Strong Pseudorandom Functions in $\mathsf{AC}^0[2]$ in the Bounded-Query Setting

Marshall Ball [*] Clément Ducros [†] Saroja Erabelli [‡] Lisa Kohl [§]
Nicolas Resch [¶] Peter Scholl [‖]

September 16, 2026

Abstract

Understanding the minimal computational power needed to realize a pseudorandom function (PRF) is a long-standing question in cryptography. By the Razborov–Smolensky polynomial approximation method, it is known that $\mathsf{AC}^0[2]$ cannot support strong pseudorandom functions with subexponential security, since any such function can be distinguished from random with quasipolynomially many samples.

In this work, we initiate the study of low-complexity strong PRFs under a refined framework that separates adversary query complexity from running time and distinguishing advantage, and observe that distinguishing algorithms for $\mathsf{AC}^0[2]$ do not apply if the number of queries is below the threshold implied by the Razborov–Smolensky approximation bound. In this setting, we prove that strong PRFs cannot exist in the smaller class $\mathsf{MOD}_2 \circ \mathsf{AC}^0$, showing that every such function family can be efficiently distinguished from random using polynomially many queries. Together with the conjecture that functions in $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ have heavy Fourier coefficients (ITCS 2014), this motivates considering $\mathsf{AC}^0[2]$ as the next natural candidate class.

We propose the first candidate strong PRF in $\mathsf{AC}^0[2]$, which we conjecture to offer subexponential security against adversaries limited to a fixed quasipolynomial number of queries. Furthermore, our candidate is in the class $\mathsf{MOD}_2 \circ \mathsf{AC}^0 \circ \mathsf{MOD}_2$, arguably the minimal class given the above barriers. We show that our candidate lacks heavy Fourier coefficients, resists many settings of linear attacks, is non-sparse over $\mathbb{F}_2$ with high probability, has high rational degree, is inapproximable by low-degree polynomials over $\mathbb{F}_2$, and has low correlation with fixed function families.

As a final contribution, we show that if any strong PRF exists in $\mathsf{AC}^0[2]$ (or a superclass), then we can construct a universal PRF, i.e., a single, fixed function which is guaranteed to be a strong PRF in the same class.

1 Introduction

A pseudorandom function (PRF) is a family of efficiently computable functions that, when keyed with a uniform secret value, is indistinguishable from a truly random function to any efficient adversary. In this work we focus on standard (also referred to as *strong*) PRFs, where the adversary gets chosen-query access to the function oracle. The question of how *computationally simple* a PRF can be is a fundamental question at the intersection of cryptography, complexity theory, and learning theory. In modern cryptography, PRFs not only enable direct constructions of encryption and authentication schemes, but they are also central for the efficiency of high-end cryptographic protocols, where a PRF is evaluated within some secure computation protocol [ARS+15, MJSC16, CCF+16, BIP+18, BHI+20, BCG+20]. Furthermore, understanding the minimal computational complexity required to realize PRFs sheds light on fundamental questions in complexity theory. Specifically, if a PRF exists in a complexity class, then any proof technique meeting the Razborov-Rudich "natural proof" criteria cannot be used to establish circuit lower bounds in that class [RR94,MV12,RST15]. Additionally, PRFs help to understand

[*]New York University, marshall.ball@cs.nyu.edu
[†]CISPA, Helmholtz Center for Information Security, clement.ducros@cispa.de
[‡]New York University, saroja.erabelli@gmail.com
[§]Cryptology Group, CWI Amsterdam, lisa.kohl@cwi.nl
[¶]Informatics Institute, University of Amsterdam, n.a.resch@uva.nl
[‖]Aarhus University, peter.scholl@cs.au.dk

when and how derandomization techniques can be applied [NW88, Wil13]. Moreover, the existence of PRFs is closely tied to questions in computational learning theory. In particular, any complexity class that admits efficient learning with membership queries cannot contain secure PRFs. Constructing concrete PRF candidates provides an explicit target for learning algorithms, offering a more tangible challenge than addressing the full generality of learning an entire complexity class.

To be considered secure, a PRF is typically required to admit at least subexponential security, i.e., there exists a constant $c > 0$ such that all adversaries running in time $T = 2^{\kappa^c}$ have distinguishing advantage at most $1/T$, where κ denotes the key length. [1] While results in learning theory usually do not distinguish between the number of samples and the computational time required to *learn* a function, in cryptography these resources are often very different: an adversary may have access to only a small number of queries, while having access to substantial total computational resources to break a pseudorandom function. Motivated by this, we initiate the systematic study of low-complexity PRFs in the bounded-query setting, where the adversary has access to a quasipolynomial number of chosen queries (i.e., $\kappa^{\log^c \kappa}$ for some fixed constant $c > 0$), but may run in subexponential time, or it only required to distinguish with subexponential advantage. We additionally allow the adversary to obtain subexponentially many *random* queries. In other words, as a *weak* PRF, we would like the candidate to offer subexponential security.

At the lower end of the complexity spectrum, it is known that any function in AC^0 (the class of constant-depth, polynomial-size circuits over AND, OR and NOT gates with unbounded fan-in) can be distinguished from a random function with quasipolynomial advantage using only two queries [LMN89]. A natural next candidate is $\mathsf{MOD}_2 \circ \mathsf{AC}^0$, namely parity of AC^0 circuits. However, in this work we show that strong PRFs cannot exist even in this class; in particular, with polynomially many chosen queries we obtain constant distinguishing advantage (this argument naturally also applies to AC^0 circuits). As for $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ (i.e., circuits with parities only applied to the inputs), Akavia et al. [ABG+14], conjecture that every such function has a heavy Fourier coefficient (i.e., of inverse quasipolynomial size). As we point out below (although this is largely implicit in prior work), this conjecture would result in a distinguisher with quasipolynomial advantage with only a constant number of chosen queries.

Thus, we choose to consider the "next" class $\mathsf{AC}^0[2]$, which augments AC^0 with XOR gates. Here, it is known that any function can be distinguished from random [Raz87,Smo87,KL01] and even learned [CIKK16] in quasipolynomial time. However, this crucially requires access to a sufficiently large quasipolynomial number of queries. In this work, we propose the following question:

Does there exist a PRF in $\mathsf{AC}^0[2]$ *that is plausibly secure against any adversary running in time* $T = 2^{\kappa^c}$ *using* $\kappa^{\log^c \kappa}$ *chosen queries for some fixed* $c > 0$*?*

We note that this in particular captures any (a-priori unbounded) polynomial number of queries. We explicitly phrase the setting as quasipolynomial to avoid ambiguity with the setting where the queries are limited to a *fixed* polynomial number $q(\kappa)$, as this would be trivially achievable (even information-theoretically) via a family of $(q(\kappa) + 1)$-wise independent hash functions.

1.1 Our Contributions

Impossibility of PRFs in $\mathsf{MOD}_2 \circ \mathsf{AC}^0$. On the negative side, we first show that no pseudorandom function family can be computed by $\mathsf{MOD}_2 \circ \mathsf{AC}^0$ circuits. More precisely, let $\mathcal{F}_{n,c,d}$ denote the class of functions given by a top parity of at most n^c many AC^0 summands, where each summand has size at most n^c and depth at most d.

Theorem 1.1 *For every fixed constants $c, d \geq 1$, there is a randomized nonadaptive algorithm $A = A_{c,d}$ making $n^{O_{c,d}(1)}$ queries such that A distinguishes $\mathcal{F}_{n,c,d}$ from a uniformly random function with advantage at least $\frac{1}{2} - n^{-9}$.*

Thus, even allowing polynomially many queries, functions in $\mathsf{MOD}_2 \circ \mathsf{AC}^0$ cannot be pseudorandom. This identifies $\mathsf{AC}^0[2]$ as the next natural candidate class for PRFs secure against quasipolynomially many queries.

[1] In our PRF constructions, the key length, PRF input length, and security parameter are polynomially related; see Section 2.2 for our parameter conventions.

Strong PRF Candidate in $\mathsf{AC}^0[2]$. We propose the first strong PRF candidate in $\mathsf{AC}^0[2]$ with subexponential security when the adversary is restricted to quasipolynomially many queries. The secret key consists of a matrix $\mathbf{A} \in \{0,1\}^{N\times N}$, sampled uniformly at random subject to being invertible over $\mathbb{F}_2$, together with a uniformly random vector $\mathbf{b} \in \{0,1\}^N$, where $N = \ell mn$.

$$\begin{aligned} &F_{\mathbf{A},\mathbf{b}} : \{0,1\}^N \to \{0,1\} \\ &F_{\mathbf{A},\mathbf{b}}(x) = g(\mathbf{A}x \oplus \mathbf{b}), \text{ where } g(\mathbf{y}) = \bigoplus_{i=1}^{\ell} \bigvee_{j=1}^{m} \bigwedge_{k=1}^{n} (\mathbf{y}_i)_{j,k} \ , \end{aligned} \tag{1}$$

and $\mathbf{y}_i \in \{0,1\}^{mn}$ denotes the vector $\mathbf{y}_i := (y_{mn\cdot(i-1)}, \ldots, y_{mn\cdot i-1})$.

Conjecture 1.2 *For ℓ, m, n, such that $\ell = \Theta(\lambda)$, $m = \Theta(\lambda/\log\lambda)$ and $n = \Theta(\log\lambda)$ there exists a $c > 0$ such that for any adversary running in time $T = 2^{\kappa^c}$ getting access to $\kappa^{\log^c \kappa}$ chosen samples, the adversary has advantage at most $1/T$ in distinguishing $F_{\mathbf{A},\mathbf{b}}$ as defined in Equation 1 from random (where κ denotes the key length $\kappa = N^2 + N$ for $N = \ell mn = \Theta(\lambda^2)$).*

Key Structure and Efficiency. Two downsides of our construction are the fact that the key is an invertible matrix, rather than an unstructured random string, and further that it has a large size $\Theta(N^2)$. Having a structured key does not affect the circuit complexity since, as is common in the literature, we treat the key as hardwired into the circuit. Nevertheless, we conjecture that a variant of our candidate with a uniformly random key is also subexponentially secure. Indeed, while a uniformly random $N \times N$ matrix over $\mathbb{F}_2$ is only full-rank with constant probability, it does have rank $\geq (1-\varepsilon)N$ with probability $1 - 2^{-\varepsilon N}$. We do not see clear reasons why the security of our PRF candidate would be meaningfully harmed by choosing a uniform key, and believe that (natural adaptations of) our arguments should still apply even with a slight rank deficiency.

In this variant with a uniform key, by assuming subexponential hardness of the *learning parity with noise (LPN)* problem, we can additionally improve *efficiency* by applying the trapdoored matrix construction of Vaikuntanathan and Zamir [VZ25] (related LPN-based trapdoored-matrix constructions were obtained independently by Braverman and Newman [BN25]). A basic implementation (with unstructured $\mathbf{A}$) evaluates our PRF via naïve matrix–vector multiplication in time $\Theta(N^2)$, and the key size is $\Theta(N^2)$. With the LPN-based trapdoor of [VZ25], matrix–vector multiplication, and hence evaluation of our PRF, can be performed in $\tilde{O}(N)$ time with key size $\tilde{O}(N)$. Even more ambitiously, one could obtain evaluation time and key size $O(N)$ under additional heuristic assumptions, using the random matrices of Druk and Ishai [DI14] or the linear-size trapdoored-matrix candidates proposed by Benhamouda et al. [BCH+25]. Note that these speedups also apply to the strong and weak PRF candidates of [ABG+14], [BIP+18] and [BCG+21] whose keys include a random matrix.

Distributed Setting. While the complexity of PRFs is usually measured relative to a hardwired key, for advanced cryptographic primitives – where the key may be secret-shared between parties – it is often essential to consider complexity as a function of the key. We want to highlight that our candidate falls into $\mathsf{AC}^0[2]$ even if considered as a function of the input *and key*. Further, we observe that for certain settings of distributed computation the speed-up to $\tilde{O}(N)$ (or $O(N)$) key size and evaluation time outlined in the previous paragraph carries over, as the transformation can be applied to each secret-share of the key individually (i.e., each matrix secret share can be replaced by a trapdoored LPN matrix).

Cryptanalysis. We show that our candidate exhibits the following security properties:

- **Fourier-analytic resilience:** The function lacks heavy Fourier coefficients, resisting adaptive learning attacks that identify such coefficients [KM91].
- **Linear attacks:** We also generalize the linear test framework popular in code-based cryptography [CRR21] to include chosen (but non-adaptive) samples. In fact, we show that attacks based on finding heavy Fourier coefficients fall within this framework. Briefly, the adversary may choose certain evaluation points and then take the XOR of the outputs, and succeeds if this XOR is noticeably biased. The main challenge is controlling the bias when the query points satisfy non-trivial linear dependencies: nonetheless, among other results, we show that with constantly many non-adaptive queries such linear attacks cannot succeed.

- **Non-sparsity over $\mathbb{F}_2$:** The function is not sparse with high probability when represented as a polynomial over $\mathbb{F}_2$, and thus evades property testing attacks that learn sparse polynomials [DLM+07].
- **Resistance to algebraic attacks:** The function has high rational degree, namely $\Omega(\kappa^c)$ for any $c < 1/4$, and therefore resists rational-degree attacks using up to 2^{κ^c} queries (see Section 2.3). This additionally suggests subexponential security even in the weak PRF setting, that is, it plausibly remains secure even against adversaries with access to a subexponential number of random queries.
- **Preservation of prior guarantees:** Our construction retains resistance to the same set of attacks demonstrated in the work of [ABG+14], including low correlation with fixed function families and inapproximability by low-degree polynomials over $\mathbb{F}_2$.

Our cryptanalysis suggests that Conjecture 1.2 should hold for any $c < \frac{1}{4}$. The motivation for this threshold is the following. To resist attacks based on Fourier coefficients (including linear attacks) and rational degree, we require 2^ℓ to be large. To resist attacks based on rational degree and sparsity, we additionally require 2^{mn} to be large. We therefore choose parameters so that $\ell = mn = \lambda$. Under this choice,

$$N = \ell mn = \lambda^2, \qquad \text{and hence} \qquad \kappa = N^2 + N = \Theta(\lambda^4).$$

Thus, attacks with complexity 2^λ correspond to running time $2^{\kappa^{1/4}}$, which corresponds to the choice $c < \frac{1}{4}$.

For the query complexity, the bound $\kappa^{\log^c \kappa}$ should be viewed as conservative. Indeed, attacks arising from the Razborov–Smolensky framework suggest that security may extend to as many as $\kappa^{\log^2 \kappa}$ queries. Lastly, while concrete cryptanalysis is not our focus, we mention that for $\lambda = 128$, it is consistent with our cryptanalysis that an adversary running in time $T = 2^{128}$ could have success probability at most $1/T$ given access to 2^{30} chosen queries. We provide these numbers not as a concrete security guarantee, but rather as a target for improvement for future cryptanalysis.

Universal PRFs from Strong PRFs in $\mathsf{AC}^0[2]$. Furthermore, we show that if a strong PRF in $\mathsf{AC}^0[2]$ exists, then we can construct a *universal PRF* in $\mathsf{AC}^0[2]$, a single, fixed function in $\mathsf{AC}^0[2]$ which serves as a PRF. This is notable because it yields a single target function for building distinguishers against, which is potentially easier to obtain than a full natural property. Moreover, if such a universal distinguisher could be transformed into a natural property, then the absence of PRFs in $\mathsf{AC}^0[2]$ would imply that $\mathsf{AC}^0[2]$ is learnable, addressing the long-standing question on the dichotomy between the existence of PRFs and learning for this class.

Theorem 1.3 *If there exists a PRF $f_n : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ secure against $\mathsf{poly}(n)$ time and query adversaries in $\mathsf{AC}^0[2]$ for some $b \geq 1$, then there exists a fixed, universal function $UG : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ in $\mathsf{AC}^0[2]$, which is a PRF secure against $\mathsf{poly}(n)$ time and query adversaries.*

Theoretical Implications. Our candidate's conjectured security would imply the hardness of learning $\mathsf{AC}^0[2]$ circuits using a fixed quasipolynomial number of membership queries. We formally state this implication, extended to all primes p,[2] in the conjecture below:

Conjecture 1.4 *There exists a $c > 0$ such that $\mathsf{AC}^0[p]$ circuits cannot be learned with $\kappa^{\log^c \kappa}$ many membership queries and 2^{κ^c} time.*

Limitations of Natural Property-Based Learners for $\mathsf{AC}^0[p]$. Razborov and Smolensky showed that every circuit in $\mathsf{AC}^0[p]$ of size m and depth d can be approximated by a polynomial over $\mathbb{F}_p$ of degree $O((p-1)^d \log^d m)$ with inverse-polynomial error [Raz87, Smo87]. Combined with testers for low-degree polynomials [KL01, BKS+10], this yields a *natural property* for $\mathsf{AC}^0[p]$, namely a distinguishing algorithm that separates functions in the class from random functions with nontrivial advantage [RR94].

A classical lower bound from coding theory [JPRZ09] shows that testing whether a function is a degree-k polynomial over $\mathbb{F}_p$ or ε-far from such a polynomial requires at least $\Omega\left(\frac{1}{\varepsilon} + p^{\frac{k+1}{p-1}}\right)$ queries.[3]

[2] That is, in an $\mathsf{AC}^0[p]$ circuit one may use MOD_p gates, which output 0 iff the number of input wires carrying a 1 is 0 mod p; else a 1.

[3] This lower bound already holds when interacting directly with an exact low-degree polynomial, rather than merely an approximation; we discuss this distinction further later.

Circuit Class	Reference	Lower Bounds
AC^0 $\mathsf{AC}^0 + O(1)$ XOR,MAJ	[LMN89] [Vio13]	*No* poly-*query PRF*
$\mathsf{DNF} \circ \mathsf{MOD2}$	[ABG+14]	*No* poly-*query PRF with* subexp-*security*
$\mathsf{AC}^0 \circ \mathsf{MOD2}$	[ABG+14]	*No* poly-*query PRF with* subexp-*security (conjectured)*
$\mathsf{AC}^0[p]$ for prime p	[RR94, KL01, CIKK16]	*No* subexp-*query PRF with* subexp-*security*
$\mathsf{MOD2} \circ \mathsf{AC}^0$	**This work**	*No* poly-*query PRF with* poly-*security*

Circuit Class	Reference	Query	Runtime	Assumption
$\mathsf{AC}^0[2]$	[Kha93, NR97, BPR12, Vio13]	quasipoly	quasipoly	Factor, DDH, LWE
	This work	quasipoly	subexp	Heuristic
almost-$\mathsf{AC}^0[2]$	[YS16]	quasipoly	subexp	Low-noise LPN
$\mathsf{CC}^0[pq]$ for prime p, q	[BIP+18, CJL25a]	exp	exp	Heuristic
quasilinear-TC^0	[MV12]	exp	exp	Heuristic
TC^0	[NR97, NRR00, BPR12] [AR16]	subexp	subexp	Factor, DDH, LWE, Goldreich's OWF

Table 1: Comparison of negative (top table) and positive (bottom table) results for low-depth PRFs (where we only consider strong PRFs, i.e., PRFs secure against chosen-query attacks). We consider the complexity of computing the output for any fixed key, where security level in terms of query complexity (Query) and adversarial runtime (Runtime) is with respect to the key length. Here, poly denotes an a-priori unbounded polynomial number of queries. We write AC^0 to denote the class of constant-depth circuits with polynomially many AND, OR, and NOT gates with unbounded fan-in. We write DNF to denote the subclass of depth-2 circuits. We write $\mathsf{DNF} \circ \mathsf{MOD2}$ to denote the complexity class DNF with mod 2 gates at the bottom (=input) layer (and equivalently for $\mathsf{AC}^0 \circ \mathsf{MOD2}$). We write $\mathsf{AC}^0[m]$ to denote the class AC^0 with mod m gates. We write $\mathsf{CC}^0[m]$ to denote the class of constant-depth circuits with polynomially many NOT gates with a single input and mod m gates with unbounded fan-in. We denote by TC^0 the class of constant-depth circuits with polynomially many AND, OR, NOT, and MAJORITY (or, equivalently, threshold) gates with unbounded fan-in. Factor refers to the hardness of factoring, DDH to the decisional Diffie-Hellman assumption, LWE refers to Learning with Errors, and LPN refers to Learning Parity with Noise.

