the matrix M_A:

$$M_A = \begin{pmatrix} 0 & s_2 \\ s_2 & 0 \end{pmatrix}, \qquad \det M_A = -s_2^2.$$

Thus the lowest homogeneous component is preserved, although the full determinant changes.

Lemma 2 (Boundary-controlled permutation)**.** *Let $O \subseteq Q_n$, and let b_i count the edges in direction i with exactly one endpoint in O. Suppose that for each $x \in O$ a set $A_x \subseteq [n]$ satisfies $x^S \in O$ for every $S \subseteq A_x$. There is a permutation $\tau : O \to O$ such that*

$$\tau(x)_i \neq x_i \quad (x \in O,\ i \in A_x), \qquad \#\{x \in O : \tau(x)_i = x_i\} \leq b_i \quad (i \in [n]).$$

Proof. The case $O = \varnothing$ is immediate. Fix an ordering of O and use it for both the rows and the columns of every matrix below. All entries lie in the polynomial ring $\mathbb{Z}[s_1, \ldots, s_n]$, which is an integral domain.

Step 1: encode agreements. Define

$$M[x, z] = \prod_{i : x_i = z_i} s_i.$$

The determinant is $\det M = \sum_\tau \operatorname{sgn}(\tau) \prod_{x \in O} M[x, \tau(x)]$, the sum running over all permutations τ of O, and the term of τ equals $\pm \prod_i s_i^{a_i(\tau)}$ with

$$a_i(\tau) = \#\{x \in O : \tau(x)_i = x_i\},$$

the number of agreements of τ in coordinate i. The identity permutation is the only one with $a_i(\tau) = |O|$ for every i, so it alone contributes the monomial $(s_1 \cdots s_n)^{|O|}$, which has the largest total degree $n|O|$ among all terms. Thus $\det M \neq 0$.

Step 2: control the lowest-degree terms. Fix i, and let e_i count the edges in direction i with both endpoints in O. For each such edge, subtract s_i times the row of its endpoint with ith coordinate -1 from the row of its endpoint with ith coordinate $+1$. The modified rows (those with $x_i = +1$) and the rows used to modify them (those with $x_i = -1$) form disjoint sets, so these operations do not interfere with one another and leave the determinant unchanged. Consider a modified row x and a column z, and let $m' = \prod_{j \neq i :\, x_j = z_j} s_j$. If $z_i = +1$, then $M[x, z] = s_i m'$ and $M[x^i, z] = m'$, so the new entry is $s_i m' - s_i m' = 0$. If $z_i = -1$, then $M[x, z] = m'$ and $M[x^i, z] = s_i m'$, so the new entry is $(1 - s_i^2)m' = (1 - s_i^2)M[x, z]$. Hence each of the e_i modified rows is $(1 - s_i^2)$ times a row with entries in $\mathbb{Z}[s_1, \ldots, s_n]$, and

$$\det M = (1 - s_i^2)^{e_i} R_i,$$

where R_i is the determinant of the matrix obtained by dividing each modified row by $(1 - s_i^2)$. Since $\det M \neq 0$, we have $R_i \neq 0$. Each entry of M has degree at most one in s_i, so $\deg_{s_i} \det M \leq |O|$. Every point of O is an endpoint of exactly one edge in direction i, so $|O| = 2e_i + b_i$. Since degrees add in the integral domain $\mathbb{Z}[s_1, \ldots, s_n]$,

$$\deg_{s_i} R_i = \deg_{s_i} \det M - 2e_i \leq |O| - 2e_i = b_i.$$

The polynomial $(1 - s_i^2)^{e_i}$ has constant term one, so by Lemma 1 the lowest homogeneous component of $\det M$ equals that of R_i. Every monomial in that component thus has exponent at most b_i in s_i.

This holds for every i simultaneously, because the component itself does not depend on i. The row operations are performed separately for each fixed i; they are not a single sequence of operations over all coordinates.

Step 3: enforce the prescribed flips. Let M_A be the matrix obtained from M by zeroing the entries forbidden by the prescribed flips: $M_A[x,z] = M[x,z]$ if $x_i \neq z_i$ for all $i \in A_x$, and $M_A[x,z] = 0$ otherwise. We claim the coordinatewise factorization

$$\sum_{S \subseteq A_x} \Big(\prod_{i \in S}(-s_i)\Big) M[x^S, z] = \Big(\prod_{i \in A_x}(1 - s_i^2)\Big) M_A[x,z].$$

All rows x^S on the left belong to O, by the hypothesis on A_x. To verify the claim, note that x^S and x agree with z in the same coordinates outside A_x, so

$$M[x^S, z] = \prod_{j \notin A_x :\, x_j = z_j} s_j \cdot \prod_{i \in A_x} \sigma_i(S), \qquad \text{where } \sigma_i(S) = \begin{cases} s_i & \text{if } (x^S)_i = z_i, \\ 1 & \text{otherwise.} \end{cases}$$

Since $\sigma_i(S)$ depends only on whether $i \in S$, the sum over $S \subseteq A_x$ factorizes coordinate by coordinate: the left side equals $\prod_{j \notin A_x :\, x_j = z_j} s_j$ times $\prod_{i \in A_x} \big(\sigma_i(\varnothing) - s_i\, \sigma_i(\{i\})\big)$. For $i \in A_x$, this factor is $s_i - s_i \cdot 1 = 0$ if $x_i = z_i$, and $1 - s_i \cdot s_i = 1 - s_i^2$ if $x_i \neq z_i$. If some $i \in A_x$ has $x_i = z_i$, both sides of the claim are zero; otherwise $M_A[x,z] = M[x,z] = \prod_{j \notin A_x :\, x_j = z_j} s_j$, and the two sides agree.

