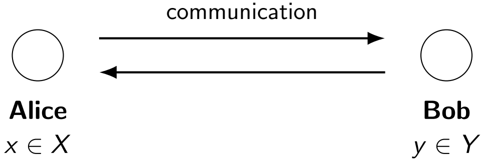
Progress on the Log-Rank Conjecture
AND and XOR Functions

Yogesh Dahiya
UCSD

How does low algebraic complexity force large combinatorial structure?

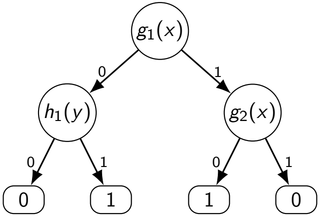
Communication Model



$$F : X \times Y \to \{0, 1\}, \qquad M_F[x, y] = F(x, y).$$

$D^{cc}(F)$ is the minimum number of bits needed in the worst case.

Protocols and Rectangles



A deterministic protocol partitions $X \times Y$ into monochromatic rectangles.

$$X \times Y = \bigsqcup_i R_i, \qquad R_i = A_i \times B_i.$$

If the protocol has cost c, then the number of rectangles is at most 2^c.

Rank: The First Lower Bound Technique

Suppose M_F is partitioned into 2^c monochromatic rectangles $R_i = A_i \times B_i$.

Then

$$M_F = \sum_{i=1}^{2^c} a_i \, \mathbf{1}_{A_i} \mathbf{1}_{B_i}^{\mathsf{T}}, \qquad a_i \in \{0, 1\}.$$

Each summand has rank at most 1. Hence $\operatorname{rank}(M_F) \leq 2^c$. Therefore

$$\boxed{\log \operatorname{rank}(M_F) \ \leq \ D^{cc}(F)}$$

Mehlhorn–Schmidt (1979)

Remark. The rank bound holds over any field. To maximize rank, choose rank over $\mathbb{R}$.

The Log-Rank Conjecture

Lovász–Saks (1988)

Log-rank conjecture. For some $c = O(1)$, every Boolean function F,

$$D^{cc}(F) \leq (\log \operatorname{rank}(M_F))^c.$$

Equivalently, every low-rank Boolean matrix should admit a partition into $2^{\mathrm{poly}(\log \operatorname{rank}(M_F))}$ monochromatic rectangles.

Algebraic simplicity should imply combinatorial simplicity.

Many Faces of the Log-Rank Conjecture

The conjecture has several equivalent formulations.

Communication Complexity

[Lovász–Saks '88]

$D^{cc}(M) \leq \mathrm{qpoly}(\mathsf{rank}(M))$

Graph Theory

[Nuffelen '76; Fajtlowicz '88]

$\chi(G) \leq \mathrm{qpoly}(\mathsf{rank}(A_G))$

Algebra

[Yannakakis '91]

$\mathsf{rank}_+(M) \leq \mathrm{qpoly}(\mathsf{rank}(M))$

Discrete Geometry

[Singer–Sudan '22]

Point–hyperplane arrangements

Low rank forces combinatorial structure.

Signed vs. Unsigned Rectangle Decompositions

Low rank gives a signed rectangle decomposition.

[Hambardzumyan–Lovett–Shirley '26]

$$M = \sum_{i=1}^{r'} \alpha_i R_i, \qquad \alpha_i \in \{\pm 1\}, \qquad r' = O(r \log r).$$

A protocol gives an unsigned rectangle decomposition.

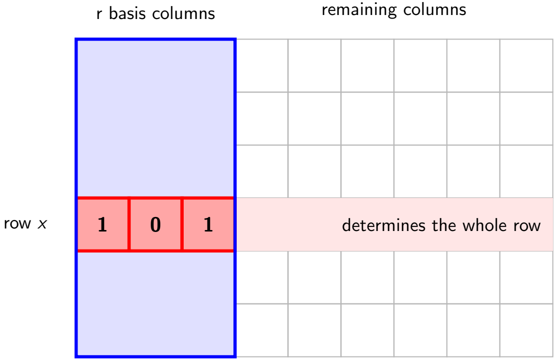
$$M = \sum_{i=1}^{s} R_i, \qquad s = 2^{D^{cc}(M)}.$$

Can we remove the signs with only quasipolynomial blow-up?

How much power does cancellation buy us?

A Trivial Upper Bound

$$D^{cc}(F) \leq \mathrm{rank}(M_F) + 1$$



Lovett's Square-Root Rank Bound

Theorem [Lovett '14]

$$D^{cc}(F) \leq O\left(\sqrt{\mathrm{rank}(M_F)}\,\log\mathrm{rank}(M_F)\right)$$

Essentially best-known general bound.

Nisan–Wigderson Formulation

If the log-rank conjecture holds, then every rank-r Boolean matrix contains a monochromatic rectangle of relative size $2^{-\operatorname{poly}(\log r)}$.

