A Fourier-Analytic Switching Lemma over $\mathbb{F}_p$ and the AC^0 Lower Bound for Generalized Parity

Yipin Wang
yipinw2@illinois.edu

Abstract

We prove a switching lemma for constant-depth circuits over the alphabet $\mathbb{F}_p$ with generalized AND/OR gates (where AND tests that all inputs are nonzero and OR tests that some input is nonzero), extending Tal's Fourier-analytic approach from the Boolean setting. This circuit model is the natural $\mathbb{F}_p$-analogue of Boolean AC^0; it does *not* include MOD_p gates and is thus distinct from $\mathrm{AC}^0[p]$. The key new ingredient is a direct computation of the L_1 Fourier mass of AND/OR gates over $\mathbb{F}_p$, which yields an exact closed-form expression for the expected high-degree Fourier mass after a random restriction. Combined with a Markov inequality argument, this gives a switching lemma with an explicit, prime-independent structure. As a consequence, we obtain that for any prime p, constant-depth circuits of sub-exponential size over $\mathbb{F}_p$ with generalized AND/OR gates cannot compute $\mathbf{1}[\sum_i x_i \equiv 0 \pmod p]$.

Keywords: switching lemma, AC^0 lower bounds, Fourier analysis over finite fields, circuit complexity, random restrictions, generalized parity

1 Introduction

Håstad's switching lemma [1] is a cornerstone of circuit complexity, establishing that random restrictions dramatically simplify constant-depth Boolean circuits. Tal [2] gave a Fourier-analytic proof that replaces the combinatorial core of Håstad's argument with an L_1 inequality: after a random restriction, the high-degree Fourier mass of a bounded-fan-in gate concentrates, which, combined with a lower bound on the L_1 mass of functions with large decision tree depth, yields the switching lemma via Markov's inequality.

In this paper we extend Tal's approach to circuits over the prime-field alphabet $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$. The generalization requires two ingredients:

(1) An upper bound on $\mathbb{E}_\rho[L_1^{\geq s}(f|_\rho)]$ for gates under $\mathbb{F}_p$-valued random restrictions.

(2) A lower bound on $L_1^{\geq s}(g)$ for AND/OR gates g of fan-in $\geq s$.

For (2), the Fourier coefficients of the generalized AND gate $\mathrm{AND}_k(x) = \prod_{i=1}^{k} \mathbf{1}[x_i \neq 0]$ are given by an explicit product formula (Proposition 3.1), from which the lower bound follows immediately.

For (1), we exploit the structural observation that random restrictions preserve AND/OR gates (Observation 2.4) to derive an exact closed-form expression for $\mathbb{E}_\rho[L_1^{\geq s}(\mathrm{AND}_K|_\rho)]$ as a weighted binomial tail (Theorem 4.1), giving a self-contained proof that avoids the general L_1 machinery.

A notable consequence of the direct computation is that the switching lemma incurs no prime-dependent penalty factor $\gamma_p < 1$: the lower bound $((p-1)/p)^s$ holds exactly for AND/OR gates, while the upper bound on expected L_1 mass is controlled by a binomial tail that admits standard Chernoff-type estimates.

Context and prior work. The question of proving AC^0 lower bounds over non-Boolean alphabets has a substantial history. Razborov [3] and Smolensky [4] established that MOD_q gates cannot be computed by $\mathrm{AC}^0[\mathrm{MOD}_p]$ circuits when $p \nmid q$; their approach uses approximation by low-degree polynomials over $\mathbb{F}_p$. Barrington, Straubing, and Thérien [5] studied circuit complexity over non-Boolean alphabets from a semigroup-theoretic perspective, showing that the computational power of constant-depth circuits depends on the algebraic structure of the gate operations. Beigel and Tarui [6] proved that ACC circuits can be simulated by depth-two circuits with symmetric gates, placing ACC inside a small circuit class.

Our contribution is complementary to these works: rather than using polynomial approximation or algebraic methods, we extend the Fourier-analytic switching lemma to the $\mathbb{F}_p$ setting. This approach provides quantitative switching bounds for the specific gate basis $\{\mathrm{AND}, \mathrm{OR}\}$ over $\mathbb{F}_p$, where $\mathrm{AND}_k(x) = \mathbf{1}[\text{all } x_i \neq 0]$ and $\mathrm{OR}_k(x) = \mathbf{1}[x \neq \mathbf{0}]$. To the best of our knowledge, the explicit Fourier computation for these generalized gates (Proposition 3.1) and the resulting exact decay formulas (Theorems 4.1 and 4.2) are new.

A note on the circuit model. We emphasize that the circuit model studied here is *not* $\mathrm{AC}^0[p]$. Our circuits use generalized AND/OR gates over the alphabet $\mathbb{F}_p$ (Definition 2.2), which test whether inputs are zero or nonzero—they do not include MOD_p gates or any other counting gates. The model is a natural $\mathbb{F}_p$-alphabet analogue of standard Boolean AC^0: the gates generalize Boolean conjunction and disjunction to the non-Boolean setting, but the computational power remains that of AC^0 (constant-depth, polynomial-size, with AND/OR operations), not of $\mathrm{AC}^0[p]$ or ACC^0. In particular, the Razborov–Smolensky lower bounds for $\mathrm{AC}^0[p]$ [3, 4] address a different and incomparable circuit model.