Let $\Pi[x,z] = \prod_{i \in \Delta(x,z)}(-s_i)$ if $\Delta(x,z) \subseteq A_x$, and $\Pi[x,z] = 0$ otherwise; let U be the diagonal matrix with $U[x,x] = \prod_{i \in A_x}(1 - s_i^2)$. Row x of ΠM is $\sum_{S \subseteq A_x} \prod_{i \in S}(-s_i)\, M[x^S, \cdot\,]$, so the claim says precisely that $\Pi M = U M_A$. Setting all $s_i = 0$ gives $\Pi(0) = U(0) = I$, so $\det \Pi$ and $\det U$ have constant term one. Taking determinants,

$$\det \Pi \, \det M = \det U \, \det M_A.$$

The left side is nonzero, because $\det \Pi \neq 0$, $\det M \neq 0$, and the polynomial ring has no zero divisors; hence $\det M_A \neq 0$. By Lemma 1, the lowest homogeneous component on the left is that of $\det M$, and the lowest homogeneous component on the right is that of $\det M_A$. Equality therefore shows that these two components coincide. This matrix identity enforces all the prescribed flips at once, even when the sets A_x vary with x.

Step 4: extract a permutation. Choose a monomial with nonzero coefficient in this component of $\det M_A$. Since $\det M_A = \sum_\tau \operatorname{sgn}(\tau) \prod_x M_A[x, \tau(x)]$ and the coefficient of the monomial is nonzero, at least one permutation τ contributes a nonzero term equal to $\pm$ this monomial (terms of different permutations may cancel, but not all of them). All entries $M_A[x, \tau(x)]$ are nonzero, so $\tau(x)_i \neq x_i$ for every $i \in A_x$. Its term is $\pm \prod_i s_i^{a_i(\tau)}$, so the exponent of s_i in the monomial is $a_i(\tau)$, and Step 2 bounds it by b_i. □

The next estimate translates this permutation into an inequality about certificates. On the left, a pair of opposite-valued inputs contributes the reciprocal of the weighted overlap between their rows on the coordinates where they differ. On the right, b_i counts boundary edges in direction i. The positive weights w_i allow the coordinates to be reweighted; they will be chosen from the average rows in Section 4.

Lemma 3 (All-pairs inequality)**.** *Let $g : Q_n \to \{0,1\}$ have a certificate row λ^x at every input x. Set $O = g^{-1}(1)$, write $O^c = Q_n \setminus O$, and let b_i count the boundary edges of O in direction i, that is, the edges in direction i with exactly one endpoint in O. For all positive reals $w_1, \dots, w_n$,*

$$\sum_{x \in O,\, y \notin O} \frac{1}{\sum_{i \in \Delta(x,y)} w_i \lambda_i^x \lambda_i^y} \leq 2^{n-1} \sum_i \frac{b_i}{w_i}, \tag{1}$$

where the sum runs over ordered pairs $(x, y) \in O \times O^c$, and every denominator on the left is positive.

Proof. If g is constant, the pair sum is empty and every b_i is zero, so the conclusion holds. This also handles $n = 0$; hence assume g is nonconstant and $n \geq 1$.

Step 1: certificates at opposite-valued inputs overlap. For $g(x) \neq g(y)$, let $S = \{i \in \Delta(x, y) : \lambda_i^x \leq \lambda_i^y\}$ and $z = x^S$. If $g(z) \neq g(x)$, the row at x gives $\lambda^x(S) \geq 1$, since $\Delta(x, z) = S$. Otherwise $g(z) = g(x) \neq g(y)$, and the row at y gives $\lambda^y(\Delta(x, y) \setminus S) \geq 1$, since $\Delta(z, y) = \Delta(x, y) \setminus S$. On S the smaller of λ_i^x, λ_i^y is λ_i^x, and on $\Delta(x, y) \setminus S$ it is λ_i^y; in either case

$$\sum_{i \in \Delta(x,y)} \min(\lambda_i^x, \lambda_i^y) \geq 1. \tag{2}$$

In particular, some product $\lambda_i^x \lambda_i^y$ with $i \in \Delta(x, y)$ is positive, which proves the claim about the denominators.

Step 2: prove the estimate for antipodal pairs. Call $-u$ the *antipode* of u; then $\Delta(u, -u) = [n]$. Write $Z = \{u \in O : -u \notin O\}$ and $\beta_i = 1/w_i$. For $u \in Z$, abbreviate $\lambda = \lambda^u$, $\mu = \lambda^{-u}$ and put

$$\omega_u = \Big(\sum_i \lambda_i \mu_i / \beta_i\Big)^{-1} = \Big(\sum_i w_i \lambda_i \mu_i\Big)^{-1}, \qquad A_u = \{i : \beta_i < \omega_u \mu_i\}.$$

The sum is positive by Step 1 applied to the pair $(u, -u)$, so ω_u is well defined and positive; it is the term of the pair $(u, -u)$ on the left of (1). The set A_u collects the coordinates where β_i is small compared with μ_i; we shall force the permutation to flip them. Since $\sum_i \omega_u \lambda_i \mu_i / \beta_i = 1$ and $\omega_u \mu_i / \beta_i > 1$ for $i \in A_u$,

$$\lambda(A_u) = \sum_{i \in A_u} \lambda_i \leq \sum_{i \in A_u} \omega_u \lambda_i \mu_i / \beta_i \leq 1,$$

and the first inequality is strict as soon as some λ_i with $i \in A_u$ is positive; if instead $\lambda_i = 0$ for all $i \in A_u$, then $\lambda(A_u) = 0$. In all cases $\lambda(A_u) < 1$. Consequently $\lambda^u(S) < 1$ for every $S \subseteq A_u$, so the certificate row at u forces $g(u^S) = g(u) = 1$: every subset of A_u can be flipped at u without leaving O.