So this structural statement is clearly necessary.

Nisan–Wigderson: it is also sufficient.

If every rank-r Boolean matrix contains such a rectangle, then

$$D^{cc}(F) \leq \operatorname{poly}(\log r).$$

Nisan–Wigderson Formulation

Thm. If every rank-r Boolean matrix contains a monochromatic rectangle of relative size $2^{-\operatorname{poly}(\log r)}$, then the log-rank conjecture holds.

Proof.

Nisan–Wigderson Formulation

Log-rank conjecture:

Every rank r Boolean matrix contains a monochromatic rectangle R

$$\frac{|R|}{|M|} \geq 2^{-\operatorname{poly}(\log r)}$$

Lovett's Theorem

Every rank r Boolean matrix $M \in \{0,1\}^{N \times N}$ contains a monochromatic rectangle R with

$$\frac{|R|}{|M|} \geq 2^{-O(\sqrt{r} \log r)}.$$

Proof Idea

Low rank	$\Longrightarrow$ low rank forces $\mathrm{disc}(M) \gtrsim 1/\sqrt{r}$	**Noticeable discrepancy**
Noticeable discrepancy	$\Longrightarrow$ amplify the bias by recursively reweighting and restricting	**Almost monochromatic**
Almost monochromatic	$\Longrightarrow$ use low rank to turn approximate structure into exact structure	**Monochromatic**

Discrepancy

Let $M \in \{0,1\}^{X \times Y}$, and let μ be a distribution on its entries.

For a rectangle $R \subseteq X \times Y$, define
$$\mathrm{disc}_\mu(R) := \left|\mu(R \cap M^{-1}(1)) - \mu(R \cap M^{-1}(0))\right|.$$

The discrepancy of M is
$$\mathrm{disc}(M) := \min_\mu \ \max_R \mathrm{disc}_\mu(R).$$

Thus, if $\mathrm{disc}(M) \geq \delta$, then for every distribution μ there is a rectangle R with $\mathrm{disc}_\mu(R) \geq \delta$.

Low Rank Forces Noticeable Discrepancy.

$$\mathrm{rank}(M) = r \quad \Longrightarrow \quad \mathrm{disc}(M) \geq \Omega\left(\frac{1}{\sqrt{r}}\right)$$

Linial–Mendelson–Schechtman–Shraibman ’07

Idea: find a biased rectangle and recurse inside it, amplifying the bias at each step.

Two Obstacles

1. **Controlling the direction.**
   Discrepancy gives a rectangle biased in *some* direction. How do we ensure that every step increases the bias toward the same sign?

2. **Controlling the size loss.**
   Every amplification step shrinks the rectangle. How do we avoid recursing all the way until the bias reaches 1?

Controlling the Direction

For a rectangle R, define its bias by $\mathrm{bias}(R) := \frac{|R^{-1}(1)| - |R^{-1}(0)|}{|R|}$.

Suppose the current rectangle R has positive bias $\alpha \geq 0$.

Choose a distribution μ that reweights entries of R so that $\mathrm{bias}_\mu(R) = 0$.

By discrepancy, some rectangle $A \times B \subset R$ satisfies
$\left|\mathrm{bias}_\mu(A \times B)\right| \geq \Omega(1/\sqrt{r})$.

$A \times B$	$A \times \overline{B}$
$\overline{A} \times B$	$\overline{A} \times \overline{B}$

If $A \times B$ is biased in the wrong direction, then—since $\mathrm{bias}_\mu(R) = 0$—one of the other three rectangels is biased in the desired direction.

Amplifying Bias

Claim. Let discrepancy be δ. Suppose the current rectangle R has bias $\alpha \geq 0$, then R contains a rectangle H such that

$$\mathrm{bias}(H) \geq \alpha + \frac{\delta}{3}(1 - \alpha^2)\frac{|R|}{|H|}.$$

There is a tradeoff: a larger loss in size forces a larger increase in bias.

Choose the Distribution

Suppose $\mathrm{bias}(R) = \alpha$. Then

$$|R^{-1}(1)| = \frac{1+\alpha}{2}|R|, \qquad |R^{-1}(0)| = \frac{1-\alpha}{2}|R|.$$

Define a distribution μ on R by giving each 1-entry weight $\frac{1}{|R|(1+\alpha)}$ and each 0-entry weight $\frac{1}{|R|(1-\alpha)}$. The total weight of the 1's and the 0's is $1/2$ each. Hence

$$\mathrm{bias}_\mu(R) = 0.$$

Controlling the Direction

Since the discrepancy is at least δ, there is a rectangle $G = A \times B \subseteq R$ with

$$|\mathrm{bias}_\mu(G)| = |\mu(R^{-1}(1) - \mu(R^{-1}(0))| \geq \delta.$$

$A \times B$	$A \times \overline{B}$
$\overline{A} \times B$	$\overline{A} \times \overline{B}$

If G is biased toward 1, we are done.

