Proof of the Aaronson–Ambainis Conjecture: An Exposition

Guy Blanc *MIT*

Jordan Docter *Stanford*

Carmen Strassle *Stanford*

Li-Yang Tan *Stanford*

November 7, 2026

Abstract

We give an exposition of a recent proof of the Aaronson–Ambainis conjecture, which states that every degree-d polynomial $p : \{0,1\}^n \to [0,1]$ has a variable of influence at least $\mathrm{poly}(\mathrm{Var}(p)/d)$. This implies that quantum query algorithms for decision problems can be efficiently simulated classically on most inputs. It also yields a near-equivalence between the status of the BPP vs. BQP problem in the random oracle and unrelativized worlds.

We organize the proof around four main steps and discuss how each connects to the literature on the analysis of boolean functions and probabilistic combinatorics. To bring out the key ideas, we do not optimize parameters, and our presentation yields weaker exponents than the original.

Authors' note. The proof was discovered by David Turturean using a research system developed at Physical Superintelligence (PSI), in combination with a non-public Anthropic model. The proof was then shared with the authors, who worked to understand it and prepare this exposition. Anthropic's public models were used in this process.

The authors did not receive financial compensation from Anthropic or PSI. This article is not intended for publication at a conference or in a journal.

Contents

1 Introduction **1**
- 1.1 This article . . . 2
- 1.2 Prior work . . . 3

2 Technical Overview **3**
- 2.1 Fractional certificates . . . 5
- 2.2 Certificate-influence inequality . . . 5
  - 2.2.1 How the AA conjecture follows . . . 6
  - 2.2.2 Discussion: Average-case, median, and worst-case versions . . . 6
  - 2.2.3 Discussion: Comparison with a conjecture of Lovett and Zhang . . . 7
- 2.3 Weighted all-pairs inequality . . . 9
  - 2.3.1 Discussion: Relationship to the Poincaré and dual BKR inequalities . . . 9
- 2.4 Routing lemma . . . 10
  - 2.4.1 Discussion: Poincaré for distances . . . 11
  - 2.4.2 Routing lemma ⇒ Weighted all-pairs inequality . . . 12
- 2.5 Antipodal routing lemma . . . 13
  - 2.5.1 Discussion: Simple proof of the faceless antipodal routing lemma . . . 14
  - 2.5.2 Two matrices and their determinants . . . 15
  - 2.5.3 The determinant equivalence lemma . . . 16
  - 2.5.4 Discussion: Relationship to Reimer's butterfly lemma . . . 18

3 Proof of the antipodal routing lemma **18**
- 3.1 Proof of Lemma 3.1 using Lemma 2.16 . . . 19
- 3.2 Proof of Lemma 2.16 . . . 20

4 Proof of the routing lemma **22**

5 Proof of the weighted all-pairs inequality **23**

6 Proof of the certificate-influence inequality **24**

7 Proof of the AA conjecture **25**
- 7.1 Proof of Theorem 2 . . . 26

8 Acknowledgments **27**

1 Introduction

Quantum computing research seeks to discover and delineate the power of quantum computers over classical ones. Simon's and Shor's algorithms [Sim97, Sho97] provided early demonstrations of quantum advantage. On the other hand, despite decades of effort, NP-complete problems and the lattice problems that underlie post-quantum cryptography appear resistant to efficient quantum and classical algorithms alike.

An oft-repeated mantra is that quantum speedups exploit *structure* in the inputs. This intuition is sharpest in the query model, where many quantum algorithms operate. The crux of Shor's factoring algorithm is an algorithm for Period-Finding, where one is given black-box query access to a string of integers that is promised to be periodic and is asked to find the period. Likewise, in Simon's problem one is given black-box query access to $f : \{0,1\}^n \to \{0,1\}^n$ that is promised to be two-to-one, with $f(x) = f(x \oplus s)$ for some nonzero s, and is asked to find s. In contrast, Grover's algorithm [Gro96] for searching an unordered list of n items does so with $O(\sqrt{n})$ queries, achieving only a quadratic speedup over classical algorithms, a speedup that Bennett, Bernstein, Brassard, and Vazirani [BBBV97] showed is optimal.

The simulation conjecture. Common to Shor's and Simon's algorithms, and many superpolynomial speedups that followed, is that they rely on a stringent promise on the input—one that is only satisfied by a tiny fraction of all inputs. The following conjecture formalizes a sense in which such a promise is necessary:

> **Conjecture 1** (Simulation conjecture)**.** *Let $\mathcal{A}$ be a quantum algorithm that makes t queries to an input $x \in \{0,1\}^n$. For every $\varepsilon, \delta > 0$ there is a classical algorithm that makes $\mathrm{poly}(t, 1/\varepsilon, 1/\delta)$ queries to x and approximates the acceptance probability of $\mathcal{A}$ to within an additive ε on a $1-\delta$ fraction of inputs.*

This conjecture first appeared in the works of Aaronson [Aar05, Aar08a] and of Aaronson and Ambainis [AA14], where it is attributed to folklore.

The conjecture has an equivalent formulation in the language of partial functions. Suppose $F : \{0,1\}^n \to \{0,1,*\}$ admits a t-query quantum algorithm. If $S \coloneqq F^{-1}(\{0,1\})$ and $|S| \geq \delta 2^n$, the conjecture asserts the existence of a classical algorithm that makes $\mathrm{poly}(t, 1/\delta, 1/\alpha)$ queries and computes F on a $1-\alpha$ fraction of S. For total functions where $S = \{0,1\}^n$, an extreme case of entirely unstructured inputs, Beals, Buhrman, Cleve, Mosca, and de Wolf [BBC+01] showed that their quantum and classical query complexities are indeed polynomially related.

The Aaronson–Ambainis (AA) conjecture. Aaronson and Ambainis [AA14] reduced the simulation conjecture to a question about bounded low-degree polynomials:

> **Conjecture 2.** *Let $p\colon \{0,1\}^n \to [0,1]$ be a polynomial of degree d. Then some variable $i \in [n]$ satisfies $\mathrm{Inf}_i(p) \geq \mathrm{poly}(\mathrm{Var}(p)/d)$.*

It is not hard to see why the simulation conjecture would follow from such a conjecture. Beals et al. [BBC+01] showed that the acceptance probability of every t-query quantum algorithm is a

polynomial p of degree at most $2t$. A classical simulator can query the most influential variable of p, restrict p accordingly, and recurse on the restricted polynomial until it becomes close to constant. The lower bound on maximum influence asserted by the AA conjecture translates into a bound on the query complexity of this classical simulator.

Implications for BPP vs. BQP relative to a random oracle. Further motivation for the AA conjecture comes from the relativized BPP versus BQP problem. Bernstein and Vazirani [BV97], in the paper that defined BQP, gave an oracle relative to which $\mathsf{BPP} \neq \mathsf{BQP}$. Fortnow and Rogers [FR99] asked whether the same holds relative to a *random* oracle. This question remains open [Aar05, For21].

Aaronson and Ambainis showed that their conjecture, if true, would provide formal evidence for the difficulty of [FR99]'s question: Assuming the AA conjecture, proving $\mathsf{BQP} \not\subseteq \mathsf{HeurP}$ relative to a random oracle would require proving the *unrelativized* separation $\mathsf{P} \neq \mathsf{P}^{\#\mathsf{P}}$.[1] This was recently strengthened by Blanc, Docter, Strassle, and Tan [BDST26], who showed that it would require proving not only $\mathsf{P} \neq \mathsf{P}^{\#\mathsf{P}}$, but even $\mathsf{PromiseBPP} \neq \mathsf{PromiseBQP}$:[2]

Claim 1.1. *Assume the AA conjecture. If* $\mathsf{BQP} \not\subseteq \mathsf{HeurP}$ *relative to a random oracle then* $\mathsf{PromiseBPP} \neq \mathsf{PromiseBQP}$.

[BDST26]'s motivation for Claim 1.1 came from the fact that $\mathsf{BPP} \neq \mathsf{BQP}$ implies $\mathsf{BPP} \neq \mathsf{BQP}$ relative to a random oracle [BG81, FR99]. This "easy direction," where an unrelativized separation implies a random oracle separation, holds for many pairs of complexity classes. Claim 1.1 says that the AA conjecture implies an approximate converse, one going in the "hard direction," and hence an approximate equivalence between the status of the BPP vs. BQP problem in the random oracle and unrelativized worlds (approximate because of the difference between BPP and HeurP, and the distinction between promise and non-promise classes). In other words, the AA conjecture implies a sense in which the random oracle hypothesis [BG81], while false in general [Kur82, Har85, CCG+94], does hold for the BPP vs. BQP problem.

1.1 This article

We give an exposition of a recent proof of the AA conjecture [Tur26]. The resolution of this conjecture implies the simulation conjecture and makes Claim 1.1 unconditional. We show:

Theorem 1 (AA is true)**.** *Let* $p\colon \{0,1\}^n \to [0,1]$ *be a polynomial of degree* d*. Then*

$$\max_{i\in[n]} \mathrm{Inf}_i(p) \;\geq\; \Omega\left(\frac{\mathrm{Var}(p)^8}{d^{12}}\right).$$

[1] [AA14] state their assumption as $\mathsf{BQP} \not\subseteq \mathsf{AvgP}$ relative to a random oracle, where by AvgP they mean the languages decided by a polynomial-time algorithm on a $1 - o(1)$ fraction of the inputs. We state their result with HeurP instead of AvgP since the latter is often reserved for errorless algorithms.

[2] [BDST26] state their assumption as $\mathsf{PromiseBQP} \not\subseteq \mathsf{PromiseHeurBPP}$ relative to a random oracle, which is implied by $\mathsf{BQP} \not\subseteq \mathsf{HeurBPP}$ relative to a random oracle. By a standard argument [BG81], $\mathsf{HeurBPP} = \mathsf{HeurP}$ relative to a random oracle.

The proof in [Tur26] achieves stronger parameters: $\max_i \mathrm{Inf}_i(p) \geq \Omega(\mathrm{Var}(p)^{5/2}/d^6)$. All the key ideas in our presentation are from the original proof. We strive to present the proof as simply as possible, to highlight its main conceptual steps, and to place each step in the context of the broader literature on the analysis of boolean functions and probabilistic combinatorics. As a result, our technical overview is fairly long, but the proof that we present is short.

1.2 Prior work

The Aaronson–Ambainis conjecture. Dinur, Friedgut, Kindler, and O'Donnell [DFKO07], in a paper that predates [AA14], showed that every bounded degree-d polynomial is ε-close to a junta on $2^{O(d)}/\varepsilon^2$ variables, which gives a variable of influence $\mathrm{poly}(\mathrm{Var}(p)/2^d)$. This had remained the best bound for general polynomials, and has been rederived by O'Donnell and Zhao [OZ16] and by Defant, Mastyło, and Pérez [DMP19]. As [AA14] also noted in their paper, for *boolean-valued* polynomials, the conjecture follows as a corollary of the "OSSS inequality" of O'Donnell, Saks, Schramm, and Servedio [OSSS05] and bounds on decision tree depth in terms of polynomial degree [Nis91, NS94]. We discuss the OSSS inequality in more detail in the technical overview.

The conjecture has since been proven for several special cases: polynomials whose nonzero Fourier coefficients have the same magnitude [Mon12, Iva26, STV26]; completely bounded block-multilinear forms [BSdW22], a class which contains the acceptance polynomials of k-fold *Forrelation* [ABP19]; homogeneous Fourier completely bounded polynomials [Esc24]; and random restrictions of arbitrary bounded polynomials [Bha25]. O'Donnell and Zhao [OZ16] showed that it suffices to prove the conjecture for one-block decoupled polynomials, while Girish, Mittal, Nehoran, and Raz [GMNR26] showed that variants of the conjecture for certain non-uniform distributions are equivalent to the original one.

Most relevant for us, Lovett and Zhang [LZ23] connected the problem to fractional certificates and proved the first bound on the fractional certificate complexity of bounded low-degree functions. We elaborate on the connections between the proof that we present and [LZ23] in the technical overview.

The simulation conjecture. The simulation conjecture concerns decision problems, and its natural analogue for *search* problems is false. Yamakawa and Zhandry [YZ24] gave a search problem for which, relative to a random oracle $\mathcal{O} : \{0,1\}^n \to \{0,1\}^n$, there is a quantum algorithm that uses polynomially many queries and succeeds with high probability, whereas every classical algorithm requires exponentially many queries.

Recent works study the simulation conjecture for quantum algorithms with limited rounds of adaptivity [BDST26, ESP26, LM26].

2 Technical Overview

Roadmap. Figure 1 summarizes the structure of the proof and of this technical overview. We will present the proof in a top-down fashion, devoting a subsection to each step. As the figure also indicates, each step draws on and extends ideas in the literature, which we discuss in the corresponding subsections. This section therefore doubles as an extended survey of related work. For a more succinct technical overview, a reader can skip all subsections titled "Discussion: [...]."

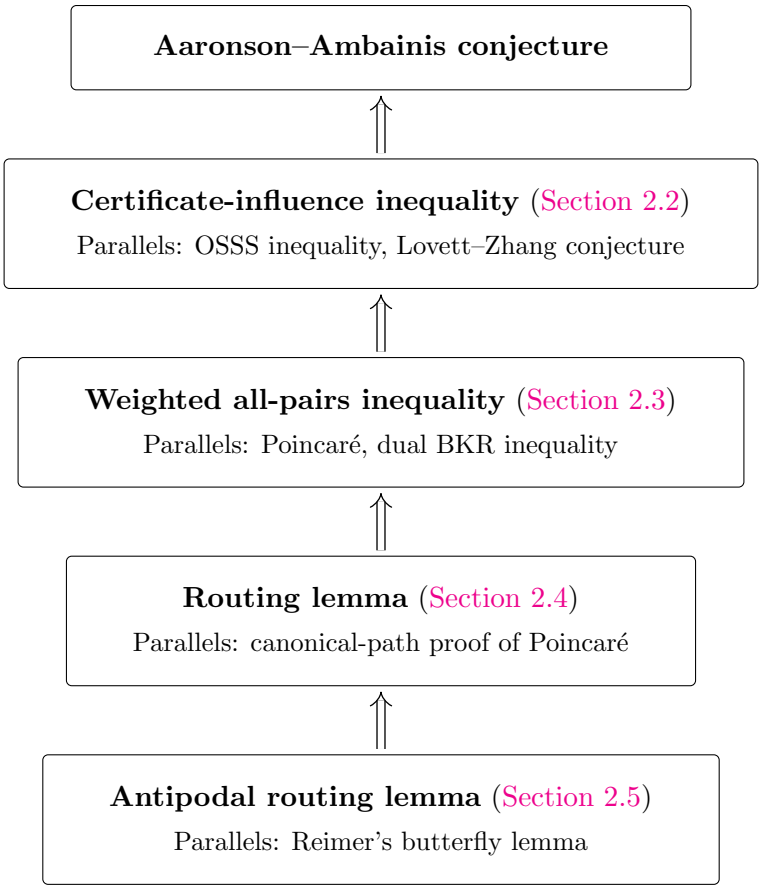


Figure 1: Structure of the proof.