Main results.

Theorem 1.1 (Switching lemma over $\mathbb{F}_p$)**.** *Let p be a prime and let $f : \mathbb{F}_p^n \to \{0,1\}$ be a generalized AND or OR gate of fan-in K. Under a random restriction ρ that independently keeps each variable alive with probability $q \le 1/(p-1)$ and fixes dead variables uniformly in $\mathbb{F}_p$,*

$$\Pr_\rho\big[\mathrm{DT}(f|_\rho) \ge s\big] \le \left(\frac{ep}{p-1} \cdot \frac{qK}{s}\right)^s$$

for all $s \ge 1$. In particular, for any constant $\alpha > 0$, setting $q = \alpha s(p-1)/(epK)$ gives $\Pr[\mathrm{DT}(f|_\rho) \ge s] \le \alpha^s$.

Corollary 1.2 (Parity $\notin \mathrm{AC}^0$ over $\mathbb{F}_p$)**.** *For any prime p, constant d, and $\epsilon > 0$, circuits of depth d and size 2^{n^ϵ} over the alphabet $\mathbb{F}_p$ with generalized AND/OR gates cannot compute $\mathbf{1}[\sum_i x_i \equiv 0 \pmod{p}]$.*

2 Preliminaries

2.1 Fourier analysis on $\mathbb{F}_p^n$

Let $\omega = e^{2\pi i/p}$ be a primitive p-th root of unity. The characters of the group $\mathbb{F}_p^n$ are $\chi_\alpha(x) = \omega^{\langle \alpha, x\rangle}$ for $\alpha \in \mathbb{F}_p^n$, where $\langle \alpha, x\rangle = \sum_i \alpha_i x_i \pmod{p}$. Every function $f : \mathbb{F}_p^n \to \mathbb{C}$ has a unique Fourier expansion

$$f(x) = \sum_{\alpha \in \mathbb{F}_p^n} \hat{f}(\alpha)\chi_\alpha(x), \qquad \hat{f}(\alpha) = \frac{1}{p^n}\sum_{x \in \mathbb{F}_p^n} f(x)\,\overline{\chi_\alpha(x)}.$$

Definition 2.1 (Fourier degree and L_1 norms)**.** *The degree of a character* χ_α *is* $|\alpha| = \#\{i : \alpha_i \neq 0\}$. *The Fourier degree of* f *is* $\mathrm{fdeg}(f) = \max\{|\alpha| : \hat{f}(\alpha) \neq 0\}$. *The* L_1 *Fourier norm at degree* $\geq s$ *is* $L_1^{\geq s}(f) = \sum_{|\alpha| \geq s} |\hat{f}(\alpha)|$.

2.2 Decision trees and gates over $\mathbb{F}_p$

A decision tree on $\mathbb{F}_p^n$ is a rooted tree where each internal node queries some variable x_i and branches into p children (one for each value in $\mathbb{F}_p$), and each leaf is labeled with an output value. The depth $\mathrm{DT}(f)$ is the minimum depth of a decision tree computing f.

Definition 2.2 (Generalized AND/OR gates)**.** *The generalized AND gate of fan-in* k *is*

$$\mathrm{AND}_k(x_1, \ldots, x_k) = \prod_{i=1}^{k} \mathbf{1}[x_i \neq 0] = \begin{cases} 1 & \text{if } x_i \neq 0 \text{ for all } i, \\ 0 & \text{otherwise.} \end{cases}$$

The generalized OR gate of fan-in k *is*

$$\mathrm{OR}_k(x_1, \ldots, x_k) = \mathbf{1}[x \neq \mathbf{0}] = \begin{cases} 1 & \text{if } x_i \neq 0 \text{ for some } i, \\ 0 & \text{if } x = \mathbf{0}. \end{cases}$$

Remark 2.3. *For* $p = 2$*, these reduce to the standard Boolean AND and OR. For general* p*, the AND gate outputs 1 iff all inputs lie in* $\mathbb{F}_p \setminus \{0\}$*, and the OR gate outputs 1 iff at least one input is nonzero.*

2.3 Random restrictions

A random restriction ρ on $\mathbb{F}_p^n$ with parameter $q \in (0, 1)$ independently sets each variable x_i to be alive (unfixed) with probability q, or dead (fixed to a uniformly random value in $\mathbb{F}_p$) with probability $1 - q$.