Apply Lemma 2 to O with these sets A_u for $u \in Z$ and with $A_x = \varnothing$ for $x \in O \setminus Z$; the hypothesis of the lemma has just been verified. Let τ be the resulting permutation. For $u \in Z$, let $G_u = \{i : u_i = \tau(u)_i\}$ be the set of coordinates where $\tau(u)$ agrees with u. Then $G_u \cap A_u = \varnothing$, because τ flips every coordinate of A_u; and $(-u)^{G_u} = \tau(u)$, because $-u$ differs from u in every coordinate while $\tau(u)$ differs from u exactly outside G_u. Since $\tau(u) \in O$ and $-u \notin O$, the certificate row at $-u$ gives $\mu(G_u) \geq 1$. On G_u we have $\beta_i \geq \omega_u \mu_i$, hence

$$\sum_{i \in G_u} \beta_i \geq \omega_u \, \mu(G_u) \geq \omega_u.$$

Summing over $u \in Z$ and using the agreement bounds of Lemma 2,

$$\sum_{u \in Z} \omega_u \leq \sum_{u \in Z} \sum_{i \in G_u} \beta_i = \sum_i \beta_i \, \#\{u \in Z : \tau(u)_i = u_i\} \leq \sum_i \beta_i b_i,$$

that is,

$$\sum_{u \in Z} \frac{1}{\sum_i w_i \lambda_i^u \lambda_i^{-u}} \leq \sum_i \frac{b_i}{w_i}.$$

This is (1) restricted to antipodal pairs, without the factor 2^{n-1}.

Step 3: pass from antipodal pairs to all pairs. For example, in Q_2 the points $(-1,-1)$ and $(1,-1)$ are antipodal on the edge with second coordinate fixed at -1. They are not antipodal in the whole square. Their differing coordinates determine the unique face on which they are antipodal; the next argument uses this fact for every pair.

For a nonempty set $T \subseteq [n]$ of *free* coordinates and an assignment $v \in \{-1,1\}^{[n]\setminus T}$ to the remaining coordinates, define the *face*

$$F(T,v) = \{x \in Q_n : x_j = v_j \text{ for every } j \notin T\}.$$

Each face occurs for exactly one pair (T,v) and is a copy of $\{-1,1\}^T$, or $Q_{|T|}$ after relabeling the coordinates. Restrict g to $F = F(T,v)$, and restrict each row λ^x with $x \in F$ to the coordinates in T. The restricted rows are certificate rows of $g|_F$, because two inputs in F differ only in coordinates of T. Two points of F are antipodal in F when they differ in every coordinate of T. Applying Step 2 inside F gives

$$\sum_{\substack{x,y\in F:\, g(x)=1,\, g(y)=0 \\ \Delta(x,y)=T}} \frac{1}{\sum_{i\in T} w_i \lambda_i^x \lambda_i^y} \le \sum_{i\in T} \frac{b_i(F)}{w_i},$$

where $b_i(F)$ counts the edges in direction i inside F with exactly one endpoint in O. If $g|_F$ is constant, both sides are zero and the same inequality holds directly. For every pair in the sum, $\Delta(x,y) = T$, so its summand is exactly the term of (x,y) in (1).

Now sum once over every face $F(T,v)$ of positive dimension. Every ordered pair $(x,y) \in O \times O^c$ is antipodal in exactly one such face: its free coordinates must be $T = \Delta(x,y)$, and its fixed assignment must be $v = (x_j)_{j\notin T} = (y_j)_{j\notin T}$. Thus the left sides add up to the left side of (1). A boundary edge $\{x, x^i\}$ of O lies in exactly those faces whose free set T contains i and whose fixed coordinates agree with x outside T. The set $T\setminus\{i\}$ can be any subset of the other $n-1$ coordinates, and the face is then determined, so the edge lies in exactly 2^{n-1} faces, and it is a boundary edge of $O \cap F$ in each of them. Conversely, every edge counted by $b_i(F)$ is a boundary edge of O. Hence $\sum_{F:\, i\in T(F)} b_i(F) = 2^{n-1} b_i$, where $T(F)$ denotes the free coordinates of F. The right sides therefore add up to $2^{n-1}\sum_i b_i/w_i$, proving (1). □

For $n = 1$, the function $g(x) = (1+x_1)/2$ with both rows equal to (1) gives equality. There is one ordered pair in $O \times O^c$, its denominator is w_1, and $b_1 = 1$, so both sides of (1) equal $1/w_1$.

3 Low degree gives inexpensive certificates

For a real-valued $f : Q_n \to \mathbb{R}$, a certificate row at x is a vector $a^x \in [0,\infty)^n$ satisfying

$$|f(x) - f(y)| \le \sum_{i\in\Delta(x,y)} a_i^x \qquad \text{for every } y \in Q_n.$$

Its *cost* is $\sum_i a_i^x$. For a Boolean function this is the definition of Section 2: opposite outputs differ by one, while equal outputs impose no additional condition. A row controls the change from x under an arbitrary set of flips, not just a single flip. Our aim is to bound its cost using only the degree, independently of the dimension.

Example 2 (A row and its dual packing)**.** Let $f(x_1,x_2) = (2+x_1+x_2)/4$. At $x = (1,1)$, the changes caused by flipping $\{1\}$, $\{2\}$, and $\{1,2\}$ are respectively $1/2$, $1/2$, and 1. A row must therefore satisfy

$$a_1 \ge \tfrac{1}{2}, \qquad a_2 \ge \tfrac{1}{2}, \qquad a_1 + a_2 \ge 1.$$

The two singleton constraints already force cost at least one, and $(a_1, a_2) = (1/2, 1/2)$ satisfies all three constraints, so the least cost is exactly one. In the dual packing problem below, this lower bound is certified by giving weight one to each singleton block and zero to the other blocks: the weighted total demand is $1/2 + 1/2 = 1$, and each coordinate is used exactly once. This is the covering–packing relationship used below.