Since the Razborov–Smolensky approximation has degree roughly $\log^d m$, any natural property based on low-degree testing must therefore use quasipolynomially many queries. Thus, obtaining a polynomial-query natural property for $\mathsf{AC}^0[p]$ would require fundamentally new techniques.

This quasipolynomial query complexity barrier is especially significant in light of the connection between natural properties and learning algorithms established by Carmosino, Impagliazzo, Kabanets, and Kolokolova (CIKK) [CIKK16]. Their learning algorithm shows how to turn the Razborov–Smolensky natural property [Raz87, Smo87] into a quasipolynomial-query learning algorithm for $\mathsf{AC}^0[p]$. More generally, their framework transforms any natural property for a complexity class $\mathcal{C}$ containing $\mathsf{AC}^0[p]$, defined for n-input circuits of size $s(n)$, denoted $\mathcal{C}[s(n)]$, into a learning algorithm for any function $f \in \mathcal{C}[s(n)]$. The learner outputs a circuit C such that, with high probability, $C(x) = f(x)$ for a large fraction of inputs $x \in \{0,1\}^n$.

Subsequent work of Oliveira and Santhanam [OS16] showed that natural property based learning algorithms can sometimes be substantially accelerated, yielding faster running times and lower query complexity. This raises a natural question:

Can the Razborov–Smolensky natural property be "bootstrapped" through the CIKK framework to obtain a significantly lower-query learner for $\mathsf{AC}^0[p]$*?*

We show that the answer is negative. Specifically, in Section 7, we prove that any learning algorithm obtained through the CIKK framework necessarily inherits the query complexity of the underlying natural property, up to the complexity of the pseudorandom generator used in the reduction.

Theorem 1.5 *Let $\mathcal{C}$ be a circuit class. Suppose $\mathcal{C}[s(n)]$ has a natural property making $g(m)$ queries, where m is the size of the target circuit. Then any CIKK learner for $\mathcal{C}$ requires at least $g(t)$ queries, where t denotes the size of the pseudorandom generator circuit used in the reduction.*

Consequently, no known CIKK-style approach can learn $\mathsf{AC}^0[p]$ using substantially fewer queries than those required by the Razborov–Smolensky natural property itself. In particular, overcoming the quasipolynomial query barrier would require either fundamentally new natural properties or qualitatively different learning techniques.

1.2 Related Work

Table 1 provides a comparison of our candidate with existing constructions, which we describe in more detail below.

PRF Candidates from LPN. Yu and Steinberger [YS16] introduce a PRF candidate based on the LPN assumption with noise rate $\mu = n^{-c}$ for any constant $0 < c < 1$. Their construction is almost in $\mathsf{AC}^0[2]$: it uses only parity and unbounded fan-in AND/OR gates, but has super-constant depth $\omega(1)$, where $\omega(1)$ can be any superconstant function.

They show that if there exists a natural property against $\mathsf{AC}^0[2]$ circuits of depth d making $O(n^{d/3})$ queries in time T with advantage $O(dn^{d/3}\epsilon)$, then there is an algorithm for solving the LPN problem with noise rate n^{-c} in αn queries for any constant $\alpha > 1$, time T and advantage ϵ. Note that one can set d to be constant to obtain a PRF in $\mathsf{AC}^0[2]$, but this would limit the adversary to a fixed polynomial number of queries. In contrast, our candidate lies fully within $\mathsf{AC}^0[2]$ (with depth just 3) and we conjecture that it remains secure against $p(\lambda)$ queries for *any* polynomial p.

In related work, Ding et al. construct a weak PRF in NC^1 based on strong LPN and propose multiple strong PRFs in NC^1 based on variants of LPN [DJK25]. Their constructions additionally satisfy key homomorphism, a property useful for many symmetric cryptographic applications. Since NC^1 strictly contains $\mathsf{AC}^0[2]$, these results highlight known separations between the two classes.

PRF Candidates from Factoring, DDH, and Ring LWE. Candidate PRFs in TC^0 have been constructed based on standard cryptographic assumptions such as Factoring, the Diffie–Hellman problem [NR97], and Ring-LWE [BPR12]. Through a generic reduction by Viola [Vio13] from subexponentially secure PRFs in TC^0 to quasipolynomially secure PRFs in $\mathsf{AC}^0[p]$, these constructions yield candidate PRFs in $\mathsf{AC}^0[p]$ that are secure against quasipolynomial-time adversaries, assuming the subexponential hardness of the underlying assumptions. However, the reductions to these assumptions do not hold in our bounded query setting. Viola's construction takes the form $f_s(h_t(x))$, where f_s is the TC^0 candidate on $\mathsf{poly}\log n$-bit inputs and h_t is a hash function. In the bounded query setting, this candidate admits a distinguisher with inverse quasipolynomial advantage by searching the key space. Because we aim for security against subexponential adversaries with non-negligible advantage, this motivates the need for a new candidate construction.

Other Low-Complexity PRF Candidates. Two weak PRF candidates in $\mathsf{AC}^0[2]$ that achieve subexponential security are proposed in [BCG+21]. However, neither of these candidates is suitable as a strong PRF. The first candidate lies in AC^0 layered on top of parity gates and has *heavy Fourier coefficients* (i.e., of weight inverse quasipolynomial). This makes it distinguishable from random with quasipolynomial advantage, even just given a constant number of queries: see Corollary 5.6 and Remark 5.7 below.The second candidate is a *sparse polynomial* over $\mathbb{F}_2$, which can be efficiently learned with membership queries using algorithms from [DLM+07].

Separately, Boneh et al. [BIP+18] introduce a strong PRF candidate with exponential security in $\mathsf{CC}^0[6]$, and conjecture that similar constructions may exist for $\mathsf{CC}^0[m]$ where $m = pq$ for distinct primes p and q. Inspired by this candidate, Chen et al. [CJL25b] formalize the Learning with Alternating Moduli (LAM) Assumption which they show to be as hard as LWE for certain large moduli. They propose candidate weak PRFs based on the LAM assumption in $\mathsf{CC}^0[m]$ where $m = pq$ for distinct primes p and q with $2^{O(n/\log n)}$ security. Note that $\mathsf{AC}^0[p]$ does not contain $\mathsf{CC}^0[pq]$ since MOD_q gates cannot be computed in $\mathsf{AC}^0[p]$ [Smo87].

Applebaum and Raykov [AR16] also propose a weak PRF candidate in AC^0 with quasipolynomial security and a strong PRF candidate in TC^0 with subexponential security based on Goldreich's one-way function [Gol00]. In addition, Alwen et al. [AKPW13] prove hardness of learning with rounding (LWR) under the learning with errors (LWE) assumption for any prescribed polynomial sample bound, yielding a bounded-sample weak PRF.

Cryptanalysis in the Bounded Query Setting Existing cryptanalytic techniques that derive additional queries from a low number of queries include Lyubashevsky's sample amplification for LPN [Lyu05], XL [CKPS00], and Gröbner-basis methods [CG20]. However, such techniques cannot circumvent the information-theoretic query barrier for attacks based solely on low-degree approximability. An attack relying solely on the existence of a degree-d approximation must also handle exact degree-d polynomials.

Since a uniformly random such polynomial is $(2^{d+1}-1)$-wise independent, no generic distinguisher for this approximation class can use fewer than 2^{d+1} queries, regardless of subsequent sample amplification.

1.3 Open Questions

We highlight several directions that remain open.

- **General linear tests.** Proving resistance of our candidate against *all* linear tests remains open. In particular, extending our result against a constant number of nonadaptive queries to support an arbitrary polynomial number of queries seems difficult. In addition, extending our result against polynomially many linearly independent queries to polynomially many linearly dependent queries is challenging because our current analysis leads to a signed sum for which a termwise application of the triangle inequality gives an insufficient bound. Obtaining a sharper bound would require showing many cancellations among the terms, which we do not currently know how to do. Note that even in the weak-PRF regime under non-adaptive queries, establishing resistance to linear tests for the second PRF candidate of [BCG+21] is still unresolved.
- **Dichotomy between Learning and PRFs.** Can we, in general, build PRFs in $\mathsf{AC}^0[2]$ assuming only the non-learnability of the class in the quasipolynomial-query regime?
- **Circuit-size lower bound for Nisan–Wigderson.** Our query lower bound in Theorem 1.5 for CIKK depends on the circuit size of the Nisan–Wigderson (NW) function(Definition 7.4). Establishing a circuit-size lower bound for the NW function would therefore yield a correspondingly stronger statement on CIKK's query complexity.

1.4 Technical Overview

Impossibility of PRFs in $\mathsf{MOD}_2 \circ \mathsf{AC}^0$. Our distinguishing algorithm combines Håstad's switching lemma with a simple parity property of shallow decision trees. We start by writing $f \in \mathsf{MOD}_2 \circ \mathsf{AC}^0$ as

$$f = \bigoplus_{j=1}^{m} C_j,$$

where each C_j is a polynomial-size constant-depth AC^0 circuit. The main challenge is that while individual AC^0 circuits are well understood via Fourier concentration and switching-lemma arguments, it is not immediate how to extend these techniques to the XOR of polynomially many such circuits.

Our approach is inspired by the Linial–Mansour–Nisan (LMN) learning algorithm [LMN93]. As in LMN, we repeatedly apply random restrictions to simplify each AC^0 summand into a shallow decision tree. Concretely, after composing d independent random restrictions, each depth-d summand C_j simplifies, with high probability, to a decision tree of depth $O(\log n)$, while still leaving many variables unset.

At this point, however, our argument diverges from LMN. Instead of exploiting Fourier concentration, the distinguisher selects a random affine subcube Q whose dimension is slightly larger than the resulting decision-tree depth, and queries the function on every point of Q. The key observation is that any Boolean function computable by a decision tree of depth strictly smaller than the dimension of the cube must have even parity on that cube. Thus, each restricted summand C_j has parity zero on Q, and therefore so does their XOR:

$$\bigoplus_{x \in Q} f(x) = 0.$$

By contrast, for a uniformly random Boolean function, the parity on any fixed nonempty affine cube is an unbiased random bit. Consequently, testing the parity of the function on a random logarithmic-dimensional affine cube yields a polynomial-query nonadaptive distinguisher for $\mathsf{MOD}_2 \circ \mathsf{AC}^0$.

Strong PRF Candidate in $\mathsf{AC}^0[2]$. Our construction builds on the weak PRF candidate of Akavia et al. [ABG+14], which is based on the TRIBES function

$$\mathsf{TRIBES}_{m,n}(x) = \bigvee_{j=1}^{m} \bigwedge_{k=1}^{n} x_{j,k},$$

where the input $x \in \{0,1\}^M$ is partitioned into m blocks of size n, with $M = mn$. For suitable choices of m and n,[4] the function has roughly equal probability of outputting 0 and 1.

The candidate of [ABG+14] uses a secret binary matrix $\mathbf{A}$ to linearly transform the input before applying TRIBES. To remove the residual bias of the TRIBES function, the construction additionally XORs in one output bit of the linear transformation:

$$\begin{aligned} &F_{\mathbf{A}} : \{0,1\}^{M+1} \to \{0,1\}, \\ &F_{\mathbf{A}}(x) = g(\mathbf{A}x), \qquad g(y,z) = \mathsf{TRIBES}_{m,n}(y) \oplus z, \end{aligned} \tag{2}$$

where arithmetic is over $\mathbb{F}_2$.

Although this candidate was later shown vulnerable to quasipolynomial-time rational-degree attacks [BR17], it still plausibly resists subexponential attacks in the bounded-query setting. A natural idea would therefore be to conjecture that the construction remains secure even against quasipolynomially many *chosen* queries. However, this turns out to be false.

The issue is that the candidate always possesses a heavy Fourier coefficient. Indeed, Akavia et al. [ABG+14] showed that their construction has a Fourier coefficient of magnitude at least $1/\kappa^{\log^c \kappa}$ for some constant $c > 0$. Consequently, the function is learnable in quasipolynomial time using the Kushilevitz–Mansour algorithm [KM91]. More importantly, distinguishing a function with a heavy Fourier coefficient from a uniformly random function requires only a *constant* number of chosen queries.[5] Thus, avoiding heavy Fourier coefficients is essential for achieving subexponential security in the bounded-query chosen-sample setting.

Motivated by this obstacle, we propose the following candidate strong PRF. Let $N = \ell mn$. The secret key consists of an invertible matrix $\mathbf{A} \in \{0,1\}^{N \times N}$ together with a uniformly random shift vector $\mathbf{b} \in \{0,1\}^N$. Define

$$\begin{aligned} &F_{\mathbf{A},\mathbf{b}} : \{0,1\}^N \to \{0,1\}, \\ &F_{\mathbf{A},\mathbf{b}}(x) = g(\mathbf{A}x \oplus \mathbf{b}), \qquad g(\mathbf{y}) = \bigoplus_{i=1}^{\ell} \mathsf{TRIBES}_{m,n}(\mathbf{y}_i), \end{aligned} \tag{3}$$

where the vector $\mathbf{y}$ is partitioned into ℓ blocks $\mathbf{y}_1, \ldots, \mathbf{y}_\ell \in \{0,1\}^{mn}$.

Design Rationale. The random shift $\mathbf{b}$ is necessary to avoid trivial distinguishers, such as evaluating the function at the zero input. More fundamentally, our design is motivated by the apparent limitations of smaller subclasses of AC^0[2].

A conjecture of Akavia et al. [ABG+14], later reiterated in a slightly stronger form in [BCG+21], suggests that every function in $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ has a heavy Fourier coefficient. Since functions with heavy Fourier coefficients are efficiently distinguishable in the chosen-query setting, this conjecture would rule out PRFs in $\mathsf{AC}^0 \circ \mathsf{MOD}_2$. This complements our negative result described above on the impossibility of PRFs in $\mathsf{MOD}_2 \circ \mathsf{AC}^0$, where parity gates are only used at the output layer.

Our candidate is designed to escape both obstructions by using the class $\mathsf{MOD}_2 \circ \mathsf{AC}^0 \circ \mathsf{MOD}_2$, featuring both an inner and an outer layer of parities. The outer XOR layer eliminates the heavy Fourier coefficients present in the Akavia et al. candidate, while still remaining within AC^0[2]. In this sense, the construction can be viewed as the simplest plausible candidate beyond the currently known barriers.

Cryptanalysis. To provide evidence for our conjectured security, we analyze our candidate against several classes of attacks, with particular focus on linear attacks and sparsity-based attacks, as these appear to require new techniques beyond prior work. The remaining guarantees, such as resistance to low-degree polynomial approximation and low correlation with fixed function families, follow more directly from known analyses of the TRIBES function and its variants [ABG+14].

Linear attacks. A major obstacle in constructing strong PRFs from low-complexity functions is resistance to linear tests. Informally, a linear attack chooses query points $x^{(1)}, \ldots, x^{(Q)}$ together with coefficients $u_1, \ldots, u_Q \in \{0,1\}$, and attempts to detect a nontrivial bias in

$$\bigoplus_{i=1}^{Q} u_i F(x^{(i)}).$$

[4] Namely, one takes $n \approx \log M$ and $m \approx M/\log M$, which makes $\mathsf{TRIBES}_{m,n}$ approximately balanced.

[5] This observation is implicit in the literature [O'D14, Theorem 1.30]; we discuss it explicitly in Section 5.2.

Such attacks generalize classical linear cryptanalysis and Fourier-based distinguishers, and are ubiquitous in code-based cryptanalysis.

Our first observation is that the Akavia et al. candidate [ABG+14] is inherently vulnerable to this type of attack because it possesses a heavy Fourier coefficient. Indeed, finding a large Fourier coefficient can itself be viewed as a special case of a linear test. This motivates our introduction of the outer XOR layer, whose role is precisely to destroy the large Fourier character present in prior candidates.

To further analyze our construction, we prove resistance against several structured classes of linear tests. One of the more challenging classes to deal with is where the adversary may choose highly correlated query sets, making the resulting bias difficult to bound. This kind of attack captures, for instance, our subcube attack that rules out PRFs in $\mathsf{MOD}_2 \circ \mathsf{AC}_0$, as well as more general forms of differential cryptanalysis. We show that any such linear attack based on $Q = O(1)$ arbitrary static queries can only succeed with negligible probability.[6] Boosting this to cover a polynomial number of queries appears to be a challenging open problem: in particular, we cannot just rely on all of the Fourier coefficients being small in magnitude, but would need to somehow argue that positive and negative contributions largely cancel out.

We then consider different attack models where the adversary's queries may be *adaptive*, but are restricted in other ways. We first show resistance to attackers who query up to N adaptively chosen, linearly independent input vectors. Here, we rely on the observation that for a matrix $\mathbf{X}$ of linearly independent queries, the attacker's adaptivity can be absorbed by the key matrix $\mathbf{A}$, and $\mathbf{AX}$ is always a uniformly random set of linearly independent vectors. Secondly, we consider a hybrid attack that contains an additional phase of an arbitrary number of random queries following the bounded adaptive queries, and then defines a final test query by summing up the previous responses. Here, we show that no fixed test function applied to the previous responses can be used to predict the final response. Both of these proofs rely solely on the property that our PRF is an XOR of functions of the form $h_i(\mathbf{A}_i x \oplus \mathbf{b}_i)$, where the functions h_i have Fourier coefficients bounded in magnitude by some $\rho < 1$, or decaying as $o(1)$. The results therefore apply equally well if TRIBES is swapped out for another function with a sufficiently small spectrum.

Sparsity attacks. Another notable attack comes from sparse polynomial representations over $\mathbb{F}_2$. Prior work [DLM+07] showed that sparse Boolean functions can be efficiently learned with membership queries, and the second candidate of [BCG+21] is not a strong PRF precisely because it admits a sparse polynomial representation.

In contrast, we show that our candidate has subexponentially many monomials with high probability. The proof proceeds by analyzing the maximal-degree coefficients of the polynomial representation. The key observation is that each such coefficient can be expressed as the determinant of a structured random submatrix derived from the secret matrix $\mathbf{A}$. This converts the sparsity question into a problem about the behavior of random determinants over $\mathbb{F}_2$. We show that these determinants are nonzero with noticeable probability and, crucially, that the corresponding events exhibit only weak correlation. This allows us to prove concentration bounds showing that subexponentially many maximal-degree monomials survive simultaneously with high probability.

The remaining challenge is to bound the covariance between different determinant events. Since distinct coefficients may depend on heavily overlapping sets of matrix entries, the corresponding vanishing events are highly dependent, and standard independence-based tail bounds do not apply. To overcome this, we condition on all but one column, after which the remaining determinant becomes a linear function of the unconditioned column. Whenever an associated submatrix is full rank, this linear function is unbiased and therefore exhibits complete cancellation. Consequently, the covariance is bounded by the probability that ℓ random independent $M \times (M-1)$-submatrices all fail to have full rank, which is exponentially small. Combining this covariance bound with Chebyshev's inequality yields the result, implying that the polynomial representation contains a subexponential number of monomials with overwhelming probability and therefore evades known sparse-polynomial learning attacks.