Observation 2.4 (Restriction preserves AND/OR structure)**.** *Let* $f = \mathrm{AND}_K$ *and let* ρ *be a random restriction. If any dead variable is fixed to* 0*, then* $f|_\rho \equiv 0$*. Otherwise, every dead variable is fixed to some* $v \in \{1, \ldots, p-1\}$*, contributing* $\mathbf{1}[v \neq 0] = 1$ *to the product, so* $f|_\rho = \mathrm{AND}_J$ *where* J *is the number of alive variables. Similarly, for* $f = \mathrm{OR}_K$*: if any dead variable is fixed to a nonzero value, then* $f|_\rho \equiv 1$*; otherwise, all dead variables are fixed to* 0 *and* $f|_\rho = \mathrm{OR}_J$*. In both cases,* $f|_\rho$ *is either constant or a gate of the same type on fewer variables. In particular,* $\mathrm{DT}(f|_\rho) \geq s$ *if and only if* $f|_\rho$ *is a gate of the same type on* $J \geq s$ *variables.*

3 Fourier Analysis of AND/OR Gates

This section contains the key new computation.

Proposition 3.1 (Fourier transform of AND_k)**.** *Let* $f = \mathrm{AND}_k : \mathbb{F}_p^k \to \{0, 1\}$*. For any* $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{F}_p^k$*,*

$$\hat{f}(\alpha) = \frac{1}{p^k} \prod_{i=1}^{k} \theta_{\alpha_i}, \quad \text{where } \theta_a = \sum_{v=1}^{p-1} \omega^{-av} = \begin{cases} p - 1 & \text{if } a = 0, \\ -1 & \text{if } a \neq 0. \end{cases}$$

In particular, $|\hat{f}(\alpha)| = p^{-k}(p-1)^{k-|\alpha|}$.

Proof. Since $\mathrm{AND}_k(x) = \prod_{i=1}^{k} \mathbf{1}[x_i \neq 0]$ and the variables factor in the sum,

$$\hat{f}(\alpha) = \frac{1}{p^k} \sum_{x_1,\dots,x_k} \prod_{i=1}^{k} \mathbf{1}[x_i \neq 0]\, \omega^{-\alpha_i x_i} = \frac{1}{p^k} \prod_{i=1}^{k} \left(\sum_{v=1}^{p-1} \omega^{-\alpha_i v} \right).$$

If $a = 0$, the inner sum is $\sum_{v=1}^{p-1} 1 = p-1$. If $a \neq 0$, then $\sum_{v=1}^{p-1} \omega^{-av} = \sum_{v=0}^{p-1} \omega^{-av} - 1 = 0 - 1 = -1$, since the full sum of all p-th roots of unity vanishes. Hence $|\theta_a| = p - 1$ if $a = 0$ and $|\theta_a| = 1$ if $a \neq 0$, giving $|\hat{f}(\alpha)| = p^{-k}(p-1)^{k-|\alpha|}$. □

Corollary 3.2 (Lower bound for AND gates)**.** *For $f = \mathrm{AND}_k$ with $k \geq s$,*

$$L_1^{\geq s}(f) = \sum_{|\alpha| \geq s} |\hat{f}(\alpha)| = \left(\frac{p-1}{p} \right)^k \sum_{j=s}^{k} \binom{k}{j}.$$

In particular, when $k = s$: $L_1^{\geq s}(\mathrm{AND}_s) = \left(\frac{p-1}{p}\right)^s$.

Proof. There are $\binom{k}{j}(p-1)^j$ characters of degree exactly j. Each has $|\hat{f}(\alpha)| = p^{-k}(p-1)^{k-j}$ by Proposition 3.1. The total L_1 at degree j is $\binom{k}{j}(p-1)^j \cdot p^{-k}(p-1)^{k-j} = \binom{k}{j}(p-1)^k/p^k$. Summing over $j \geq s$ gives the result. □

Remark 3.3 (The OR gate)**.** *For $\mathrm{OR}_k(x) = \mathbf{1}[x \neq \mathbf{0}]$, we have $\mathrm{OR}_k(x) = 1 - \mathbf{1}[x = \mathbf{0}]$, so the Fourier coefficients are $\hat{f}(\mathbf{0}) = 1 - p^{-k}$ and $\hat{f}(\alpha) = -p^{-k}$ for all $\alpha \neq \mathbf{0}$. Hence*

$$L_1^{\geq s}(\mathrm{OR}_J) = \frac{1}{p^J} \sum_{j=s}^{J} \binom{J}{j} (p-1)^j = \Pr\left[\mathrm{Bin}\left(J, \tfrac{p-1}{p}\right) \geq s\right].$$

Since $\mathrm{Bin}(J,(p-1)/p)$ has mean $J(p-1)/p \geq s(p-1)/p \geq s/2$, the lower bound $L_1^{\geq s}(\mathrm{OR}_J) \geq \left(\frac{p-1}{p}\right)^s$ holds for all $J \geq s$: at $J = s$ the only term is $j = s$ giving exactly $((p-1)/p)^s$, and the probability $\Pr[\mathrm{Bin}(J,(p-1)/p) \geq s]$ is non-decreasing in J.

Remark 3.4 (All Fourier coefficients are nonzero)**.** *A notable feature of Proposition 3.1 is that $|\hat{f}(\alpha)| > 0$ for every $\alpha \in \mathbb{F}_p^k$. In particular, AND_k has Fourier degree exactly k. This is not true for general $\{0,1\}$-valued functions on $\mathbb{F}_p^k$: as we discuss in Section 7, there exist functions with decision tree depth s but Fourier degree $< s$.*

Remark 3.5 (Boolean comparison)**.** *For $p = 2$, Corollary 3.2 gives $L_1^{\geq s}(\mathrm{AND}_s) = (1/2)^s$, matching the standard Boolean computation. The lower bound $((p-1)/p)^s$ holds for all primes with the same structural form.*

4 Exact Formulas for Expected Fourier Decay

The following theorems provide exact closed-form expressions for the expected high-degree L_1 mass of AND and OR gates after a random restriction.