Lemma 4. *If $f : Q_n \to [-1, 1]$ has degree at most d, it has a certificate row of cost at most $36d^2$ at every input.*

Proof. If f is constant, the zero vector is a row of cost zero at every input. This includes $n = 0$ and $d = 0$, so assume henceforth that f is nonconstant and $n, d \geq 1$.

The proof first bounds the sum of single-coordinate changes of a bounded polynomial. Linear-programming duality then turns the row problem into a bound on weighted block changes. Finally, a Chebyshev composition converts those block changes into single-coordinate changes of another bounded polynomial.

Step 1: bound the sum of single-coordinate changes. Let $H : \{-1, 1\}^k \to [-1, 1]$ have degree at most D, and let $\widetilde{H}$ be its multilinear extension, the unique multilinear polynomial agreeing with H on the cube. Explicitly, for $u \in [-1, 1]^k$,

$$\widetilde{H}(u) = \sum_{z \in \{-1,1\}^k} H(z) \prod_{j=1}^{k} \frac{1 + z_j u_j}{2}.$$

The weights are nonnegative and sum to

$$\prod_{j=1}^{k} \left(\frac{1 + u_j}{2} + \frac{1 - u_j}{2} \right) = 1,$$

so this is a convex combination of vertex values and $|\widetilde{H}| \leq 1$ on $[-1, 1]^k$. The displayed polynomial is multilinear and agrees with H at every vertex, so by uniqueness it is the polynomial in the Fourier representation of H; hence it has total degree at most D. Since $\widetilde{H}$ is affine in each variable,

$$\frac{\partial \widetilde{H}}{\partial u_b}(u) = \frac{\widetilde{H}(u^{(b,+)}) - \widetilde{H}(u^{(b,-)})}{2},$$

where $u^{(b,\pm)}$ denotes u with its bth coordinate replaced by ± 1; at $u = \mathbf{1}$ this equals $(H(\mathbf{1}) - H(\mathbf{1}^b))/2$.

The endpoint Markov inequality (proved in Appendix A) says that a polynomial q of degree at most D with $|q| \leq 1$ on $[0, 1]$ satisfies $|q'(0)| \leq 2D^2$. For $R \subseteq [k]$, let $u(t)$ be the point whose coordinates in R equal $1 - 2t$ and whose other coordinates equal one, and put $p(t) = \widetilde{H}(u(t))$. Then p is a polynomial of degree at most D, with $|p| \leq 1$ on $[0, 1]$ because $u(t) \in [-1, 1]^k$ there, and by the chain rule

$$p'(0) = \sum_{b \in R} (-2) \frac{\partial \widetilde{H}}{\partial u_b}(\mathbf{1}) = -\sum_{b \in R} \big(H(\mathbf{1}) - H(\mathbf{1}^b) \big).$$

For the set R where $H(\mathbf{1}) - H(\mathbf{1}^b) > 0$, the absolute value of this derivative is the sum of the positive changes, so that sum is at most $2D^2$. Taking instead the set where the changes are negative bounds the sum of their absolute values by $2D^2$. Adding gives

$$\sum_{b=1}^{k} |H(\mathbf{1}) - H(\mathbf{1}^b)| \leq 4D^2. \tag{3}$$

If $k = 0$ or $D = 0$, the left side is zero, so the same estimate holds.

Step 2: express the row problem as a packing problem. Fix x, and set $t_S = |f(x) - f(x^S)|$ for $S \subseteq [n]$; thus $0 \le t_S \le 2$ and $t_\varnothing = 0$. Since every $y \in Q_n$ is x^S for exactly one S, a row at x is exactly a vector $a \ge 0$ with $\sum_{i \in S} a_i \ge t_S$ for every S. The empty set imposes only $0 \ge 0$, so omit it. In the rest of this proof, block sums range over nonempty $S \subseteq [n]$. The covering problem and its dual packing problem are

$$C_x = \min \left\{ \sum_i a_i : a_i \ge 0,\ \sum_{i \in S} a_i \ge t_S \text{ for every nonempty } S \right\},$$

$$P_x = \max \left\{ \sum_S \kappa_S t_S : \kappa_S \ge 0,\ \sum_{S \ni i} \kappa_S \le 1 \text{ for every } i \right\}.$$

Here κ_S is the weight assigned to block S, and each coordinate has total usage at most one. The covering problem is feasible, since $a_i = 2$ works, and its objective is bounded below by zero. Its minimum is attained: we may restrict to the nonempty compact set of feasible vectors with $\sum_i a_i \le 2n$. The packing problem is feasible at $\kappa = 0$, and its feasible set is compact, since $0 \le \kappa_S \le 1$ for every nonempty S. Thus both optima exist, and finite linear-programming duality gives $C_x = P_x$.

For intuition, any feasible row and packing satisfy

$$\sum_S \kappa_S t_S \le \sum_S \kappa_S \sum_{i \in S} a_i = \sum_i a_i \sum_{S \ni i} \kappa_S \le \sum_i a_i.$$

Duality says that their optimal values agree. It therefore suffices to bound the packing objective by $36d^2$ for every feasible κ.