Preliminaries. For a real number ξ we write $\xi_+ = \max(\xi, 0)$. We write $\mathbb{1}[\cdot]$ for the indicator of a condition. Random variables are written in boldface (e.g. $\boldsymbol{x} \sim \{0,1\}^n$), and unless otherwise stated, all probabilities and expectations are with respect to the uniform distribution. For a vector $\lambda \in \mathbb{R}^n_{\geq 0}$ we write $\lambda(i)$ for its i-th entry, $\lambda(S) = \sum_{i \in S} \lambda(i)$ for $S \subseteq [n]$, and $\|\lambda\|_1 = \lambda([n])$.

For $x \in \{0,1\}^n$ and $S \subseteq [n]$, we write $x^{\oplus S}$ to denote x with the coordinates in S flipped. For $i \in [n]$ we write $x^{\oplus i}$ for $x^{\oplus \{i\}}$. We write $\Delta(x,y) = \{i \in [n] : x_i \neq y_i\}$ for the set of coordinates where x and y differ. For $x, y \in \{0,1\}^n$, we write

$$\mathrm{cube}(x,y) \coloneqq \left\{z \in \{0,1\}^n : z_i = x_i = y_i \text{ for every } i \notin \Delta(x,y)\right\}$$

to denote the subcube spanned by x and y, i.e. the subcube with x and y as antipodal points.

We reserve f for boolean-valued functions and use p for real-valued ones. The variance of a function $p : \{0,1\}^n \to \mathbb{R}$ is $\mathrm{Var}(p) \coloneqq \mathbb{E}[(p(\boldsymbol{x}) - \mathbb{E}[p(\boldsymbol{x})])^2]$. For boolean-valued functions $f : \{0,1\}^n \to \{0,1\}$ we have $\mathrm{Var}(f) = \Pr[f(\boldsymbol{x}) = 1] \Pr[f(\boldsymbol{x}) = 0]$.

Definition 2.1 (Influences)**.** Let $p\colon \{0,1\}^n \to \mathbb{R}$ and $i \in [n]$. The *influence* of coordinate i on p is

$$\mathrm{Inf}_i(p) \coloneqq \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\left(p(\boldsymbol{x}) - p(\boldsymbol{x}^{\oplus i})\right)^2\right].$$

The total influence of p is $\mathrm{Inf}(p) \coloneqq \sum_{i=1}^n \mathrm{Inf}_i(p)$.

2.1 Fractional certificates

An important notion in this proof is that of fractional certificates. Fractional certificates for boolean-valued functions were first studied by Tal [Tal13] and Gilmer, Saks, and Srinivasan [GSS16]. Up to constant factors, they are equivalent to randomized certificates, introduced by Aaronson [Aar08b].

We will need an analogue for real-valued functions:

Definition 2.2 (Fractional certificates)**.** Let $p\colon \{0,1\}^n \to \mathbb{R}$ and $x \in \{0,1\}^n$. A *fractional certificate for p at x* is a vector $\lambda_x \in \mathbb{R}^n_{\geq 0}$ such that

$$\lambda_x\big(\Delta(x,y)\big) \geq |p(x) - p(y)| \qquad \text{for every } y \in \{0,1\}^n.$$

Its *cost* is $\|\lambda_x\|_1$. For a boolean-valued function $f : \{0,1\}^n \to \{0,1\}$, this is equivalent to:

$$\lambda_x\big(\Delta(x,y)\big) \geq 1 \qquad \text{for all } y \text{ such that } f(y) \neq f(x).$$

The $\{0,1\}$-valued vectors λ_x satisfying this are exactly the indicator vectors of ordinary certificates of f at x.

Definition 2.3 (Fractional certificate families)**.** A *fractional certificate family* for $p\colon \{0,1\}^n \to \mathbb{R}$ is an indexed collection $\Lambda = (\lambda_x)_{x\in\{0,1\}^n}$ in which λ_x is a fractional certificate for p at x, for every $x \in \{0,1\}^n$. Its *cost* is $\max_x \|\lambda_x\|_1$.

The *fractional certificate complexity* of p at x, written $\mathrm{FC}(p,x)$, is the minimum cost of a fractional certificate for p at x, and the *fractional certificate complexity* of p is $\mathrm{FC}(p) = \max_x \mathrm{FC}(p,x)$, the minimum cost of a fractional certificate family for p.

Agarwal and Ben-David [ABD26], building on [ABDK21, Tal13], proved the following bound on the fractional certificate complexity of bounded low-degree polynomials:

Fact 2.4. *Every $p\colon \{0,1\}^n \to [0,1]$ of degree at most d has $\mathrm{FC}(p) = O(d^2)$.*

Earlier, Lovett and Zhang [LZ23] proved that every such p has $\mathrm{FC}_\varepsilon(p) = O(d^8 \log^{16}(n)/\varepsilon^4)$. Here $\mathrm{FC}_\varepsilon(p)$ is an analogue of $\mathrm{FC}(p)$ in which x's certificate must satisfy $\lambda_x(\Delta(x,y)) \geq 1$ for every y with $|p(x) - p(y)| \geq \varepsilon$, and is unconstrained for other y's. A fractional certificate of cost K yields one of cost K/ε in their sense, and so Fact 2.4 implies $\mathrm{FC}_\varepsilon(p) = O(d^2/\varepsilon)$. We will use the stronger notion $\mathrm{FC}(p)$.

2.2 Certificate-influence inequality

Fact 2.4 is the only step in the proof that uses the degree assumption of the AA conjecture. With this fact in hand, we show that every bounded function with small fractional certificate complexity, regardless of degree, has an influential variable. This claim in turn reduces to the following statement about boolean-valued functions:

Lemma 2.5 (Certificate-influence inequality)**.** *Let $f\colon \{0,1\}^n \to \{0,1\}$ be nonconstant, with a fractional certificate family $\Lambda = (\lambda_x)_{x\in\{0,1\}^n}$. Then for uniform $\boldsymbol{x}$ and $\boldsymbol{y}$,*

$$\max_{i\in[n]} \mathrm{Inf}_i(f) \geq 4 \cdot \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{\mathbb{1}[f(\boldsymbol{x}) = 1]}{\|\lambda_{\boldsymbol{x}}\|_1}\right] \cdot \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{\mathbb{1}[f(\boldsymbol{y}) = 0]}{\|\lambda_{\boldsymbol{y}}\|_1}\right].$$

Note that the RHS is well defined: Since f is nonconstant, every input requires a certificate of positive cost. The inequality is exactly tight for the dictator function $f(x) = \mathbb{1}[x_1 = 1]$ with $\lambda_x = e_1$ for every x. To our knowledge, even the special case for ordinary (non-fractional) certificates was not known, though we need the stronger fractional version to prove the AA conjecture.

2.2.1 How the AA conjecture follows

Let $p : \{0,1\}^n \to [0,1]$ be a bounded polynomial and $\Lambda = (\lambda_x)_{x \in \{0,1\}^n}$ be a fractional certificate family for p of cost $\mathrm{FC}(p)$. For $t \in [0,1]$, we define the boolean-valued threshold function

$$p_t(x) \coloneqq \mathbb{1}[p(x) > t].$$

The high-level idea is that the certificate-influence inequality applied to p_t yields an AA-like statement for p_t, which, for an appropriate choice of t, we can then transfer to p. See Figure 2.

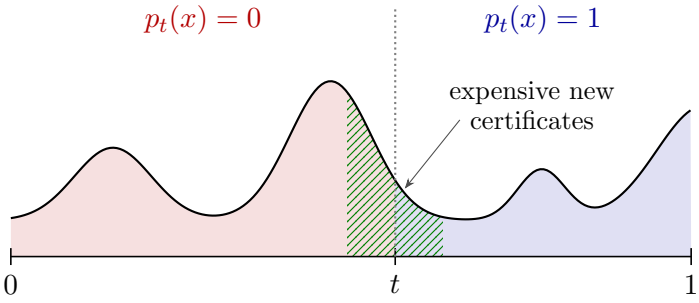


Figure 2: The curve depicts the distribution of $p(\boldsymbol{x})$ for a uniform $\boldsymbol{x} \sim \{0,1\}^n$. We booleanize p by thresholding it at a value $t \in [0,1]$. Since the certificates for p_t are the certificates for p scaled by $1/|p(x) - t|$, inputs with $p(x)$ close to t (shaded green) have expensive new certificates. Such inputs contribute little to the RHS of the certificate-influence inequality.

It can be checked that

$$\left(\frac{\lambda_x}{|p(x) - t|}\right)_{x \in \{0,1\}^n} \tag{1}$$

is a fractional certificate family for p_t. Applying the certificate-influence inequality to p_t and these new certificates, we get that

$$\max_{i \in [n]} \mathrm{Inf}_i(p_t) \geq 4 \cdot \frac{\mathbb{E}\left[(p(\boldsymbol{x}) - t)_+\right] \mathbb{E}\left[(t - p(\boldsymbol{y}))_+\right]}{\mathrm{FC}(p)^2}.$$

Fact 2.4 upper bounds $\mathrm{FC}(p)$ by $O(\deg(p)^2)$, while $\mathbb{E}[(p(\boldsymbol{x}) - t)_+]$ and $\mathbb{E}[(t - p(\boldsymbol{y}))_+]$ are each at least $\Omega(\mathrm{Var}(p))$ as long as t is sufficiently close to $\mathbb{E}[p(\boldsymbol{x})]$. Finally, for $\boldsymbol{t}$ randomly chosen within this admissible interval, an elementary argument upper bounds $\mathbb{E}_{\boldsymbol{t}}[\mathrm{Inf}_i(p_{\boldsymbol{t}})]$ in terms of $\mathrm{Inf}_i(p)$ and $\mathrm{Var}(p)$. Putting these together yields the AA conjecture.

2.2.2 Discussion: Average-case, median, and worst-case versions

The certificate-influence inequality is an average-case bound: The RHS averages, over 1-inputs and 0-inputs separately, the reciprocals of the costs of their certificates. This implies a weaker version where the lower bound is expressed in terms of the *median* costs of 1- and 0-certificates:

Corollary 2.6 (Median version of the certificate-influence inequality)**.** *Let f and Λ be as in Lemma 2.5. For $b \in \{0,1\}$, let m_b be a median of the cost $\|\lambda_x\|_1$ over the inputs x with $f(x) = b$. Let $\mu \coloneqq \Pr[f(\boldsymbol{x}) = 1]$. Then*

$$\max_{i \in [n]} \mathrm{Inf}_i(f) \geq \frac{\mu(1-\mu)}{m_0 \cdot m_1}.$$

Our reduction from the AA conjecture can be modified so that this median version suffices, though we do not know of a proof of it that is substantially simpler than that of the average-case version.

Worst-case version and the OSSS inequality. An even weaker version of the inequality is one where the lower bound is expressed in terms of the *maximum* costs of the certificates. This worst-case version, in contrast, does not suffice for our reduction from the AA conjecture. In our reduction, the inequality (or its median version) is applied to the new fractional certificates (1). The costs of these certificates blow up whenever $p(x)$ is close to t. An average-case or median version allows us to argue that this does not happen too often, but a worst-case version does not.

We also remark that, up to polynomial factors, such a worst-case version already follows from the OSSS inequality:

$$\max_{i \in [n]} \mathrm{Inf}_i(f) \geq \frac{4\mu(1-\mu)}{D(f)} \qquad \text{where } D(f) \coloneqq \text{decision tree complexity of } f$$

along with bounds on $D(f)$ in terms of $\mathrm{FC}(f)$ via [Nis91, BBC+01]. The proof of Lemma 2.5 is rather different from, and quite a bit more involved than, the various known proofs of the OSSS inequality.

2.2.3 Discussion: Comparison with a conjecture of Lovett and Zhang

Lovett and Zhang [LZ23] proposed a plan for the AA conjecture whose main step is a conjectured L_∞ analogue of Talagrand's convex-distance inequality [Tal96]. We now state and discuss their conjecture, describe a counterexample, and compare their conjecture to the certificate-influence inequality.

[LZ23]'s conjecture in the language of fractional certificates. Roughly speaking, [LZ23]'s conjecture says that if $X, Y \subseteq \{0,1\}^n$ are two disjoint sets such that every $x \in X$ is far in L_∞ from $\mathrm{conv}(Y)$, then there must be a variable that is influential with respect to X or to Y. See their paper for the precise formulation of their conjecture and the definition of this more general notion of influence. Their reduction from the AA conjecture requires this level of generality, but we state here the special case of their conjecture where X and Y partition the cube:

Conjecture 3 (Special case of [LZ23, Conjecture 37])**.** *Let $f \colon \{0,1\}^n \to \{0,1\}$ be nonconstant, $A \coloneqq f^{-1}(1)$, and $\mu \coloneqq \Pr[f(\boldsymbol{x}) = 1] = |A|/2^n$. Then*

$$\max_{i \in [n]} \mathrm{Inf}_i(f) \geq \mu(1-\mu) \cdot \mathrm{poly}\Big(\min_{x \in A} d_\infty\big(x, \mathrm{conv}(\overline{A})\big)\Big).$$

The distance $d_\infty(x, \mathrm{conv}(\overline{A}))$ is exactly captured by fractional certificates:

Fact 2.7. *$d_\infty\big(x, \operatorname{conv}(\overline{A})\big) = 1/\mathrm{FC}(f,x)$ for $x \in A$.*

One direction of this identity ($\geq$) is proved in [LZ23, Lemma 28], while the other ($\leq$) follows from LP duality.[3] Thanks to this identity, Conjecture 3 can be equivalently stated as follows:

Conjecture 4 (Conjecture 3 in the language of fractional certificates)**.** *Let $f : \{0,1\}^n \to \{0,1\}$ be nonconstant and $\mu \coloneqq \Pr[f(\boldsymbol{x}) = 1]$. Then*

$$\max_{i\in[n]} \mathrm{Inf}_i(f) \geq \frac{\mu(1-\mu)}{\mathrm{poly}(\mathrm{FC}_1(f))},$$

where $\mathrm{FC}_b(f) \coloneqq \max\{\mathrm{FC}(f,x) \colon x \in f^{-1}(b)\}$.