If G is biased toward 0, then since $\mathrm{bias}_\mu(R) = 0$, one of the other three rectangles has 1-bias at least $\delta/3$. Thus there is always $H \subseteq R$ with

$$\mathrm{bias}_\mu(H) \geq \delta/3.$$

Back to Ordinary Bias

Let $H_1 := |H^{-1}(1)|$ and $H_0 := |H^{-1}(0)|$.

By the definition of μ,

$$\begin{aligned}
\mathrm{bias}_\mu(H) &= \frac{1}{|R|}\left(\frac{H_1}{1+\alpha} - \frac{H_0}{1-\alpha}\right) \\
&= \frac{(H_1 - H_0) - \alpha(H_1 + H_0)}{|R|(1-\alpha^2)} \\
&= \frac{|H|}{|R|(1-\alpha^2)}\bigl(\mathrm{bias}(H) - \alpha\bigr).
\end{aligned}$$

Since $\mathrm{bias}_\mu(H) \geq \delta/3$, rearranging gives

$$\mathrm{bias}(H) \geq \alpha + \frac{\delta}{3}(1-\alpha^2)\frac{|R|}{|H|}.$$

Halving the Error

Suppose the current rectangle R has $\mathrm{bias}(R) = \alpha = 1 - \beta$.

Claim. There is a subrectangle $H \subseteq R$ such that

$$\mathrm{bias}(H) \geq 1 - \frac{\beta}{2}, \quad \text{while} |H| \geq 2^{-O(1/\delta)} |R|.$$

We can halve the distance from monochromaticity at a cost of only $2^{-O(1/\delta)}$.

Halving the Error: Proof

Repeatedly apply the previous claim:

$$R = R_0 \supset R_1 \supset \cdots \supset R_T,$$

and write $\alpha_i := \mathrm{bias}(R_i)$.

Let T be the first index for which $\alpha_T \geq 1 - \beta/2$.

For every $i < T$, we have $\alpha_i < 1 - \beta/2$, and therefore

$$1 - \alpha_i^2 \geq 1 - \left(1 - \frac{\beta}{2}\right)^2 = \beta - \frac{\beta^2}{4} \geq \frac{\beta}{2}.$$

Hence at every step,

$$\alpha_{i+1} - \alpha_i \geq \frac{\beta\delta}{6} \frac{|R_i|}{|R_{i+1}|}.$$

Controlling the Size Loss

Summing over the T steps,

$$\alpha_T - \alpha_0 \geq \frac{\beta\delta}{6} \sum_{i=0}^{T-1} \frac{|R_i|}{|R_{i+1}|}.$$

Since $\alpha_0 = 1 - \beta$ and $\alpha_T \leq 1$, therefore $\sum_{i=0}^{T-1} \frac{|R_i|}{|R_{i+1}|} \leq \frac{6}{\delta}$.

The total size loss is

$$\frac{|R_0|}{|R_T|} = \prod_{i=0}^{T-1} \frac{|R_i|}{|R_{i+1}|}.$$

By AM–GM,

$$\frac{|R_0|}{|R_T|} \leq \left(\frac{6}{T\delta}\right)^T = \left[\left(\frac{6}{T\delta}\right)^{T\delta/6}\right]^{6/\delta} = 2^{O(1/\delta)}.$$

Halving the Error

Claim. Suppose disc$(M) \geq \delta$ and R is a rectangle with $\mathrm{bias}(R) = 1 - \beta$. Then there exists a subrectangle $H \subseteq R$ such that

$$\mathrm{bias}(H) \geq 1 - \frac{\beta}{2}, \qquad |H| \geq 2^{-O(1/\delta)}|R|.$$

The error is halved at a cost of $2^{-O(1/\delta)}$ in size.

From Discrepancy to Almost Monochromatic

Starting with constant error, repeatedly halve it:

$$1 \longrightarrow \frac{1}{2} \longrightarrow \frac{1}{4} \longrightarrow \cdots \longrightarrow \varepsilon.$$

This requires $O(\log(1/\varepsilon))$ rounds, each costing $2^{-O(1/\delta)}$ in size.

Therefore

$$\boxed{\mathrm{disc}(M) \geq \delta \quad \Longrightarrow \quad \exists R : \begin{cases} \mathrm{bias}(R) \geq 1 - \varepsilon, \\ \dfrac{|R|}{|M|} \geq 2^{-O(\delta^{-1} \log(1/\varepsilon))}. \end{cases}}$$

Almost Monochromatic Is Enough

Lemma. [Gavinsky–Lovett '13]:

Let R be a Boolean matrix of rank at most r with $\mathrm{bias}(R) \geq 1 - \frac{1}{100r}$.
Then R contains a 1-monochromatic rectangle of constant relative size.