Theorem 4.1 (Exact formula for AND)**.** *Let $f = \mathrm{AND}_K : \mathbb{F}_p^K \to \{0,1\}$ and let ρ be a random restriction with parameter q. Then*

$$\mathbb{E}_\rho\left[L_1^{\geq s}(f|_\rho)\right] = \left(\frac{p-1}{p} \right)^K \sum_{j=s}^{K} \binom{K}{j} q^j. \tag{1}$$

Theorem 4.2 (Exact formula for OR)**.** *Let $f = \mathrm{OR}_K : \mathbb{F}_p^K \to \{0,1\}$ and let ρ be a random restriction with parameter q. Then*

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = \frac{1}{p^K}\sum_{j=s}^{K}\binom{K}{j}\big((p-1)q\big)^j. \tag{2}$$

Remark 4.3. *For $p = 2$, the two formulas coincide: $(1/2)^K \sum_{j=s}^{K}\binom{K}{j}q^j$. For $p \geq 3$, they differ because the AND gate's survival condition (all dead variables nonzero) and the OR gate's survival condition (all dead variables zero) have different probabilities.*

Proof of Theorem 4.1. By Observation 2.4, the restricted function $f|_\rho$ is either identically zero (if any dead variable is fixed to 0) or AND_J on the J alive variables (if all dead variables are nonzero). In the former case, $L_1^{\geq s}(f|_\rho) = 0$.

Each variable independently falls into one of three categories: alive (probability q), dead and fixed to 0 (probability $(1-q)/p$), or dead and fixed to a nonzero value (probability $(1-q)(p-1)/p$). The gate survives (is not killed to 0) precisely when no dead variable is fixed to 0.

For a specific alive set $A \subseteq [K]$ with $|A| = J$, the probability that exactly these variables are alive and all $K - J$ dead variables are nonzero is $q^J \cdot ((1-q)(p-1)/p)^{K-J}$. The resulting function is AND_J, so by Corollary 3.2, $L_1^{\geq s}(\mathrm{AND}_J) = ((p-1)/p)^J \sum_{j=s}^{J}\binom{J}{j}$. Summing over all choices of alive set:

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = \sum_{J=s}^{K}\binom{K}{J}q^J\left(\frac{(1-q)(p-1)}{p}\right)^{K-J}\cdot\left(\frac{p-1}{p}\right)^J\sum_{j=s}^{J}\binom{J}{j}. \tag{3}$$

We exchange the order of summation: for fixed j (the "degree" index), J ranges from j to K. Using $\binom{K}{J}\binom{J}{j} = \binom{K}{j}\binom{K-j}{J-j}$ and substituting $m = J - j$:

$$(3) = \sum_{j=s}^{K}\binom{K}{j}\left(\frac{q(p-1)}{p}\right)^j\sum_{m=0}^{K-j}\binom{K-j}{m}\left(\frac{q(p-1)}{p}\right)^m\left(\frac{(1-q)(p-1)}{p}\right)^{K-j-m}. \tag{4}$$

The inner sum is a binomial expansion: $\sum_{m=0}^{K-j}\binom{K-j}{m}(q(p-1)/p)^m((1-q)(p-1)/p)^{K-j-m} = ((p-1)/p)^{K-j}$. Substituting:

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = \sum_{j=s}^{K}\binom{K}{j}\left(\frac{q(p-1)}{p}\right)^j\left(\frac{p-1}{p}\right)^{K-j} = \left(\frac{p-1}{p}\right)^K\sum_{j=s}^{K}\binom{K}{j}q^j. \qquad \square$$

Proof of Theorem 4.2. By Observation 2.4, $f|_\rho$ is either identically 1 (if any dead variable is nonzero) or OR_J on J alive variables (if all dead variables are zero). The constant case contributes $L_1^{\geq s}(1) = 0$ for $s \geq 1$.

For a specific alive set A with $|A| = J$, the probability that exactly these variables are alive and all $K - J$ dead variables are zero is $q^J \cdot ((1-q)/p)^{K-J}$. The resulting function is OR_J, so by Remark 3.3, $L_1^{\geq s}(\mathrm{OR}_J) = p^{-J}\sum_{j=s}^{J}\binom{J}{j}(p-1)^j$. Summing:

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = \sum_{J=s}^{K}\binom{K}{J}q^J\left(\frac{1-q}{p}\right)^{K-J}\cdot\frac{1}{p^J}\sum_{j=s}^{J}\binom{J}{j}(p-1)^j.$$