More generally, the covariance argument is not specific to TRIBES. It applies more broadly to candidates of the form $F_{\mathbf{A},\mathbf{b}}(x) = g(\mathbf{A}x + \mathbf{b})$ whenever the relevant polynomial coefficients of $g(\mathbf{A}x + \mathbf{b})$ can be represented as determinants of random submatrices of $\mathbf{A}$.

Resistance to algebraic attacks. As a byproduct of adding the outer XOR-layer in our candidate, we can show that the candidate has high rational degree $\Omega(\kappa^c)$ for any $c < 1/4$ and therefore provides

[6] In fact, while we do not attemp to optimize this, an inspection of the proof shows we can handle roughly $Q = o(\sqrt{\log \lambda})$ queries.

security against rational-degree attacks using 2^{κ^c} queries (see Section 2.3 for a more detailed discussion). Note that this does not help to boost the PRF security in the bounded-query setting (as here, even [ABG+14] resists rational-degree attacks). However, unlike [ABG+14], our candidate PRF plausibly attains full subexponential security as a weak PRF – that is, it remains secure even against adversaries with access to a subexponential number of random input/output pairs (as distinguishing algorithms based on Razborov–Smolensky require access to chosen samples). A similar result was achieved in [BCG+21]. Their candidate weak PRF, however, falls in $\mathsf{AC}^0 \circ \mathsf{MOD2}$ and therefore does not provide any security guarantees in the chosen sample setting for the same reason as [ABG+14].

Universal PRFs from Strong PRF in $\mathsf{AC}^0[2]$. We show that the existence of any strong PRF in $\mathsf{AC}^0[2]$ implies the existence of a *universal* PRF in the same class, namely a single fixed function family that is itself pseudorandom.

The first step is to show that if a PRF exists in $\mathsf{AC}^0[2]$, then there also exists one constructable in bounded polynomial time. Concretely, by combining domain-shrinking and domain-extension arguments, we show that any PRF in $\mathsf{AC}^0[2]$ can be transformed into another PRF in the same class computable by circuits constructable in time $O(n^2)$. This allows us to effectively enumerate all candidate PRF families by enumerating Turing machines running for at most n^2 steps and interpreting their outputs as $\mathsf{AC}^0[2]$ circuits.

Since any explicit PRF family has a constant-size description, some index in a $\mathsf{poly}(n)$-size enumeration for large enough n must correspond to a secure PRF. We therefore define a universal candidate by XORing together the first $\mathsf{poly}(n)$ enumerated circuit families:

$$UF(k_1, \ldots, k_t, x) = \bigoplus_{i=1}^{t} F_i(k_i, x),$$

where $F_1, \ldots, F_t$ are the first enumerated candidates.

If the combined function UF were distinguishable from random, then, one could isolate an index j such that replacing F_j with a truly random function changes the distinguisher's acceptance probability by a non-negligible amount. Since at least one enumerated family F_j is itself a secure PRF, this yields a contradiction.

Finally, we apply the same domain-shrinking and domain-extension technique used earlier to restore the key and input lengths to the desired parameters while preserving membership in $\mathsf{AC}^0[2]$. This yields a fixed universal PRF family in $\mathsf{AC}^0[2]$.

Impossibility of Learning Speedup. In Section 7, we argue that existing natural-property-based techniques cannot yield polynomial-query distinguishers against $\mathsf{AC}^0[p]$.

Our starting point is the learning framework of Carmosino, Impagliazzo, Kabanets, and Kolokolova (CIKK) [CIKK16], which converts natural properties into learning algorithms using the Nisan–Wigderson (NW) generator. At a high level, the learner embeds the target function into an NW pseudorandom generator and then applies the natural property distinguisher to the resulting NW function.

A natural hope is that one could substantially reduce the query complexity of this approach using known "speedups" for CIKK-style learners, such as those of Oliveira and Santhanam [OS16], since the input length of the NW function can be much smaller than the original input length. We show that this intuition breaks down once the circuit-size dependence of the natural property is taken into account. Although the NW function has small input length, its circuit size can still be large. Moreover, for classes such as $\mathsf{AC}^0[p]$, any meaningful natural property must necessarily depend on circuit size rather than input length alone, since every Boolean function on n bits can be represented by an exponential-size depth-2 DNF.

Formally, we prove that every CIKK learner inherits the query complexity of the underlying natural property evaluated at the circuit size of the NW function. Consequently, since the best known natural properties for $\mathsf{AC}^0[p]$ require quasipolynomially many queries as a function of circuit size, any known CIKK-style learner must also use quasipolynomially many queries unless one substantially improves the circuit complexity of NW generators.

2 Preliminaries

Throughout this work, we denote by λ the security parameter. We say a function $f(\lambda)$ is negligible in the security parameter λ if $f(\lambda) = \lambda^{-\omega(1)}$. We denote this by writing $f(\lambda) = \mathsf{negl}(\lambda)$. We write

$\mathsf{poly}(\lambda)$ to denote a function that is bounded by a fixed polynomial in λ. We say an algorithm is efficient if it runs in probabilistic polynomial time (PPT) in the length of its input. For a positive integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. For a set S, $x \leftarrow S$ denotes uniformly sampling from S; additionally, $x_1, \ldots, x_L \leftarrow S$ for $L \in \mathbb{N}$ denotes that $x_1, \ldots, x_L$ are all independent and uniformly sampled from S. We use $\mathcal{U}_n$ to denote the uniform distribution over $\{0,1\}^n$. For a vector $x \in \{0,1\}^N$, we let $|x| := |\{i \in [N] : x_i = 1\}|$ denote the (Hamming) weight of the vector.

We will use the following standard result on the rank of a random matrix over $\mathbb{F}_2$.

Proposition 2.1 (Rank of a random binary matrix) *Let $0 \le Q \le N$, and let $\mathcal{R}_{N\times Q}$ denote the set of full-rank $N \times Q$ matrices over $\mathbb{F}_2$. It holds that*

$$c_{N,Q} := \Pr_{V \leftarrow \mathbb{F}_2^{N\times Q}}[V \in \mathcal{R}_{N\times Q}] = \prod_{i=0}^{Q-1}(1 - 2^{i-N}) \ge \prod_{i=1}^{\infty}(1 - 2^{-i}) > 0.288.$$

Moreover,

$$1 - c_{N,Q} \le \sum_{i=0}^{Q-1} 2^{i-N} < 2^{Q-N}.$$

Proof. The formula for $c_{N,Q}$ follows by sampling each column of V one-by-one, multiplying the probabilities, for each i, that column $i+1$ falls outside of the span of the previous i columns, of size 2^i. The upper bound for $1 - c_{N,Q}$ follows by expanding the product formula. □

2.1 Fourier Analysis of Boolean Functions

In this section (and this work), we will regularly identify $\{0,1\}$ with $\mathbb{F}_2$, the finite Galois field with two elements.

Characters and Fourier basis. For $s \in \{0,1\}^N$, define the *character*

$$\chi_s(x) \;=\; (-1)^{\langle s,x\rangle} \qquad (x \in \{0,1\}^N)\ .$$

The set $\{\chi_s : s \in \mathbb{F}_2^N\}$ forms an orthonormal basis for real-valued functions on the Boolean hypercube $\{0,1\}^N$ under the inner product

$$\langle f, g\rangle \;=\; \mathbb{E}_x\big[f(x)\, g(x)\big], \qquad f, g : \{0,1\}^N \to \mathbb{R}\ .$$

Here and elsewhere, the notation $\mathbb{E}_x$ means that this expectation is over a uniformly random $x \in \{0,1\}^N$. For this inner-product, the characters form an orthonormal basis.

Lemma 2.2 (Orthonormality) *For all $s, t \in \{0,1\}^N$,*

$$\mathbb{E}_x\big[\chi_s(x)\chi_t(x)\big] = \mathbf{1}_{s=t}.$$

Fourier coefficients and expansion. For $f : \{0,1\}^N \to \mathbb{R}$, the *Fourier coefficient* of f at $s \in \{0,1\}^N$ is

$$\widehat{f}(s) \;=\; \mathbb{E}_x\big[f(x)\,\chi_s(x)\big]\ .$$

Due to Lemma 2.2, every f admits the Fourier expansion

$$f(x) \;=\; \sum_{s\in\{0,1\}^N} \widehat{f}(s)\,\chi_s(x)\ .$$

Lemma 2.2 implies this change-of-basis transformation is orthogonal, at least when the norms are suitably defined. That is, we have the following.

Theorem 2.3 (Plancherel's Theorem) *For all $f, g : \{0,1\}^N \to \mathbb{R}$,*

$$\mathbb{E}_x[f(x)g(x)] \;=\; \sum_{s\in\{0,1\}^N} \widehat{f}(s)\widehat{g}(s).$$

When $f = g$, we have the following special case.

Theorem 2.4 (Parseval's Theorem) *For all $f : \{0,1\}^N \to \mathbb{R}$,*

$$\mathbb{E}_x[f(x)^2] \;=\; \sum_{s\in\{0,1\}^N} \widehat{f}(s)^2.$$

In particular, when the range of f is $\{-1,1\}$, then as clearly $\mathbb{E}_x\big[f(x)^2\big] = 1$, we have

$$\sum_{s\in\{0,1\}^N} \widehat{f}(s)^2 \;=\; 1 \;.$$

Lastly, we record what it means for a function to have a concentrated Fourier spectrum.

Definition 2.5 (Fourier ε-concentration) *Let $\mathcal{F} \subseteq \mathbb{F}_2^N$ be a set. We say that $f : \{0,1\}^N \to \{-1,1\}$ has Fourier mass ε-concentrated on $\mathcal{F}$ if*

$$\sum_{s\in\mathcal{F}} \widehat{f}(s)^2 \;\geq\; 1-\varepsilon \;.$$

See [O'D14] for further background.[7]

2.2 Pseudorandom Functions

We give the definitions of weak pseudorandom functions (WPRFs) and of standard pseudorandom functions (PRFs, or *strong* PRFs). We note that, as is standard, the adversary's query complexity, running time, and success probability are all measured with respect to the *key length* of the PRF.

Definition 2.6 ((Weak) Pseudorandom Function) *Let $\lambda \in \mathbb{N}$ be the security parameter, and let $N = N(\lambda)$ and $\kappa = \kappa(\lambda)$ denote polynomially bounded functions specifying the input length and key length, respectively.*

A (weak) pseudorandom function is a family of functions $\mathcal{F} = \{f_\lambda : \{0,1\}^\kappa \times \{0,1\}^N \to \{0,1\}\}$, where $f_\lambda(k,x)$ can be evaluated in polynomial time given (k,x). We write $f_k(x)$ instead of $f_\lambda(k,x)$. For $Q = Q(\kappa)$, $T = T(\kappa)$ and $\varepsilon = \varepsilon(\kappa)$, we say that $\mathcal{F}$ is a (Q,T,ε)-secure (strong) pseudorandom function (PRF) *if for every $\lambda \in \mathbb{N}$ and every oracle circuit $\mathcal{A}$ of size $T(\kappa)$,*

$$\left| \Pr_{k\leftarrow\{0,1\}^\kappa}[\mathcal{A}^{f_k(\cdot)} = 1] - \Pr_{R}[\mathcal{A}^{R(\cdot)} = 1] \right| \leq \varepsilon(\kappa),$$

where $R : \{0,1\}^N \to \{0,1\}$ denotes a uniformly random function. We call $\mathcal{F}$ a (Q,T,ε)-secure weak PRF (WPRF) *if the same condition holds when $\mathcal{A}$ is only given input-output pairs $(x_i, f_k(x_i))$, where each $x_i \leftarrow \{0,1\}^n$ is sampled uniformly and independently. For brevity, we also say $\mathcal{F}$ is a* (Q,T)-secure (strong/weak) PRF *if it is $(Q,T,1/T)$-secure (in either the strong or weak sense).*

Remark 2.7 *We remark that it is a common observation that one can trade running time for success probability. That is, to increase success probability, one can run the adversary multiple times with independent random tapes, and take a majority vote. However, this amplification also requires the adversary to use additional queries. Hence, if a PRF is (Q,T,ε)-secure with $Q = \mathsf{poly}(\kappa)$ but $\varepsilon = 2^{-\kappa^{\Omega(1)}}$, amplification to constant success would demand exponentially many queries, it is unclear how to obtain an attack with constant success probability while maintaining polynomial query complexity.*

Remark 2.8 *In all PRF constructions considered in this paper, the security parameter λ, input length N, and key length κ are polynomially related. Unless stated otherwise, we express security bounds in terms of λ. Polynomial resource bounds, negligibility, and polynomial circuit-size bounds are invariant under switching between these parameters. Since circuit depth is unaffected by this reparameterization, membership in $\mathsf{AC}^0[2]$ is also unchanged.*

[7] We remark that there, Fourier characters/coefficients use a domain of $\{-1,1\}$.

2.3 Algebraic Attacks and Rational Degree

Bogdanov et al. [BR17] used an algebraic attack to break Akavia et al.'s weak PRF candidate [ABG+14]. The attack requires the candidate to have a low *rational degree*. We define rational degree below and later in Section 5.5 prove our candidate has high rational degree.

Definition 2.9 ((Primal/Dual) Rational Degree) *Let $F : \mathbb{F}_2^N \to \mathbb{F}_2$ be a boolean function, which we consider as an element of the Boolean ring $\mathcal{R}_N := \mathbb{F}_2[x_1, \ldots, x_N]/I$, where $I = \langle x_1^2 - x_1, \ldots, x_N^2 - x_N \rangle$ denotes the Boolean ideal. The* primal rational degree *of F is*

$$\min\{d : \exists G \in \mathcal{R}_N \text{ with } \deg(G) = d \text{ s.t. } GF \equiv 0 \pmod{I}\} .$$

The dual rational degree *of F is*

$$\min\{d : \exists G \in \mathcal{R}_N \text{ with } \deg(G) = d \text{ s.t. } G(F+1) \equiv 0 \pmod{I}\} .$$

The rational degree $\mathsf{rdeg}(F)$ *is the minimum of its primal and dual rational degree.*

If a function has rational degree d (either primal or dual), one can look to interpolate a degree-d multilinear polynomial G witnessing $GF \equiv 0 \pmod{I}$ (in the primal case) or $G(F+1) \equiv 0 \pmod{I}$ (in the dual case). The interpolation is done by viewing the $\binom{N+1}{d}$ coefficients of G as unknowns, and solving for them via Gaussian elimination. To ensure that there are enough equations, one must query $\Theta\left(\binom{N+1}{d}\right)$ samples. Note that random samples are sufficient.

This attack demonstrates that, if F has rational degree d, then it is *not* a $\big(N^{O(d)}, N^{O(d)}, \Omega(1)\big)$-weak PRF. Note that in the weak PRF setting, in order for our candidate to satisfy subexponential security *if the adversary is also given subexponentially many samples*, it is necessary for us to argue that our candidates rational degree is $\geq \kappa^{\Omega(1)}$ (where we recall κ denotes key-length). This is what we do in Section 5.5 (in fact, we show the rational degree is $\Omega\big(\kappa^{1/6}\big)$).

2.4 Linear Attacks

Linear attacks have been implicitly considered in much of the cryptanalytic literature for many years (especially in code-based cryptography), but was only recently formalized (see, e.g., [CRR21]). The framework has become popular as it encompasses many practical attacks on code-based assumptions, especially those based on information set decoding (ISD) [Pra62, Ste89, FS09, BLP11, MMT11, BJMM12, MO15, EKM17, BM18].

To provide context, we describe how linear attacks apply to a dual-LPN style assumption which states that given a matrix $\mathbf{A}$ sampled according to some distribution, $(\mathbf{A}, \mathbf{Ae})$ is indistinguishable from $(\mathbf{A}, \mathbf{b})$ where $\mathbf{b}$ is a uniformly random vector, even if $\mathbf{e}$ is random sparse vector. The linear attack proceeds by identifying a nonzero *attack vector* $\mathbf{u}$ such that $\langle \mathbf{u}, \mathbf{y} \rangle$ is biased when $\mathbf{y} = \mathbf{Ae}$ for a randomly chosen sparse vector $\mathbf{e}$. Equivalently, $\mathbf{uA}$ must have low Hamming weight; i.e., $\mathbf{u}$ corresponds to a message whose codeword under $\mathbf{A}$ has low weight. Detecting such a bias in $\mathbf{y}$ is already enough to distinguish $(\mathbf{A}, \mathbf{Ae})$ from uniform.

This generalizes to an attack on a weak PRF in the following way. Given randomly sampled inputs $x^{(1)}, \ldots, x^{(Q)}$, before seeing the values $F(k, x^{(1)})$, $\ldots$, $F(k, x^{(Q)})$, the adversary must choose an attack vector $\mathbf{u} \in \{0,1\}^Q \setminus \{\mathbf{0}\}$, and the adversary is said to succeed if the bias

$$\left| \mathbb{E}\left[(-1)^{\sum_{i=1}^{Q} u_i F_k(x^{(i)})} \right] \right|$$

is non-negligible, where the randomness is over the key k, the values $x^{(1)}, \ldots, x^{(Q)}$, and any other internal randomness of the adversary.[8]

A contribution of our work is to define a generalized notion of linear tests that even *strong PRF's* must satisfy. Essentially, we allow the adversary to non-adaptively choose a set of query points $x^{(1)}, \ldots, x^{(Q)}$. In particular, they need not be independent and uniform. Again, before submitting the queries to the oracle, the adversary must choose the attack vector $\mathbf{u}$, and the attack is said to succeed if the adversaries determines a non-trivial bias in the inner product of the PRF values and the vector $\mathbf{u}$. The formal definition follows.

[8] Since we typically will not be explicitly giving a running time bound on the adversaries, we will be able to assume the adversary is deterministic. But in full generality, we allow the adversary to toss random coins, and this is useful when we generalize to the non-adaptive setting.

Definition 2.10 (Non-adaptive linear test for Strong PRF) *Let $\{f_k : \{0,1\}^n \to \{0,1\} \mid k \leftarrow \{0,1\}^\kappa\}$ be a family of keyed functions, and let $Q \in \mathbb{N}$ be a query bound. Let $x^{(1)}, \ldots, x^{(Q)} \in \{0,1\}^n$ be distinct inputs* chosen non-adaptively *by the adversary, and let $\mathbf{u} = (u_1, \ldots, u_Q) \in \{0,1\}^Q \setminus \{\mathbf{0}\}$ be a vector* chosen by the adversary *(also known as the* attack vector*) prior to making any queries. The adversary is said to* succeed *in a linear attack if*

$$\mathrm{bias}(\mathbf{u}) := \left|\mathbb{E}\left[(-1)^{\sum_{i=1}^{Q} u_i f_k(x^{(i)})}\right]\right| \geq \varepsilon(n)$$

for some non-negligible function $\varepsilon(n)$, where the randomness is over the choice of the key k and any internal randomness of the adversary.

In this setting of non-adaptive linear tests, it in fact suffices to consider the attack vector $\mathbf{u} = (1, 1, \ldots, 1)$. This is because the adversary must decide the attack vector $\mathbf{u}$ based on her queries $x^{(1)}, \ldots, x^{(Q)}$, and if she wishes to set $u_i = 0$, she can simply not submit the query $x^{(i)}$. (This implicitly assumes the adversary need not fix ahead of time exactly how many queries it will make to the PRF, but we consider that reasonable.) In Section 5.2, all the non-adaptive linear attacks we consider have attack vector $\mathbf{u} = (1, 1, \ldots, 1)$.

Additionally, we will even consider a natural *adaptive* attack that can be viewed as a linear attack (see Proposition 5.10). However, defining adaptive linear attacks in full generality appears to be quite unreasonable (at least, any "natural" definition seems to be impossible to fulfill, as it would imply information-theoretic security of the PRF).