1. Keep at least half the rows, each containing at most $1/(10r)$ fraction of 0's.
2. Choose r of these rows that span all remaining rows.
3. Delete at most $1/10$ of the columns, making all r spanning rows identically 1.
4. Every remaining row is constant (all 0 or all 1); it cannot be an all-0 row.

Low rank turns approximate structure into exact structure.

Putting Things Together

Low rank $\Longrightarrow$ **Noticeable discrepancy**

$\mathrm{rank}(M) = r \implies \mathrm{disc}(M) \gtrsim 1/\sqrt{r}$

Noticeable discrepancy $\Longrightarrow$ **Almost monochromatic**

$\mathrm{disc}(M) \geq \delta$ gives R with

$$\frac{|R|}{|M|} \geq 2^{-O(\delta^{-1}\log(1/\varepsilon))},\ \mathrm{bias}(R) \geq 1-\varepsilon$$

Almost monochromatic $\Longrightarrow$ **Monochromatic**

bias $1 - O(1/r)$ suffices

$$\frac{|R_{\mathrm{mono}}|}{|M|} \geq 2^{-O(\sqrt{r}\log r)}.$$

Lovett's Square-Root Rank Bound

By the Nisan–Wigderson formulation, a monochromatic rectangle of relative size $2^{-O(\sqrt{r}\log r)}$ gives

$$D^{cc}(F) \leq O(\sqrt{r}\log r).$$

Structured Communication Matrices

Block-Composed Functions

Let $f : \{0,1\}^n \to \{0,1\}$ and let $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$.

Define

$$F = f \circ g^n, \qquad F(x,y) = f(g(x_1,y_1),\ldots,g(x_n,y_n)).$$

We want to understand log-rank for natural subclasses of this form.

Rich Gadgets Reduce to Query Complexity

If g contains both AND_2 and OR_2 as submatrices, then communication model reduces to query model.

$$g = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 1 \end{array}$$

AND_2

OR_2

$$D^{cc}(f \circ g^n) \approx D^{dt}(f), \qquad \log \operatorname{rank}(M_{f \circ g^n}) \approx \deg(f).$$

Thus log-rank reduces to

$$D^{dt}(f) \overset{?}{\leq} \operatorname{poly}(\deg(f)),$$

which we know.

Two Natural Exceptional Gadgets

The two natural cases that do not reduce to ordinary query complexity are

$$f \circ \mathrm{AND}_2 \qquad\qquad f \circ \mathrm{XOR}_2.$$

AND Functions

AND Functions

For $f : \{0,1\}^n \to \{0,1\}$, define

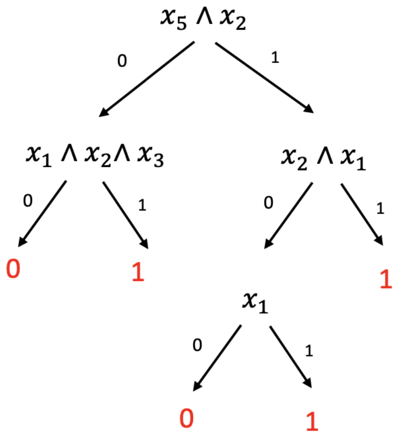
$$F(x, y) = f(x_1 \wedge y_1, \ldots, x_n \wedge y_n).$$

Ordinary decision trees immediately give

$$D^{cc}(f \circ \mathrm{AND}_2) \leq O(D^{dt}(f)).$$

Can be very weak. But for AND_2 we can simulate a stronger query model.

AND Decision Tree



Minimum depth computing f is denoted $D^{\wedge\text{-dt}}(f)$.

Can Simulate AND-Decision Trees

An AND-query asks for

$$\bigwedge_{i \in S} z_i .$$

For $z_i = x_i \wedge y_i$,

$$\bigwedge_{i \in S} z_i = \bigwedge_{i \in S} (x_i \wedge y_i) = \left(\bigwedge_{i \in S} x_i \right) \wedge \left(\bigwedge_{i \in S} y_i \right) .$$

So Alice and Bob can answer one AND-query using 2 bits.

$$D^{cc}(f \circ \mathrm{AND}_2) \leq 2D^{\wedge\text{-}dt}(f).$$

Reducing to AND Decision Trees

$$D^{cc}(f \circ \mathrm{AND}_2) \leq 2\, D^{\wedge\text{-dt}}(f).$$

Thus, to prove log-rank for AND-functions, it is enough to show

$$D^{\wedge\text{-dt}}(f) \leq \mathrm{poly}\big(\log \mathrm{rank}(M_{f \circ \mathrm{AND}_2})\big).$$

Rank Becomes Sparsity

Write the multilinear polynomial of f over the reals:

$$f(z) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} z_i .$$

Substitute $z_i = x_i y_i$:

$$(f \circ \mathrm{AND}_2)(x, y) = \sum_{S \subseteq [n]} a_S \left(\prod_{i \in S} x_i \right) \left(\prod_{i \in S} y_i \right) .$$

The monomials $\prod_{i \in S} x_i$ are linearly independent, thus

$$\mathsf{rank}(f \circ \mathrm{AND}_2) = \mathsf{spar}(f).$$

where $\mathsf{spar}(f)$ is the number of nonzero monomials in the real polynomial of f.