Exchanging summation using the same identity $\binom{K}{J}\binom{J}{j} = \binom{K}{j}\binom{K-j}{J-j}$ and substituting $m = J-j$:

$$= \sum_{j=s}^{K} \binom{K}{j} \frac{(p-1)^j q^j}{p^j} \sum_{m=0}^{K-j} \binom{K-j}{m} \frac{q^m}{p^m} \left(\frac{1-q}{p}\right)^{K-j-m}.$$

The inner sum equals $(q/p + (1-q)/p)^{K-j} = p^{-(K-j)}$. Hence

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = \sum_{j=s}^{K} \binom{K}{j} \frac{((p-1)q)^j}{p^j} \cdot \frac{1}{p^{K-j}} = \frac{1}{p^K} \sum_{j=s}^{K} \binom{K}{j} \big((p-1)q\big)^j. \qquad \square$$

Remark 4.4 (Exactness)**.** *Both formulas are exact, not merely upper bounds. For instance, when* $K = s$ *and* $q = 1$ *(no restriction), Theorem 4.1 gives* $((p-1)/p)^s$*, matching Corollary 3.2.*

5 The Switching Lemma

Proof of Theorem 1.1. The argument combines the exact formulas with the Fourier lower bound via Markov's inequality.

Step 1 (Lower bound). Suppose $\mathrm{DT}(f|_\rho) \geq s$. By Observation 2.4, $f|_\rho = \mathrm{AND}_J$ (or OR_J) for some $J \geq s$. By Corollary 3.2 and Remark 3.3,

$$L_1^{\geq s}(f|_\rho) \geq \left(\frac{p-1}{p}\right)^s.$$

(For $J > s$ in the AND case, note that $\mathrm{AND}_J = \mathrm{AND}_s \otimes \mathrm{AND}_{J-s}$, so $L_1^{\geq s}(\mathrm{AND}_J) \geq L_1^{=s}(\mathrm{AND}_s) \cdot L_1(\mathrm{AND}_{J-s}) = ((p-1)/p)^s \cdot (2(p-1)/p)^{J-s} \geq ((p-1)/p)^s$, since $2(p-1)/p \geq 1$ for $p \geq 2$.)

Step 2 (Upper bound on expected L_1). By Theorems 4.1 and 4.2, the expected L_1 mass has the form

$$\mathbb{E}_\rho\Big[L_1^{\geq s}(f|_\rho)\Big] = P^K \sum_{j=s}^{K} \binom{K}{j} Q^j,$$

where $(P, Q) = ((p-1)/p,\ q)$ for the AND gate and $(P, Q) = (1/p,\ (p-1)q)$ for the OR gate.

To bound the binomial tail without introducing a correction factor, write

$$\sum_{j=s}^{K} \binom{K}{j} Q^j = (1+Q)^K \Pr\Big[\mathrm{Bin}\Big(K, \tfrac{Q}{1+Q}\Big) \geq s\Big],$$

and apply the Chernoff bound via the moment generating function. For $X \sim \mathrm{Bin}(K, p')$ with $p' = Q/(1+Q)$, and any $t > 0$:

$$\Pr[X \geq s] \leq \frac{\mathbb{E}[e^{tX}]}{e^{ts}} = \frac{(1 - p' + p'e^t)^K}{e^{ts}}.$$

Setting $e^t = s/(Kp')$ (optimal for the leading term):

$$\Pr[X \geq s] \leq \frac{(1 - p' + s/K)^K}{(s/(Kp'))^s} \leq \frac{e^{s - Kp'}}{(s/(Kp'))^s} = \left(\frac{eKp'}{s}\right)^s \cdot e^{-Kp'} \leq \left(\frac{eKp'}{s}\right)^s,$$

where we used $1+x \le e^x$ and $e^{-Kp'} \le 1$. Substituting $Kp' = KQ/(1+Q)$:

$$\mathbb{E}_\rho\Big[L_1^{\ge s}(f|_\rho)\Big] \le \big(P(1+Q)\big)^K \cdot \left(\frac{eQK}{(1+Q)s}\right)^s.$$

The prefactor $\big(P(1+Q)\big)^K$ is at most 1 in both cases:

- *AND gate:* $P(1+Q) = (p-1)(1+q)/p \le 1$ if and only if $q \le 1/(p-1)$, which holds by hypothesis.
- *OR gate:* $P(1+Q) = (1+(p-1)q)/p \le 1$ for all $q \le 1$.

Dropping the prefactor and using $1+Q \ge 1$:

$$\mathbb{E}_\rho\Big[L_1^{\ge s}(f|_\rho)\Big] \le \left(\frac{eQK}{s}\right)^s.$$

Step 3 (Markov's inequality).

$$\Pr_\rho[\mathrm{DT}(f|_\rho) \ge s] \le \frac{\mathbb{E}[L_1^{\ge s}(f|_\rho)]}{((p-1)/p)^s} \le \left(\frac{epQK}{(p-1)s}\right)^s.$$