First take a rational feasible κ and a common positive integer denominator N of its entries. The zero packing has objective zero, so assume some $\kappa_S > 0$. Make a list $B_1, \dots, B_k$ of nonempty sets containing $N\kappa_S$ copies of each S; repetitions are allowed. Then $k \ge 1$, and each coordinate i occurs in exactly $N \sum_{S \ni i} \kappa_S \le N$ listed blocks. For instance, the packing on two coordinates with $\kappa_{\{1\}} = \kappa_{\{2\}} = 1/2$, all other weights zero, and denominator $N = 2$ gives the list $\{1\}, \{2\}$. This conversion replaces fractional weights by an ordinary list of blocks with bounded overlap.

Step 3: simulate block flips by single flips. Set $m = \lceil 2\sqrt{N} \rceil$. We seek a function G on the k block-selection variables with three properties: (i) $|G| \le 1$ on the cube; (ii) its degree is at most dm; and (iii) $G(\mathbf{1}) = f(x)$ and $G(\mathbf{1}^b) = f(x^{B_b})$ for every block B_b. The single-coordinate bound (3) will then control all the listed block changes at once.

For each original coordinate, we need a bounded polynomial that equals $+1$ when no containing block is selected and -1 when exactly one containing block is selected. With several blocks selected, only boundedness matters; we do not require the polynomial to reproduce successive block flips. Chebyshev polynomials provide such a construction with degree of order $\sqrt{N}$. Squaring this degree in (3) will then cancel the packing denominator N.

The Chebyshev polynomials are defined by $T_0(z) = 1$, $T_1(z) = z$, and $T_{m+1}(z) = 2zT_m(z) - T_{m-1}(z)$. The recurrence gives $\deg T_m = m$, and the cosine addition formula gives $T_m(\cos\theta) = \cos(m\theta)$, hence $|T_m(z)| \le 1$ on $[-1, 1]$. Choose

$$c = 1 - \cos(\pi/m), \qquad r_i(w) = \sum_{b : i \in B_b} \frac{1 - w_b}{2} \quad (i \in [n]),$$

where $w \in \{-1,1\}^k$, and call block B_b selected when $w_b = -1$. Thus $r_i(w)$ counts the selected blocks containing i. Let $\widetilde{f}$ be the multilinear extension of f and define, for $w \in \{-1,1\}^k$,

$$G(w) = \widetilde{f}\big(x_1 T_m(1 - c\, r_1(w)), \dots, x_n T_m(1 - c\, r_n(w))\big).$$

Three properties make this construction useful.

- G takes values in $[-1,1]$. Indeed $0 \le r_i(w) \le N$ and

  $$Nc \le \frac{N\pi^2}{2m^2} \le \frac{\pi^2}{8} < 2,$$

  using $1 - \cos u = 2\sin^2(u/2) \le u^2/2$ and $m^2 \ge 4N$. Since $c > 0$, each $1 - c\, r_i(w)$ lies in $[-1,1]$. Thus each $x_i T_m(1 - c\, r_i(w))$ lies in $[-1,1]$, and $|\widetilde{f}| \le 1$ on $[-1,1]^n$ as in Step 1.

- G has degree at most dm as a function on $\{-1,1\}^k$. Each $r_i(w)$ is affine in w, so each substituted coordinate has degree at most m, and $\widetilde{f}$ has total degree at most d. The displayed composition need not be multilinear in w. On the cube, however, replacing each power w_b^q by 1 for even q and by w_b for odd q preserves its values and cannot increase total degree. The resulting multilinear representation therefore has degree at most dm.

- $G(\mathbf{1}) = f(x)$ and $G(\mathbf{1}^b) = f(x^{B_b})$. At $w = \mathbf{1}$ every r_i is zero and $T_m(1) = 1$. At $w = \mathbf{1}^b$ we have $r_i = 1$ for $i \in B_b$ and $r_i = 0$ otherwise, while $T_m(1 - c) = T_m(\cos(\pi/m)) = \cos \pi = -1$, so the argument of $\widetilde{f}$ is x^{B_b}.

Applying (3) to $H = G$ with $D = dm$, and using $m \le 2\sqrt{N} + 1 \le 3\sqrt{N}$ (as $N \ge 1$), yields

$$N \sum_S \kappa_S t_S = \sum_{b=1}^{k} \big|f(x) - f(x^{B_b})\big| = \sum_{b=1}^{k} |G(\mathbf{1}) - G(\mathbf{1}^b)| \le 4d^2 m^2 \le 36 d^2 N.$$

Dividing by N bounds the objective of every rational feasible packing by $36d^2$. For an arbitrary real feasible packing, take integers $q \ge 1$ and set $\kappa_S^{(q)} = \lfloor q\kappa_S \rfloor / q$. These rational packings remain feasible because $0 \le \kappa_S^{(q)} \le \kappa_S$ for every S, and they converge coordinatewise to κ. There are finitely many blocks, so their objective values converge to $\sum_S \kappa_S t_S$. The same bound therefore holds for every feasible packing, hence for $P_x = C_x$. A minimizing row at the arbitrary input x has cost at most $36d^2$, proving the lemma. □

4 From certificates to influence

We now combine the Boolean all-pairs inequality with two properties of f: its values lie in $[0,1]$, and its certificate rows have uniformly bounded cost. The degree enters only through the row bound already proved. To see why thresholds are useful, note that for $a, b \in [0,1]$,

$$|a - b| = \int_0^1 \big|\mathbf{1}_{\{a > t\}} - \mathbf{1}_{\{b > t\}}\big|\, dt.$$

Apart from the endpoints, the integrand is one precisely for thresholds lying between a and b. Thus integrating Boolean boundary information can recover differences of real values.

The argument below follows that idea in four steps: scale each real certificate row to obtain a row for the threshold function; apply the Boolean all-pairs inequality at each nonexceptional threshold; integrate over the threshold; and choose weights using the average rows. Cauchy–Schwarz then turns the resulting inequality into a variance and influence bound.