Notably, the lower bound is one-sided in the sense that it only depends on the fractional certificate complexities of the 1-inputs, not those of the 0-inputs.

A counterexample. [LZ23] showed that their conjecture would imply the AA conjecture. Unfortunately, their conjecture is already false in the special case described above. Consider the Tribes function with tribes of width $w \approx \log n - \log\ln n$, for which $\mu \approx \frac{1}{2}$. Every $x \in \textsc{Tribes}^{-1}(1)$ has a fractional certificate (in fact, an ordinary certificate) of cost $O(w) = O(\log n)$, and so Conjecture 4 asserts a lower bound of $1/\operatorname{polylog}(n)$ on maximum influence. This is false, since every variable has influence $O(\log(n)/n)$.

How the certificate-influence inequality differs from Conjecture 4. The certificate-influence inequality is two-sided: It depends on the certificate costs of both the 1-inputs and the 0-inputs. This is what lets it evade the counterexample, since every 0-input of Tribes requires fractional certificate complexity $\Omega(n/w) = \Omega(n/\log n)$.

Two-sidedness alone, however, is not the new ingredient. A two-sided worst-case statement,

$$\max_{i\in[n]} \mathrm{Inf}_i(f) \geq \frac{\mu(1-\mu)}{\mathrm{poly}(\mathrm{FC}_0(f), \mathrm{FC}_1(f))},$$

already follows from the OSSS inequality as discussed in Section 2.2.2. What is new is the average-case nature of the certificate-influence inequality, which, as discussed in the same section, is what our reduction from the AA conjecture needs.

These differences aside, our reduction from the AA conjecture is similar in spirit to [LZ23]'s. For a polynomial $p : \{0,1\}^n \to [0,1]$, they apply their conjecture to the sets $X = \{x \colon p(x) \geq t + \theta\}$ and $Y = \{y \colon p(y) \leq t - \theta\}$, while we apply the certificate-influence inequality to the function $p_t(x) = \mathbb{1}[p(x) > t]$. In both cases, the AA conjecture for a real-valued function is reduced to an AA-like statement for a boolean function (a partial boolean function in their case).

[3] Fact 2.7 gives a particularly nice reformulation of the certificate-influence inequality. It can be stated as

$$\max_{i\in[n]} \mathrm{Inf}_i(f) \geq 4 \cdot \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[d_\infty(\boldsymbol{x}, \operatorname{conv}(f^{-1}(0)))\right] \cdot \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[d_\infty(\boldsymbol{y}, \operatorname{conv}(f^{-1}(1)))\right].$$

2.3 Weighted all-pairs inequality

Overlap. We call a pair (x, y) of inputs *bichromatic* if $f(x) \neq f(y)$. We will be interested in understanding the amount of "overlap" between certificates of bichromatic pairs:

$$\mathrm{overlap}_\Lambda(x, y) \coloneqq \sum_{i \in \Delta(x,y)} \lambda_x(i)\lambda_y(i),$$

which is the inner product of λ_x and λ_y restricted to the coordinates on which x and y differ.

The ordinary (non-fractional) certificates of bichromatic pairs have overlap at least 1: Viewed as sets, these certificates must intersect within $\Delta(x, y)$, since otherwise there will be an input that is consistent with both certificates. For fractional certificates, however, overlap can be much smaller. Consider the n-bit MAJORITY function, $x = 1^n$ and $y = 0^n$, and fractional certificates:

$$\lambda_x = \lambda_y = \left(\tfrac{2}{n}, \ldots, \tfrac{2}{n}\right).$$

The overlap between these two certificates is only $O(\frac{1}{n})$.

To prove the certificate-influence inequality, we will need a more general definition where overlap is measured with respect to positive weights $w \in \mathbb{R}^n_{>0}$:

$$\mathrm{overlap}_{w,\Lambda}(x, y) \coloneqq \sum_{i \in \Delta(x,y)} w_i \lambda_x(i)\lambda_y(i).$$

We can now state the weighted all-pairs inequality:

Lemma 2.8 (Weighted all-pairs inequality)**.** *Let $f\colon \{0,1\}^n \to \{0,1\}$ be nonconstant, with a fractional certificate family $\Lambda = (\lambda_x)_{x \in \{0,1\}^n}$. Let $w \in \mathbb{R}^n_{>0}$. For independent and uniform $\boldsymbol{x}$ and $\boldsymbol{y}$,*

$$\sum_{i=1}^n \frac{\mathrm{Inf}_i(f)}{w_i} \geq 4 \cdot \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}} \left[\frac{\mathbb{1}[f(\boldsymbol{x}) = 1, f(\boldsymbol{y}) = 0]}{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x}, \boldsymbol{y})}\right].$$

The ratio on the RHS is taken to be 0 when the numerator is 0. When the numerator is 1 (i.e. when (x, y) is bichromatic), it can be checked that the overlap must be positive.[4]

The certificate-influence inequality follows quite straightforwardly from the weighted all-pairs inequality, via an appropriate choice of weights and an application of Cauchy–Schwarz. (This is why we need the weighted version.)

2.3.1 Discussion: Relationship to the Poincaré and dual BKR inequalities

Consider the *unweighted* version of the all-pairs inequality where $w = 1^n$, which already captures essentially all the complexity of the proof. It says that:

$$\mathrm{Inf}(f) \geq 4 \cdot \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}} \left[\frac{\mathbb{1}[f(\boldsymbol{x}) = 1, f(\boldsymbol{y}) = 0]}{\mathrm{overlap}_\Lambda(\boldsymbol{x}, \boldsymbol{y})}\right].$$

[4] In fact, fractional certificates for bichromatic pairs satisfy $\sum_{i \in \Delta(x,y)} \min\{\lambda_x(i), \lambda_y(i)\} \geq 1$. See [KT16, Lemma 6.2].

Poincaré inequality. Unlike the certificate-influence inequality, the unweighted all-pairs inequality is known and elementary in the case of *ordinary* certificates. Since the overlap between ordinary certificates of bichromatic pairs is at least 1, the inequality follows from the classic Poincaré inequality over the hypercube:

$$\textbf{Poincaré:} \quad \mathrm{Inf}(f) \geq 4\mu(1-\mu), \quad \mu \coloneqq \Pr[f(\boldsymbol{x}) = 1].$$

The challenge in proving the all-pairs inequality, even the unweighted version, stems from the fact that the overlap between *fractional* certificates can be much smaller than 1.

The overlap can also be larger than 1, and so the all-pairs inequality is incomparable to Poincaré. That said, en route to the all-pairs inequality we will prove a generalization of Poincaré. See Section 2.4.

Dual BKR inequality. One can consider a more general notion of ordinary certificates for two sets $X, Y \subseteq \{0,1\}^n$, extending the standard setting where X and Y partition the cube. For each $x \in X$, let $C_x \subseteq [n]$ certify x's membership in X: all z's that agree with x on C_x also belong to X. Likewise, we assign to each $y \in Y$ a certificate $D_y \subseteq [n]$ and let Λ be the family of all certificates.

If X and Y are disjoint, then the overlap between the certificates for any pair $(x, y) \in X \times Y$ must be at least 1 (i.e. their corresponding subcubes must be disjoint), for the same reason as in the standard setting. Now consider the setting where X and Y intersect, in which case these certificates can have 0 overlap (i.e. their corresponding subcubes may intersect). The dual van den Berg–Kesten–Reimer (BKR) inequality of Kahn, Saks, and Smyth [KSS11] bounds the probability of this event:

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}\left[\boldsymbol{x} \in X,\ \boldsymbol{y} \in Y,\ \mathrm{overlap}_{\Lambda}(\boldsymbol{x},\boldsymbol{y}) = 0\right] \ \leq\ \Pr[\boldsymbol{z} \in X \cap Y].$$

(Their actual inequality is stronger: A pair (x, y) counts on the LHS as long as there exist certificates for them that have 0 overlap. That is, one does not have to fix the certificate family beforehand.) [KSS11] used the dual BKR inequality to prove Rudich's conjecture [Rud88], and Smyth [Smy01, Smy02] used it to resolve Tardos's conjecture [Tar89]. We refer the reader to these papers for details, mentioning here that both Rudich's and Tardos's conjectures are motivated by questions concerning random oracles, in the same way that the AA conjecture is connected to the question of whether $\mathsf{BPP} \neq \mathsf{BQP}$ relative to a random oracle via Claim 1.1.

While the statement of the all-pairs inequality is quite different from that of the dual BKR inequality, their proofs are related. The key ingredient in the proof of the dual BKR inequality is Reimer's butterfly lemma [Rei00]. The proof of our antipodal routing lemma yields a generalization of Reimer's butterfly lemma. See Section 2.5.4.

2.4 Routing lemma

For the rest of this overview, it will be convenient to view $f : \{0,1\}^n \to \{0,1\}$ as the set $A \coloneqq f^{-1}(1)$. A pair (x, y) is therefore bichromatic if exactly one of them lies in A. We write $\mathrm{Inf}_i(A)$ for $\mathrm{Inf}_i(f)$, and likewise $\mathrm{Inf}(A)$ for $\mathrm{Inf}(f)$.

Definition 2.9 (Monochromatic faces)**.** Fix $A \subseteq \{0,1\}^n$. For $x \in A$ and $T \subseteq [n]$, we call T a *monochromatic face of* x if $\mathrm{cube}(x, x^{\oplus T})$ is monochromatic. Equivalently, $x^{\oplus S} \in A$ for every $S \subseteq T$.

We are now ready to state the routing lemma:

Lemma 2.10 (Routing lemma)**.** *Let $A \subseteq \{0,1\}^n$. For every bichromatic pair $(x,y) \in A \times \overline{A}$, let $T = T(x,y) \subseteq \Delta(x,y)$ be a monochromatic face of x. Then one can choose, for every such pair (x,y) and monochromatic face T, a point $z = z(x,y,T) \in A$ such that*

(a) $z \in \mathrm{cube}(x^{\oplus T}, y)$.

(b) For every $i \in [n]$, when $\boldsymbol{x}, \boldsymbol{y}$ are uniform and independent on $\{0,1\}^n$,

$$\Pr_{\boldsymbol{x},\boldsymbol{y}}\left[\boldsymbol{x} \in A,\ \boldsymbol{y} \notin A,\ \boldsymbol{z}_i \neq \boldsymbol{y}_i\right] \ \leq\ \tfrac{1}{4} \cdot \mathrm{Inf}_i(A).$$

See Figure 3 for an illustration of the lemma.

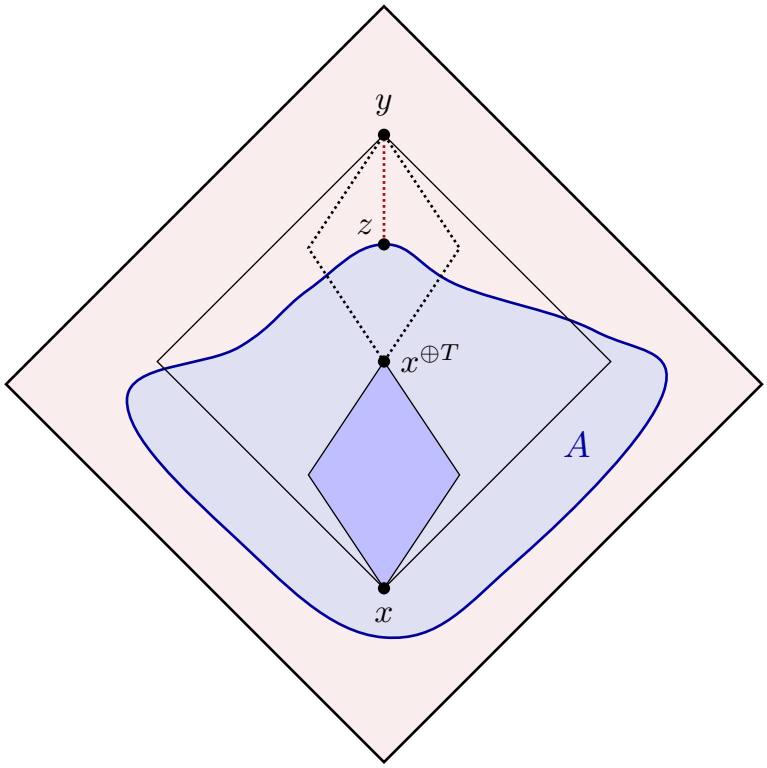


Figure 3: An illustration of the routing lemma. The points x and y form a bichromatic pair, and T is a monochromatic face of x: the shaded diamond $\mathrm{cube}(x, x^{\oplus T})$ lies in A. The dotted diamond is $\mathrm{cube}(x^{\oplus T}, y)$. The routing lemma says that we can assign to (x,y) and T a point $z \in A \cap \mathrm{cube}(x^{\oplus T}, y)$ that is "close" to y—close, on average over all such (x,y) pairs, in the coordinate-wise sense of (b).