Since $Q \le (p-1)q$ in both cases, this gives

$$\Pr_\rho[\mathrm{DT}(f|_\rho) \ge s] \le \left(\frac{C_p qK}{s}\right)^s$$

with $C_p = ep/(p-1)$, completing the proof. □

Remark 5.1 (The condition $q \le 1/(p-1)$)**.** *The hypothesis $q \le 1/(p-1)$ ensures that the prefactor $((p-1)(1+q)/p)^K$ arising in the AND case is at most 1. For $p=2$, this is $q \le 1$, which is always satisfied. For $p \ge 3$, it is a mild restriction: in the AC^0 application (Section 6), we set $q = O(s/K) \ll 1$, well within this range. The OR case requires no condition on q.*

Remark 5.2 (Comparison with Håstad's switching lemma)**.** *Håstad's switching lemma [1] gives $\Pr[\mathrm{DT}(f|_\rho) \ge s] \le (CqK)^s$ for width-K DNFs or CNFs with any number of terms; this bound is tight for DNFs. For a* single *AND or OR gate (a width-K DNF/CNF with one term), the event $\mathrm{DT}(f|_\rho) \ge s$ is equivalent to a binomial tail (Observation 2.4), and the Chernoff bound gives the tighter estimate $(C_p qK/s)^s$—stronger than $(CqK)^s$ by a factor of s^{-s}.*

This additional s^{-s} factor is genuine, not an artifact: it reflects the exact combinatorics of a single gate and allows one to set q to be a small constant (rather than $O(s/K)$) while still ensuring that all bottom-level gates simplify to decision trees of depth $O(\log K)$ with high probability. However, the bound applies only to individual gates, not to DNFs with many terms and shared variables. The DNF switching lemma of [7], which gives the bound $(2pwq/(1-q))^s$ independently of the number of terms, is needed for the subsequent rounds of the iterative argument (see Section 6).

Remark 5.3 (No γ_p penalty)**.** *In earlier versions of this work, the switching lemma was conditional on a conjecture that $c_p(s) \ge D_p \cdot \gamma_p^s$ for all $\{0,1\}$-valued functions of decision tree depth $\ge s$, where $\gamma_p < 1$. The AND/OR gate computation eliminates this entirely: the lower bound $((p-1)/p)^s$ is exact and applies to the specific functions appearing as circuit gates. There is no need to lower-bound the L_1 mass of arbitrary $\{0,1\}$-valued functions, which as we show in Section 7 would indeed require a weaker bound.*

6 Application: Parity $\notin$ AC^0 over $\mathbb{F}_p$

Proof of Corollary 1.2. Let C be a depth-d circuit of size M over $\mathbb{F}_p$ computing $\mathrm{Parity}_p(x) = \mathbf{1}[\sum_i x_i \equiv 0 \pmod{p}]$. We assume without loss of generality that C has alternating levels of AND and OR gates (any constant-depth circuit can be put in this form with at most a constant factor increase in depth). We describe the iterative depth reduction in detail.

Step 1 (Switching the bottom level). Suppose the bottom level (level 1) consists of AND gates $g_1, \ldots, g_M$ of fan-in at most K, and the next level (level 2) consists of OR gates. We apply a random restriction ρ with survival probability q. By Theorem 1.1, each bottom gate satisfies $\Pr[\mathrm{DT}(g_i|_\rho) \geq s] \leq (C_p qK/s)^s$. By a union bound over all M gates, with high probability every bottom gate has $\mathrm{DT}(g_i|_\rho) < s$.

Step 2 (Flattening: how two levels become one). Each simplified bottom gate $g_i|_\rho$ is now computed by a decision tree of depth $< s$. Over $\mathbb{F}_p$, such a tree has at most p^s root-to-leaf paths, and each accepting path (leading to output 1) corresponds to a conjunction of at most s conditions of the form $x_j = v$. Using the identity $\mathbf{1}[x_j = v] = 1 - \mathbf{1}[x_j \neq v]$ and expanding, each accepting path can be rewritten as a width-$\leq s$ AND of generalized literals. Hence $g_i|_\rho$ is equivalent to a DNF of width $\leq s$ with at most p^s terms.

Now consider a parent OR gate at level 2, say $h = \mathrm{OR}(g_{i_1}, \ldots, g_{i_r})$. After substitution, $h|_\rho = g_{i_1}|_\rho \vee \cdots \vee g_{i_r}|_\rho$. Since each $g_{i_j}|_\rho$ is itself a DNF, the OR of these DNFs is again a DNF: one simply takes the union of all terms. The resulting DNF has width $\leq s$ (each term has width $< s$) and up to $r \cdot p^s \leq M \cdot p^s$ terms, whose variables may overlap across terms.

Crucially, the two-level subcircuit (AND gates at level 1, OR gate at level 2) has been replaced by a single DNF at what was level 2. The circuit depth has decreased by 1. The analogous argument applies with AND and OR interchanged: if the bottom level consists of OR gates and the next level of AND gates, each simplified OR gate becomes a width-$\leq s$ CNF, and the parent AND of CNFs is again a CNF.

Step 3 (Subsequent switching rounds). After Step 2, the new bottom level consists of width-$\leq s$ DNFs or CNFs—not single AND/OR gates. These are multi-term formulas whose terms share variables, and Theorem 1.1 does not apply to them. To continue the iterative depth reduction, we invoke the M-independent DNF switching lemma over $\mathbb{F}_p$ proved in [7]:

$$\Pr_\rho[\mathrm{DT}(f|_\rho) \geq s] \leq \left(\frac{2pwq}{1-q}\right)^s$$

for any width-w DNF f with arbitrarily many terms, where the bound is independent of the number of terms. The analogous statement holds for CNFs by duality.