3 Impossibility of PRFs in $\mathsf{MOD}_2 \circ \mathsf{AC}^0$

Since AC^0 is learnable, a natural next candidate class for pseudorandom functions is $\mathsf{MOD}_2 \circ \mathsf{AC}^0$, namely XORs of AC^0 circuits. In this section, we show that this class is still too weak: polynomial-size XORs of AC^0 circuits can be distinguished from uniform random functions using only polynomially many nonadaptive queries. Together with a conjecture of Akavia et al. [ABG+14], later reiterated in a slightly stronger form [BCG+21], that every function in $\mathsf{AC}^0 \circ \mathsf{MOD}_2$ has a heavy Fourier coefficient, this suggests that the most minimal remaining candidate class for PRFs is $\mathsf{AC}^0[2]$.

We now make this argument formal. For a restriction ρ, the notation $f \restriction \rho$ means the restricted function. For $0 \leq p \leq 1$, R_p denotes the usual random restriction distribution in which each variable is left unset with probability p, and otherwise is fixed to a uniformly random bit, independently across variables. Let $\mathsf{DT}(f)$ be the least possible depth of a decision tree computing f.

Fix constants $c, d \geq 1$. Let

$$\mathcal{F}_{n,c,d} = \left\{ f : \{0,1\}^n \to \{0,1\} \;\middle|\; \begin{array}{l} f = \bigoplus_{j=1}^{m} C_j, \quad m \leq n^c\,, \\ \text{and each } C_j \text{ is a depth-}d\ \mathsf{AC}^0 \text{ circuit of size at most } n^c \end{array} \right\}.$$

Thus $\mathcal{F}_{n,c,d}$ is the class of top-parity-of-AC^0 functions with polynomially many summands and polynomial-size depth-d summands. The goal is to distinguish $\mathcal{F}_{n,c,d}$ from a uniformly random Boolean function on n bits by making only polynomially many queries.

The first step is the switching-lemma, which simplifies each AC^0 summand to a shallow decision tree under random restrictions. The second step is proving a parity property of shallow decision trees on affine subcubes. We begin by stating Håstad's switching lemma and deriving the decision-tree formulation that will be used below.

Lemma 3.1 (Håstad's switching lemma) *Let F be a CNF or a DNF of width w. Then for every $0 \leq p \leq 1/2$ and every $t \in \mathbb{N}$,*

$$\Pr_{\rho \sim \mathcal{R}_p}\left[\mathsf{DT}(F \restriction \rho) \geq t\right] \leq (5pw)^t.$$

Corollary 3.2 (Decision-tree form of the switching lemma) *Let F be the AND or OR of a finite family $\{F_i\}_{i \in I}$ of Boolean functions. Suppose that $\mathsf{DT}(F_i) \leq w$ for every $i \in I$. Then for every $0 \leq p \leq 1/2$ and every $t \in \mathbb{N}$,*

$$\Pr_{\rho \sim \mathcal{R}_p}\left[\mathsf{DT}(F \restriction \rho) \geq t\right] \leq (5pw)^t.$$

Proof. Every Boolean function of decision-tree depth at most w can be written both as a w-DNF and as a w-CNF: take one term for each accepting leaf, or one clause excluding each rejecting leaf. Therefore, if F is an OR of the F_i's, then F is a w-DNF. If F is an AND of the F_i's, then F is a w-CNF. The claim follows from Lemma 3.1. □

We prove the parity lemma about shallow decision trees below.

Lemma 3.3 (Parity lemma for decision trees) *Let $g : \{0,1\}^k \to \{0,1\}$. If $\mathsf{DT}(g) < k$, then*

$$\bigoplus_{x \in \{0,1\}^k} g(x) = 0.$$

Equivalently, on every affine subcube of dimension k, a Boolean function of decision-tree depth $< k$ has even parity.

Proof. We prove the first statement by induction on k. The case $k = 1$ is immediate. Suppose $k \geq 2$.

If g is constant, then the parity of its truth table on $\{0,1\}^k$ is 0 because 2^k is even. Otherwise, let the root of an optimal decision tree for g query some variable, say x_i. Write

$$g_0 := g \restriction (x_i = 0), \qquad g_1 := g \restriction (x_i = 1).$$

Then $\mathsf{DT}(g_b) \leq \mathsf{DT}(g) - 1 < k - 1$ for each $b \in \{0,1\}$. By the induction hypothesis,

$$\bigoplus_{y \in \{0,1\}^{k-1}} g_0(y) = 0 \qquad \text{and} \qquad \bigoplus_{y \in \{0,1\}^{k-1}} g_1(y) = 0.$$

Therefore

$$\bigoplus_{x \in \{0,1\}^k} g(x) = \left(\bigoplus_{y \in \{0,1\}^{k-1}} g_0(y) \right) \oplus \left(\bigoplus_{y \in \{0,1\}^{k-1}} g_1(y) \right) = 0.$$

The affine-subcube version follows by identifying any k-dimensional affine subcube with $\{0,1\}^k$. □

Finally, we prove the main result of this section, a distinguisher for all $f \in \mathcal{F}_{n,c,d}$.

Theorem 3.4 (Main Result) *For every fixed constants $c, d \geq 1$, there is a randomized nonadaptive oracle algorithm $A = A_{c,d}$ making $n^{O_{c,d}(1)}$ queries such that for every sufficiently large n,*

$$\Pr[A^f = 1] \geq 1 - n^{-9} \qquad \text{for every } f \in \mathcal{F}_{n,c,d},$$

while for a uniformly random Boolean function $F : \{0,1\}^n \to \{0,1\}$,

$$\Pr[A^F = 1] = \tfrac{1}{2}.$$

Hence A distinguishes $\mathcal{F}_{n,c,d}$ from uniform random with advantage at least $\frac{1}{2} - n^{-9}$.

Proof. Choose an integer $B = B(c,d)$ so large that $B \log_2 20 \geq 2c + 20$. Set

$$\ell := \lceil B \log_2 n \rceil, \qquad p := \frac{1}{100\ell}, \qquad k := \ell + 1.$$

The algorithm A is as follows.

1. Sample independent restrictions $\rho_1, \ldots, \rho_d \sim R_p$.
2. Let $\sigma := \rho_1 \rho_2 \cdots \rho_d$ be their composition.
3. Let $L(\sigma) \subseteq [n]$ be the set of variables left unset by σ.
4. If $|L(\sigma)| < k$, output a fair coin.
5. Otherwise, choose uniformly random $S \subseteq L(\sigma)$ with $|S| = k$, and choose a uniformly random assignment to the variables in $L(\sigma) \setminus S$.
6. This defines a k-dimensional affine subcube $Q \subseteq \{0,1\}^n$. Query the oracle on all 2^k points of Q.

7. Accept iff

$$\bigoplus_{x\in Q} f(x) = 0.$$

Since $\ell = \Theta_{c,d}(\log n)$, the algorithm will query at most

$$2^k \le 2^{\ell+1} \le 4n^B$$

points, which is polynomial in n.

We analyze the two cases.

1. $f \in \mathcal{F}_{n,c,d}$*:* Fix

$$f = \bigoplus_{j=1}^{m} C_j \in \mathcal{F}_{n,c,d}, \qquad m \le n^c,$$

where each C_j is a depth-d AC^0 circuit of size at most n^c.

Fix one summand $C = C_j$. For a gate g of C, define the *height* of g to be the number of AND/OR gates on a longest input-to-g path. Thus literals have height 0 and the output gate has height at most d.

For $r = 0, 1, \ldots, d$, let $\sigma_r := \rho_1 \cdots \rho_r$ (with σ_0 the empty restriction), and let $\mathcal{E}_r(C)$ be the event that every gate of height at most r in $C \restriction \sigma_r$ computes a function of decision-tree depth at most ℓ.

The event $\mathcal{E}_0(C)$ holds deterministically, since literals have decision-tree depth $0 \le \ell$.

Now fix $r \in \{1, \ldots, d\}$ and condition on $\mathcal{E}_{r-1}(C)$. Let g be a gate of height exactly r. Under σ_{r-1}, each child of g computes a function of decision-tree depth at most ℓ. Since g is an AND or OR of its children, Corollary 3.2 gives

$$\Pr\big[\mathsf{DT}(g \restriction \sigma_r) > \ell \mid \sigma_{r-1}, \mathcal{E}_{r-1}(C)\big] \le (5p\ell)^{\ell+1} = 20^{-\ell}.$$

The circuit C has at most n^c gates, so by a union bound over all gates of height exactly r,

$$\Pr\big[\neg\mathcal{E}_r(C) \mid \mathcal{E}_{r-1}(C)\big] \le n^c \cdot 20^{-\ell}.$$

Iterating the inequality

$$\Pr[\neg\mathcal{E}_{i+1}(C)] \le \Pr[\neg\mathcal{E}_i(C)] + \Pr[\neg\mathcal{E}_{i+1}(C) \mid \mathcal{E}_i(C)]$$

and using that $\mathcal{E}_0(C)$ holds deterministically, we get

$$\Pr\big[\neg\mathcal{E}_d(C)\big] \le \sum_{i=0}^{d-1} \Pr\big[\neg\mathcal{E}_{i+1}(C) \mid \mathcal{E}_i(C)\big] \le dn^c\, 20^{-\ell}.$$

Union-bounding over the at most $m \le n^c$ summands, we obtain

$$\Pr\big[\exists j \in [m] : \neg\mathcal{E}_d(C_j)\big] \le dn^{2c}\, 20^{-\ell} \le n^{-10}$$

for all sufficiently large n, by the choice of B.

Hence, with probability at least $1 - n^{-10}$ over the random restrictions, every restricted summand $C_j \restriction \sigma$ has decision-tree depth at most ℓ.

Next we show that with overwhelming probability there are enough free variables left to define the k-dimensional cube Q. Since the d restrictions are independent, each variable remains free after their composition with probability exactly p^d, independently of the other variables. Thus

$$|L(\sigma)| \sim \mathrm{Bin}\big(n, p^d\big) \qquad \text{and} \qquad \mu := \mathbb{E}[|L(\sigma)|] = np^d = \frac{n}{(100\ell)^d}.$$

Because $\ell = \Theta(\log n)$ and d is constant, we have $\mu \gg \log n$, and in particular $\mu \ge 32k$ for all sufficiently large n. A Chernoff bound therefore gives

$$\Pr\big[|L(\sigma)| < k\big] \le e^{-31^2\mu/(2\cdot 32^2)} \le n^{-10}$$

for all sufficiently large n.

Combining the last two bounds, with probability at least $1 - 2n^{-10}$, both of the following hold:

- every summand $C_j \restriction \sigma$ has decision-tree depth at most ℓ; and
- $|L(\sigma)| \geq k$.

On this event, the algorithm defines the cube Q by fixing all but k free variables. Decision-tree depth cannot increase under further restriction, so for every j, $\mathsf{DT}(C_j \restriction Q) \leq \ell < k$. Applying Lemma 3.3 to the k-dimensional cube Q, we get

$$\bigoplus_{x \in Q} C_j(x) = 0 \qquad \text{for every } j \in [m].$$

Therefore

$$\bigoplus_{x \in Q} f(x) = \bigoplus_{x \in Q} \Big(\bigoplus_{j=1}^{m} C_j(x) \Big) = \bigoplus_{j=1}^{m} \Big(\bigoplus_{x \in Q} C_j(x) \Big) = 0,$$

so the algorithm accepts.

Thus

$$\Pr[A^f = 1] \geq 1 - 2n^{-10} \geq 1 - n^{-9}$$

for all sufficiently large n.

2. Random case: Let $F : \{0,1\}^n \to \{0,1\}$ be uniformly random.

Condition on the internal randomness of the algorithm up to the point where the cube Q is determined. If $|L(\sigma)| < k$, then the algorithm outputs a fair coin by definition. Otherwise the algorithm queries all points of a fixed k-dimensional affine cube Q. Since F is uniformly random, the values $\{F(x) : x \in Q\}$ are independent unbiased bits, and hence

$$\Pr\left[\bigoplus_{x \in Q} F(x) = 0 \;\middle|\; Q \right] = \frac{1}{2}.$$

Therefore in all cases,

$$\Pr[A^F = 1] = \frac{1}{2}.$$

This proves the theorem. □

4 Strong PRFs in $\mathsf{AC}^0[2]$

Let λ denote the security parameter. We first define an auxiliary function $G : \{0,1\}^N \to \{0,1\}$ as the product of ℓ independent TRIBES functions on disjoint input blocks:

$$G(x) = \bigoplus_{i=1}^{\ell} \mathsf{TRIBES}_{m,n}\big(x^{(i)}\big), \tag{4}$$

where $x^{(i)}$ denotes the i^{th} M-bit block of x, and where each $\mathsf{TRIBES}_{m,n}$ function is given by

$$\mathsf{TRIBES}_{m,n}(x) = \bigvee_{j=1}^{m} \bigwedge_{k=1}^{n} x_{j,k}.$$

where $x_{j,k}$ denotes the k^{th} bit in the j^{th} sub-block of x. Here:

- $n = \log \lambda - \log \ln \lambda + O(1)$ and $m = \ln(2) 2^n + O(1)$ are the *balanced tribes* parameters, chosen so that $mn = \lambda$,
- $M = mn$ is the input length to each TRIBES function and $N = \ell mn$ is the input length to G, where $\ell = \lambda$.

These choices of $n = \Theta(\log \lambda)$ and $m = \Theta(\lambda / \log \lambda)$ guarantee that $\mathsf{TRIBES}_{m,n}$ is fairly *unbiased*, i.e., so that $\Pr_x[\mathsf{TRIBES}_{m,n}(x) = 1] \approx 1/2$. Specifically, following [O'D14, Proposition 4.12] (and the surrounding discussion), these parameters ensure that $\Pr_x[\mathsf{TRIBES}_{m,n}(x) = 1] = 1 - (1 - 2^{-n})^m = 1/2 - O\left(\frac{\log \lambda}{\lambda}\right)$.

We now define our candidate strong PRF. Let the PRF key $\mathbf{A} \in \{0,1\}^{N \times N}$ be a randomly sampled invertible matrix over $\mathbb{F}_2$ and $\mathbf{b} \in \{0,1\}^N$ be sampled uniformly at random. We remark that the key length is $\kappa(\lambda) = \lambda^4 + \lambda^2$.

Candidate Strong PRF. Our candidate strong PRF $F_{\mathbf{A},\mathbf{b}} : \{0,1\}^N \to \{0,1\}$ is defined as

$$F_{\mathbf{A},\mathbf{b}}(x) \;=\; G\big(\mathbf{A}x \oplus \mathbf{b}\big)\ . \tag{5}$$

Throughout the rest of the paper, we use F and G as defined above. We remark as well that $F_{\mathbf{I},\mathbf{0}} = G$, i.e., when the key is set to $(\mathbf{I},\mathbf{0})$ the PRF is this function G.

At certain points, it will be useful for us to analyze the version of our function where the range is $\{-1,1\}$ rather than $\{0,1\}$. That is, we will sometimes consider the function $f_{\mathbf{A},\mathbf{b}} : \{0,1\}^N \to \{-1,1\}$ defined as

$$f_{\mathbf{A},\mathbf{b}}(x) := (-1)^{F_{\mathbf{A},\mathbf{b}}(x)}\ .$$

We similarly, we define

$$g(x) := (-1)^{G(x)}, \qquad\qquad \mathsf{TRIBES}^{\pm}_{m,n}(x) := (-1)^{\mathsf{TRIBES}_{m,n}(x)}.$$

In particular, we then have $g(x) = \prod_{i=1}^{\ell} \mathsf{TRIBES}^{\pm}_{m,n}(x^{(i)})$ and $f_{\mathbf{A},\mathbf{b}}(x) = g(\mathbf{A}x \oplus \mathbf{b})$, where as above $x^{(i)}$ denotes the i^{th} M-bit block of x.

5 Rationales for Security

In this section, we demonstrate that our candidate withstands several classes of attacks. Specifically, we show that it is resilient to Fourier-analytic attacks, linear tests, and sparse polynomial testing. In addition, we establish that our candidate inherits the same guarantees as the construction of Akavia et al. [ABG+14], including inapproximability by low-degree polynomials over finite fields and the absence of correlation with any fixed family of functions. In each subsection, we prove resistance against attacks running in time $O(2^{\lambda})$. This implies that the security is $2^{O(\kappa^{1/4})}$, sub-exponential in the key length κ (recall $\kappa(\lambda) = \lambda^4$).

5.1 Absence of Large Fourier Characters

We analyze the Fourier spectrum of our candidate function $f_{\mathbf{A},\mathbf{b}}$ as defined in Equation (3). Note here that we are explicitly looking at the $\{\pm 1\}$-valued version of the Boolean function. We prove that all Fourier coefficients are exponentially small in the security parameter λ.

First, before continuing, we record the Fourier-characters for the version of the TRIBES function with range $\{\pm 1\}$.

Proposition 5.1 (Fourier Coefficients of TRIBES [O'D14], Proposition 4.14) *We have*

$$\widehat{\mathsf{TRIBES}^{\pm}}_{m,n}(0) = 2(1-2^{-n})^m - 1\ .$$

Else, if $z = (z^{(1)},\ldots,z^{(m)}) \neq 0$ *where* $z^{(i)}$ *denotes the* i^{th} n*-bit block of* $z \in \{0,1\}^M$*, if* $p = |\{i \in [m] : z^{(i)} \neq 0\}|$*, then*

$$\widehat{\mathsf{TRIBES}^{\pm}}_{m,n}(z) = 2(-1)^{p+|z|}2^{-pn}(1-2^{-n})^{m-p}\ .$$

Note that, for a Boolean function $f : \{0,1\}^N \to \{-1,1\}$,

$$\hat{f}(0) = \mathbb{E}_x[f(x)\chi_0(x)] = \Pr[f(x)=1] - \Pr[f(x)=-1]\ ;$$

hence, showing that $|\hat{f}(0)|$ is small is enough to argue that f is *unbiased*. Thus, the following proposition is a strengthening of the claim that our PRF candidate is unbiased.

Proposition 5.2 (No Large Fourier Coefficients)

$$\max_{z\in\{0,1\}^N}\Big\{\big|\widehat{f_{\mathbf{A},\mathbf{b}}}(z)\big|\Big\} \le O(2^{-\lambda}).$$

Proof. Our proof relies on two lemmas.

Lemma 5.3 (Product Lemma for Disjoint Inputs) *Let ℓ, M, N be positive integers such that $N = \ell M$. Let $h_1, \ldots, h_\ell : \{0,1\}^M \to \{-1,1\}$ be Boolean functions, and define $f : \{0,1\}^N \to \{-1,1\}$ by*

$$f(x) = \prod_{i=1}^{\ell} h_i(x^{(i)}),$$

where $x = (x^{(1)}, \ldots, x^{(\ell)})$ is partitioned into ℓ disjoint blocks of size M. Then

$$\widehat{f}(z) = \prod_{i=1}^{\ell} \widehat{h}_i(z^{(i)}) \quad \textit{and} \quad \max_{z \in \widehat{O}^N} |\widehat{f}(z)| = \prod_{i=1}^{\ell} \max_{z^{(i)} \in \{-1,1\}^M} |\widehat{h}_i(z^{(i)})|.$$

Proof. For $z = (z^{(1)}, \ldots, z^{(\ell)})$,

$$\begin{aligned}
\widehat{f}(z) &= \frac{1}{2^N} \sum_{x \in \{0,1\}^N} \prod_{i=1}^{\ell} h_i(x^{(i)}) \cdot (-1)^{\langle x, z \rangle} \\
&= \prod_{i=1}^{\ell} \left[\frac{1}{2^M} \sum_{x^{(i)} \in \{0,1\}^M} h_i(x^{(i)}) \cdot (-1)^{\langle x^{(i)}, z^{(i)} \rangle} \right] \\
&= \prod_{i=1}^{\ell} \widehat{h}_i(z^{(i)}).
\end{aligned}$$

Taking maxima over z yields the second claimed equality. □

In our construction, the blocks of inputs to the ℓ tribes functions are mixed by $\mathbf{A}$ (and shifted by $\mathbf{b}$), rather than being disjoint. The following lemma shows that an invertible linear transformation preserves the maximum magnitude of a Fourier coefficient.