2.4.1 Discussion: Poincaré for distances

To gain some intuition for (b), consider the statement that one gets by summing it over all $i \in [n]$:

(b') $\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\mathbb{1}[\boldsymbol{x}\in A,\boldsymbol{y}\notin A]\cdot \mathrm{dist}(\boldsymbol{z},\boldsymbol{y})\right]\le \frac{1}{4}\cdot \mathrm{Inf}(A)$,

where $\mathrm{dist}(\cdot,\cdot)$ denotes Hamming distance. This says the Hamming distance from z to y, averaged over all bichromatic (x,y) pairs, is upper bounded in terms of $\mathrm{Inf}(A)$. The original statement (b) says that this closeness even holds in a more stringent coordinate-wise manner, where for each $i\in[n]$, the disagreement in the i-th coordinate is upper bounded in terms of $\mathrm{Inf}_i(A)$.

The summed version (b') alone, without (a), is already of independent interest. It implies:

$$\textbf{Poincaré for distances:}\quad \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\mathbb{1}[\boldsymbol{x}\in A,\ \boldsymbol{y}\notin A]\cdot \mathrm{dist}(\boldsymbol{y},A)\right]\le \frac{1}{4}\cdot \mathrm{Inf}(A).$$

This strengthens Poincaré, which can be written as:

$$\textbf{Poincaré:}\quad \Pr_{\boldsymbol{x},\boldsymbol{y}}\left[\boldsymbol{x}\in A,\ \boldsymbol{y}\notin A\right]\le \frac{1}{4}\cdot \mathrm{Inf}(A).$$

While Poincaré bounds the fraction of bichromatic pairs in terms of total influence, Poincaré for distances says that the same upper bound holds even when we take their distances to A into account.

2.4.2 Routing lemma ⇒ Weighted all-pairs inequality

Poincaré for restricted distances. Taking (a) into consideration allows us to further strengthen Poincaré for distances, by restricting which paths from y to A we are allowed to take. Since $T\subseteq\Delta(x,y)$, requiring $z\in\mathrm{cube}(x^{\oplus T},y)$ is equivalent to requiring $\Delta(z,y)\subseteq\Delta(x,y)\setminus T$. Thus, we let

$$\mathrm{dist}_{\Delta(x,y)\setminus T}(y,A)\coloneqq \mathrm{dist}\big(y,\ A\cap\mathrm{cube}(x^{\oplus T},y)\big)$$

be the length of the shortest path from y to A that walks only in directions in $\Delta(x,y)\setminus T$. This is at least $\mathrm{dist}(y,A)$. Combining (a) with (b') yields

$$\textbf{Poincaré for restricted distances:}\quad \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\mathbb{1}[\boldsymbol{x}\in A,\ \boldsymbol{y}\notin A]\cdot \mathrm{dist}_{\Delta(\boldsymbol{x},\boldsymbol{y})\setminus T}(\boldsymbol{y},A)\right]\le \frac{1}{4}\cdot \mathrm{Inf}(A).$$

Recall that T has to be a monochromatic face of x, and we can ensure this by ensuring $\lambda_x(T)<1$. Therefore, as long as a set of coordinates has λ_x-mass less than 1, we may freeze those coordinates at their values in y and force the path to only walk in the other directions.

Routing lemma ⇒ unweighted all-pairs inequality. We now sketch how the routing lemma implies the unweighted all-pairs inequality; the weighted version is similar. For every bichromatic pair (x,y), we will define a monochromatic face $T=T(x,y)\subseteq\Delta(x,y)$ of x such that

$$\mathrm{dist}_{\Delta(x,y)\setminus T}(y,A)\ \ge\ \frac{1}{\mathrm{overlap}_\Lambda(x,y)}. \tag{2}$$

The inequality then follows by averaging and applying Poincaré for restricted distances.

Our goal is to define T to make the restricted distance on the LHS of Equation (2) large. Consider an arbitrary $z\in A$. Since (y,z) is bichromatic, $\lambda_y(\Delta(y,z))\ge 1$. Thus, if every coordinate $i\in\Delta(y,z)$ is "light," in the sense that $\lambda_y(i)\le\tau$ for some $\tau>0$, then

$$|\Delta(y,z)|\ge\frac{1}{\tau},$$

since at least $1/\tau$ coordinates of weight at most τ are needed to reach total weight 1. This suggests that we should force the path from y to A to only walk along light directions, which we can do by freezing the heavy ones. Indeed, by letting

$$T = \{i \in \Delta(x, y) \colon \lambda_y(i) > \tau\},$$

we get

$$\operatorname{dist}_{\Delta(x,y)\setminus T}(y, A) \geq \frac{1}{\tau}.$$

It remains to choose τ. A smaller τ gives a better bound, but the routing lemma needs T to be a monochromatic face of x. Every $i \in T$ has $\lambda_y(i) > \tau$, so by Markov's inequality for the function λ_y with respect to the measure λ_x,

$$\lambda_x(T) \;\leq\; \frac{1}{\tau}\sum_{i\in T} \lambda_x(i)\,\lambda_y(i) \;\leq\; \frac{1}{\tau}\sum_{i\in\Delta(x,y)} \lambda_x(i)\,\lambda_y(i) \;=\; \frac{\operatorname{overlap}_\Lambda(x, y)}{\tau},$$

where the first inequality is strict unless $\lambda_x(T) = 0$. We therefore choose $\tau = \operatorname{overlap}_\Lambda(x, y)$ to guarantee $\lambda_x(T) < 1$. With this choice, we get Equation (2).

See Figure 4 for an illustration of this argument.

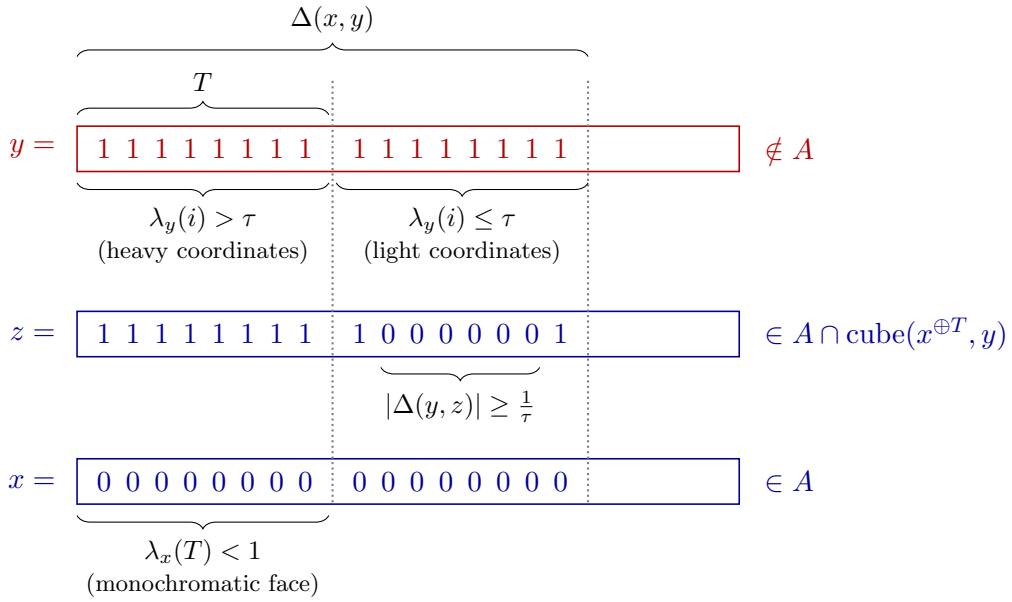


Figure 4: How the routing lemma implies the (unweighted) all-pairs inequality. The three strings agree outside $\Delta(x, y)$, and x and y are drawn as all-0 and all-1 on $\Delta(x, y)$ for concreteness. The heavy coordinates T of λ_y form a monochromatic face of x if $\tau = \operatorname{overlap}_\Lambda(x, y)$, and z is the point given by the routing lemma.

2.5 Antipodal routing lemma

The crux of the proof of the routing lemma is the following analogue for *antipodal* pairs $(x, \bar{x})$ where $\bar{x} \coloneqq (\bar{x}_1, \ldots, \bar{x}_n)$:

Lemma 2.11 (Antipodal routing lemma)**.** *Let $A \subseteq \{0,1\}^n$ and for every $x \in A$, let $T = T(x) \subseteq [n]$ be a monochromatic face of x. Then one can choose, for every such x and monochromatic face T, a point $z = z(x,T) \in A$ such that*

(a) $z \in \mathrm{cube}(x^{\oplus T}, \bar{x})$.

(b) *For every $i \in [n]$, when $\boldsymbol{x}$ is uniform on $\{0,1\}^n$,*

$$\Pr_{\boldsymbol{x}}\left[\boldsymbol{x} \in A,\ \boldsymbol{z}_i \neq \bar{\boldsymbol{x}}_i\right] \ \leq\ \tfrac{1}{2} \cdot \mathrm{Inf}_i(A).$$

The routing lemma follows from this antipodal version by an averaging argument: Every bichromatic pair (x,y) is an antipodal pair in $\mathrm{cube}(x,y)$, and we apply the antipodal routing lemma in every such subcube.

2.5.1 Discussion: Simple proof of the faceless antipodal routing lemma

The main difficulty in proving the antipodal routing lemma, and hence the routing lemma, lies in the constraint imposed by the monochromatic faces: the requirement (a) that $z \in \mathrm{cube}(x^{\oplus T}, \bar{x})$. We sketch a short and elementary proof for the special case where $T(x) = \varnothing$ for all x:

Proposition 2.12 (Faceless antipodal routing lemma)**.** *Let $A \subseteq \{0,1\}^n$. One can choose, for each $x \in A$, a point $z = z(x) \in A$ such that*

$$\Pr_{\boldsymbol{x}}\left[\boldsymbol{x} \in A,\ \boldsymbol{z}_i \neq \bar{\boldsymbol{x}}_i\right] \ \leq\ \tfrac{1}{2} \cdot \mathrm{Inf}_i(A) \quad \textit{for all } i \in [n].$$

Proof sketch. For $i \in [n]$, let $\pi_i \colon A \to A$ swap x and $x^{\oplus i}$ (i.e. flip their i-th coordinates) iff both lie in A, and let $\pi \coloneqq \pi_n \circ \cdots \circ \pi_1$. Our map is $x \mapsto \pi(x)$, which we furthermore claim is in fact a permutation of A. A straightforward induction shows that for all $m \in [n]$, the map $\pi_{\leq m} \coloneqq \pi_m \circ \cdots \circ \pi_1$ is indeed a permutation of A and satisfies:

$$\left|\{x \in A \colon \pi_{\leq m}(x)_i \neq \bar{x}_i\}\right| \ =\ \left|\{x \in A \colon x^{\oplus i} \notin A\}\right| \quad \text{for all } i \in [m].$$

The $m = n$ case of the identity above, divided by 2^n on both sides, yields the proposition. □

Relationship to the canonical-path proof of Poincaré. Recall the following standard proof of the Poincaré inequality: For every pair (x,y), walk from x to y by flipping the coordinates of $\Delta(x,y)$ in the fixed order $1 \to 2 \to \cdots \to n$. If the pair is bichromatic, this path must cross a bichromatic *edge* of A, and we charge the pair to the first such edge. Each edge of the cube lies on these "canonical paths" of exactly 2^n of the 4^n pairs, so the number of bichromatic pairs is at most 2^n times the number of bichromatic edges, which is the Poincaré inequality.

The proof of Proposition 2.12 also walks toward a target, the antipodal point $\bar{x}$, flipping coordinates in the fixed order $1 \to 2 \to \cdots \to n$, and charges bichromatic edges direction by direction. It differs in two ways. First, it never leaves A: It skips any flip that would exit A (i.e. it skips any bichromatic edge), and ends at the point $z = \pi(x) \in A$ rather than at $\bar{x}$. Second, every skipped flip is charged, not only the first: If the walk from x is at w and skips direction i, then x is charged the bichromatic edge $\{w, w^{\oplus i}\}$. Thus x incurs one charge for each coordinate on which z and $\bar{x}$ differ. Since each step π_i is a permutation of A, every bichromatic edge in direction i is charged by exactly one x, which is where the equality in the proof comes from.

The actual antipodal routing lemma with faces. The face constraint (*a*) requires z to flip every coordinate $i \in T(x)$, which is not satisfied by this proof, and we know of no argument along these lines.Its proof, which we discuss next, is linear algebraic, and entirely different from any proof of the Poincaré inequality that we know.

2.5.2 Two matrices and their determinants

For a map $\pi : A \to A$, we let

$$\text{errors}_i(\pi) := \sum_{x \in A} \mathbb{1}[\pi(x)_i \neq \bar{x}_i] = 2^n \cdot \Pr_{\boldsymbol{x}} \left[\boldsymbol{x} \in A,\ \pi(\boldsymbol{x})_i \neq \bar{\boldsymbol{x}}_i\right].$$

Consider the following natural approach. First construct a map π satisfying (*b*) (as we did in the proof of the faceless version), and then show that *any* such map can be "repaired" so that it furthermore satisfies the face constraints (*a*) while not increasing $\text{errors}_i(\pi)$ for any direction $i \in [n]$ (thereby preserving (*b*)). Unfortunately, there are counterexamples showing that not all such maps can be repaired this way. One can then hope that any *permutation* of A can be repaired this way; after all, our proof of the faceless version constructs a permutation. But there are counterexamples to this as well.[5]

Nevertheless, we show that a suitable modification of this high-level approach succeeds. The magical step is to take a *signed combination of all possible permutations*. This makes some of them "cancel" with each other. We show that *any* permutation that survives this cancellation can indeed be repaired to satisfy the face constraints without increasing $\text{errors}_i(\pi)$ for any direction $i \in [n]$.

In more detail, we first define a matrix M in formal variables $\zeta_1, \ldots, \zeta_n$ such that

$$\det M = \sum_{\text{permutations } \pi \text{ of } A} \text{sign}(\pi) \prod_{i \in [n]} \zeta_i^{\text{errors}_i(\pi)}.$$

We show that $\det M$ (viewed as a polynomial in the ζ_i's) has a "good" monomial, one whose degrees, suitably normalized, are upper bounded in terms of the influences of A. Here and throughout, when we refer to a monomial of a polynomial, we mean one with nonzero coefficient:

Claim 2.13 (There is a good monomial)**.** *The polynomial* $\det M$ *has a monomial* $\zeta_1^{k_1} \cdots \zeta_n^{k_n}$ *such that*

$$\frac{k_i}{2^n} \leq \tfrac{1}{2} \cdot \text{Inf}_i(A) \qquad \textit{for every } i \in [n]. \tag{3}$$

Since a monomial $\zeta_1^{k_1} \cdots \zeta_n^{k_n}$ of $\det M$ has a nonzero coefficient, it arises from at least one permutation π with $\text{errors}_i(\pi) = k_i$ for every i, and $\text{errors}_i(\pi)/2^n$ is exactly the LHS of (*b*) for $z = \pi$. So (3) tells us that this π satisfies (*b*).