We apply $d-2$ further rounds of random restriction, each time using the DNF/CNF switching lemma of [7] to simplify the current bottom level and then flattening as in Step 2 to reduce depth by 1. At each round, the width parameter is at most s (set by the previous round's simplification).

Step 4 (Union bound). After all $d-1$ rounds, the total number of gates encountered across all rounds is at most $M \cdot p^{(d-1)s}$. By setting $q = \alpha s(p-1)/(epK_{\max})$ for a suitable $\alpha < 1/2$ in the first round, and $q = c/(pw)$ for a small constant c in subsequent rounds (as in [7]), the probability of failure for each gate is at most exponentially small in s. Setting $s = c' \log n$ for large enough c':

$$M \cdot p^{(d-1)s} \cdot (\text{failure probability per gate}) \leq 2^{n^\epsilon} \cdot p^{(d-1)c' \log n} \cdot n^{-c' \log(1/\beta)} \to 0$$

for an appropriate constant $\beta < 1$.

Step 5 (Contradiction). With positive probability, all rounds succeed and the circuit is reduced to depth 1 with $\mathrm{DT} < s$. The number of surviving variables satisfies $|A| = n^{\Omega(1)}$ (since each round preserves a $q = \Omega(s/K_{\max})$ fraction). But Parity restricted to the surviving set depends on all $|A|$ variables: changing any single x_i by 1 changes $\sum x_i$ modulo p. Hence $\mathrm{DT}(\mathrm{Parity}_p|_A) = |A| \gg s$, contradicting the simplified circuit.

Quantitative bound. Setting $q = \alpha s/(C_p K_{\max})$ with $K_{\max} \leq M \leq 2^{n^\epsilon}$ and $s = c \log n$, the number of surviving variables after $d-1$ rounds is at least

$$|A| \geq n \cdot q^{d-1} = n \cdot \left(\frac{\alpha c \log n}{C_p \cdot 2^{n^\epsilon}} \right)^{d-1}.$$

For $|A| > s = c \log n$ to hold, we need $n^{1-\epsilon(d-1)} \gg \log n$, which holds for $\epsilon < 1/(d-1)$. This yields the exponential lower bound $M \geq 2^{n^{\epsilon'}}$ for $\epsilon' > 0$ depending on d and p. $\square$

Remark 6.1 (Role of the two switching lemmas). *The iterative argument uses two different switching lemmas for two different purposes. Theorem 1.1 handles the first round: each bottom gate is a single AND or OR gate on independent inputs, and the exact Fourier formulas of Sections 3–4 (which exploit the product structure via Observation 2.4) yield the bound* $(C_p q K/s)^s$. *The DNF/CNF switching lemma of [7] handles all subsequent rounds: after flattening, the bottom level consists of multi-term DNFs/CNFs with shared variables, for which the single-gate analysis does not apply. The CDT argument yields the M-independent bound* $(2pwq/(1-q))^s$*, which is essential for the union bound across rounds since flattening can increase the number of terms exponentially.*

7 The Decision Tree – Fourier Degree Gap

We record an observation that arose during our investigation and is of independent interest.

Proposition 7.1. *For any* $f : \mathbb{F}_p^s \to \{0,1\}$,

$$\mathrm{fdeg}(f) \leq \mathrm{DT}(f) \leq \mathrm{rel}(f) \leq s,$$

where $\mathrm{rel}(f)$ *denotes the number of relevant variables.*

Proof. A decision tree of depth d writes f as a sum of products of at most d single-variable indicators $\mathbf{1}[x_i = v]$. Over $\mathbb{F}_p$, each indicator $\mathbf{1}[x_i = v] = \frac{1}{p}\sum_{a=0}^{p-1} \omega^{a(x_i - v)}$ has Fourier degree ≤ 1. A product of d such terms involves characters with at most d nonzero coordinates, so $\mathrm{fdeg}(f) \leq d = \mathrm{DT}(f)$. The bound $\mathrm{DT}(f) \leq \mathrm{rel}(f)$ holds because querying all relevant variables determines f. $\square$

Observation 7.2 (Both inequalities can be strict). *Both* $\mathrm{fdeg} < \mathrm{DT}$ *and* $\mathrm{DT} < \mathrm{rel}$ *can occur, even for* $p = 2$.