Lemma 5.4 *Let f and g be functions such that $f(x) = g(\mathbf{A}x + \mathbf{b})$ with $\mathbf{A} \in \{0,1\}^{N \times N}$ invertible over $\mathbb{F}_2$ and $\mathbf{b} \in \{0,1\}^N$. Then*

$$\max_{z \in \{0,1\}^N} |\widehat{f}(z)| = \max_{z \in \{0,1\}^n} |\widehat{g}(z)|.$$

Proof. Let $y = \mathbf{A}x + \mathbf{b}$, so $x = \mathbf{A}^{-1}(y - \mathbf{b})$. Then

$$\begin{aligned}
\widehat{f}(z) &= \frac{1}{2^N} \sum_{x \in \{0,1\}^N} g(\mathbf{A}x + \mathbf{b}) \cdot (-1)^{\langle x, z \rangle} \\
&= \frac{1}{2^N} \sum_{y \in \{0,1\}^N} g(y) \cdot (-1)^{\langle \mathbf{A}^{-1}(y - \mathbf{b}), z \rangle} \\
&= (-1)^{\langle z, -\mathbf{A}^{-1}\mathbf{b} \rangle} \cdot \frac{1}{2^N} \sum_{y \in \{0,1\}^N} g(y) \cdot (-1)^{\langle y, (\mathbf{A}^{-1})^\top z \rangle} \\
&= (-1)^{\langle z, -\mathbf{A}^{-1}\mathbf{b} \rangle} \cdot \widehat{g}\big((\mathbf{A}^{-1})^\top z\big),
\end{aligned}$$

which implies $|\widehat{f}(z)| = |\widehat{g}((\mathbf{A}^{-1})^\top z)|$ and hence the maxima are equal. □

Now, we apply the lemmas to our candidate. From Proposition 5.1, we see that for our choice of parameters,

$$\max_z |\widehat{\mathsf{TRIBES}}^{\pm}(z)| \leq O\left(\tfrac{\log \lambda}{\lambda}\right) .$$

Indeed, following the discussion in Section 4, we see that or choice of m and n guarantees that $(1 - 2^{-n})^m \leq \frac{1}{2} + O\left(\frac{\log \lambda}{\lambda}\right)$, and so

$$|\widehat{\mathsf{TRIBES}}^{\pm}(0)| = \left|2(1 - 2^{-n})^m - 1\right| \leq O\left(\tfrac{\log \lambda}{\lambda}\right) .$$

Otherwise, for $z = (z^{(1)}, \ldots, z^{(m)}) \neq 0$ (where as in Proposition 5.1, $z^{(i)}$ denotes the i^{th} n-bit block of $z \in \{0,1\}^M$), we have $p = |\{i \in [m] : z^{(i)} \neq 0\}| \geq 1$, implying

$$|\widehat{\mathsf{TRIBES}}^{\pm}(z)| = \left|2(-1)^{p+|z|} 2^{-pn} (1 - 2^{-n})^{m-p}\right| \leq 2 \cdot 2^{-pn} = O\left(\tfrac{1}{\lambda^p}\right) .$$

Therefore, by Lemma 5.3,

$$\max_z |\widehat{G}(z)| \le \left(O\left(\tfrac{\log\lambda}{\lambda}\right)\right)^{\ell}.$$

Since $\ell = \lambda$, this is certainly at most $O(2^{-\lambda})$, as desired. Finally, $f_{\mathbf{A},\mathbf{b}}(x) = g(\mathbf{A}x \oplus \mathbf{b})$, so Lemma 5.4 implies that $\max_z |\widehat{f}(z)| = \max_z |\widehat{g}(z)|$, completing the proof. □

This shows that the largest Fourier coefficient of $f_{\mathbf{A},\mathbf{b}}$ is exponentially small in λ, and thus $f_{\mathbf{A},\mathbf{b}}$ has no non-negligibly large Fourier characters.

5.2 Resistance Against Linear Tests

We show that our candidate $F_{\mathbf{A},\mathbf{b}}$ as defined in Equation (3) resists a specific set of non-adaptive linear attacks as defined in Definition 2.10. Specifically, we show resistance against a simple linear-sum attack with polynomially many queries and all non-adaptive linear attacks with $O(1)$ queries. Additionally, recalling the discussion that followed the definition, we will always assume the attack vector $\mathbf{u} = (1, 1, \ldots, 1)$.

Again in this section we find it easier to work with $\{-1, 1\}$-valued versions of the functions, namely, $f_{\mathbf{A},\mathbf{b}}$ and g.

Resisting a Linear-Sum Attack We study the following natural *linear-sum* distinguishing attack, inspired by the Blum–Luby–Rubinfeld (BLR) linearity test. Fix an integer $Q \ge 3$. The distinguisher samples $x^{(1)}, \ldots, x^{(Q-1)} \leftarrow \mathbb{F}_2^N$ and sets

$$x^{(Q)} \;:=\; \sum_{i=1}^{Q-1} x^{(i)}.$$

It then outputs the result of $\sum_{i=1}^{Q} F(x^{(i)})$. The parity bias of this distinguisher on F is equivalent to the following:

$$\Delta_Q(f) \;:=\; \mathbb{E}\Big[\prod_{i=1}^{Q} f\big(x^{(i)}\big)\Big] \;=\; \mathbb{E}\Big[(-1)^{\sum_{i=1}^{Q} F\left(x^{(i)}\right)}\Big] \;,$$

where the expectation is over the random choice of the $x^{(i)}$'s. Equivalently, if the distinguisher accepts when $\sum_{i=1}^{Q} F(x^{(i)}) = 0 \pmod 2$, then its acceptance probability is $\Pr[\text{accept}] = \frac{1+\Delta_Q(f)}{2}$. Thus, against a uniformly random function, for which this parity is unbiased, the distinguishing advantage is $\frac{|\Delta_Q(f)|}{2}$.

We obtain a closed form Fourier sum for the bias of the linear sum attack in Lemma 5.5.

Lemma 5.5 (Bias as a Fourier sum) *For any $Q \ge 3$ and any $h : \mathbb{F}_2^N \to \{-1, 1\}$,*

$$\Delta_Q(h) \;=\; \sum_{s \in \mathbb{F}_2^N} \widehat{h}(s)^{\,Q}.$$

Proof. Expand each $h(x^{(i)})$ into its Fourier series:

$$\prod_{i=1}^{Q} h(x^{(i)}) \;=\; \sum_{s_1,\ldots,s_Q \in \mathbb{F}_2^N} \Big(\prod_{i=1}^{Q} \widehat{h}(s_i)\Big)(-1)^{\sum_{i=1}^{Q}\langle s_i, x^{(i)}\rangle}.$$

Taking expectation over the jointly distributed $x^{(1)}, \ldots, x^{(Q)}$ with the constraint $x^{(Q)} = \sum_{i=1}^{Q-1} x^{(i)}$ gives

$$\mathbb{E}\Big[(-1)^{\sum_{i=1}^{Q}\langle s_i, x^{(i)}\rangle}\Big] \;=\; \prod_{i=1}^{Q-1} \mathbb{E}_{x^{(i)}}\Big[(-1)^{\langle s_i + s_Q,\, x^{(i)}\rangle}\Big],$$

which equals 1 iff $s_i = s_Q$ for every $i \in [Q-1]$, and equals 0 otherwise. Thus only the diagonal terms $s_1 = \cdots = s_Q = s$ survive, yielding $\Delta_Q(h) = \sum_{s \in \mathbb{F}_2^N} \widehat{h}(s)^Q$. □

We can first use this observation to argue that functions with their Fourier mass concentrated on a small set (as in Definition 2.5) are distinguished from uniform via a linear attack.

Corollary 5.6 *[Distinguisher for functions with concentrated Fourier mass] If $h : \{0,1\}^N \to \{-1,1\}$ has its Fourier coefficients ε-concentrated on a set $\mathcal{F}$, then the linear-sum attack distinguishes g from a random function with advantage $\frac{(1-\varepsilon)^2}{2|\mathcal{F}|}$.*

Proof. If h is ε-concentrated on a set $\mathcal{F}$, then by Cauchy-Schwarz, setting $Q = 4$ in Lemma 5.5 yields

$$\Delta_4(h) \geq \sum_{s\in\mathcal{F}} \widehat{h}(s)^4 \;\geq\; \frac{1}{|\mathcal{F}|}\left(\sum_{s\in\mathcal{F}} \widehat{h}(s)^2\right)^2 \;\geq\; \frac{(1-\varepsilon)^2}{|\mathcal{F}|} \;.$$

□

Remark 5.7 *The Kushilevitz–Mansour algorithm [KM91] learns any Boolean function whose Fourier mass is ε-concentrated on a set $\mathcal{F}$ of size K in time* $\mathsf{poly}(K, n, 1/\varepsilon)$. *For distinguishing a PRF from random, however, full learning is unnecessary. Under the same spectral-concentration assumption, a much simpler* distinguisher *suffices. In particular, one can phrase the attack as a linear test that queries the function on a constant number of inputs and checks the corresponding correlation. This highlights the strength of linear tests. Consequently, establishing resistance against* all *linear tests would, in particular, rule out Fourier-based attacks that rely on detecting heavy coefficients. But we emphasize that attacks detecting heavy Fourier coefficients constitute just a subset of all possible linear attacks.*

We now prove the upper bound that we use for our candidates.

Proposition 5.8 (Small Fourier spectrum ⇒ negligible bias) *Let $h : \mathbb{F}_2^N \to \{-1, 1\}$ and suppose $\max_{s\in\mathbb{F}_2^N} |\widehat{h}(s)| \leq \tau$. Then for every $Q \geq 3$,*

$$|\Delta_Q(h)| \;\leq\; \tau^{\,Q-2} \;.$$

In particular, if $\tau = 2^{-\Omega(\lambda)}$ and $Q \geq 3$, then $|\Delta_Q(h)| = 2^{-\Omega(\lambda)}$.

Proof. The statement follows from Lemma 5.5, the triangle inequality, and Parseval through the series of inequalities below.

$$|\Delta_Q(h)| \;=\; \left|\sum_s \widehat{h}(s)^{\,Q}\right| \;\leq\; \sum_s |\widehat{h}(s)|^Q \;\leq\; \tau^{\,Q-2}\sum_s \widehat{h}(s)^2 \;=\; \tau^{\,Q-2} \;.$$

□

Applying the upper bound to $g(x)$ which Proposition 5.2 establishes has exponentially small Fourier coefficients, we obtain Corollary 5.9.

Corollary 5.9 *For $x^{(1)}, \ldots, x^{(Q-1)} \leftarrow \mathbb{F}_2^N$ and $x^{(Q)} = \sum_{i=1}^{Q-1} x^{(i)}$, conditioned on all Q queries being distinct,*

$$\Delta_Q(g) \;=\; 2^{-\Omega(\lambda)}.$$

Finally, we show that our candidate $F_{\mathbf{A},\mathbf{b}}$ resists the linear sum attack.

Proposition 5.10 *Let $x^{(1)}, \ldots, x^{(Q-1)} \leftarrow \mathbb{F}_2^N$ and $x^{(Q)} = \sum_{i=1}^{Q-1} x^{(i)}$ conditioned on all Q queries being distinct. Define*

$$\delta \;:=\; \mathbb{E}\Big[(-1)^{\sum_{i=1}^{Q} F_{\mathbf{A},\mathbf{b}}(x^{(i)})}\Big] \;=\; \Delta_Q(f) \;.$$

Then $|\delta| \;=\; 2^{-\Omega(\lambda)}$.

Proof. Let $y^{(i)} := \mathbf{A}x^{(i)} + \mathbf{b}$. Because $\mathbf{A}$ is invertible and $x^{(i)}$ are uniform and independent, the vectors $y^{(1)}, \ldots, y^{(Q-1)}$ are uniform and independent. Moreover

$$\begin{aligned} y^{(Q)} = \mathbf{A}x^{(Q)} + \mathbf{b} \;&=\; \mathbf{A}\Big(\sum_{i=1}^{Q-1} x^{(i)}\Big) + \mathbf{b} \\ &= \sum_{i=1}^{Q-1} (\mathbf{A}x^{(i)} + \mathbf{b}) + (Q \bmod 2)\,\mathbf{b} \\ &= \sum_{i=1}^{Q-1} y^{(i)} \;+\; (Q \bmod 2)\,\mathbf{b} \;. \end{aligned}$$

Hence:

- If Q is even, then $y^{(Q)} = \sum_{i=1}^{Q-1} y^{(i)}$, and by Lemma 5.5 (applied to g evaluated at the $y^{(i)}$'s) we get $\delta = \sum_s \widehat{g}(s)^Q$.
- If Q is odd, then $y^{(Q)} = \sum_{i=1}^{Q-1} y^{(i)} + \mathbf{b}$. Conditioning on $\mathbf{b}$ and repeating the same computation yields $\delta = \sum_s \widehat{g}(s)^Q (-1)^{\langle s,\mathbf{b}\rangle}$. In particular, $|\delta| \le \sum_s \left|\widehat{g}(s)\right|^Q$.

In both cases, $|\delta| \le \sum_s |\widehat{g}(s)|^Q \le \left(\max_s |\widehat{g}(s)|\right)^{Q-2}$ by Proposition 5.8. Using the bound $\max_s |\widehat{g}(s)| \le 2^{-\Omega(\lambda)}$ from Proposition 5.2 completes the proof. □

Lemma 5.5 shows that the attack's bias is upper bounded by a spectral sum of $\widehat{g}^{Q-2}$. Thus, whenever the spectrum has a *heavy* coefficient or is ε-concentrated on a small set $\mathcal{F}$, the bias is non-negligible. In addition, if all coefficients are small (as we established for $f_{\mathbf{A},\mathbf{b}}$ in Proposition 5.2), then the bias is negligible by Proposition 5.8. This shows why linear attacks subsume the usual "heavy Fourier" attacks and why our candidate resists them.

Resisting Linear Tests for $Q = O(1)$ Queries We prove that our candidate resists every fixed non-adaptive linear test making $Q = O(1)$ distinct queries. We briefly summarize the proof strategy. The actual PRF samples $\mathbf{A} \leftarrow \mathcal{GL}_N(\mathbb{F}_2)$, but this distribution is slightly inconvenient for the Fourier calculation. We therefore first analyze a relaxed test in which $\mathbf{A}$ is sampled uniformly from all matrices in $\mathbb{F}_2^{N\times N}$. Let Ξ_g^{inv} denote the bias in the actual test and Ξ_g^{all} the bias in the relaxed all-matrices test. At the end of the proof, we show that for $Q = O(1)$,

$$|\Xi_g^{\mathrm{inv}}| \le O_Q(|\Xi_g^{\mathrm{all}}|) + O_Q(2^{-N}),$$

because a uniformly random matrix is injective on the span of the Q queries except with probability $O_Q(2^{-N})$. It remains to bound Ξ_g^{all}. Expanding in the Fourier basis and applying the triangle inequality reduces this to bounding an absolute constrained Fourier sum $S_g(V)$, where the subspace $V \subseteq \mathbb{F}_2^Q$ encodes the linear relations among the query points. We bound $S_g(V)$ piece by piece: first for the function $u = 1 - \mathsf{AND}_n$, then for $q = 1 - \mathsf{TRIBES}_{m,n}$, then for the signed TRIBES block $h = (-1)^{\mathsf{TRIBES}_{m,n}}$, and finally for the full product $g = \prod_{i=1}^{\ell} h_i$.

The key insight is the following. The subspace $V^\perp$ is the set of Q-bit column patterns that are orthogonal to the constraints determined in V. Thus, $V^\perp$ determines which collections of Fourier coefficients can contribute nonzero terms to the sum. For the function $u = 1 - \mathsf{AND}_n$, the Fourier coefficient at 0 is large, namely $1 - 2^{-n}$, while every nonzero Fourier coefficient has magnitude only 2^{-n}. Hence terms in which many of the rows $\xi_1, \ldots, \xi_Q$ are zero can potentially make large contributions, whereas each nonzero row incurs a small factor of 2^{-n}. The distinctness of the queries rules out low-weight nonzero column patterns: every nonzero vector in $V^\perp$ has Hamming weight at least 4. Applying the Singleton bound to the column span of C gives a lower bound on the number of nonzero rows of C, not just on the number of nonzero columns. Consequently, every nonzero contribution to the constrained sum must involve several nonzero rows, each of which contributes a small nonzero Fourier coefficient. These forced nonzero rows provide enough factors of 2^{-n} to make the constrained Fourier sum small. This argument is useful only for $Q = o(\sqrt{\log \lambda})$: the loss from summing over subspaces of $\mathbb{F}_2^Q$, as well as the hidden constants in the constrained sum bounds, can grow exponentially in Q. Thus the proof only shows security against every fixed constant-query linear test. We leave the full proof of Theorem 5.11 to Appendix A.

Theorem 5.11 *For every fixed $Q = O(1)$ and every set of distinct non-adaptive queries $x^{(1)}, \ldots, x^{(Q)} \in \mathbb{F}_2^N$,*

$$\left| \mathop{\mathbb{E}}_{A \leftarrow \mathcal{GL}_N(\mathbb{F}_2),\, b} \left[\prod_{r=1}^{Q} g(Ax^{(r)} + b) \right] \right| \le 2^{-\Omega(\lambda \log \lambda)}.$$

Equivalently, every fixed non-adaptive linear test making $Q = O(1)$ distinct queries has negligible parity bias against the candidate.

Remark 5.12 *For every fixed constant t, this implies that our candidate is statistically almost t-wise independent. Indeed, by Vazirani's XOR lemma, bounding all possible subset sums of a vector by δ bounds its statistical distance from uniform by $\frac{1}{2}\sqrt{2^t - 1}\,\delta$, yielding distance $2^{-\Omega(\lambda \log \lambda)}$.*

5.3 Resistance to Adaptive Tests

We now turn to two different classes of attack strategies that support a larger number of queries that can even be chosen *adaptively*, albeit under linear independence conditions.

Adaptive Linearly Independent Queries. We show exponential security against linear tests for any number of adaptively chosen, linearly independent queries. Concretely, the adversary can adaptively query up to N linearly independent vectors $x^{(i)} \in \mathbb{F}_2^N$ and performs a linear test on the responses.

Theorem 5.13 *Let an adversary interact with $F_{A,b}(x) = G(\mathbf{A}x \oplus \mathbf{b})$ by adaptively issuing $Q \leq N$ queries $x^{(1)}, \ldots, x^{(Q)}$, subject to the constraint $x^{(i)} \notin \mathsf{Span}(x^{(1)}, \ldots, x^{(i-1)})$ for each i. For any fixed attack vector $u \in \mathbb{F}_2^Q \setminus \{0\}$ (chosen independently of the transcript), the bias*

$$\delta_Q \;=\; \left| \mathbb{E}\left[(-1)^{\sum_{i=1}^{Q} u_i \, F_{\mathbf{A},\mathbf{b}}(x^{(i)})}\right]\right|$$

is unconditionally bounded by $2^{-\Omega(\lambda \log \lambda)}$.

Proof. Without loss of generality, we fix our nonzero attack vector $u = (1, \ldots, 1)$, since the argument applies identically to any nonzero support.