Reducing to AND Decision Trees

For AND-functions, to show log-rank:

$$D^{cc}(f \circ \mathrm{AND}_2) \overset{?}{\leq} \mathsf{poly}(\log \mathsf{rank}(f \circ \mathrm{AND}_2)).$$

it suffices to prove

$$D^{\wedge\text{-}dt}(f) \overset{?}{\leq} \mathsf{poly}(\log \mathsf{spar}(f)).$$

It's a question about AND decision-tree complexity.

Log-Rank for AND-Functions

Knop–Lovett–McGuire–Yuan (2021)

For every total Boolean function f,

$$D^{\wedge\text{-dt}}(f) = O\big((\log \operatorname{spar}(f))^5 \log n\big)\,.$$

Consequently,

$$D^{cc}(f \circ \mathrm{AND}_2) \leq O\big((\log \operatorname{rank}(f \circ \mathrm{AND}_2))^5 \log n\big).$$

Log-rank for AND-functions, up to an extra $\log n$ factor.

Simplifying the Target

Let $D^{\text{0-dt}}(f)$ denote the minimum depth of an ordinary decision tree when we count only 0-answers along a path.

$$D^{\text{0-dt}}(f) \;\leq\; D^{\wedge\text{-dt}}(f) \;\leq\; O\Big(D^{\text{0-dt}}(f)\log n\Big)\,.$$

So, up to a logarithmic factor, it is enough to build a small 0-depth decision tree.

Simplifying the Target

Goal. Show

$$D^{\wedge\text{-dt}}(f) = O\big((\log \mathrm{spar}(f))^5 \log n\big) \, .$$

Use zero-decision trees:

$$D^{0\text{-dt}}(f) \le D^{\wedge\text{-dt}}(f) \le D^{0\text{-dt}}(f) \lceil \log(n+1) \rceil .$$

So it is enough to prove

$$D^{0\text{-dt}}(f) = O\big((\log \mathrm{spar}(f))^5\big) \, .$$

What Would a Small 0-Depth Tree Need?

Suppose along the all-0 path we query variables in H.

If we set all variables in H to 0 that makes f constant.
Which means that every monomial of f must contain some variable from H.

Thus H is a hitting set for the monomials of f.

$$\text{small 0-depth tree} \quad \Longrightarrow \quad \text{small hitting sets for monomials.}$$

We will show that this necessary condition is also sufficient.

Hitting Set Complexity

$$f(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i, \qquad \mathcal{M}(f) := \{S \neq \emptyset : a_S \neq 0\}.$$

A set $H \subseteq [n]$ is a hitting set for f if $H \cap S \neq \emptyset$ for every $S \in \mathcal{M}(f)$.

For a restriction $\rho \in \{0, 1, *\}^n$, let

$$\mathrm{HSC}(f, \rho) := \text{minimum hitting-set size for } f|_\rho.$$

$$\mathrm{HSC}(f) := \max_{\rho \in \{1,*\}^n} \mathrm{HSC}(f, \rho).$$

It is enough to consider $\{1, *\}$-restrictions: setting variables to 0 only deletes monomials.

Small Hitting Set Complexity Give Small Zero-Depth Decision Trees

Assume $\mathrm{HSC}(f) \leq h$.

At the current restriction, choose a hitting set H of size at most h.

Some variable in H appears in at least a $1/h$ fraction of the current monomials.

Query that variable.

$x_i = 0 \quad \Longrightarrow \quad$ kill at least a $1/h$ fraction of monomials.

$x_i = 1 \quad \Longrightarrow \quad$ sparsity does not increase.

Therefore

$$D^{\text{0-dt}}(f) \leq O(h \log \mathrm{spar}(f)).$$

So What Remains?

The target becomes

$$\mathrm{HSC}(f) \leq O\left((\log \mathsf{spar}(f))^{O(1)}\right).$$

So the problem becomes a question about the monomial structure of Boolean functions.

Fractional Version of Hitting Sets

Instead of choosing a hitting set, choose a distribution q over variables.