Next, we define an analogous matrix $M_{\mathcal{T}}$ that takes the monochromatic faces $\mathcal{T} := (T(x))_{x \in A}$ into account. Note that the face constraint $\pi(x) \in \text{cube}(x^{\oplus T}, \bar{x})$ is equivalent to the requirement that $\pi(x)$ agrees with $\bar{x}$ on all coordinates in T:

$$\det M_{\mathcal{T}} = \sum_{\text{permutations } \pi \text{ of } A} \underbrace{\mathbb{1}[\pi(x)_i = \bar{x}_i \text{ for all } x \in A, i \in T(x)]}_{\text{face constraint (a)}} \cdot \text{sign}(\pi) \prod_{i \in [n]} \zeta_i^{\text{errors}_i(\pi)}.$$

We show that the existence of a good monomial in $\det M$ implies the existence of one in $\det M_{\mathcal{T}}$:

[5] Counterexamples for n as small as 3 can be easily found with your AI system of choice.

Claim 2.14 (Good monomial ⇒ Good face-respecting monomial)**.** *Let $\mathcal{T} = (T(x))_{x\in A}$ specify monochromatic faces for each $x \in A$. For any monomial $\zeta_1^{k_1}\cdots\zeta_n^{k_n}$ of $\det M$ there is a monomial $\zeta_1^{k'_1}\cdots\zeta_n^{k'_n}$ of $\det M_{\mathcal{T}}$ such that*

$$k'_i \le k_i \qquad \text{for every } i \in [n].$$

Applying Claim 2.14 to the monomial given by Claim 2.13 yields a monomial of $\det M_{\mathcal{T}}$, and hence a permutation satisfying (a) and (b). This is the antipodal routing lemma.

2.5.3 The determinant equivalence lemma

The key lemma underlying both Claims 2.13 and 2.14 is that $\det M$ and $\det M_{\mathcal{T}}$ have exactly the same *minimal* monomials, defined as follows:

Definition 2.15 (Pareto-minimal component of a polynomial)**.** We say a monomial $\zeta_1^{k_1}\cdots\zeta_n^{k_n}$ is *Pareto-minimal* in a polynomial p if there is no other monomial $\zeta_1^{k'_1}\cdots\zeta_n^{k'_n}$ of p for which $k'_i \le k_i$ for all $i \in [n]$. We then denote the *Pareto-minimal component* of p as

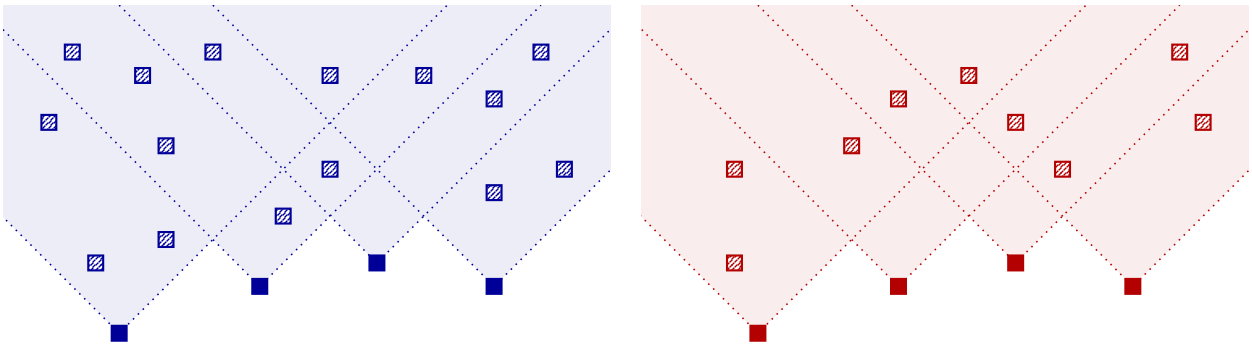
$$\text{minimal}(p) \coloneqq \sum_{\text{monomials } m \text{ in } p} \mathbb{1}[m \text{ is Pareto-minimal in } p] \cdot c_m\, m,$$

where c_m is m's coefficient in p.

Lemma 2.16 (Determinant equivalence lemma)**.** *Let $\mathcal{T} = (T(x))_{x\in A}$ specify monochromatic faces for each $x \in A$. Then*

$$\text{minimal}(\det M) = \text{minimal}(\det M_{\mathcal{T}}).$$

See Figure 5 for an illustration of the lemma.



monomials of $\det M$ monomials of $\det M_{\mathcal{T}}$

Figure 5: An illustration of the determinant equivalence lemma. The Pareto-minimal monomials of $\det M$, shaded solid blue, are the same as the Pareto-minimal monomials of $\det M_{\mathcal{T}}$, shaded solid red. Their other monomials may differ.

Claim 2.14 follows directly from the determinant equivalence lemma. Indeed, it suffices to consider only Pareto-minimal monomials of $\det M$, and the determinant equivalence lemma says that every such monomial is also present in $\det M_{\mathcal{T}}$.

We also use the determinant equivalence lemma to prove Claim 2.13. However, instead of applying it with the monochromatic faces specified by the antipodal routing lemma, we apply it with monochromatic faces that we define. We actually prove the stronger statement that *every* Pareto-minimal monomial of $\det M$ satisfies the degree bound of Equation (3). The advantage of this stronger *for all* statement is that we can separately show, for each $i \in [n]$, that the degree bound is satisfied. For any i, we choose monochromatic faces $\mathcal{T}^{(i)}$ so that the ζ_i-degree of every monomial in $\det M_{\mathcal{T}^{(i)}}$ satisfies the bound of Equation (3). By the determinant equivalence lemma, this must also be true of every Pareto-minimal monomial of $\det M$. Applying this argument n times, once for each $i \in [n]$, yields Claim 2.13.

Proof of the determinant equivalence lemma. The key step in the proof of Lemma 2.16 is a linear-algebraic identity expressing $M_{\mathcal{T}}$ in terms of M. First, we define the two matrices, both with rows and columns indexed by elements of A:

$$M[x,z] \coloneqq \prod_{i\in[n]} \zeta_i^{\mathbb{1}[x_i=z_i]} \qquad \text{and} \qquad M_{\mathcal{T}}[x,z] \coloneqq \begin{cases} 0 & \text{if } x_i = z_i \text{ for any } i \in T(x) \\ M[x,z] & \text{otherwise.} \end{cases}$$

Thus $M_{\mathcal{T}}$ is M with certain entries zeroed out. Using the fact that each $T(x) \in \mathcal{T}$ is a monochromatic face of x, we show that there is in fact a closer relationship between the two matrices: Every row of $M_{\mathcal{T}}$, appropriately scaled, can be written as a linear combination of rows in M.

$$\left(\prod_{i\in T(x)} (1-\zeta_i^2)\right) \cdot \operatorname{row}_x(M_{\mathcal{T}}) = \sum_{S\subseteq T(x)} \left(\prod_{i\in S} (-\zeta_i)\right) \cdot \operatorname{row}_{x^{\oplus S}}(M). \tag{4}$$

Note that the RHS is only well defined because $T(x)$ is a monochromatic face: The rows of M are indexed by elements of A, so $x^{\oplus S}$ only appears as a row in M if $x^{\oplus S} \in A$.

See Figure 6 for an illustration of Equation (4).

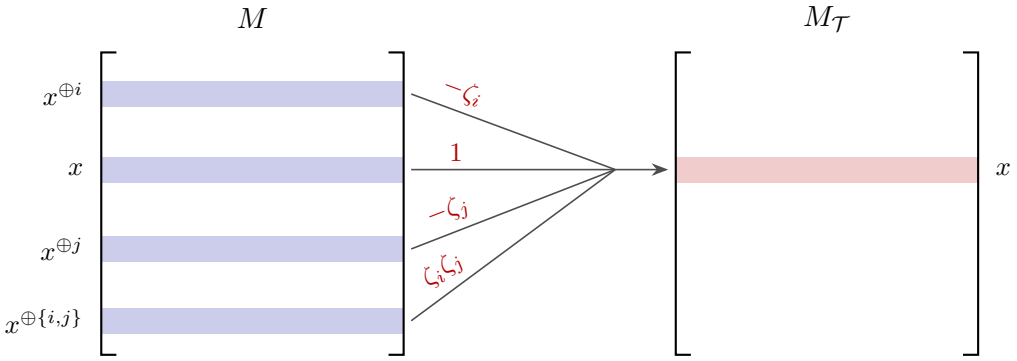


Figure 6: An illustration of Equation (4) for $T(x) = \{i, j\}$. Up to the scaling factor $(1-\zeta_i^2)(1-\zeta_j^2)$, the highlighted row x of $M_{\mathcal{T}}$ is a linear combination of the highlighted rows of M.

Using Equation (4) for each row of $M_{\mathcal{T}}$, we can write

$$DM_{\mathcal{T}} = FM,$$

where D is a diagonal matrix capturing the scaling factors $\prod_{i \in T(x)}(1-\zeta_i^2)$ and each row of F corresponds to the linear combination on the right-hand side of Equation (4). Taking determinants, we obtain

$$\det D \det M_{\mathcal{T}} = \det F \det M.$$

We then show that each of $\det D$ and $\det F$ evaluates to 1 when $\zeta_1 = \cdots = \zeta_n = 0$, which, combined with the following elementary fact, proves the determinant equivalence lemma.

Proposition 2.17. *For any polynomials p, q, if $p(0,\ldots,0) = 1$ then* $\text{minimal}(pq) = \text{minimal}(q)$.

2.5.4 Discussion: Relationship to Reimer's butterfly lemma

The antipodal routing lemma does not require $x \mapsto z(x)$ to be a permutation of A, but our proof produces one. We do not need this property, but we also do not know of a simpler proof that does not give a permutation. With the permutation requirement included, condition (a) alone is equivalent to Reimer's *butterfly lemma* [Rei00] (see also [KSS11]).[6] The name is explained by Figure 3: the point $x^{\oplus T}$ is the *body* of a butterfly, and the subcubes $\text{cube}(x, x^{\oplus T})$ and $\text{cube}(x^{\oplus T}, y)$, which meet only at the body, are its *wings*. Our proof of the antipodal routing lemma therefore strengthens the butterfly lemma by adding the coordinate-wise bound (b).

Reimer used the butterfly lemma to prove a conjecture of van den Berg and Kesten [vdBK85]. They had shown that two monotone events occur disjointly with probability at most the product of their probabilities (the BK inequality), and conjectured that this holds for all events. Reimer's theorem is now known as the van den Berg–Kesten–Reimer (BKR) inequality. The *dual* BKR inequality of Kahn, Saks, and Smyth [KSS11], discussed in Section 2.3.1, is also proved using the butterfly lemma.

Reimer's proof is also linear-algebraic, but entirely different from ours. Each point x is assigned a vector in $\mathbb{F}_2^{2^n}$ that is a tensor product over the coordinates, the factor at coordinate i depending on x_i and on whether $i \in T$. The lemma follows from the linear independence of these vectors. See [BCR99] for details.

Table 1 summarizes how the antipodal routing lemma relates to versions where either (a) or (b) is dropped.

Conditions	**Statement**	**Proof**
(a) + (b)	Antipodal routing lemma	determinants over $\mathbb{Z}[\zeta]$; gives a permutation
(a)	Trivial	$z(x) = x^{\oplus T}$
(a) + permutation	Reimer's butterfly lemma	linear algebra over $\mathbb{F}_2$
(b)	Faceless version (Proposition 2.12)	elementary; gives a permutation

Table 1: The antipodal routing lemma and related statements.

3 Proof of the antipodal routing lemma

In this section, we prove the antipodal routing lemma (Lemma 2.11), restated here in an equivalent but more convenient form:

[6] Without the permutation requirement, condition (a) alone is easily achieved via the map $x \mapsto x^{\oplus T}$.

Lemma 3.1 (Antipodal routing lemma)**.** *Let $A \subseteq \{0,1\}^n$ and for every $x \in A$, let $T(x) \subseteq [n]$ be a monochromatic face of x. Write $\mathcal{T} \coloneqq (T(x))_{x\in A}$. Then there is $\pi = \pi_{\mathcal{T}} : A \to A$ satisfying*

(a) *$\pi(x)_i \neq x_i$ for every $x \in A$ and $i \in T(x)$.*

(b) *For every $i \in [n]$,*

$$\operatorname{errors}_i(\pi) \leq E_i$$

where $\operatorname{errors}_i(\pi) \coloneqq \sum_{x\in A} \mathbb{1}[\pi(x)_i = x_i]$ *and* $E_i \coloneqq \sum_{x\in A} \mathbb{1}[x^{\oplus i} \notin A]$.

We recall the definitions of the two matrices M and $M_{\mathcal{T}}$ from the technical overview (Section 2.5.2):

$$M[x,z] \coloneqq \prod_{i\in[n]} \zeta_i^{\mathbb{1}[x_i=z_i]} \quad \text{and} \quad M_{\mathcal{T}}[x,z] \coloneqq \prod_{i\in[n]} \zeta_i^{\mathbb{1}[x_i=z_i]} \cdot \prod_{i\in T(x)} \mathbb{1}[x_i \neq z_i].$$

These satisfy

$$\det M = \sum_{\text{permutations } \pi \text{ of } A} \operatorname{sign}(\pi) \prod_{i\in[n]} \zeta_i^{\operatorname{errors}_i(\pi)}$$

$$\det M_{\mathcal{T}} = \sum_{\text{permutations } \pi \text{ satisfying (a)}} \operatorname{sign}(\pi) \prod_{i\in[n]} \zeta_i^{\operatorname{errors}_i(\pi)}.$$

The key lemma is the following, restated for convenience.