Fix the adversary's random tape. Thus, each query $x^{(i)}$ is a deterministic function of the previous transcript. Let $v^{(i)} = \mathbf{A}x^{(i)}$. We first show that, despite the adaptivity, the tuple $\mathbf{V} = (v^{(1)}, \ldots, v^{(Q)})$ is distributed as a uniformly random full-rank $N \times Q$ matrix.

Claim 5.14 *Conditioned on any fixed value of $\mathbf{b}$, the matrix $\mathbf{V} = (\mathbf{A}x^{(1)}, \ldots, \mathbf{A}x^{(Q)})$ is uniformly distributed over the set*

$$R_{N\times Q} = \left\{\mathbf{V} \in \mathbb{F}_2^{N\times Q} : \operatorname{rank}(\mathbf{V}) = Q\right\}.$$

In particular, $\mathbf{V}$ is independent of $\mathbf{b}$.

Proof. We prove the claim by induction on the number of queries. Suppose that after $i-1$ queries, we have fixed a transcript satisfying

$$\mathbf{A}X_{<i} = V_{<i},$$

where

$$X_{<i} = (x^{(1)}, \ldots, x^{(i-1)}) \qquad \text{and} \qquad V_{<i} = (v^{(1)}, \ldots, v^{(i-1)}).$$

By the linear-independence condition imposed on the adversary, $x^{(i)} \notin \mathsf{Span}(X_{<i})$.

We claim that, conditioned on $\mathbf{A}X_{<i} = V_{<i}$, the next image $\mathbf{A}x^{(i)}$ is uniform over $\mathbb{F}_2^N \setminus \mathsf{Span}(V_{<i})$. Indeed, fix any two vectors

$$y, y' \in \mathbb{F}_2^N \setminus \mathsf{Span}(V_{<i}).$$

There exists an invertible linear map $T \in \mathcal{GL}_N(\mathbb{F}_2)$ that fixes $\mathsf{Span}(V_{<i})$ pointwise and maps y to y'. Hence the map

$$\mathbf{A} \mapsto T\mathbf{A}$$

is a bijection from the set of invertible matrices satisfying $\mathbf{A}X_{<i} = V_{<i}$ and $\mathbf{A}x^{(i)} = y$ to the set of invertible matrices satisfying $\mathbf{A}X_{<i} = V_{<i}$ and $\mathbf{A}x^{(i)} = y'$. Therefore all possible images outside $\mathsf{Span}(V_{<i})$ occur with equal probability.

It follows inductively that $(v^{(1)}, \ldots, v^{(Q)})$ is generated by choosing $v^{(1)}$ uniformly from $\mathbb{F}_2^N \setminus \{0\}$, then choosing $v^{(i)}$ uniformly from

$$\mathbb{F}_2^N \setminus \mathsf{Span}(v^{(1)}, \ldots, v^{(i-1)})$$

for each $i > 1$. This is exactly the uniform distribution over $R_{N\times Q}$. Since $\mathbf{A}$ is sampled independently of $\mathbf{b}$, the same distribution holds conditioned on every fixed value of $\mathbf{b}$, and therefore $\mathbf{V}$ is independent of $\mathbf{b}$. □

The following claim allows us to compute the bound when sampling $\mathbf{V}$ from a uniform distribution, rather than full rank matrices.

Claim 5.15

$$\delta_Q \leq 3.47 \cdot \mathop{\mathbb{E}}_{\mathbf{V}\leftarrow\mathbb{F}_2^{N\times Q}} \left[\left|\mathop{\mathbb{E}}_{\mathbf{b}}\left[\prod_{i=1}^{Q} g(v^{(i)} \oplus \mathbf{b})\right]\right|\right].$$

Proof. By Proposition 2.1, the uniformly sampled $\mathbf{V}$ has full-rank with probability $c_{N,Q} > 0.288$, and hence $\frac{1}{c_Q} < 3.47$. Therefore,

$$\begin{aligned}\delta_Q &= \frac{1}{c_Q}\left|\mathop{\mathbb{E}}_{\mathbf{V}\leftarrow\mathbb{F}_2^{N\times Q}}\left[\mathbf{1}(\mathbf{V}\in R_{N\times Q})\mathop{\mathbb{E}}_{\mathbf{b}}\left[\prod_{i=1}^{Q} g(v^{(i)}\oplus\mathbf{b})\right]\right]\right| \\ &\le \frac{1}{c_Q}\mathop{\mathbb{E}}_{\mathbf{V}\leftarrow\mathbb{F}_2^{N\times Q}}\left[\mathbf{1}(\mathbf{V}\in R_{N\times Q})\left|\mathop{\mathbb{E}}_{\mathbf{b}}\left[\prod_{i=1}^{Q} g(v^{(i)}\oplus\mathbf{b})\right]\right|\right] \\ &\le 3.47\cdot\mathop{\mathbb{E}}_{\mathbf{V}\leftarrow\mathbb{F}_2^{N\times Q}}\left[\left|\mathop{\mathbb{E}}_{\mathbf{b}}\left[\prod_{i=1}^{Q} g(v^{(i)}\oplus\mathbf{b})\right]\right|\right].\end{aligned}$$

This proves the claim. □

Claim 5.16 *With μ defined by*

$$\mu = \mathop{\mathbb{E}}_{\mathbf{V}_w\leftarrow\mathbb{F}_2^{M\times Q}}\left[\left|\mathop{\mathbb{E}}_{\mathbf{b}_w}\prod_{i=1}^{Q} h(v_w^{(i)}\oplus\mathbf{b}_w)\right|\right],$$

we have

$$\delta_Q \le 3.47\cdot\mu^\ell.$$

Proof. By Claim 5.15, it remains to bound

$$\mathop{\mathbb{E}}_{\mathbf{V}\leftarrow\mathbb{F}_2^{N\times Q}}\left[\left|\mathop{\mathbb{E}}_{\mathbf{b}}\prod_{i=1}^{Q} g(v^{(i)}\oplus\mathbf{b})\right|\right].$$

Since $g(x)=\prod_{w=1}^{\ell} h(x_w)$, and since the block submatrices $\mathbf{V}_1,\ldots,\mathbf{V}_\ell$ and block shifts $\mathbf{b}_1,\ldots,\mathbf{b}_\ell$ are mutually independent, the expectation factors over the blocks. Thus the above quantity is exactly μ^ℓ, giving

$$\delta_Q \le 3.47\cdot\mu^\ell.$$

□

Claim 5.17 *The block bias is bounded by the largest Fourier coefficient of h. Namely,*

$$\mu \le \max_{s\in\mathbb{F}_2^M}|\widehat{h}(s)|.$$

Proof. Define

$$\widetilde{h}(\mathbf{V}_w) = \mathop{\mathbb{E}}_{\mathbf{b}_w}\left[\prod_{i=1}^{Q} h(v_w^{(i)}\oplus\mathbf{b}_w)\right].$$

By Cauchy–Schwarz,

$$\mu = \mathop{\mathbb{E}}_{\mathbf{V}_w}\left[\left|\widetilde{h}(\mathbf{V}_w)\right|\right] \le \sqrt{\mathop{\mathbb{E}}_{\mathbf{V}_w}\left[\widetilde{h}(\mathbf{V}_w)^2\right]}.$$

Expanding the square gives

$$\mathop{\mathbb{E}}_{\mathbf{V}_w}\left[\widetilde{h}(\mathbf{V}_w)^2\right] = \mathop{\mathbb{E}}_{\mathbf{b},\mathbf{b}'}\mathop{\mathbb{E}}_{\mathbf{V}_w}\left[\prod_{i=1}^{Q} h(v_w^{(i)}\oplus\mathbf{b})h(v_w^{(i)}\oplus\mathbf{b}')\right].$$

Since the columns $v_w^{(1)},\ldots,v_w^{(Q)}$ are independent and uniformly distributed in $\mathbb{F}_2^M$, this equals

$$\mathop{\mathbb{E}}_{\mathbf{b},\mathbf{b}'}\left[\left(\mathop{\mathbb{E}}_{z\leftarrow\mathbb{F}_2^M}\left[h(z\oplus\mathbf{b})h(z\oplus\mathbf{b}')\right]\right)^Q\right].$$

Set $\Delta=\mathbf{b}\oplus\mathbf{b}'$. Then the inner expectation is the autocorrelation $C_h(\Delta)=\mathbb{E}_{z\leftarrow\mathbb{F}_2^M}\left[h(z)h(z\oplus\Delta)\right]$. Thus

$$\mathop{\mathbb{E}}_{\mathbf{V}_w}\left[\widetilde{h}(\mathbf{V}_w)^2\right] = \mathop{\mathbb{E}}_{\Delta\leftarrow\mathbb{F}_2^M}\left[C_h(\Delta)^Q\right].$$

We now bound this autocorrelation moment. For $h_\Delta(x) = h(x \oplus \Delta)$, we have $\widehat{h}_\Delta(s) = (-1)^{\langle \Delta, s\rangle}\widehat{h}(s)$. By Plancherel,

$$C_h(\Delta) = \mathop{\mathbb{E}}_x[h(x)h(x\oplus\Delta)] = \sum_s \widehat{h}(s)\widehat{h}_\Delta(s) = \sum_s \widehat{h}(s)^2(-1)^{\langle s,\Delta\rangle}.$$

Therefore,

$$\begin{aligned}
\mathop{\mathbb{E}}_\Delta\left[C_h(\Delta)^Q\right] &= \mathop{\mathbb{E}}_\Delta\left[\left(\sum_s \widehat{h}(s)^2(-1)^{\langle s,\Delta\rangle}\right)^Q\right] \\
&= \sum_{s_1,\ldots,s_Q}\prod_{i=1}^{Q}\widehat{h}(s_i)^2 \mathop{\mathbb{E}}_\Delta\left[(-1)^{\langle s_1+\cdots+s_Q,\Delta\rangle}\right] \\
&= \sum_{s_1+\cdots+s_Q=0}\prod_{i=1}^{Q}\widehat{h}(s_i)^2.
\end{aligned}$$

This transforms the Q-th power of the autocorrelation into a simple collision probability: viewing $p(s) := \widehat{h}(s)^2$ as a probability distribution (since by Parseval, $\sum_s p(s) = 1$), the above sum is exactly the probability that $S_1 \oplus \cdots \oplus S_Q = 0$, where each S_i is drawn independently from the distribution p. This collision probability is upper-bounded by the maximum probability of the distribution, since for every fixed choice of $s_1, \ldots, s_{Q-1}$, there is a unique value of s_Q satisfying $s_1 + \cdots + s_Q = 0$. Hence,

$$\begin{aligned}
\sum_{s_1+\cdots+s_Q=0}\prod_{i=1}^{Q}p(s_i) &\le \max_s p(s) \sum_{s_1,\ldots,s_{Q-1}}\prod_{i=1}^{Q-1}p(s_i) \\
&= \max_s p(s)\left(\sum_s p(s)\right)^{Q-1} \\
&= \max_s p(s).
\end{aligned}$$

Therefore, $\mathbb{E}_\Delta\left[C_h(\Delta)^Q\right] \le \max_s \widehat{h}(s)^2$. Combining this with Cauchy–Schwarz gives

$$\mu \le \sqrt{\mathop{\mathbb{E}}_\Delta\left[C_h(\Delta)^Q\right]} \le \max_s |\widehat{h}(s)|.$$

□

Finally, by Proposition 5.1, $\max_s |\widehat{h}(s)| \le O\left(\frac{\log\lambda}{\lambda}\right)$. Since $\ell = \Theta(\lambda)$, Claim 5.16 and Claim 5.17 imply

$$\delta_Q \le 3.47 \cdot O\left(\frac{\log\lambda}{\lambda}\right)^{\Theta(\lambda)} = 2^{-\Omega(\lambda\log\lambda)}.$$

This completes the proof.

□

Hybrid Attack: Adaptive Linearly Independent Queries, Random Queries Plus One Adaptive Sum-Query. Next, we consider a more complex form of "adaptive sum-query" attack. Here, the adversary first makes a batch of linearly independent queries as in the previous game, followed by a batch of an arbitrary number of random queries. Finally, the adversary chooses a subset of the previous query inputs to sum up, to obtain the final sum-query. An arbitrary test function is then applied to the first $Q-1$ queries, which tries to predict the result of the final query.

The game is formalized below.

Definition 5.18 (Hybrid adaptive game) *Let $Q_0 < N$, Q_1 be any function of λ and $Q = Q_0 + Q_1$. The game proceeds as follows.*

1. ***Adaptive, linearly independent queries.*** *The adversary adaptively chooses Q_0 queries $x^{(1)}, \ldots, x^{(Q_0)} \in \mathbb{F}_2^N$, subject to the constraint $x^{(i)} \notin \mathsf{Span}(x^{(1)}, \ldots, x^{(i-1)})$ for each i, receiving the outputs $c_i = g(\mathbf{A}x^{(i)} \oplus \mathbf{b}), i \in [Q_0]$.*
2. *The adversary fixes a subset selection function $\sigma : \{-1,1\}^{Q-1} \to \{S \subseteq [Q-1] : |S| \ge 2\}$ and a test function $T : \{-1,1\}^{Q-1} \to \{-1,1\}$.*
3. ***Random queries.*** *Sample Q_1 additional queries $x^{(j)}, c_j$, where $x^{(j)} \leftarrow \mathbb{F}_2^N$ and $c_j = g(\mathbf{A}x^{(j)} \oplus \mathbf{b})$.*

4. ***Adaptive sum-query and test.*** *Define the subset* $S = \sigma(c_{<Q}) \subseteq [Q-1]$ *with* $|S| \geq 2$*, and the subset sum query* $x^{(Q)} = \sum_{i \in S} x^{(i)}$*. Compute the response* $c_Q = g(\mathbf{A}x^{(Q)} \oplus \mathbf{b})$.

Finally, the test T *is applied to the first* $Q-1$ *responses. The advantage of the adversary is its correlation with the final answer,*

$$\left| \mathop{\mathbb{E}}_{\mathbf{A},\mathbf{b}} [T(c_{<Q}) \cdot c_Q] \right|.$$

Below, we show that any attack of this form has exponentially small advantage provided Q_0 is not too large. Note that, despite the additional random queries and arbitrary test function, this result not subsume Theorem 5.13 for two reasons: the sum-query test does not capture the simple linear test from Theorem 5.13, and even if it did, Theorem 5.13 gives a stronger bound that works for all Q up to N.

Theorem 5.19 *Let* $\tau = \max_s |\widehat{g}(s)| = 2^{-\Omega(\lambda \log \lambda)}$*. For any adversary in the hybrid game of Definition 5.18, with bounded test function* $T : \{-1,1\}^{Q-1} \to [-1,1]$*, the advantage*

$$\delta = \left| \mathop{\mathbb{E}}_{\mathbf{A},\mathbf{b}} [T(c_{<Q}) \cdot c_Q] \right|$$

is bounded by

$$\delta \leq \tau + \frac{\tau^2}{(1-\tau)^3} + 2^{Q_0+1-N}.$$

In particular, if $Q_0 < N - \omega(\log \lambda)$*, then*

$$\delta = \mathsf{negl}(\lambda).$$

Proof. We write

$$v^{(i)} = \mathbf{A}x^{(i)} \qquad \text{and} \qquad y^{(i)} = v^{(i)} \oplus \mathbf{b}.$$

Let

$$\mathbf{V}_0 = (v^{(1)}, \ldots, v^{(Q_0)}), \qquad \mathbf{V}_1 = (v^{(Q_0+1)}, \ldots, v^{(Q-1)}).$$

By Claim 5.14, $\mathbf{V}_0$ is distributed uniformly over $\mathcal{R}_{N \times Q_0}$, independent of $\mathbf{b}$. Furthermore, the random query block $\mathbf{V}_1$ is uniform over $\mathbb{F}_2^{N \times Q_1}$, and independent of $(\mathbf{V}_0, \vec{b})$.

We now relax the distribution on $\mathbf{V}_0$ to be uniform over all matrices.

Claim 5.20 *For every function* $F : \mathbb{F}_2^{N \times Q_0} \to [-1,1]$*,*

$$\left| \mathop{\mathbb{E}}_{\mathbf{V}_0 \leftarrow \mathcal{R}_{N \times Q_0}} [F(\mathbf{V}_0)] - \mathop{\mathbb{E}}_{\mathbf{V}_0 \leftarrow \mathbb{F}_2^{N \times Q_0}} [F(\mathbf{V}_0)] \right| \leq 2^{Q_0+1-N}.$$

Proof. By Proposition 2.1, the probability that a uniformly random $N \times Q_0$ matrix fails to have full rank is at most 2^{Q_0-N}, which bounds the statistical distance between the two distributions of $\mathbf{V}_0$. Since F is bounded in $[-1,1]$, applying F changes the expectation by at most twice the statistical distance. □

For a matrix $Y = (y^{(1)}, \ldots, y^{(Q-1)})$, let

$$c(Y) = \left(g(y^{(1)}), \ldots, g(y^{(Q-1)})\right)$$

denote the transcript induced by Y. Given this transcript, the adversary's selection function σ defines a subset

$$S(Y) \subseteq [Q-1], \qquad |S(Y)| \geq 2.$$

Write

$$P(S(Y)) = (|S(Y)| + 1) \bmod 2.$$

For the selected subset $S(Y)$, the adaptive query is $x^{(Q)} = \sum_{i \in S(Y)} x^{(i)}$. Therefore its shifted internal state is

$$\begin{aligned} \mathbf{A}x^{(Q)} \oplus \mathbf{b} &= \sum_{i \in S(Y)} \mathbf{A}x^{(i)} \oplus \mathbf{b} = \sum_{i \in S(Y)} (y^{(i)} \oplus \mathbf{b}) \oplus \mathbf{b} \\ &= \sum_{i \in S(Y)} y^{(i)} \oplus (|S(Y)| + 1)\mathbf{b} = \sum_{i \in S(Y)} y^{(i)} \oplus P(S(Y))\mathbf{b}. \end{aligned}$$

For fixed $\mathbf{b}$ and $\mathbf{V}_1$, define a function of $\mathbf{V}_0$ by:

$$F_{\mathbf{b},\mathbf{V}_1}(\mathbf{V}_0) = T(c(Y)) \cdot g\left(\sum_{i \in S(Y)} y^{(i)} \oplus P(S(Y))\mathbf{b}\right)$$

where $\mathbf{V} = [\mathbf{V}_0\, \mathbf{V}_1]$ and $Y = \mathbf{V} \oplus \mathbf{b}\mathbf{1}^\top$, and note that the adversary's advantage is given by:

$$\delta = \left| \mathop{\mathbb{E}}_{\substack{\mathbf{b},\mathbf{V}_1 \leftarrow \mathcal{U}, \\ \mathbf{V}_0 \leftarrow \mathcal{R}_{N \times Q_0}}} [F_{\mathbf{b},\mathbf{V}_1}(\mathbf{V}_0)] \right|$$

By Claim 5.20, for fixed $\mathbf{b}, \mathbf{V}_1$, we have

$$\left| \mathop{\mathbb{E}}_{\mathbf{V}_0 \leftarrow \mathcal{R}_{N \times Q_0}} [F_{\mathbf{b},\mathbf{V}_1}(\mathbf{V}_0)] - \mathop{\mathbb{E}}_{\mathbf{V}_0 \leftarrow \mathbb{F}_2^{N \times Q_0}} [F_{\mathbf{b},\mathbf{V}_1}(\mathbf{V}_0)] \right| \leq 2^{Q_0+1-N}. \tag{6}$$

Defining the ideal advantage

$$\delta_{\text{ideal}} = \left| \mathop{\mathbb{E}}_{\substack{\mathbf{b},\mathbf{V}_1 \leftarrow \mathcal{U}, \\ \mathbf{V}_0 \leftarrow \mathbb{F}_2^{N \times Q_0}}} [F_{\mathbf{b},\mathbf{V}_1}(\mathbf{V}_0)] \right|$$

and averaging (6) over $\mathbf{b}, \mathbf{V}_1$, we obtain

$$\delta \leq \delta_{\text{ideal}} + 2^{Q_0+1-N}. \tag{7}$$

The ideal distribution, where $\mathbf{V} \leftarrow \mathbb{F}_2^{N \times (Q-1)}$ and $Y = \mathbf{V} \oplus \mathbf{b}\mathbf{1}^\top$, can equivalently be rewritten as sampling $Y \leftarrow \mathbb{F}_2^{N \times (Q-1)}$, independent of $\mathbf{b}$. We use this form in what follows.