For a restriction $\rho \in \{1, *\}^n$, define the best hitting probability:

$$\mathrm{HP}(f, \rho) := \max_q \min_{S \in \mathcal{M}(f|_\rho)} \sum_{i \in S} q_i.$$

Then define

$$\mathrm{FHSC}(f, \rho) := \frac{1}{\mathrm{HP}(f, \rho)}, \qquad \mathrm{FHSC}(f) := \max_\rho \mathrm{FHSC}(f, \rho).$$

A hitting set of size h gives hitting probability $1/h$, so

$$\mathrm{FHSC}(f, \rho) \leq \mathrm{HSC}(f, \rho).$$

Fractional to Integral

Fractional hitting sets are easier to analyze, and they are enough.

$$\mathrm{HSC}(f) \le O(\mathrm{FHSC}(f) \log \mathsf{spar}(f))\,.$$

Idea. Sample $O(\mathrm{FHSC}(f) \log \mathsf{spar}(f))$ variables from the fractional distribution.

A union bound over all monomials shows that every monomial is hit.

So it remains to prove

$$\mathrm{FHSC}(f) \le \mathsf{poly}(\log \mathsf{spar}(f)).$$

What We Plan to Do

We want to show

$$\mathrm{FHSC}(f) \leq \mathsf{poly}(\log \mathsf{spar}(f)).$$

We will do following

$$\text{large } \mathrm{FHSC}(f) \implies \text{large Promise-OR structure} \implies \text{large } \mathsf{spar}(f).$$

From Monomials to Sensitive Blocks

For $z \in \{0,1\}^n$, consider only upward moves

$$z \longrightarrow z \vee w, \qquad w \subseteq [n] \setminus \mathrm{supp}(z).$$

Define

$$W(f,z) := \{\, w : \; f(z) \neq f(z \vee w) \,\}.$$

The elements of $W(f,z)$ are the sensitive 0-blocks at z.

For $\rho \in \{1,*\}^n$, let $z_\rho \in \{0,1\}^n$ be the input with $(z_\rho)_i = 1 \iff \rho_i = 1$.

Key observation. The minimal monomials of $f|_\rho$ are exactly the minimal elements of $W(f, z_\rho)$.

So hitting the monomials of $f|_\rho$ is the same as hitting the minimal sensitive 0-blocks at z_ρ.

Equivalent Fractional Hitting Problem

For $\rho \in \{1, *\}^n$,

$$\mathrm{HP}(f, \rho) = \max_{q} \min_{S \in \mathcal{M}(f|_\rho)} \sum_{i \in S} q_i,$$

where q is a distribution over the free coordinates, is same as

$$\mathrm{HP}(f, \rho) = \max_{q} \min_{B \in W_{\min}(f, z_\rho)} \sum_{i \in B} q_i.$$

So the same quantity can be viewed as the best probability of hitting every minimal sensitive 0-block at z_ρ.

The Dual Viewpoint

Fix z_ρ, and write $\mathcal{W} := W_{\min}(f, z_\rho)$.

Fractional hitting

$$\begin{aligned} \text{maximize} \quad & p \\ \text{subject to} \quad & \sum_{i \in B} w_i \geq p \qquad \forall B \in \mathcal{W} \\ & \sum_i w_i = 1 \\ & w_i \geq 0 \end{aligned}$$

w: distribution over the 0-coordinates of z_ρ.

Dual packing

$$\begin{aligned} \text{minimize} \quad & p \\ \text{subject to} \quad & \sum_{\substack{B \in \mathcal{W} \\ i \in B}} \lambda_B \leq p \qquad \forall i \\ & \sum_{B \in \mathcal{W}} \lambda_B = 1 \\ & \lambda_B \geq 0 \end{aligned}$$

λ: distribution over minimal sensitive 0-blocks.

By LP duality, both programs have the same optimum.

Fractional Monotone Block Sensitivity

The dual chooses a distribution λ over sensitive 0-blocks so that no coordinate is too popular. Indeed,

$$\sum_{\substack{B \in \mathcal{W} \\ i \in B}} \lambda_B = \Pr_{B \sim \lambda}[\, i \in B \,].$$

Thus small p means that the sensitive blocks are *fractionally almost disjoint*.

Define $\mathrm{FMBS}(f, z_\rho) := \frac{1}{p}$. By LP duality,

$$\mathrm{FHSC}(f, z_\rho) = \mathrm{FMBS}(f, z_\rho).$$

The integral version is

$$\mathrm{MBS}(f, z_\rho) := \max\{\text{number of pairwise disjoint sensitive 0-blocks at } z_\rho\}.$$

Hence

$$\mathrm{MBS}(f, z_\rho) \leq \mathrm{FMBS}(f, z_\rho) = \mathrm{FHSC}(f, z_\rho).$$

Four Measures

For every z,

$$\mathrm{MBS}(f, z) \leq \mathrm{FMBS}(f, z) = \mathrm{FHSC}(f, z) \leq \mathrm{HSC}(f, z).$$

The pointwise gap between fractional and integral versions can be large.