Lemma 5. *Let $f : Q_n \to [0,1]$ have variance $V > 0$, and suppose that at every input x there is a certificate row a^x of cost at most K, where $K > 0$. Then*

$$\max_i J_i(f) \geq \frac{4V^4}{9K^4}.$$

Proof. Since $V > 0$, the function is nonconstant and $n \geq 1$. Fix one such row a^x for each input x, and let X, Y be independent uniform inputs. Define

$$L_i = \mathbb{E}|f(X) - f(X^i)|, \qquad L^* = \max_i L_i.$$

We will prove $L^* \geq 2V^2/(3K^2)$; the influence bound will then follow from $J_i(f) \geq L_i^2$, which is the Cauchy–Schwarz inequality $\mathbb{E}W^2 \geq (\mathbb{E}|W|)^2$.

Step 1: apply the Boolean inequality at every level. Fix positive numbers $w_1, \ldots, w_n$; they do not depend on the threshold and will be chosen in Step 2. Set

$$s_w(x,y) = \sum_{i \in \Delta(x,y)} w_i a_i^x a_i^y.$$

Call a threshold $t \in [0,1]$ *exceptional* if it is a value of f; there are finitely many such thresholds. For a nonexceptional t, define

$$g_t(x) = \mathbf{1}_{\{f(x) > t\}}, \qquad \lambda_i^x(t) = \frac{a_i^x}{|f(x) - t|}.$$

Every denominator here is positive. If $g_t(x) \neq g_t(y)$, then t lies strictly between $f(x)$ and $f(y)$, so the original certificate inequality gives

$$\sum_{i \in \Delta(x,y)} \lambda_i^x(t) \geq \frac{|f(x) - f(y)|}{|f(x) - t|} \geq 1.$$

Thus $\lambda^x(t)$ is a certificate row of g_t at x. For example, if $t = 1/2$ is nonexceptional and $f(x) = 3/4$, the scaled row is $4a^x$: crossing the threshold from x changes f by at least $1/4$.

We also need positivity of s_w whenever the endpoint values differ. If $f(x) \neq f(y)$, their open interval contains a nonexceptional threshold, since only finitely many thresholds are excluded. Apply (2) to the scaled rows at that threshold. Some $i \in \Delta(x,y)$ has $\lambda_i^x(t)\lambda_i^y(t) > 0$, hence $a_i^x a_i^y > 0$ and $s_w(x,y) > 0$. In particular, $s_w(x,y) = 0$ forces $f(x) = f(y)$. For quotients with denominator s_w, we henceforth assign value zero whenever the numerator is zero, including the case $0/0$.

For any Boolean g, $b_i(g) = 2^{n-1} \Pr[g(X) \neq g(X^i)]$, because each boundary edge in direction i gives exactly two inputs x with $g(x) \neq g(x^i)$. Apply (1) to g_t with the scaled rows: the term of an ordered pair (x,y) with $f(x) > t > f(y)$ is $(f(x) - t)(t - f(y))/s_w(x,y)$. Dividing the pair sum by $|Q_n|^2 = 4^n$ turns it into an expectation over (X,Y). On the right, the normalization is $2^{n-1}2^{n-1}/4^n = 1/4$, so

$$\mathbb{E}\frac{\mathbf{1}_{\{f(X) > t > f(Y)\}}(f(X) - t)(t - f(Y))}{s_w(X,Y)} \leq \frac{1}{4}\sum_i \frac{\Pr[g_t(X) \neq g_t(X^i)]}{w_i},$$

with the zero-quotient convention just stated. The denominator is positive whenever the numerator is positive. Integrate over $0 \leq t \leq 1$. The finitely many exceptional thresholds do not affect the

integrals, and the expectations are finite sums, so integration and expectation may be exchanged. For $f(x) > f(y)$, the indicator is one exactly for $t \in (f(y), f(x))$, and

$$\int_{f(y)}^{f(x)} (f(x) - t)(t - f(y))\, dt = \frac{(f(x) - f(y))^3}{6},$$

as the substitution $t = f(y) + r(f(x) - f(y))$ and $\int_0^1 r(1-r)\, dr = 1/6$ show. The cube of the difference has a useful interpretation: two factors come from scaling the endpoint certificates, and the third comes from integrating over the interval between the endpoint values. Also $\int_0^1 \Pr[g_t(X) \neq g_t(X^i)]\, dt = \mathbb{E} \int_0^1 |\mathbf{1}_{\{f(X)>t\}} - \mathbf{1}_{\{f(X^i)>t\}}|\, dt = L_i$, by the identity displayed before the lemma. Hence, after multiplying by 6,

$$\mathbb{E} \frac{(f(X) - f(Y))_+^3}{s_w(X, Y)} \leq \frac{3}{2} \sum_i \frac{L_i}{w_i}, \tag{4}$$

where $r_+ = \max\{r, 0\}$ and, as before, the quotient is zero when its numerator is zero.

Step 2: choose weights from the average certificates. Put $\bar{a}_i = \mathbb{E} a_i^X$ and $\bar{a} = \sum_i \bar{a}_i = \mathbb{E} \sum_i a_i^X \leq K$. If $\bar{a} = 0$, then every row is zero and the certificate inequality makes f constant, contradicting $V > 0$; thus $\bar{a} > 0$. If $\bar{a}_i = 0$, then $a_i^x = 0$ at every input, since the entries are nonnegative and every input has positive uniform probability. The certificate inequality with $y = x^i$ then gives $|f(x) - f(x^i)| \leq a_i^x = 0$, so $L_i = 0$.