Claim 5.21 *In the ideal distribution,*

$$\delta_{\text{ideal}} \leq \tau + \Delta,$$

where

$$\Delta = \left| \mathop{\mathbb{E}}_{Y} \left[T(c(Y)) \mathbf{1}_{P(S(Y))=0} \sum_{\alpha \neq 0} \widehat{g}(\alpha) (-1)^{\left\langle \alpha, \sum_{i \in S(Y)} y^{(i)} \right\rangle} \right] \right|.$$

Proof. In the ideal distribution, the columns $y^{(1)}, \ldots, y^{(Q-1)}$ are independent and uniformly distributed in $\mathbb{F}_2^N$, and they are independent of $\mathbf{b}$. Since $c(Y)$ depends only on Y, the test value $T(c(Y))$ and subset $S(Y)$ are also fixed once Y is fixed.

For fixed Y, define

$$Z(Y) = \mathop{\mathbb{E}}_{\mathbf{b}} \left[g\left(\sum_{i \in S(Y)} y^{(i)} \oplus P(S(Y))\mathbf{b} \right) \right].$$

Then

$$\delta_{\text{ideal}} = \left| \mathop{\mathbb{E}}_{Y} [T(c(Y)) Z(Y)] \right|.$$

Expanding g in the Fourier basis,

$$g(z) = \sum_{\alpha \in \mathbb{F}_2^N} \widehat{g}(\alpha) (-1)^{\langle \alpha, z \rangle},$$

we get

$$Z(Y) = \sum_{\alpha \in \mathbb{F}_2^N} \widehat{g}(\alpha) (-1)^{\left\langle \alpha, \sum_{i \in S(Y)} y^{(i)} \right\rangle} \mathop{\mathbb{E}}_{\mathbf{b}} \left[(-1)^{P(S(Y)) \langle \alpha, \mathbf{b} \rangle} \right].$$

If $P(S(Y)) = 1$, equivalently $|S(Y)|$ is even, then the expectation over $\mathbf{b}$ vanishes for every $\alpha \neq 0$ and equals 1 for $\alpha = 0$. If $P(S(Y)) = 0$, equivalently $|S(Y)|$ is odd, then the expectation over $\mathbf{b}$ equals 1 for every α. Therefore

$$Z(Y) = \widehat{g}(0) + \mathbf{1}_{P(S(Y))=0} \sum_{\alpha \neq 0} \widehat{g}(\alpha) (-1)^{\left\langle \alpha, \sum_{i \in S(Y)} y^{(i)} \right\rangle}.$$

Substituting this expression for $Z(Y)$ into δ_{ideal} and applying the triangle inequality gives

$$\delta_{\text{ideal}} \le \left| \mathop{\mathbb{E}}_{Y}[T(c(Y))]\widehat{g}(0) \right| + \Delta \le \tau + \Delta,$$

where we used $|T(c(Y))| \le 1$ and $|\widehat{g}(0)| \le \tau$. □

Claim 5.22 *Assume $\tau \le 1/2$. Then*

$$\Delta \le \frac{\tau^2}{(1-\tau)^3}.$$

Proof. We condition on a transcript $c = (c_1, \dots, c_{Q-1}) \in \{-1,1\}^{Q-1}$. Conditioned on this transcript, the subset $S = S(c)$ and the value $T(c)$ are fixed. Moreover, since the event $g(y^{(i)}) = c_i$ depends only on the ith column, and since the columns are independent in the ideal distribution, the random variables $y^{(1)}, \dots, y^{(Q-1)}$ remain mutually independent after conditioning on c.

We first compute the conditional expectation of a Fourier character. For $\alpha \in \mathbb{F}_2^N$ and $c_i \in \{-1,1\}$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{y|g(y)=c_i}\left[(-1)^{\langle \alpha, y\rangle}\right] &= \frac{\mathbb{E}_y\left[(-1)^{\langle \alpha, y\rangle}\mathbf{1}[g(y)=c_i]\right]}{\Pr[g(y)=c_i]} \\
&= \frac{\mathbb{E}_y\left[(-1)^{\langle \alpha, y\rangle}\frac{1+c_i g(y)}{2}\right]}{\mathbb{E}_y\left[\frac{1+c_i g(y)}{2}\right]} \\
&= \frac{\delta_{\alpha,0} + c_i\widehat{g}(\alpha)}{1 + c_i\widehat{g}(0)},
\end{aligned}$$

where $\delta_{\alpha,0}$ denotes the Kronecker delta: $\delta_{\alpha,0} = \begin{cases} 1 & \text{if } \alpha = 0, \\ 0 & \text{otherwise.} \end{cases}$ Thus, for every $\alpha \ne 0$,

$$\left| \mathop{\mathbb{E}}_{y|g(y)=c_i}\left[(-1)^{\langle \alpha, y\rangle}\right] \right| \le \frac{|\widehat{g}(\alpha)|}{1-\tau}.$$

Now fix a transcript c. If $P(S(c)) = 1$, then the indicator $\mathbf{1}_{P(S(c))=0}$ makes the residual contribution vanish. So it suffices to consider the case $P(S(c)) = 0$. In this case $|S(c)|$ is odd. Since the game requires $|S(c)| \ge 2$, we have $|S(c)| \ge 3$.

Using conditional independence, for every fixed transcript c with $P(S(c)) = 0$,

$$\begin{aligned}
&\left| \mathop{\mathbb{E}}_{Y|c}\left[\sum_{\alpha\ne 0}\widehat{g}(\alpha)(-1)^{\langle \alpha, \sum_{i\in S(c)} y^{(i)}\rangle}\right] \right| \\
&\quad\le \sum_{\alpha\ne 0}|\widehat{g}(\alpha)| \prod_{i\in S(c)} \left| \mathop{\mathbb{E}}_{y^{(i)}|g(y^{(i)})=c_i}\left[(-1)^{\langle \alpha, y^{(i)}\rangle}\right] \right| \\
&\quad\le \sum_{\alpha\ne 0}|\widehat{g}(\alpha)| \left(\frac{|\widehat{g}(\alpha)|}{1-\tau}\right)^{|S(c)|} \\
&\quad= \sum_{\alpha\ne 0}\frac{|\widehat{g}(\alpha)|^{|S(c)|+1}}{(1-\tau)^{|S(c)|}}.
\end{aligned}$$

Let $k = |S(c)| \ge 3$. Since $\max_\alpha |\widehat{g}(\alpha)| \le \tau$, we have

$$|\widehat{g}(\alpha)|^{k+1} \le \tau^{k-1}\widehat{g}(\alpha)^2.$$

Therefore, by Parseval's identity,

$$\begin{aligned}
\sum_{\alpha\ne 0}\frac{|\widehat{g}(\alpha)|^{k+1}}{(1-\tau)^k} &\le \frac{\tau^{k-1}}{(1-\tau)^k}\sum_{\alpha\ne 0}\widehat{g}(\alpha)^2 \\
&\le \frac{\tau^{k-1}}{(1-\tau)^k}.
\end{aligned}$$

Since $\tau \le 1/2$,

$$\frac{\tau^{k-1}}{(1-\tau)^k} = \frac{\tau^2}{(1-\tau)^3}\left(\frac{\tau}{1-\tau}\right)^{k-3} \le \frac{\tau^2}{(1-\tau)^3}.$$

Thus, for every transcript c, the conditional residual contribution is at most $\tau^2/(1-\tau)^3$. Since $|T(c)| \le 1$, averaging over c gives

$$\Delta \le \frac{\tau^2}{(1-\tau)^3}.$$

□

We now combine the claims. By (7),

$$\delta \le \delta_{\text{ideal}} + 2^{Q_0 - N}.$$

Since $\tau = 2^{-\Omega(\lambda \log \lambda)}$, we have $\tau \le 1/2$ for all sufficiently large λ. Hence, we can apply Claim 5.21 and Claim 5.22, giving

$$\delta_{\text{ideal}} \le \tau + \frac{\tau^2}{(1-\tau)^3}.$$

Therefore

$$\delta \le \tau + \frac{\tau^2}{(1-\tau)^3} + 2^{Q_0 - N}.$$

Finally, since $\tau = 2^{-\Omega(\lambda \log \lambda)}$ and $Q_0 < N - \omega(\log \lambda)$, all three terms are negligible in λ. This proves the theorem. □

5.4 Non-sparsity over $\mathbb{F}_2$

We show that, with overwhelming probability over the choice of the key, the multilinear expansion of our candidate contains subexponentially many nonzero monomials.

Throughout this section, let

$$U := [\ell] \times [m] \times [n], \qquad M := mn, \qquad N := |U| = \ell M.$$

Our candidate can be written over $\mathbb{F}_2$ as

$$F_{\mathbf{A},\mathbf{b}}(X) = \sum_{w=1}^{\ell} \left(1 + \prod_{j=1}^{m} \left(1 + \prod_{k=1}^{n} (\mathbf{A}X + \mathbf{b})_{(w,j,k)}\right)\right).$$

The degree of $F_{\mathbf{A},\mathbf{b}}$ is at most M.

Theorem 5.23 *The multilinear expansion of $F_{\mathbf{A},\mathbf{b}}$ contains $\Omega\left(\binom{N}{M}\right)$ nonzero monomials of degree exactly M with probability $1 - 2^{-\Omega(\lambda)}$. In particular, under our parameter setting, the number of nonzero monomials is at least $2^{\Omega(\lambda \log \lambda)}$ with probability $1 - 2^{-\Omega(\lambda)}$.*

Proof. We first analyze the simpler case in which $\mathbf{A}$ is sampled uniformly from all matrices in $\mathbb{F}_2^{N \times N}$. We then transfer the result to the true key distribution by conditioning on the event that $\mathbf{A}$ is invertible.

For every set $T \subseteq [N]$ of size M, let $C_T(\mathbf{A})$ denote the coefficient of the monomial $X^T := \prod_{v \in T} X_v$ in the degree-M homogeneous part of $F_{\mathbf{A},\mathbf{b}}$. Define

$$S(\mathbf{A}) := \sum_{\substack{T \subseteq [N] \\ |T| = M}} C_T(\mathbf{A}),$$

where the coefficients are viewed as elements of $\{0,1\}$. Thus $S(\mathbf{A})$ is the number of nonzero monomials of degree exactly M.

Claim 5.24 *For every $T \subseteq [N]$ with $|T| = M$,*

$$C_T(\mathbf{A}) = \sum_{w=1}^{\ell} \det(\mathbf{A}_{w,T}) \pmod 2,$$

where $\mathbf{A}_{w,T}$ denotes the $M \times M$ submatrix of $\mathbf{A}$ formed by the M rows in block w and the columns indexed by T.

Proof. The degree-M homogeneous part of $F_{\mathbf{A},\mathbf{b}}$ is independent of $\mathbf{b}$. Indeed, choosing any shift term $\mathbf{b}_{(w,j,k)}$ from one of the affine forms $(\mathbf{A}X+\mathbf{b})_{(w,j,k)}$ lowers the degree.

Fix a block $w \in [\ell]$. The unique degree-M contribution from the wth block comes from choosing the product over all m tribes and all n variables inside each tribe:

$$\prod_{j=1}^{m}\prod_{k=1}^{n}(\mathbf{A}X)_{(w,j,k)}.$$

Equivalently, after indexing the $M = mn$ rows inside block w by $r \in [M]$, this product has the form

$$\prod_{r=1}^{M}\left(\sum_{v=1}^{N}\mathbf{A}_{w,r,v}X_v\right).$$

The coefficient of X^T in this product is obtained by choosing, from each of the M linear forms, one variable, and requiring that the set of chosen variables is exactly T. Hence the coefficient contributed by block w is

$$\sum_{\sigma:[M]\to T \text{ bijection}}\prod_{r=1}^{M}\mathbf{A}_{w,r,\sigma(r)}.$$

This is the permanent of the $M \times M$ submatrix $\mathbf{A}_{w,T}$. Over $\mathbb{F}_2$, the permanent equals the determinant, because the signs in the determinant are all equal to 1 in characteristic two. Summing over blocks gives the claimed formula. □

Claim 5.25 *Suppose* $\mathbf{A}$ *is sampled uniformly from* $\mathbb{F}_2^{N\times N}$. *Then*

$$\mathbb{E}[S(\mathbf{A})] = \binom{N}{M}\cdot\frac{1-\delta_M^\ell}{2},$$

where

$$\delta_M := 1-2p_M \qquad \textit{and} \qquad p_M := \prod_{i=1}^{M}(1-2^{-i}).$$

In particular, for all sufficiently large λ, $\mathbb{E}[S] \geq \frac{1}{4}\binom{N}{M}$.

Proof. Fix $T \subseteq [N]$ with $|T| = M$. By Claim 5.24, $C_T(\mathbf{A})$ is the parity of the bits $\det(\mathbf{A}_{1,T}), \ldots, \det(\mathbf{A}_{\ell,T})$. Under the uniform matrix distribution, the matrices $\mathbf{A}_{1,T}, \ldots, \mathbf{A}_{\ell,T}$ are independent uniform $M \times M$ matrices. Therefore, by Proposition 2.1,

$$\Pr[\det(\mathbf{A}_{w,T}) = 1] = \prod_{i=1}^{M}(1-2^{-i}) = p_M \in [0.288, 0.289].$$

Let $Y_T := (-1)^{C_T(\mathbf{A})}$. Since C_T is the parity of ℓ independent determinant bits,

$$\mathbb{E}[Y_T] = (1-2p_M)^\ell = \delta_M^\ell.$$

Because $C_T = \frac{1-Y_T}{2}$, we obtain $\mathbb{E}[C_T] = \frac{1-\delta_M^\ell}{2}$. Summing over all $T \subseteq [N]$ of size M gives

$$\mathbb{E}[S] = \binom{N}{M}\cdot\frac{1-\delta_M^\ell}{2}.$$

Finally, from the bounds on the quantity p_M, it holds that $\delta_M = 1-2c_{M,M}$ lies strictly between zero and one. Since $\ell = \Theta(\lambda)$, we have $\delta_M^\ell = 2^{-\Omega(\lambda)}$. In particular, for all sufficiently large λ, $\delta_M^\ell \leq 1/2$, and hence

$$\mathbb{E}[S] \geq \frac{1}{4}\binom{N}{M}.$$

□

Claim 5.26 *Suppose* $\mathbf{A}$ *is sampled uniformly from* $\mathbb{F}_2^{N\times N}$. *For any distinct* $T, T' \subseteq [N]$ *with* $|T| = |T'| = M$,

$$|\operatorname{Cov}(C_T, C_{T'})| \leq \frac{1}{2}\delta_M^\ell.$$

Proof. It is convenient to first bound the covariance of the sign variables

$$Y_T := (-1)^{C_T(\mathbf{A})} \qquad \text{and} \qquad Y_{T'} := (-1)^{C_{T'}(\mathbf{A})}.$$

Fix one block $w \in [\ell]$ and define $\rho_{T,T'} := \mathbb{E}\left[(-1)^{\det(\mathbf{A}_{w,T})+\det(\mathbf{A}_{w,T'})}\right]$. We claim that $|\rho_{T,T'}| \le \delta_M$.

Since $T \neq T'$, choose a column $c \in T \setminus T'$. Condition on all columns in $(T \cup T') \setminus \{c\}$ of the wth row-block. This fixes $\det(\mathbf{A}_{w,T'})$. It also fixes the $M \times (M-1)$ matrix $\mathbf{B} := \mathbf{A}_{w,T\setminus\{c\}}$. As a function of the remaining unconditioned column $a_c \in \mathbb{F}_2^M$, the determinant $\det(\mathbf{A}_{w,T})$ is a linear form: $\det(\mathbf{A}_{w,T}) = v^\top a_c$, where $v \in \mathbb{F}_2^M$ is the vector of cofactors determined by $\mathbf{B}$.

If $\mathsf{rank}(\mathbf{B}) = M-1$, then $v \neq 0$, and therefore $v^\top a_c$ is a uniform bit over the choice of a_c. In this case, the conditional expectation of $(-1)^{\det(\mathbf{A}_{w,T})+\det(\mathbf{A}_{w,T'})}$ is zero. If $\mathsf{rank}(\mathbf{B}) < M-1$, then the conditional expectation has absolute value at most 1. Hence $|\rho_{T,T'}| \le \Pr[\mathsf{rank}(\mathbf{B}) < M-1]$.

From Proposition 2.1, a random $M \times (M-1)$ matrix over $\mathbb{F}_2$ has full rank with probability $p_{M,M-1} = \prod_{j=0}^{M-2}(1-2^{j-M}) = \prod_{i=2}^{M}(1-2^{-i}) = 2p_M$. Therefore $\Pr[\mathsf{rank}(\mathbf{B}) < M-1] = 1 - 2p_M = \delta_M$. This proves $|\rho_{T,T'}| \le \delta_M$.

Now use independence across the ℓ row-blocks. By Claim 5.24,

$$Y_T Y_{T'} = \prod_{w=1}^{\ell} (-1)^{\det(\mathbf{A}_{w,T})+\det(\mathbf{A}_{w,T'})}.$$

Thus $|\mathbb{E}[Y_T Y_{T'}]| \le \delta_M^\ell$ and $\mathbb{E}[Y_T]\,\mathbb{E}[Y_{T'}] = \delta_M^{2\ell}$. Therefore

$$|\operatorname{Cov}(Y_T, Y_{T'})| \le \delta_M^\ell + \delta_M^{2\ell} \le 2\delta_M^\ell.$$

Since $C_T = (1-Y_T)/2$ and $C_{T'} = (1-Y_{T'})/2$, we have

$$|\operatorname{Cov}(C_T, C_{T'})| = \frac{1}{4}|\operatorname{Cov}(Y_T, Y_{T'})| \le \frac{1}{2}\delta_M^\ell,$$

as claimed. □

Claim 5.27 *Suppose* $\mathbf{A}$ *is sampled uniformly from* $\mathbb{F}_2^{N\times N}$. *Then*

$$S(\mathbf{A}) \ge \frac{1}{8}\binom{N}{M}$$

with probability $1 - 2^{-\Omega(\lambda)}$.

Proof. All sums below are over subsets $T \subseteq [N]$ of size M. By Claim 5.26,

$$\begin{aligned}\operatorname{Var}(S) &= \sum_T \operatorname{Var}(C_T) + \sum_{T\neq T'} \operatorname{Cov}(C_T, C_{T'}) \\ &\le \binom{N}{M} + \frac{1}{2}\binom{N}{M}^2 \delta_M^\ell,\end{aligned}$$

where we use the trivial bound $\operatorname{Var}(C_T) \le 1$ for each T.