But globally we can show

$$\mathrm{FMBS}(f) = O(\mathrm{MBS}(f)^2).$$

Therefore, we only need to argue that large $\mathrm{MBS}(f)$ implies large sparsity.

Monotone Block Sensitivity to Sparsity

Lemma. For every Boolean f, $\mathrm{MBS}(f) = O\big((\log \mathsf{spar}(f))^2\big)$.

Proof idea.

Suppose $\mathrm{MBS}(f) = k$, witnessed at some z by pairwise disjoint blocks $w_1, \ldots, w_k$.

Identify each block w_i into one variable and fix all other variables.

This gives a Boolean function $g : \{0,1\}^k \to \{0,1\}$ such that

$$s(g, 0^k) = k \qquad \text{and} \qquad \mathsf{spar}(g) \leq \mathsf{spar}(f).$$

So f contains a k-variable **Promise-OR structure**.

Promise-OR Has Large Sparsity

Claim. If Q agrees with OR_n on $0^n, e_1, \ldots, e_n$, then

$$\mathrm{spar}(Q) = 2^{\Omega(\sqrt{n})}.$$

Suppose, towards contradiction, that $\mathrm{spar}(Q) \leq 2^{\sqrt{n}/100}$.

- Choose a random restriction $\rho \in \{0, *\}^n$, leaving each variable free independently with probability $1/2$.
- With high probability, at least $n/4$ variables remain free.
- A degree-k monomial survives iff all its variables remain free, which happens with probability 2^{-k}.

Promise-OR Has Large Sparsity

By a union bound,

$$\Pr_{\rho}\left[\deg(Q|_{\rho}) > \frac{\sqrt{n}}{10}\right] \leq \mathrm{spar}(Q)\, 2^{-\sqrt{n}/10} \ll 1.$$

Hence there is a restriction ρ leaving $m \geq n/4$ variables free such that

$$\deg(Q|_{\rho}) \leq \frac{\sqrt{n}}{10}.$$

Thus $Q|_{\rho}$ has degree at most $\sqrt{n}/10$, while it computes $\mathrm{PromiseOR}$ on at least $n/4$ variables. This contradicts degree lower bound for $\mathrm{PromiseOR}_n$ of $\sqrt{n}$. Therefore

$$\mathrm{spar}(Q) = 2^{\Omega(\sqrt{n})}.$$

Putting Things Together

$$\mathrm{MBS}(f) = O((\log \mathsf{spar}(f))^2)$$

$$\mathrm{FHSC}(f) = O(\mathrm{MBS}(f)^2) = O((\log \mathsf{spar}(f))^4)$$

$$\mathrm{HSC}(f) = O(\mathrm{FHSC}(f) \log r) = O((\log \mathsf{spar}(f))^5)$$

Thus sparse Boolean functions have small hitting sets for their monomials.

From Hitting Sets to Communication

$$\begin{aligned}
D^{cc}(f \circ \mathrm{AND}_2) &\leq 2D^{\wedge\text{-dt}}(f) \\
&\leq O\Big(D^{0\text{-dt}}(f) \log n\Big) \\
&\leq O(\mathrm{HSC}(f) \log \mathrm{spar}(f) \log n) \\
&\leq O\big((\log \mathrm{spar}(f))^6 \log n\big) \\
&= O\big((\log \mathrm{rank}(M_{f \circ \mathrm{AND}_2}))^6 \log n\big) .
\end{aligned}$$

Step Back

For general matrices, log-rank asks:

$$\text{small rank} \quad \Longrightarrow \quad \text{large monochromatic rectangle.}$$

For AND-functions, it becomes:

$$\text{small sparsity} \quad \Longrightarrow \quad \text{large monochromatic subcube obtained by fixing only 0's,}$$

At its heart, the log-rank conjecture asks:

When does algebraic simplicity force large monochromatic structure?

XOR Functions

XOR Functions

For $f : \{0,1\}^n \to \{0,1\}$, define

$$F(x, y) = f(x \oplus y).$$

This case has a completely different algebraic flavor.

Parity Decision Trees

A parity decision tree query

$$\bigoplus_{i \in S} z_i, \qquad S \subseteq [n].$$

Minimum depth computing f is denoted $D^{\oplus\text{-dt}}(f)$.