Lemma 2.16 (Determinant equivalence lemma)**.** *Let $\mathcal{T} = (T(x))_{x\in A}$ specify monochromatic faces for each $x \in A$. Then*

$$\operatorname{minimal}(\det M) = \operatorname{minimal}(\det M_{\mathcal{T}}).$$

3.1 Proof of Lemma 3.1 using Lemma 2.16

The first step is the following.

Proposition 3.2. *Every monomial in* $\operatorname{minimal}(\det M)$ *has ζ_i-degree at most E_i for all $i \in [n]$.*

Proof. We will show this separately for each $i \in [n]$. Fixing any such i, we define the monochromatic faces $\mathcal{T}^{(i)} = (T^{(i)}(x))_{x\in A}$ by

$$T^{(i)}(x) = \begin{cases} \{i\} & \text{if } x^{\oplus i} \in A \\ \varnothing & \text{otherwise.} \end{cases}$$

Then $\det M_{\mathcal{T}^{(i)}}$ only sums over permutations where $\pi(x)_i \neq x_i$ for every x where $x^{\oplus i} \in A$. For such permutations

$$\operatorname{errors}_i(\pi) = \sum_{x\in A} \mathbb{1}[\pi(x)_i = x_i] \leq \sum_{x\in A} \mathbb{1}[x^{\oplus i} \notin A] = E_i.$$

Therefore, every monomial in $\det M_{\mathcal{T}^{(i)}}$, and hence every monomial in $\operatorname{minimal}(\det M_{\mathcal{T}^{(i)}})$, has ζ_i-degree at most E_i. By Lemma 2.16, this is also true of $\operatorname{minimal}(\det M)$. □

We next observe that:

Proposition 3.3. *The polynomial* $\text{minimal}(\det M)$ *is nonzero.*

Proof. This is equivalent to showing that $\det M$ is nonzero. For this, observe that the identity permutation $I(x) = x$ is the only permutation where $\text{errors}_i(I) = |A|$ for all $i \in [n]$. It therefore cannot cancel with any other permutation in the determinant of M. □

The antipodal routing lemma, Lemma 3.1, follows easily from these ingredients. Combining Propositions 3.2 and 3.3, we have that $\text{minimal}(\det M)$ must have a nonzero monomial with ζ_i-degree at most E_i for all $i \in [n]$. By Lemma 2.16, this is also true of $\text{minimal}(\det M_{\mathcal{T}})$, which can only occur if there is a permutation satisfying all the constraints.

3.2 Proof of Lemma 2.16

We first show the following:

Proposition 2.17. *For any polynomials* p, q, *if* $p(0, \dots, 0) = 1$ *then* $\text{minimal}(pq) = \text{minimal}(q)$.

Proof. Since $p(0) = 1$, it can be written as $1 + p'$ where every monomial of p' has degree at least 1. The Pareto-minimal monomials of $pq = q + p'q$ are exactly those of q. □

The next step is the following expression linking M and $M_{\mathcal{T}}$.

Claim 3.4. *For any monochromatic faces* $T(x)$ *for every* $x \in A$, *let* D *be the* $A \times A$ *diagonal matrix with* $D[x, x] = \prod_{i \in T(x)} (1 - \zeta_i^2)$ *and* F *be the* $A \times A$ *matrix with*

$$F[x, x^{\oplus S}] = \begin{cases} \prod_{i \in S} (-\zeta_i) & \text{if } S \subseteq T(x), \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$DM_{\mathcal{T}} = FM.$$

Proof. We wish to show that, for any $x, z \in A$,

$$(DM_{\mathcal{T}})[x, z] = (FM)[x, z].$$

Expanding matrix multiplication, this is equivalent to showing that

$$\prod_{i \in T(x)} (1 - \zeta_i^2) M_{\mathcal{T}}[x, z] = \sum_{S \subseteq T(x)} \mathbb{1}[x^{\oplus S} \in A] \cdot \left(\prod_{i \in S} (-\zeta_i) \right) M[x^{\oplus S}, z],$$

where the indicator $\mathbb{1}[x^{\oplus S} \in A]$ appears because M only has rows corresponding to elements of A. Here, we use the assumption that $T(x)$ is a monochromatic face of x, meaning this condition is always satisfied. The goal is therefore to show

$$\prod_{i \in T(x)} (1 - \zeta_i^2) M_{\mathcal{T}}[x, z] = \sum_{S \subseteq T(x)} \left(\prod_{i \in S} (-\zeta_i) \right) M[x^{\oplus S}, z]. \tag{5}$$

Warm-up: single index. First, consider the case where $T(x) = \{i\}$. Then our desired equation is equivalent to

$$(1 - \zeta_i^2) M_{\mathcal{T}}[x, z] = M[x, z] - \zeta_i M[x^{\oplus i}, z].$$

For notational convenience, we define

$$M^{-i}[x,z] \coloneqq \prod_{j\neq i} \zeta_j^{\mathbb{1}[x_j=z_j]}$$

to contain the part of $M[x,z]$ that does not depend on the i-th index. Then,

$$\begin{aligned}
&M[x,z] - \zeta_i \cdot M[x^{\oplus i},z] \\
&\quad= \zeta_i^{\mathbb{1}[x_i=z_i]} \cdot M^{-i}[x,z] - \zeta_i \cdot \zeta_i^{\mathbb{1}[x_i\neq z_i]} \cdot M^{-i}[x^{\oplus i},z] && \text{(Definitions of } M,\ M^{-i}, \text{ and } x^{\oplus i}) \\
&\quad= M^{-i}[x,z] \cdot \left(\zeta_i^{\mathbb{1}[x_i=z_i]} - \zeta_i \cdot \zeta_i^{\mathbb{1}[x_i\neq z_i]}\right) && (M^{-i}[x,\cdot] = M^{-i}[x^{\oplus i},\cdot]) \\
&\quad= M^{-i}[x,z] \cdot \mathbb{1}[x_i \neq z_i] \cdot (1-\zeta_i^2) && \text{(Casework on } x_i = z_i \text{ or } x_i \neq z_i) \\
&\quad= (1-\zeta_i^2) M_{\mathcal{T}}[x,z].
\end{aligned}$$

This verifies Equation (5) in this simple case. Underlying the above derivation is the equality

$$\zeta_i^{\mathbb{1}[x_i=z_i]} - \zeta_i \cdot \zeta_i^{\mathbb{1}[x_i\neq z_i]} = \mathbb{1}[x_i \neq z_i] \cdot (1-\zeta_i^2), \tag{6}$$

which we will reuse shortly.

Full case. For arbitrary $T = T(x)$, we will similarly define

$$M^{-T}[x,z] \coloneqq \prod_{i\notin T} \zeta_i^{\mathbb{1}[x_i=z_i]}.$$

With this notation, Equation (5) can be rewritten as

$$M^{-T}[x,z] \prod_{i\in T(x)} \mathbb{1}[x_i \neq z_i](1-\zeta_i^2) = M^{-T}[x,z] \sum_{S\subseteq T(x)} \left(\prod_{j\in S}(-\zeta_j)\right) \prod_{i\in T(x)} \zeta_i^{\mathbb{1}[(x^{\oplus S})_i=z_i]}.$$

It suffices to show this holds after dropping $M^{-T}[x,z]$ from both sides:

$$\begin{aligned}
&\prod_{i\in T(x)} \mathbb{1}[x_i \neq z_i](1-\zeta_i^2) \\
&\quad= \prod_{i\in T(x)} \left(\zeta_i^{\mathbb{1}[x_i=z_i]} - \zeta_i \cdot \zeta_i^{\mathbb{1}[x_i\neq z_i]}\right) && \text{(Equation (6))} \\
&\quad= \sum_{S\subseteq T(x)} \prod_{i\in S}\left(-\zeta_i \cdot \zeta_i^{\mathbb{1}[x_i\neq z_i]}\right) \cdot \prod_{i\in T(x)\setminus S} \left(\zeta_i^{\mathbb{1}[x_i=z_i]}\right) && \text{(Distributive law)} \\
&\quad= \sum_{S\subseteq T(x)} \left(\prod_{i\in S}(-\zeta_i)\right) \cdot \left(\prod_{i\in T(x)} \zeta_i^{\mathbb{1}[(x^{\oplus S})_i=z_i]}\right). && \text{(Collect terms by index)}
\end{aligned}$$

This is the equality we wished to show. □

We now prove the determinant equivalence lemma. We will use just two facts about the determinant:

1. For any square matrices X and Y, $\det XY = \det X \det Y$.

2. The determinant of the identity matrix is 1.

Proof of Lemma 2.16. By Claim 3.4,

$$\det D \det M_{\mathcal{T}} = \det F \det M.$$

At $\zeta_1 = \cdots = \zeta_n = 0$, we have that both D and F become the identity matrix, which has determinant 1. Applying Proposition 2.17 gives that

$$\text{minimal}(\det M_{\mathcal{T}}) = \text{minimal}(\det D \det M_{\mathcal{T}}) = \text{minimal}(\det F \det M) = \text{minimal}(\det M). \qquad \square$$

4 Proof of the routing lemma

In this section we prove the routing lemma (Lemma 2.10) using the antipodal routing lemma (Lemma 2.11).

For $S \subseteq [n]$, we say that a set $Q \subseteq \{0,1\}^n$ is an *S-subcube* if $Q = \text{cube}(u, u^{\oplus S})$ for some $u \in \{0,1\}^n$. Each subset S gives rise to $2^{n-|S|}$ different subcubes, each of which is isomorphic to $\{0,1\}^S$ and which together partition $\{0,1\}^n$. For *every* S and *every* S-subcube Q, we are going to apply the antipodal routing lemma within that subcube.

Applying the antipodal routing lemma within Q. Fix an S-subcube Q. The assumption of the routing lemma gives us, for every bichromatic pair $(x, y) \in A \times \bar{A}$, a monochromatic face $T(x, y) \subseteq \Delta(x, y)$ of x that is contained in A. We use these to define, for each $x \in Q \cap A$, a monochromatic face of x that is contained in $Q \cap A$:

$$T_Q(x) \coloneqq \begin{cases} T(x, x^{\oplus S}) & \text{if } x^{\oplus S} \notin A, \\ \varnothing & \text{otherwise.} \end{cases}$$

Since $T(x, x^{\oplus S}) \subseteq \Delta(x, x^{\oplus S}) = S$, this monochromatic face of x is indeed contained in $Q \cap A$. In the second case, the empty face is trivially monochromatic.

We can therefore apply the antipodal routing lemma to the set $A \cap Q$ in $Q \cong \{0,1\}^S$, with faces $T_Q(x)$. In Q, the antipode of x is $x^{\oplus S}$. Thus, for all $x \in Q \cap A$ the antipodal routing lemma gives a point $z_Q(x) \in Q \cap A$ such that

(a^Q) $z_Q(x) \in \text{cube}\big(x^{\oplus T_Q(x)}, x^{\oplus S}\big)$, and

(b^Q) for every $i \in S$, $\Pr_{\boldsymbol{x} \sim Q}\left[\boldsymbol{x} \in A,\ z_Q(\boldsymbol{x})_i \neq (\boldsymbol{x}^{\oplus S})_i\right] \leq \frac{1}{2} \cdot \text{Inf}_i^Q(A)$

where $\text{Inf}_i^Q(A) \coloneqq \Pr_{\boldsymbol{x} \sim Q}[\mathbb{1}[\boldsymbol{x} \in A] \neq \mathbb{1}[\boldsymbol{x}^{\oplus i} \in A]]$ is the influence of the i-th variable within the subcube Q.

Finally, we are ready to define $z(x, y, T)$ as required by the routing lemma. We associate each bichromatic pair (x, y) with the subcube in which they are antipodal. This corresponds exactly to the $\Delta(x, y)$-subcube $Q = \text{cube}(x, y)$. We let

$$z(x, y, T) \coloneqq z_Q(x) \qquad \text{for } Q = \text{cube}(x, y).$$

Note that $z(x, y, T) \in A$ as required. The remainder of the proof will show that with this choice of z, items (a) and (b) of the routing lemma hold.

Item (a) holds. Since $(x, y) \in A \times \bar{A}$ is a bichromatic pair, the first case in the definition of T_Q applies, and so $T_Q(x) = T(x, y)$. Then, with our choice of $z(x, y, T)$, (a^Q) becomes $z(x, y, T) \in \text{cube}\big(x^{\oplus T(x,y)}, y\big)$ as desired. This comes simply from plugging in $S = \Delta(x, y)$ and noting that $x^{\oplus \Delta(x,y)} = y$.

Item (b) holds. A uniform random pair $(\boldsymbol{x}, \boldsymbol{y})$ can be sampled as follows: draw $\boldsymbol{x} \sim \{0,1\}^n$ and $\boldsymbol{S} \subseteq [n]$ uniformly and independently, and set $\boldsymbol{y} \coloneqq \boldsymbol{x}^{\oplus \boldsymbol{S}}$. Then $\boldsymbol{S} = \Delta(\boldsymbol{x}, \boldsymbol{y})$, and $\boldsymbol{y}$ is the antipode of $\boldsymbol{x}$ in the $\boldsymbol{S}$-subcube containing $\boldsymbol{x}$. Each coordinate lies in $\boldsymbol{S}$ with probability $\frac{1}{2}$. We can therefore rewrite the inequality we wish to prove as:

$$\Pr_{\boldsymbol{x},\boldsymbol{S}}\left[\boldsymbol{x} \in A,\ \boldsymbol{x}^{\oplus \boldsymbol{S}} \notin A,\ \boldsymbol{z}_i \neq (\boldsymbol{x}^{\oplus \boldsymbol{S}})_i\right] \ \leq\ \tfrac{1}{4} \cdot \mathrm{Inf}_i(A).$$

Let E be the event $\{\boldsymbol{x} \in A,\ \boldsymbol{x}^{\oplus \boldsymbol{S}} \notin A,\ \boldsymbol{z}_i \neq (\boldsymbol{x}^{\oplus \boldsymbol{S}})_i\}$ whose probability we wish to bound. By the law of total probability,

$$\begin{aligned}\Pr[E] &= \Pr[i \in \boldsymbol{S}] \cdot \Pr\left[E \,\middle|\, i \in \boldsymbol{S}\right] \ +\ \Pr[i \notin \boldsymbol{S}] \cdot \Pr\left[E \,\middle|\, i \notin \boldsymbol{S}\right] \\ &= \tfrac{1}{2} \cdot \Pr\left[E \,\middle|\, i \in \boldsymbol{S}\right]. \end{aligned} \tag{7}$$

The second equality follows from the fact that if $i \notin \boldsymbol{S}$, then $\boldsymbol{z}$ and $\boldsymbol{y} = \boldsymbol{x}^{\oplus \boldsymbol{S}}$ both lie in the $\boldsymbol{S}$-subcube containing $\boldsymbol{x}$, whose points all agree outside $\boldsymbol{S}$. Hence $\boldsymbol{z}_i = \boldsymbol{y}_i$, and so $\Pr[E \mid i \notin \boldsymbol{S}] = 0$.