Over $\mathbb{F}_2$*: the function* $f(x_1, x_2, x_3) = \mathbf{1}[|x| \in \{1,2\}]$ *on* $\mathbb{F}_2^3$ *satisfies* $\mathrm{DT}(f) = 3$ *but* $\mathrm{fdeg}(f) = 2$, *since* $\hat{f}(\{1,2,3\}) = 0$ *while the degree-2 coefficients are nonzero.*

Over $\mathbb{F}_3$*: there exist subsets* $S \subset \mathbb{F}_3^4$ *with* $|S| = 6$ *such that* $\mathbf{1}_S$ *depends on all 4 variables and requires depth 4 to compute, yet has Fourier degree 3.*

This gap is why a generic lower bound of the form "$\mathrm{DT}(f) \geq s$ *implies* $L_1^{\geq s}(f) > 0$*" fails over* $\mathbb{F}_p$ *for arbitrary* $\{0,1\}$*-valued functions. The switching lemma avoids this obstacle because it applies to AND/OR gates specifically, which have* $\mathrm{fdeg} = \mathrm{DT} = \mathrm{rel}$ *(Remark 3.4).*

Remark 7.3 (Size of the gap). *In all cases we have examined computationally (exhaustive for* $\mathbb{F}_2^3$, $\mathbb{F}_2^4$, $\mathbb{F}_3^2$*; sampling for* $\mathbb{F}_3^s$ *with* $s \leq 6$*), the gap* $\mathrm{DT}(f) - \mathrm{fdeg}(f)$ *is at most 1. Whether* $\mathrm{DT} - \mathrm{fdeg}$ *can grow with the ambient dimension remains an open question.*

8 Discussion and Open Problems

8.1 Comparison with the Boolean case

The Fourier-analytic switching lemma over $\mathbb{F}_p$ has the same qualitative form as in the Boolean case, with exponential decay in s. The constant $C_p = ep/(p-1)$ depends mildly on p, with $C_2 = 2e$ and $C_p \to e$ as $p \to \infty$. For single gates, the bound $(C_p qK/s)^s$ is stronger than Håstad's $(CqK)^s$ by a factor of s^{-s}, reflecting the exact binomial structure (Remark 5.2). However, this stronger form holds only for individual AND/OR gates, not for DNFs with shared variables. Extending Tal's full character-by-character analysis to DNFs over $\mathbb{F}_p$—which would yield an M-independent bound via the Fourier-analytic route—remains an open problem (see Section 8.3).

8.2 The extremal L_1 problem

Although not needed for the switching lemma, the following question remains mathematically interesting: given $f : \mathbb{F}_p^s \to \{0,1\}$ with $\mathrm{fdeg}(f) \geq s$, what is the minimum value of $L_1^{=s}(f)/((p-1)/p)^s$? Computational evidence for $p = 3$, $s \leq 4$ reveals a rich structure: the extremal sets include lines, affine quadrics, and affine subspace cosets, with the AND gate achieving ratio exactly 1.

8.3 Open problems

1. **Fourier-analytic DNF switching lemma over $\mathbb{F}_p$.** Theorem 1.1 gives the bound $(C_p qK/s)^s$ for single gates, which is tight up to the constant C_p (it matches the Chernoff bound on the underlying binomial). The combinatorial CDT method [7] gives the M-independent bound $(2pwq/(1-q))^s$ for DNFs. Can one obtain an M-independent DNF switching lemma over $\mathbb{F}_p$ via Tal's Fourier-analytic approach? This would require extending the character-by-character L_1 bound with $\min(1, \cdot)$ truncation to the $\mathbb{F}_p$ setting.
2. **DT–Fourier degree gap.** Is $\mathrm{DT}(f) - \mathrm{fdeg}(f)$ bounded by an absolute constant for all $\{0,1\}$-valued functions on $\mathbb{F}_p^s$? Our data shows a maximum gap of 1, but this is only verified for small s.
3. **Multi-prime circuits.** Can the switching lemma be extended to circuits that mix gates modulo different primes? The L_1 approach seems promising since the Fourier structure is well-understood for each prime individually.
4. **Tight $\mathbf{AC}^0$ bounds.** Determine the optimal exponent in the exponential size lower bound for Parity over $\mathbb{F}_p$. In the Boolean case, Håstad obtained the tight bound $\exp(\Omega(n^{1/(d-1)}))$; does the same hold over $\mathbb{F}_p$?

Acknowledgements

The author thanks Makrand Sinha for his lectures on Fourier analysis over $\mathbb{F}_2$, which inspired this line of investigation. The author is also deeply grateful to his parents for their constant help and support over the years.

References

[1] J. Håstad. Almost optimal lower bounds for small depth circuits. In *Proc. 18th STOC*, pages 6–20, 1986.

[2] A. Tal. Tight bounds on the Fourier spectrum of AC^0. In *Proc. 32nd CCC*, pages 15:1–15:31, 2017.

[3] A. A. Razborov. Lower bounds on the size of bounded depth networks over a complete basis with logical addition. *Math. Notes*, 41(4):598–607, 1987.

[4] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proc. 19th STOC*, pages 77–82, 1987.

[5] D. A. M. Barrington, H. Straubing, and D. Thérien. Non-uniform automata over groups. *Inform. and Comput.*, 89(2):109–132, 1990.

[6] R. Beigel and J. Tarui. On ACC. *Comput. Complexity*, 4:350–366, 1994.

[7] Y. Wang. A switching lemma for DNFs over $\mathbb{F}_p$: the canonical decision tree approach. Preprint, 2026. Available at `https://tianyilemon.github.io/writings/cdt_switching_lemma.pdf`.