Choose

$$w_i = \begin{cases} \bar{a}/\bar{a}_i, & \bar{a}_i > 0, \\ 1, & \bar{a}_i = 0. \end{cases}$$

All these weights are finite and positive. They give

$$\sum_i \frac{L_i}{w_i} = \sum_{i:\bar{a}_i>0} \frac{\bar{a}_i}{\bar{a}} L_i \leq L^* \sum_{i:\bar{a}_i>0} \frac{\bar{a}_i}{\bar{a}} = L^*,$$
$$\mathbb{E} s_w(X, Y) \leq \sum_i w_i \mathbb{E}(a_i^X a_i^Y) = \sum_i w_i \bar{a}_i^2 = \sum_{i:\bar{a}_i>0} \bar{a}\, \bar{a}_i = \bar{a}^2 \leq K^2.$$

The first line uses $L_i = 0$ when $\bar{a}_i = 0$. The second line uses that s_w omits the coordinates outside $\Delta(X, Y)$, that all terms are nonnegative, and that X and Y are independent. Thus (4) controls a quotient, while the mean of its denominator is also controlled.

Step 3: recover variance, then influence. Set $\xi = (f(X) - f(Y))_+$, so $0 \leq \xi \leq 1$, and abbreviate $s_w = s_w(X, Y)$. Exchanging X and Y preserves their joint distribution, so the positive and negative parts of $f(X) - f(Y)$ have the same second moment. Independence gives

$$\mathbb{E}(f(X) - f(Y))^2 = 2\mathbb{E}[f(X)^2] - 2(\mathbb{E} f(X))^2 = 2V.$$

Consequently $\mathbb{E}\xi^2 = V$.

On the event $s_w = 0$, Step 1 gives $f(X) = f(Y)$ and hence $\xi = 0$. Define the random variable

$$U = \begin{cases} \xi^2/\sqrt{s_w}, & s_w > 0, \\ 0, & s_w = 0. \end{cases}$$

Then $\xi^2 = U\sqrt{s_w}$ everywhere, and $U^2 = \xi^4/s_w$ under our convention. All moments are finite because (X, Y) ranges over a finite set. Cauchy–Schwarz, the inequality $\xi^4 \leq \xi^3$ on $[0, 1]$, (4), and

Step 2 now give

$$V^2 = (\mathbb{E}[U\sqrt{s_w}])^2 \le \mathbb{E}U^2\,\mathbb{E}s_w$$
$$= \mathbb{E}\frac{\xi^4}{s_w}\,\mathbb{E}s_w \le \mathbb{E}\frac{\xi^3}{s_w}\,\mathbb{E}s_w \le \tfrac{3}{2}L^*K^2.$$

Dividing by $3K^2/2 > 0$ yields $L^* \ge 2V^2/(3K^2)$. Choose a coordinate i attaining L^* and apply Cauchy–Schwarz once more:

$$J_i(f) = \mathbb{E}(f(X) - f(X^i))^2 \ge L_i^2 = (L^*)^2 \ge \frac{4V^4}{9K^4}.$$ □

Proof of Theorem 1. Lemma 4 applies to f, since $[0,1] \subset [-1,1]$, and gives a certificate row of cost at most $K = 36d^2$ at every input. Since $d \ge 1$, this K is positive. Lemma 5 therefore gives

$$\max_i J_i(f) \ge \frac{4V^4}{9(36d^2)^4} = \frac{4V^4}{9 \cdot 36^4 d^8},$$

which is Theorem 1. □

A The endpoint Markov inequality

This appendix proves the one-variable estimate used in Lemma 4. The idea is to express an endpoint derivative as a weighted sum of polynomial values. At the Chebyshev points, the signs of the weights match the alternating values of a Chebyshev polynomial, allowing us to compute the sum of their absolute values exactly.

Lemma 6 (Endpoint Markov inequality)**.** *Let $D \ge 0$ be an integer. If a real polynomial p has degree at most D and satisfies $|p(z)| \le 1$ for every $z \in [-1,1]$, then $|p'(1)| \le D^2$. Consequently, if a real polynomial q has degree at most D and satisfies $|q(t)| \le 1$ for every $t \in [0,1]$, then $|q'(0)| \le 2D^2$.*

Proof. For $D = 0$ both polynomials are constant and both derivatives vanish. Assume $D \ge 1$.

Step 1: interpolate at the Chebyshev points. Set $\eta_j = \cos(j\pi/D)$ for $0 \le j \le D$. Since cosine is strictly decreasing on $[0,\pi]$, these points satisfy $1 = \eta_0 > \eta_1 > \cdots > \eta_D = -1$. Define the Lagrange basis polynomials

$$\ell_j(z) = \prod_{\substack{0 \le k \le D \\ k \ne j}} \frac{z - \eta_k}{\eta_j - \eta_k}.$$

Every denominator is nonzero, and $\ell_j(\eta_k)$ is one for $k = j$ and zero otherwise. Thus

$$p(z) = \sum_{j=0}^{D} p(\eta_j)\ell_j(z);$$

the identity holds because both sides have degree at most D and agree at the $D+1$ distinct nodes.

Step 2: determine the signs of the derivative weights. For $j \ge 1$, the product defining ℓ_j contains the factor $(z-1)/(\eta_j - 1)$. Only the term differentiating this factor survives at $z = 1$, giving

$$\ell_j'(1) = \frac{1}{\eta_j - 1} \prod_{\substack{1 \le k \le D \\ k \ne j}} \frac{1 - \eta_k}{\eta_j - \eta_k}.$$

All numerator factors are positive. Counting also $\eta_j - 1$, exactly j denominator factors are negative, namely those corresponding to $k = 0, \ldots, j-1$. Hence $\operatorname{sgn} \ell_j'(1) = (-1)^j$. For $j = 0$, every factor of ℓ_0 equals one at $z = 1$, so the product rule gives

$$\ell_0'(1) = \sum_{k=1}^{D} \frac{1}{1 - \eta_k} > 0.$$

Thus $|\ell_j'(1)| = (-1)^j \ell_j'(1)$ for every $0 \le j \le D$.