It remains to show that $\Pr[E \mid i \in \boldsymbol{S}] \leq \frac{1}{2}\mathrm{Inf}_i(A)$. Fix any $S \ni i$. Given $\boldsymbol{S} = S$, the point $\boldsymbol{x}$ is uniform on $\{0,1\}^n$, which is the same as drawing a uniform S-subcube $\boldsymbol{Q}$ and then a uniform $\boldsymbol{x} \sim \boldsymbol{Q}$. Conditioned on $\boldsymbol{Q} = Q$, we have $\boldsymbol{z} = z_Q(\boldsymbol{x})$ and $\boldsymbol{y} = \boldsymbol{x}^{\oplus S}$, so dropping the condition $\boldsymbol{y} \notin A$ and applying (b^Q),

$$\Pr\left[E \,\middle|\, \boldsymbol{S} = S,\ \boldsymbol{Q} = Q\right] \ \leq\ \Pr_{\boldsymbol{x} \sim Q}\left[\boldsymbol{x} \in A,\ z_Q(\boldsymbol{x})_i \neq (\boldsymbol{x}^{\oplus S})_i\right] \ \leq\ \tfrac{1}{2} \cdot \mathrm{Inf}_i^Q(A).$$

Since $i \in S$, every edge $\{x, x^{\oplus i}\}$ lies inside a single S-subcube, and so averaging over $\boldsymbol{Q}$ gives $\mathbb{E}_{\boldsymbol{Q}}\left[\mathrm{Inf}_i^{\boldsymbol{Q}}(A)\right] = \mathrm{Inf}_i(A)$. Hence $\Pr[E \mid \boldsymbol{S} = S] \leq \frac{1}{2} \cdot \mathrm{Inf}_i(A)$ for every $S \ni i$, and therefore $\Pr[E \mid i \in \boldsymbol{S}] \leq \frac{1}{2} \cdot \mathrm{Inf}_i(A)$. Combining this with Equation (7) completes the proof.

5 Proof of the weighted all-pairs inequality

We prove the weighted all-pairs inequality (Lemma 2.8) assuming the routing lemma (Lemma 2.10). The argument is identical to the unweighted case sketched in the technical overview (Section 2.4.2), suitably modified to take the weights into account.

Let $A \coloneqq f^{-1}(1)$. For any pair $x \in A$ and $y \notin A$, let $\tau = \mathrm{overlap}_{w,\Lambda}(x, y)$ and $T = \{i \in \Delta(x, y) \colon w_i \lambda_y(i) > \tau\}$ be the heavy coordinates of λ_y weighted according to w. By the same Markov argument as given in the technical overview, our choice of τ ensures $\lambda_x(T) < 1$, which in turn ensures that T is a monochromatic face of x.

Applying the routing lemma with this choice of monochromatic faces and taking a w-weighted sum over $i \in [n]$ in condition (b), we get

$$\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\mathbb{1}[\boldsymbol{x} \in A, \boldsymbol{y} \notin A] \cdot \mathrm{dist}_w(\boldsymbol{z}, \boldsymbol{y})\right] \leq \frac{1}{4}\sum_{i=1}^{n} \frac{\mathrm{Inf}_i(A)}{w_i}, \tag{8}$$

where $\mathrm{dist}_w(z,y) \coloneqq \sum_{i\in\Delta(z,y)} \frac{1}{w_i}$ is the w-weighted Hamming distance between z and y.

We lower bound the LHS of Equation (8) for each (x,y) pair. By condition (a) of the routing lemma, the point z satisfies $\Delta(z,y) \subseteq \Delta(x,y) \setminus T$, and so

$$\lambda_y(i) \le \frac{\tau}{w_i} \quad \text{for all } i \in \Delta(z,y).$$

And yet $z \in A$, so these coordinates must furthermore satisfy $\sum_{i\in\Delta(z,y)} \lambda_y(i) \ge 1$. Putting these two observations together,

$$1 \le \sum_{i\in\Delta(z,y)} \lambda_y(i) \le \sum_{i\in\Delta(z,y)} \frac{\tau}{w_i} = \tau \cdot \mathrm{dist}_w(z,y).$$

By our choice of τ, we have shown that

$$\mathrm{dist}_w(z,y) \ge \frac{1}{\mathrm{overlap}_{w,\Lambda}(x,y)}.$$

Averaging this inequality over $\boldsymbol{x}, \boldsymbol{y}$ and combining it with Equation (8) yields the weighted all-pairs inequality.

6 Proof of the certificate-influence inequality

We prove the certificate-influence inequality (Lemma 2.5) using the weighted all-pairs inequality (Lemma 2.8). Throughout this proof, let $\boldsymbol{x}$ be uniform on $f^{-1}(1)$ and $\boldsymbol{y}$ independently uniform on $f^{-1}(0)$. The certificate-influence inequality can be written as:

$$\underbrace{4\Pr[f=1]\Pr[f=0]}_{=\mathrm{Var}(f)} \cdot \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{1}{\|\lambda_{\boldsymbol{x}}\|_1}\right] \cdot \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{1}{\|\lambda_{\boldsymbol{y}}\|_1}\right] \le \max_{i\in[n]} \mathrm{Inf}_i(f) =: M. \tag{9}$$

In this notation, the weighted all-pairs inequality is

$$4\,\mathrm{Var}(f) \cdot \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\frac{1}{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}\right] \le \sum_{i\in[n]} \frac{\mathrm{Inf}_i(f)}{w_i} \le \sum_{i\in[n]} \frac{1}{w_i} \cdot M. \tag{10}$$

We will use Cauchy–Schwarz to upper bound the left side of our goal with two terms, one of which we can bound using Lemma 2.8.

$$\begin{aligned}
\left(\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\tfrac{1}{\|\lambda_{\boldsymbol{x}}\|_1}\right] \cdot \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\tfrac{1}{\|\lambda_{\boldsymbol{y}}\|_1}\right]\right)^2 &= \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\tfrac{1}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right]^2 && \text{(Independence)}\\
&= \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right)\cdot\left(\frac{1}{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}\right)\right]^2\\
&\le \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right)^2\right]\cdot\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{1}{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}\right)^2\right]\\
& && \text{(Cauchy–Schwarz)}\\
&\le \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right)^2\right]\cdot\sum_{i\in[n]}\frac{1}{w_i}\cdot\frac{M}{4\,\mathrm{Var}(f)}. && \text{(Equation (10))}
\end{aligned}$$

We now upper bound the first term in the above expression:

$$\begin{aligned}
\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right)^2\right] &= \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\frac{\sum_{i\in\Delta(\boldsymbol{x},\boldsymbol{y})} w_i\lambda_{\boldsymbol{x}}(i)\lambda_{\boldsymbol{y}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2\cdot\|\lambda_{\boldsymbol{y}}\|_1^2}\right] && \text{(Definition of overlap)}\\
&\le \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\frac{\sum_{i\in[n]} w_i\lambda_{\boldsymbol{x}}(i)\lambda_{\boldsymbol{y}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2\cdot\|\lambda_{\boldsymbol{y}}\|_1^2}\right] && \text{(Sum over all indices)}\\
&= \sum_{i\in[n]} w_i\cdot\mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\frac{\lambda_{\boldsymbol{x}}(i)\lambda_{\boldsymbol{y}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2\cdot\|\lambda_{\boldsymbol{y}}\|_1^2}\right] && \text{(Linearity of expectation)}\\
&= \sum_{i\in[n]} w_i\cdot\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{\lambda_{\boldsymbol{x}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2}\right]\cdot\mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{\lambda_{\boldsymbol{y}}(i)}{\|\lambda_{\boldsymbol{y}}\|_1^2}\right]. && \text{(Independence of } \boldsymbol{x} \text{ and } \boldsymbol{y})
\end{aligned}$$

We will choose $w_i = \left(\mathbb{E}_{\boldsymbol{x}}\left[\frac{\lambda_{\boldsymbol{x}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2}\right]\right)^{-1}$ to cancel the first term.[7] Combining these expressions, we then have that

$$\begin{aligned}
\left(\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{1}{\|\lambda_{\boldsymbol{x}}\|_1}\right]\cdot\mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{1}{\|\lambda_{\boldsymbol{y}}\|_1}\right]\right)^2 &\le \mathop{\mathbb{E}}_{\boldsymbol{x},\boldsymbol{y}}\left[\left(\frac{\sqrt{\mathrm{overlap}_{w,\Lambda}(\boldsymbol{x},\boldsymbol{y})}}{\|\lambda_{\boldsymbol{x}}\|_1\cdot\|\lambda_{\boldsymbol{y}}\|_1}\right)^2\right]\cdot\sum_{i\in[n]}\frac{1}{w_i}\cdot\frac{M}{4\,\mathrm{Var}(f)}\\
&\le \sum_{i\in[n]}\mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{\lambda_{\boldsymbol{y}}(i)}{\|\lambda_{\boldsymbol{y}}\|_1^2}\right]\cdot\sum_{i\in[n]}\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{\lambda_{\boldsymbol{x}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2}\right]\cdot\frac{M}{4\,\mathrm{Var}(f)}\\
&= \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{\|\lambda_{\boldsymbol{y}}\|_1}{\|\lambda_{\boldsymbol{y}}\|_1^2}\right]\cdot\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{\|\lambda_{\boldsymbol{x}}\|_1}{\|\lambda_{\boldsymbol{x}}\|_1^2}\right]\cdot\frac{M}{4\,\mathrm{Var}(f)} && \text{(Linearity of expectation)}\\
&= \mathop{\mathbb{E}}_{\boldsymbol{y}}\left[\frac{1}{\|\lambda_{\boldsymbol{y}}\|_1}\right]\cdot\mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\frac{1}{\|\lambda_{\boldsymbol{x}}\|_1}\right]\cdot\frac{M}{4\,\mathrm{Var}(f)}.
\end{aligned}$$

Rearranging gives Equation (9), which is what we wanted to show.

7 Proof of the AA conjecture

We now prove the AA conjecture (Theorem 1) using the certificate-influence inequality (Lemma 2.5). We in fact prove a more general statement that gives a lower bound on the maximum influence of an arbitrary function $p\colon\{0,1\}^n\to\mathbb{R}$ in terms of its fractional certificate complexity, with no assumptions on degree or boundedness:

Theorem 2 (AA-like statement for arbitrary functions)**.** *Let* $p\colon\{0,1\}^n\to\mathbb{R}$ *be nonconstant,* $\mu \coloneqq \mathbb{E}[p(\boldsymbol{x})]$*, and* $\sigma \coloneqq \mathbb{E}[|p(\boldsymbol{x})-\mu|]$*. Then*

$$\max_{i\in[n]}\mathrm{Inf}_i(p) \;\ge\; \Omega\left(\frac{\sigma^8}{\mathrm{FC}(p)^6}\right).$$

Proof of Theorem 1 (the AA conjecture) assuming Theorem 2. Let $p:\{0,1\}^n\to[0,1]$ be a bounded degree-d polynomial. By Fact 2.4, $\mathrm{FC}(p)=O(d^2)$ (this uses both assumptions of low degree and boundedness). Since p takes values in $[0,1]$, we have that $|p-\mu|\le 1$, hence $|p-\mu|\ge(p-\mu)^2$, and so $\sigma\ge\mathrm{Var}(p)$ (this uses only boundedness). Theorem 1 now follows from Theorem 2. □

[7] To avoid dividing by 0, we formally set $w_i = \left(\varepsilon+\mathbb{E}_{\boldsymbol{x}}\left[\frac{\lambda_{\boldsymbol{x}}(i)}{\|\lambda_{\boldsymbol{x}}\|_1^2}\right]\right)^{-1}$ and take the limit as $\varepsilon\to 0$. The resulting proof is identical up to bookkeeping.

7.1 Proof of Theorem 2

Fix a fractional certificate family $\Lambda = (\lambda_x)_{x\in\{0,1\}^n}$ for p of cost $\mathrm{FC}(p)$. For a threshold $t \in \mathbb{R}$, let $p_t(x) \coloneqq \mathbb{1}[p(x) > t]$.

We first claim that for each $x \in \{0,1\}^n$, the scaled vector

$$\frac{\lambda_x}{|p(x) - t|}$$

is a fractional certificate for p_t at x.[8] Indeed, if $p_t(x) \neq p_t(y)$, then $p(x)$ and $p(y)$ lie on opposite sides of t, and so

$$\lambda_x(\Delta(x,y)) \;\geq\; |p(x) - p(y)| \;\geq\; |p(x) - t|.$$

Hence, whenever p_t is nonconstant, we can apply the certificate-influence inequality to get

$$\max_{i\in[n]} \mathrm{Inf}_i(p_t) \geq 4 \cdot \frac{\mathbb{E}\left[(p(\boldsymbol{x}) - t)_+\right] \mathbb{E}\left[(t - p(\boldsymbol{y}))_+\right]}{\mathrm{FC}(p)^2}. \tag{11}$$

Since $\mathbb{E}[(p(\boldsymbol{x}) - \mu)_+] = \mathbb{E}[(\mu - p(\boldsymbol{x}))_+]$ and they sum to σ, each equals $\sigma/2$. The maps $t \mapsto \mathbb{E}[(p(\boldsymbol{x}) - t)_+]$ and $t \mapsto \mathbb{E}[(t - p(\boldsymbol{x}))_+]$ are 1-Lipschitz, so for $t \in [\mu - \frac{\sigma}{4}, \mu + \frac{\sigma}{4}]$ both are at least $\sigma/4$. In particular, for any such t we have that p_t is nonconstant and the numerator on the RHS of Equation (11) is $\Omega(\sigma^2)$.