For $z = x \oplus y$,

$$\bigoplus_{i \in S} z_i = \left(\bigoplus_{i \in S} x_i\right) \oplus \left(\bigoplus_{i \in S} y_i\right).$$

Thus Alice and Bob can simulate parity decision tree now.

$$D^{cc}(f \circ \mathrm{XOR}_2) \leq 2D^{\oplus\text{-dt}}(f).$$

Rank Becomes Fourier Sparsity

$$f(z) = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S(z), \qquad \chi_S(z) := (-1)^{\sum_{i \in S} z_i}.$$

$$\chi_S(x \oplus y) = \chi_S(x)\chi_S(y),$$

$$M_{f \circ \mathrm{XOR}_2}(x, y) = \sum_{S \subseteq [n]} \widehat{f}(S)\chi_S(x)\chi_S(y).$$

The characters $\{\chi_S\}$ are linearly independent (in fact, orthogonal under $\langle g, h \rangle := \mathbb{E}_x[g(x)h(x)]$), hence

$$\mathrm{rank}(M_{f \circ \mathrm{XOR}_2}) = \mathrm{spar}_{\pm}(f)$$

Reducing to Parity Decision Trees

For XOR-functions, to show log-rank:

$$D^{cc}(f \circ \mathrm{AND}_2) \overset{?}{\leq} \mathrm{poly}(\log \mathrm{rank}(f \circ \mathrm{AND}_2)).$$

it suffices to prove

$$D^{\oplus\text{-dt}}(f) \overset{?}{\leq} \mathrm{poly}\big(\log \mathrm{spar}_{\pm}(f)\big) .$$

It's a question about Parity decision-tree complexity.

What Would Log-Rank Need for XOR-Functions?

If the log-rank conjecture holds for $f \circ \mathrm{XOR}_2$, then f must have a monochromatic affine subspace of codimension $\mathsf{poly}\big(\log \mathsf{spar}_{\pm}(f)\big)$.

So a necessary structural condition is

small Fourier sparsity $\implies$ large affine subspace on which f is constant.

As in the previous cases, this condition is also sufficient.

Thus log-rank for XOR-functions reduces to this structural question.

Restricting One Parity

$$f(x) = \sum_{\alpha} \widehat{f}(\alpha)\chi_\alpha(x).$$

Suppose we restrict to $\chi_\gamma(x) = b$. Then for every α,

$$\chi_{\alpha+\gamma}(x) = b\chi_\alpha(x).$$

So the two frequencies α and $\alpha + \gamma$ become indistinguishable after the restriction.

Their coefficients *fold together*:

$$\widehat{f}(\alpha)\chi_\alpha + \widehat{f}(\alpha+\gamma)\chi_{\alpha+\gamma} = (\widehat{f}(\alpha) + \widehat{f}(\alpha+\gamma))\chi_\alpha.$$

for every x satisfying $\chi_\gamma(x) = b$.

Restricting Several Parities

$$\chi_{\gamma_1}(x) = b_1, \ldots, \chi_{\gamma_t}(x) = b_t, \qquad b_i \in \{\pm 1\},$$

and let

$$G := \mathrm{span}\{\gamma_1, \ldots, \gamma_t\} \subseteq \mathbb{F}_2^n.$$

If $\beta = \alpha + \gamma$ for some $\gamma \in G$, then on the restricted affine subspace

$$\chi_\beta(x) = \chi_\alpha(x)\chi_\gamma(x) = \pm\,\chi_\alpha(x).$$

Thus frequencies are identified modulo G:

$$\alpha \sim \beta \qquad \Longleftrightarrow \qquad \alpha + \beta \in G.$$

Equivalently, the frequency space is partitioned into cosets

$$\alpha + G.$$

Buckets in the Fourier Spectrum

Let $S = \mathrm{supp}(\widehat{f})$, and let $G = \{\gamma_1, \ldots, \gamma_t\}$.

Choose one representative $\alpha(C)$ from each coset $C \in \mathbb{F}_2^n/G$. Then

$$f(x) = \sum_{C \in \mathbb{F}_2^n/G} \left(\sum_{\gamma \in G} \widehat{f}(\alpha(C) + \gamma)\chi_\gamma(x) \right) \chi_{\alpha(C)}(x).$$

After restricting $\chi_{\gamma_i}(x) = b_i$, every $\chi_\gamma(x)$, $\gamma \in G$, becomes a fixed sign.

Thus each coset of G is one *bucket*, and each bucket contributes at most one Fourier coefficient after restriction.

If $f|_{H_b}$ is constant, then every bucket $C \neq G$ must fold to 0. Hence every occupied nonzero bucket contains at least two frequencies.

Log-Rank for XOR-Functions: The Open Problem

The log-rank conjecture for XOR-functions is equivalent to asking:

Does every Boolean function f of Fourier sparsity s have a monochromatic affine subspace of codimension $\mathrm{poly}(\log s)$?

Step Back

For general matrices, log-rank asks:

small rank $\implies$ large monochromatic rectangle.

For AND-functions:

small sparsity $\implies$ large monochromatic subcube obtained by fixing only 0's.

For XOR-functions:

small Fourier sparsity $\implies$ large monochromatic affine subspace.

At its heart, the log-rank conjecture asks:

When does algebraic simplicity force large monochromatic structure?

Thank You