Step 3: sum the absolute weights. Recall $T_0(z) = 1$, $T_1(z) = z$, and $T_{m+1}(z) = 2zT_m(z) - T_{m-1}(z)$ for $m \ge 1$. The recurrence gives $\deg T_m = m$, and the cosine addition formula gives $T_m(\cos\theta) = \cos(m\theta)$. In particular, $T_D(\eta_j) = (-1)^j$, so uniqueness of interpolation gives

$$T_D(z) = \sum_{j=0}^{D} (-1)^j \ell_j(z).$$

Also $T_m(1) = 1$ for every m, and $T_0'(1) = 0$, $T_1'(1) = 1$. Differentiating the recurrence and evaluating at one proves inductively that

$$T_{m+1}'(1) = 2 + 2m^2 - (m-1)^2 = (m+1)^2 \qquad (m \ge 1).$$

Therefore $T_D'(1) = D^2$. Differentiating the interpolation identities and using $|p(\eta_j)| \le 1$ now yields

$$|p'(1)| = \left| \sum_{j=0}^{D} p(\eta_j) \ell_j'(1) \right| \le \sum_{j=0}^{D} |\ell_j'(1)| = \sum_{j=0}^{D} (-1)^j \ell_j'(1) = T_D'(1) = D^2.$$

Finally, for q on $[0,1]$, set $p(z) = q((1-z)/2)$. This affine substitution preserves the degree bound and gives $|p| \le 1$ on $[-1,1]$. Since $p'(1) = -q'(0)/2$, the first bound gives

$$|q'(0)| = 2|p'(1)| \le 2D^2. \qquad \square$$

Use of AI: The author prompted Claude and ChatGPT interactively to simplify the proofs, focusing only on proving the Aaronson–Ambainis influence conjecture without optimizing the bounds. The author expanded some proofs manually and made them more human-readable and added explanations and worked examples. The author's main intent is to make the proof more accessible to researchers, teachers and students with basic knowledge of linear algebra, linear programming, probability and quantum computing.

References

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

[2] Scott Aaronson and Andris Ambainis, *The need for structure in quantum speedups*, Theory of Computing 10 (2014), 133–166. https://theoryofcomputing.org/articles/v010a006/

[3] Noam Nisan and Mario Szegedy, *On the degree of Boolean functions as real polynomials*, Computational Complexity 4 (1994), 301–313.

[4] Ryan O'Donnell, Michael Saks, Oded Schramm, and Rocco A. Servedio, *Every decision tree has an influential variable*, FOCS 2005, 31–39.

[5] Irit Dinur, Ehud Friedgut, Guy Kindler, and Ryan O'Donnell, *On the Fourier tails of bounded functions over the discrete cube*, Israel J. Math. 160 (2007), 389–412.

[6] Andreas Defant, Mieczysław Mastyło, and Antonio Pérez, *On the Fourier spectrum of functions on Boolean cubes*, Math. Ann. 374 (2019), 653–680.

[7] Ashley Montanaro, *Some applications of hypercontractive inequalities in quantum information theory*, J. Math. Phys. 53 (2012), 122206.

[8] Nikhil Bansal, Makrand Sinha, and Ronald de Wolf, *Influence in completely bounded block-multilinear forms and classical simulation of quantum algorithms*, CCC 2022, LIPIcs 234, Article 28. https://arxiv.org/abs/2203.00212

[9] Francisco Escudero Gutiérrez, *Influences of Fourier completely bounded polynomials and classical simulation of quantum algorithms*, Chicago J. Theoret. Comput. Sci. (2024), Article 2.

[10] Paata Ivanisvili, *Polynomial growth of Bohnenblust–Hille constants on the Hamming cube*, arXiv:2609.12427 (2026).

[11] Ryan O'Donnell and Yu Zhao, *Polynomial bounds for decoupling, with applications*, CCC 2016, LIPIcs 50, Article 24. https://arxiv.org/abs/1512.01603

[12] Sreejata Kishor Bhattacharya, *Aaronson–Ambainis conjecture is true for random restrictions*, arXiv:2402.13952, revised 2026; conference version: *Random restrictions of bounded low degree polynomials are juntas*, ITCS 2025, LIPIcs 325, Article 17.

[13] Guy Blanc, Jordan Docter, Carmen Strassle, and Li-Yang Tan, *Quantum speedups require structure or depth*, arXiv:2608.19158 (2026).

[14] Qipeng Liu and Saachi Mutreja, *Parallel quantum advantage with limited adaptivity requires structure*, arXiv:2608.20297 (2026).

[15] Uma Girish, Kunal Mittal, Barak Nehoran, and Ran Raz, *Distributional variants of the Aaronson–Ambainis conjecture*, arXiv:2609.35327 (2026).

[16] Shachar Lovett and Jiapeng Zhang, *Fractional certificates for bounded functions*, ITCS 2023, LIPIcs 251, Article 84. https://doi.org/10.4230/LIPIcs.ITCS.2023.84

[17] Avantika Agarwal and Shalev Ben-David, *Oracle separations for the quantum-classical polynomial hierarchy*, arXiv:2410.19062. https://arxiv.org/abs/2410.19062

[18] David Turturean, *A proof of the Aaronson–Ambainis conjecture*, Physical Superintelligence, November 2026. https://github.com/physical-superintelligence/aaronson-ambainis

[19] David Reimer, *Proof of the van den Berg–Kesten conjecture*, Combinatorics, Probability and Computing 9 (2000), 27–32.