Next, we relate $\mathrm{Inf}_i(p_t)$ and $\mathrm{Inf}_i(p)$. We first write:

$$\mathrm{Inf}_i(p_t) = 2 \Pr\left[p(\boldsymbol{x}^{\oplus i}) \leq t < p(\boldsymbol{x})\right].$$

Fix $\theta > 0$. If $p(x^{\oplus i}) \leq t < p(x)$, then either $p(x)$ lies in the window $(t, t+\theta]$, or the jump $p(x) - p(x^{\oplus i})$ exceeds θ. The probability of the latter is at most

$$\Pr\left[p(\boldsymbol{x}) - p(\boldsymbol{x}^{\oplus i}) > \theta\right] = \tfrac{1}{2}\Pr\left[|p(\boldsymbol{x}) - p(\boldsymbol{x}^{\oplus i})| > \theta\right] \leq \frac{\mathrm{Inf}_i(p)}{2\theta^2}.$$

The equality is by symmetry and the inequality is Markov's inequality applied to the quantity $(p(\boldsymbol{x}) - p(\boldsymbol{x}^{\oplus i}))^2$. Therefore,

$$\mathrm{Inf}_i(p_t) \leq 2 \Pr\left[t < p(\boldsymbol{x}) \leq t + \theta\right] + \frac{\mathrm{Inf}_i(p)}{\theta^2}.$$

Combining Equation (11) with this window-jump argument, we have shown that for every $t \in [\mu - \frac{\sigma}{4}, \mu + \frac{\sigma}{4}]$ and $\theta > 0$,

$$\frac{M}{\theta^2} + 2 \Pr\left[t < p(\boldsymbol{x}) \leq t + \theta\right] \geq \max_{i\in[n]} \mathrm{Inf}_i(p_t) \geq \Omega\left(\frac{\sigma^2}{\mathrm{FC}(p)^2}\right), \tag{12}$$

where $M \coloneqq \max_{i\in[n]} \mathrm{Inf}_i(p)$. Now let $\boldsymbol{t}$ be uniform on $[\mu - \frac{\sigma}{4}, \mu + \frac{\sigma}{4}]$, an interval of length $\sigma/2$, and note that

$$\mathop{\mathbb{E}}_{\boldsymbol{t}}\left[\Pr_{\boldsymbol{x}}\left[\boldsymbol{t} < p(\boldsymbol{x}) \leq \boldsymbol{t} + \theta\right]\right] \;=\; \mathop{\mathbb{E}}_{\boldsymbol{x}}\left[\Pr_{\boldsymbol{t}}\left[p(\boldsymbol{x}) - \theta \leq \boldsymbol{t} < p(\boldsymbol{x})\right]\right] \;\leq\; \frac{\theta}{\sigma/2} \;=\; \frac{2\theta}{\sigma}.$$

[8] Strictly speaking, these certificates are undefined when $p(x) = t$. However, since there are only finitely many x's and we will be choosing t uniformly from an interval, this bad event happens with probability 0.

Averaging Equation (12) over $\boldsymbol{t}$ therefore gives, for every $\theta > 0$,

$$\frac{M}{\theta^2} + \frac{4\theta}{\sigma} \geq \Omega\left(\frac{\sigma^2}{\mathrm{FC}(p)^2}\right).$$

Finally, choosing $\theta = (\sigma M)^{1/3}$ and rearranging, we get:

$$\max_{i\in[n]} \mathrm{Inf}_i(p) \geq \Omega\left(\frac{\sigma^8}{\mathrm{FC}(p)^6}\right).$$

8 Acknowledgments

We thank PSI, and in particular David Turturean and Jim Halverson, for sharing the proof with us and suggesting that we write this exposition. We also thank Scott Aaronson for putting the two teams in touch.

Guy, Carmen, and Li-Yang are supported by NSF awards 1942123, 2211237, 2224246, a Sloan Research Fellowship, and a Google Research Scholar Award. Jordan is supported by NSF awards 2016245 and 2440805 and by the Air Force Office of Scientific Research under grant agreement FA9550-21-1-0392. Guy is also supported by a Norbert Wiener Postdoctoral Fellowship and Omer Reingold's Simons Foundation Investigator Award. Jordan is also supported by the Shoucheng Zhang Graduate Fellowship. Carmen is also supported by a Stanford Computer Science Distinguished Fellowship, an NSF GRFP, Omer Reingold's Simons Foundation Investigator Award, and a Jump Trading PhD Fellowship. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of Jump Trading.

References

[AA14] Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. *Theory of Computing*, 10(6):133–166, 2014. Preliminary version in ICS 2011. 1, 1.2

[Aar05] Scott Aaronson. Ten semi-grand challenges for quantum computing theory, 2005. https://www.scottaaronson.com/writings/qchallenge.html. 1

[Aar08a] Scott Aaronson. How to solve longstanding open problems in quantum computing using only Fourier analysis, 2008. Lecture at the Banff International Research Station. https://www.scottaaronson.com/talks/openqc.ppt. 1

[Aar08b] Scott Aaronson. Quantum certificate complexity. *Journal of Computer and System Sciences*, 74(3):313–322, 2008. Preliminary version in CCC 2003. 2.1

[ABD26] Avantika Agarwal and Shalev Ben-David. Oracle separations for the quantum-classical polynomial hierarchy. In *Proceedings of the 17th Innovations in Theoretical Computer Science Conference (ITCS 2026)*, pages 2:1–2:22, 2026. 2.1

[ABDK21] Anurag Anshu, Shalev Ben-David, and Srijita Kundu. On query-to-communication lifting for adversary bounds. In *Proceedings of the 36th Computational Complexity Conference (CCC 2021)*, pages 30:1–30:39, 2021. 2.1

[ABP19] Srinivasan Arunachalam, Jop Briët, and Carlos Palazuelos. Quantum query algorithms are completely bounded. *SIAM Journal on Computing*, 48(3):903–925, 2019. Preliminary version in ITCS 2018. 1.2

[BBBV97] Charles Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, 1997. 1

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Preliminary version in FOCS 1998. 1, 2.2.2

[BCR99] Christian Borgs, Jennifer Chayes, and Dana Randall. The van den Berg–Kesten–Reimer inequality: A review. In *Perplexing Problems in Probability: Festschrift in Honor of Harry Kesten*, volume 44 of *Progress in Probability*, pages 159–173. Birkhäuser, 1999. 2.5.4

[BDST26] Guy Blanc, Jordan Docter, Carmen Strassle, and Li-Yang Tan. Quantum speedups require structure or depth. In *Proceedings of the 67th Annual Symposium on Foundations of Computer Science (FOCS 2026)*, 2026. 1, 2, 1.2

[BG81] Charles Bennett and John Gill. Relative to a random oracle A, $\mathsf{P}^A \neq \mathsf{NP}^A \neq \mathsf{co\text{-}NP}^A$ with probability 1. *SIAM Journal on Computing*, 10(1):96–113, 1981. 1, 2

[Bha25] Sreejata Kishor Bhattacharya. Random restrictions of bounded low degree polynomials are juntas. In *Proceedings of the 16th Innovations in Theoretical Computer Science Conference (ITCS 2025)*, pages 17:1–17:21, 2025. 1.2

[BSdW22] Nikhil Bansal, Makrand Sinha, and Ronald de Wolf. Influence in completely bounded block-multilinear forms and classical simulation of quantum algorithms. In *Proceedings of the 37th Computational Complexity Conference (CCC 2022)*, pages 28:1–28:21, 2022. 1.2

[BV97] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997. Preliminary version in STOC 1993. 1

[CCG+94] Richard Chang, Benny Chor, Oded Goldreich, Juris Hartmanis, Johan Håstad, Desh Ranjan, and Pankaj Rohatgi. The random oracle hypothesis is false. *Journal of Computer and System Sciences*, 49(1):24–39, 1994. 1

[DFKO07] Irit Dinur, Ehud Friedgut, Guy Kindler, and Ryan O'Donnell. On the Fourier tails of bounded functions over the discrete cube. *Israel Journal of Mathematics*, 160:389–412, 2007. Preliminary version in STOC 2006. 1.2

[DMP19] Andreas Defant, Mieczysław Mastyło, and Antonio Pérez. On the Fourier spectrum of functions on Boolean cubes. *Mathematische Annalen*, 374(1):653–680, 2019. 1.2

[Esc24] Francisco Escudero Gutiérrez. Influences of Fourier completely bounded polynomials and classical simulation of quantum algorithms. *Chicago Journal of Theoretical Computer Science*, 2024(2), 2024. 1.2

[ESP26] Francisco Escudero Gutiérrez, Miquel Saucedo, and Carlos Palazuelos. Optimal inequalities for completely bounded polynomials and the limitations of quantum query algorithms, 2026. arXiv:2609.05201. 1.2

[For21] Lance Fortnow. Worlds to die harder for: Open oracle questions for the 21st century. *ACM SIGACT News*, 52(3):26–36, 2021. 1

[FR99] Lance Fortnow and John Rogers. Complexity limitations on quantum computation. *Journal of Computer and System Sciences*, 59(2):240–252, 1999. Preliminary version in CCC 1998. 1

[GMNR26] Uma Girish, Kunal Mittal, Barak Nehoran, and Ran Raz. Distributional variants of the Aaronson–Ambainis conjecture, 2026. arXiv:2609.35327. 1.2

[Gro96] Lov Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual Symposium on Theory of Computing (STOC 1996)*, pages 212–219, 1996. 1

[GSS16] Justin Gilmer, Michael Saks, and Srikanth Srinivasan. Composition limits and separating examples for some Boolean function complexity measures. *Combinatorica*, 36(3):265–311, 2016. Preliminary version in CCC 2013. 2.1

[Har85] Juris Hartmanis. Solvable problems with conflicting relativizations. *Bulletin of the European Association for Theoretical Computer Science*, 27:40–49, 1985. 1

[Iva26] Paata Ivanisvili. Polynomial growth of Bohnenblust–Hille constants on the Hamming cube, 2026. arXiv:2609.12427. 1.2

[KSS11] Jeff Kahn, Michael Saks, and Clifford Smyth. The dual BKR inequality and Rudich's conjecture. *Combinatorics, Probability and Computing*, 20(2):257–266, 2011. Preliminary version in CCC 2000. 2.3.1, 2.5.4

[KT16] Raghav Kulkarni and Avishay Tal. On fractional block sensitivity. *Chicago Journal of Theoretical Computer Science*, 2016(8), 2016. 4

[Kur82] Stuart Kurtz. On the random oracle hypothesis. In *Proceedings of the 14th Annual Symposium on Theory of Computing (STOC 1982)*, pages 224–230, 1982. 1

[LM26] Qipeng Liu and Saachi Mutreja. Parallel quantum advantage with limited adaptivity requires structure, 2026. arXiv:2608.20297. 1.2

[LZ23] Shachar Lovett and Jiapeng Zhang. Fractional certificates for bounded functions. In *Proceedings of the 14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, pages 84:1–84:13, 2023. 1.2, 2.1, 2.2.3, 3

[Mon12] Ashley Montanaro. Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):122206, 2012. 1.2

[Nis91] Noam Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991. Preliminary version in STOC 1989. 1.2, 2.2.2

[NS94] Noam Nisan and Mario Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994. Preliminary version in STOC 1992. 1.2

[OSSS05] Ryan O'Donnell, Michael Saks, Oded Schramm, and Rocco Servedio. Every decision tree has an influential variable. In *Proceedings of the 46th Annual Symposium on Foundations of Computer Science (FOCS 2005)*, pages 31–39, 2005. 1.2

[OZ16] Ryan O'Donnell and Yu Zhao. Polynomial bounds for decoupling, with applications. In *Proceedings of the 31st Conference on Computational Complexity (CCC 2016)*, pages 24:1–24:18, 2016. 1.2

[Rei00] David Reimer. Proof of the van den Berg–Kesten conjecture. *Combinatorics, Probability and Computing*, 9(1):27–32, 2000. 2.3.1, 2.5.4

[Rud88] Steven Rudich. *Limits on the Provable Consequences of One-Way Functions*. PhD thesis, University of California, Berkeley, 1988. 2.3.1

[Sho97] Peter Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. Preliminary version in FOCS 1994. 1

[Sim97] Daniel Simon. On the power of quantum computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997. Preliminary version in FOCS 1994. 1

[Smy01] Clifford Smyth. *The Conjectures of Rudich, Tardos, and Kusner*. PhD thesis, Rutgers, The State University of New Jersey, 2001. 2.3.1

[Smy02] Clifford Smyth. Reimer's inequality and Tardos' conjecture. In *Proceedings of the 34th Annual Symposium on Theory of Computing (STOC 2002)*, pages 218–221, 2002. 2.3.1

[STV26] Joseph Slote, Chun-Kai Tseng, and Alexander Volberg. An $m^{2.943}$ Bohnenblust–Hille bound on the Boolean cube, 2026. arXiv:2609.21144. 1.2

[Tal96] Michel Talagrand. A new look at independence. *Annals of Probability*, 24(1):1–34, 1996. 2.2.3

[Tal13] Avishay Tal. Properties and applications of Boolean function composition. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science (ITCS 2013)*, pages 441–454, 2013. 2.1

[Tar89] Gábor Tardos. Query complexity, or why is it difficult to separate $\mathsf{NP}^A \cap \mathsf{coNP}^A$ from P^A by random oracles A? *Combinatorica*, 9(4):385–392, 1989. 2.3.1

[Tur26] David Turturean. A Proof of the Aaronson–Ambainis Conjecture, November 2026. Manuscript. https://psi.inc/research/aaronson-ambainis/. 1.1

[vdBK85] Jacob van den Berg and Harry Kesten. Inequalities with applications to percolation and reliability. *Journal of Applied Probability*, 22(3):556–569, 1985. 2.5.4

[YZ24] Takashi Yamakawa and Mark Zhandry. Verifiable quantum advantage without structure. *Journal of the ACM*, 71(3):20, 2024. Preliminary version in FOCS 2022. 1.2