By Chebyshev's inequality,

$$\Pr\left[S < \frac{1}{2}\,\mathbb{E}[S]\right] \le \frac{4\operatorname{Var}(S)}{\mathbb{E}[S]^2}.$$

Using Claim 5.25, for sufficiently large λ we have $\mathbb{E}[S] \ge \frac{1}{4}\binom{N}{M}$. Therefore

$$\Pr\left[S < \frac{1}{8}\binom{N}{M}\right] \le O\left(\frac{1}{\binom{N}{M}} + \delta_M^\ell\right).$$

Since $\delta_M < 1$ and $\ell = \Theta(\lambda)$, $\delta_M^\ell = 2^{-\Omega(\lambda)}$. Moreover, under our parameter setting, $\binom{N}{M} = 2^{\Omega(\lambda\log\lambda)}$, so $\frac{1}{\binom{N}{M}} = 2^{-\Omega(\lambda\log\lambda)}$. Hence the failure probability is $2^{-\Omega(\lambda)}$, as claimed. □

Claim 5.28 *Under the true key distribution, where* $\mathbf{A}$ *is sampled uniformly from* $\mathcal{GL}_N(\mathbb{F}_2)$, *we have*

$$S(\mathbf{A}) \ge \frac{1}{8}\binom{N}{M}$$

with probability $1 - 2^{-\Omega(\lambda)}$.

Proof. Let $\mathcal{E}$ be the bad event

$$\mathcal{E} := \left\{ S < \frac{1}{8}\binom{N}{M} \right\}.$$

By Claim 5.27, under the uniform matrix distribution, $\Pr[\mathcal{E}] \le 2^{-\Omega(\lambda)}$. Also, for a uniformly random $N \times N$ matrix over $\mathbb{F}_2$, by Proposition 2.1,

$$\Pr[\mathbf{A} \in \mathcal{GL}_N(\mathbb{F}_2)] > 0.288$$

Therefore

$$\begin{aligned}\Pr[\mathcal{E} \mid \mathbf{A} \in \mathcal{GL}_N(\mathbb{F}_2)] &\le \frac{\Pr[\mathcal{E}]}{\Pr[\mathbf{A} \in \mathcal{GL}_N(\mathbb{F}_2)]} \\ &\le \rho^{-1} \cdot 2^{-\Omega(\lambda)} = 2^{-\Omega(\lambda)}.\end{aligned}$$

This proves the claim. □

By Claim 5.28, under the true key distribution, $S(\mathbf{A}) = \Omega\Big(\binom{N}{M}\Big)$ with probability $1 - 2^{-\Omega(\lambda)}$. Thus, the total number of nonzero monomials satisfies the same lower bound. Finally, under our parameter setting, $\binom{N}{M} = 2^{\Omega(\lambda \log \lambda)}$. This proves the theorem. □

5.5 High Rational Degree

We show that our candidate has high rational degree as defined in Definition 2.9, thereby avoiding the algebraic attack that breaks the weak PRF candidate of [ABG+14]. For $N = \ell mn \in \mathbb{N}$, let

$$\mathcal{R}_N \;:=\; \mathbb{F}_2[X_1, \ldots, X_N] \big/ \big(X_1^2 + X_1, \ldots, X_N^2 + X_N\big)$$

be the ring of multilinear polynomials over $\mathbb{F}_2$. Every $F \in \mathcal{R}_N$ has a unique multilinear representative

$$F \;=\; \sum_{S \subseteq [N]} c_S \prod_{i \in S} X_i \qquad (c_S \in \mathbb{F}_2),$$

and multiplication satisfies

$$\Big(\prod_{i \in S} X_i\Big) \cdot \Big(\prod_{i \in T} X_i\Big) \;=\; \prod_{i \in S \cup T} X_i \quad \text{in } \mathcal{R}_N$$

for all $S, T \subseteq [N]$. In particular, each $F \in \mathcal{R}_N$ is idempotent: $F^2 = F$.[9]

Let I denote the ideal $(X_{ijk}^2 + X_{ijk})_{(i,j,k)}$ in $\mathbb{F}_2[X_1, \ldots, X_N]$. We use the following elementary properties about the rational degree.

Lemma 5.29 *Let $F \in \mathcal{R}_N \setminus \{0\}$.*

1. (Algebraic-degree upper bound) *The (primal) rational degree of F is at most the algebraic degree of its multilinear representative in $\mathbb{F}_2[X_1, \ldots, X_N]$.*
2. (Affine invariance) *If $\mathbf{A} \in \mathbb{F}_2^{N \times N}$ is invertible and $\mathbf{b} \in \mathbb{F}_2^N$, then the rational degree of F equals that of $F(\mathbf{AX} + \mathbf{b})$.*

Proof. (1) Since $F^2 + F \equiv 0 \pmod I$, taking $G := F + 1$ shows that $GF \equiv 0 \pmod I$ and $\deg(G) = \deg(F)$, proving the claimed upper bound.

(2) Suppose $G(\mathbf{X})\, F(\mathbf{AX} + \mathbf{b}) \equiv 0$ in $\mathcal{R}_N$. Evaluating on $\mathbb{F}_2^N$ and substituting $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ shows $G'(\mathbf{Y})\, F(\mathbf{Y}) = 0$ for all $\mathbf{Y} \in \mathbb{F}_2^N$, where $G'(\mathbf{Y}) := G(\mathbf{A}^{-1}\mathbf{Y} + \mathbf{A}^{-1}\mathbf{b})$ has the same algebraic degree as G. Thus any annihilator of $F(\mathbf{AX} + \mathbf{b})$ of degree d induces one of F of the same degree (and vice versa), for both primal and dual definitions. Hence rational degree is preserved. □

Write the input variables as $\{X_{ijk}\}_{i \in [\ell],\, j \in [m],\, k \in [n]}$ so that block i consists of m tribes, each of width n. Let

$$H_i(X_i) := \prod_{j=1}^{m} \Big(1 + \prod_{k=1}^{n} X_{ijk}\Big) \qquad \text{and} \qquad \tilde{F}(X) \;:=\; \sum_{i=1}^{\ell} H_i(X_i) \in \mathcal{R}_N.$$

[9] This follows from the fact that two multilinear polynomials agree as elements of $\mathcal{R}_N$ iff they agree pointwise on $\mathbb{F}_2^N$, and $F(x)^2 = F(x)$ for all $x \in \mathbb{F}_2^N$, as $y^2 = y$ for all $y \in \mathbb{F}_2$.

Up to adding a constant (i.e., replacing F by $F+1$), this agrees with the standard multilinearization of the TRIBES construction; rational degree is invariant under $F \mapsto F+1$ (it swaps the primal and dual notions). By Lemma 5.29, it suffices to analyze the unshifted, identity case $F_{\mathbf{A},\mathbf{b}}(\mathbf{X}) = F(\mathbf{X})$ with $\mathbf{A} = I$ and $\mathbf{b} = \mathbf{0}$. We work with $\tilde{F}$ defined above since $\mathsf{rdeg}(F_{\mathbf{A},\mathbf{b}}) = \mathsf{rdeg}(\tilde{F})$. Each H_i has algebraic degree mn, hence $\deg(\tilde{F}) = mn$. By Lemma 5.29, we have the following corollary.

Corollary 5.30 (Trivial upper bound) *For the function $\tilde{F}$ with parameters ℓ, m, n,*

$$\mathsf{rdeg}\Big(\tilde{F}\Big) \;\leq\; mn.$$

We now prove the main lower bound on the rational degree of our candidate.

Proposition 5.31 (Lower bound) *For the function $\tilde{F}$ with parameters ℓ, m, n,*

$$\mathsf{rdeg}\Big(\tilde{F}\Big) \;\geq\; \min\{\ell,\, mn\}.$$

Consequently, for any choice of invertible $\mathbf{A} \in \mathbb{F}_2^{N\times N}$ and $\mathbf{b} \in \mathbb{F}_2^N$, we also have $\mathsf{rdeg}(F_{\mathbf{A},\mathbf{b}}) \geq \min\{\ell, mn\}$.

Proof. We show that no nonzero multilinear G of degree $< t := \min\{\ell, mn\}$ can satisfy either $G\,\tilde{F} \equiv 0 \pmod I$ (primal) or $G(\tilde{F}+1) \equiv 0 \pmod I$ (dual).

Write $G = \sum_{S\in\mathcal{G}} \mathbf{X}^S$, where each monomial $\mathbf{X}^S := \prod_{(i,j,k)\in S} X_{ijk}$ is square-free and $\mathcal{G} \subseteq 2^{[\ell]\times[m]\times[n]}$. Let $S^\star \in \mathcal{G}$ be of maximum size, so $|S^\star| = \deg(G) < t$. Because $|S^\star| < \ell$, there exists an index $i^\star \in [\ell]$ such that no variable from block $i^\star$ appears in $S^\star$, i.e., $(i^\star, j, k) \notin S^\star$ for all $j \in [m], k \in [n]$. Consider the monomial

$$M^\star \;:=\; \Big(\prod_{(i,j,k)\in S^\star} X_{ijk}\Big)\cdot\Big(\prod_{j=1}^{m}\prod_{k=1}^{n} X_{i^\star jk}\Big).$$

In the expansion of

$$G\cdot\tilde{F} = \sum_{i=1}^{\ell} G\cdot H_i = \sum_{i=1}^{\ell}\sum_{T\subseteq[m]}\sum_{S\in\mathcal{G}}\Big(\prod_{(i',j',k')\in S} X_{i'j'k'}\Big)\Big(\prod_{j\in T}\prod_{k=1}^{n} X_{ijk}\Big)\;,$$

the term with $i = i^\star$ and $T = [m]$ produces $M^\star$. Because $S^\star$ is disjoint from block $i^\star$, the variables in $M^\star$ are pairwise distinct, hence

$$\deg(M^\star) \;=\; |S^\star| + mn.$$

We claim that $M^\star$ appears *exactly once* (and hence, with coefficient 1) in $G\cdot\tilde{F}$ modulo I. Indeed, suppose another triple (i, T, S) contributes the same monomial $M^\star$. Then we must have

$$S \;\cup\; \big(\{i\}\times T\times[n]\big) \;=\; S^\star \;\cup\; \big(\{i^\star\}\times[m]\times[n]\big)\;.$$

If $i \neq i^\star$, the right-hand side would force $\{i^\star\}\times[m]\times[n] \subseteq S$, implying $|S| \geq mn$, which contradicts $|S| < t \leq mn$. Thus $i = i^\star$, and the disjointness shows $T = [m]$ and $S = S^\star$. Hence $M^\star$ survives with coefficient 1 in $G\tilde{F}$, so $G\tilde{F} \not\equiv 0 \pmod I$.

The above argument also establishes the dual rational degree. Indeed, we need to argue that $G\tilde{F} \not\equiv G \pmod I$. We have just shown that even after reducing modulo I, it must be that $\deg(G\tilde{F}) = \deg(G) + \deg(\tilde{F})$, which is strictly greater than $\deg(G)$ since $\deg(\tilde{F}) = mn \geq 1$. Hence, we cannot have $G\tilde{F} \equiv G \pmod I$. We conclude that no nonzero G of degree $< t$ annihilates $\tilde{F}$ or $\tilde{F}+1$, so $\mathsf{rdeg}(\tilde{F}) \geq t$. □

Corollary 5.32 *Let $N = \ell mn$. For the candidate $F_{\mathbf{A},\mathbf{b}}$ viewed as an element of $\mathcal{R}_N$,*

$$\min\{\ell, mn\} \;\leq\; \mathsf{rdeg}(F_{\mathbf{A},\mathbf{b}}) \;\leq\; mn.$$

This implies there is no rational degree attack on our candidate with less than $2^{\Omega(\lambda)}$ queries, which is impossible in our bounded query setting.

5.6 Inapproximability by Low-Degree $\mathbb{F}_2$ Polynomials

We show that the correlation between a low degree $\mathbb{F}_2$ polynomial and our candidate is small. We measure correlation between a Boolean function and low-degree $\mathbb{F}_2$ polynomials as follows. For a function $f : \{0,1\}^M \to \{-1,1\}$ and an integer $d \geq 0$, define

$$\mathsf{Corr}_d(f) \;:=\; \max_{\substack{p:\{0,1\}^M \to \{0,1\} \\ \deg(p) \leq d}} \left| \mathbb{E}_{x \leftarrow \{0,1\}^M} \left[f(x) \cdot (-1)^{p(x)} \right] \right| .$$

We show that our candidate satisfies the same inapproximability guarantee as in [ABG+14]. The proof is in **??**.

Theorem 5.33 *Let $\alpha > 0$ be a constant and $K = K(\alpha)$. Let $M = mn$ be the input size for $\mathsf{TRIBES}_{m,n}$, and assume $\alpha n 2^n \leq M \leq (1-\alpha) n 2^n$. Then, for every $d \leq \log M - \log\big(K \log M\big)$ there exists $c = c(\alpha) > 0$ such that*

$$\mathsf{Corr}_d\big(F_{\mathbf{A},\mathbf{b}}\big) \;\leq\; 1 - c \cdot 2^{-d}.$$

In particular, taking parameters $\ell = \Theta(\lambda), m = \Theta(\lambda/\log(\lambda))$ and $n = \Theta(\log \lambda)$ satisfies the regime above, and the same bound holds uniformly for all invertible $\mathbf{A}$ and all $\mathbf{b}$.

5.7 Lack of Correlation with Fixed Function Families

Similarly to Section 5.6, we define the *correlation* of two functions (not necessarily polynomials) $g, h : \mathbb{F}_2^N \to \{-1,1\}$ as

$$\mathsf{Corr}(g,h) := \mathop{\mathbb{E}}_{x \in \mathbb{F}_2^N} [g(x)h(x)] \;,$$

where x is sampled uniformly at random from $\mathbb{F}_2^N$.

Considering any family $\mathcal{H}$ of functions $h : \mathbb{F}_2^N \to \{-1,1\}$, we show that, with high probability over the choice of secret key $(\mathbf{A}, \mathbf{b})$, $f_{\mathbf{A},\mathbf{b}}$ will not noticeably correlate with any function $h \in \mathcal{H}$, assuming $\mathcal{H}$ is not too large. We prove the following theorem, which is analogous to Corollary 17 from [ABG+14], showing that our candidate has low correlation with fixed function families. The proof is in **??**.

Theorem 5.34 *Let $\eta, \gamma > 0$ be such that $1-\gamma-2\eta > 0$. Let $\mathcal{H}$ be any set of functions $h : \mathbb{F}_2^N \to \{-1,1\}$ of size $|\mathcal{H}| \leq 2^{\gamma N}$. Let $\mu_g := \mathbb{E}_x[f_{\mathbf{I},\mathbf{0}}(x)]$ and $\mu_h := \mathbb{E}_x[h(x)]$. Then with probability at least $1 - 2^{-(1-\gamma-2\eta)N}$ over the choice of key $(\mathbf{A}, \mathbf{b})$, for all $h \in \mathcal{H}$, we have*

$$-2^{-\eta N} \leq \mathsf{Corr}(f_{\mathbf{A},\mathbf{b}}, h) - \mu_g \mu_h \leq 2^{-\eta N} \;.$$

In particular, for all $h \in \mathcal{H}$,

$$-|\mu_g| - 2^{-\eta N} \leq \mathsf{Corr}(f_{\mathbf{A},\mathbf{b}}, h) \leq |\mu_g| + 2^{-\eta N}$$

Remark 5.35 *Note that for our choice of parameters, we have $|\mu_g| \leq 2^{-\ell} = O\left(2^{-\lambda}\right)$, so the total bound on the correlation is negligible in the security parameter λ, and thus also in the key-length $\kappa(\lambda) = O(\lambda^4)$.*

6 Universal PRF from Strong PRFs in $\mathsf{AC}^0[2]$

We construct an explicit universal PRF UG in $\mathsf{AC}^0[2]$ whose circuit generator takes only 1^n and a depth bound d as inputs, under the assumption that a strong PRF of depth d exists in $\mathsf{AC}^0[2]$. Before we begin, we say that a PRF $\{f_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}\}_{n \in \mathbb{N}}$ is constructible in time $t(n)$ if there is a deterministic Turing running in $t(n)$ steps that outputs an circuit computing f_n.

We prove Theorem 6.4 by first establishing the PRF domain shrinkage (Lemma 6.1) and domain extension (Lemma 6.2) lemmas below.

Lemma 6.1 (PRF Domain Shrinkage) *Suppose $f_n : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ is a PRF secure against polynomial-time and polynomial-query adversaries, and that f_n can be represented by an $\mathsf{AC}^0[2]$ circuit of depth $d = O(1)$ constructible in time and size $O(n^c)$ for some constants $b, c \geq 1$.*

Then there exists $\epsilon \in (0,1)$ and a PRF

$$f'_n := f_{n^\epsilon} : \{0,1\}^{n^{b\epsilon}} \times \{0,1\}^{n^\epsilon} \to \{0,1\},$$

which remains secure against polynomial-time and polynomial-query adversaries and can be represented by an $\mathsf{AC}^0[2]$ circuit constructible in time and size $O(n)$ and depth d.

Proof. Set $\epsilon = 1/c$. The function family f'_n is simply the original PRF family evaluated at the smaller security parameter n^ϵ.

Suppose for sake of contradiction that f'_n admits an adversary running in time or making queries bounded by $O(n^{c_2})$ for some constant c_2. Replacing n by $n^{1/\epsilon}$ yields an attack on the original family f_n running in time or making queries bounded by $O\big((n^{1/\epsilon})^{c_2}\big) = O(n^{c_2/\epsilon})$, which is still polynomial in n. This contradicts the assumed security of f_n.

Finally, since f_n is constructible in time and size $O(n^c)$, the rescaled family f'_n is constructible in time and size $O((n^\epsilon)^c) = O(n)$. The circuit depth remains unchanged. □

Lemma 6.2 (PRF Domain Extension) *Suppose $f'_n : \{0,1\}^{n^{b\epsilon}} \times \{0,1\}^{n^\epsilon} \to \{0,1\}$ is a PRF secure against polynomial-time and polynomial-query adversaries, and that f'_n can be represented by an $\mathsf{AC}^0[2]$ circuit of depth d constructible in time and size $O(n)$ for some $b \geq 1$ and $\epsilon \in (0,1)$. Then there exists a PRF*

$$g_n : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$$

which remains secure against polynomial-time and polynomial-query adversaries and can be represented as an $\mathsf{AC}^0[2]$ circuit constructible in time and size $O(n^2)$ and depth $2d+1$.

Proof. We define $g_n(k,x)$ as follows:

1. Let k_1 be the first $n^{b\epsilon}$ bits of k and k_2 be the second $n^{b\epsilon}$ bits of k.
2. Let $l = \frac{n^\epsilon}{2}$ and $m = \frac{n}{l}$. Partition the function input x into m blocks of size l: x_1, x_2, …, x_m.
3. For each $i \in [m]$, we define $y_i \in \{0,1\}^{n^\epsilon}$ such that its j^{th} bit is $f'_n(k_1, x_i || n^\epsilon \cdot (i-1) + j)$ for $j \in [n^\epsilon]$, where $||$ denotes the concatenation. Note that $n^\epsilon \cdot (i-1) + j$ can be represented by $\log(3n) < \frac{n^\epsilon}{2}$ bits for large enough n.
4. Let $z = \bigoplus_{i=1}^{m} y_i$.
5. Output $f'_n(k_2, z)$.

First, we verify that g_n is constructible in time and size $O(n^2)$. The function g_n contains one XOR gate and $m \cdot n^\epsilon + 1 = 2n + 1$ functions f'_n, which are each constructible in time and size $O(n)$, which means g_n is constructible in time and size $O(n^2)$. The depth is $2d+1$.

Let $\mathcal{A}$ be an adversary against the PRF security of g_n. Let $q^{(1)}, q^{(2)}, \ldots, q^{(t(n))}$ are its distinct queries for some polynomial $t(n)$. Let $z^{(1)}, z^{(2)}, \ldots, z^{(t(n))}$ be the outputs of step 4 for the queries, before they are plugged into f'_n. We wish to show that these $z^{(i)}$ are all distinct with high probability. If they are distinct, since f'_n is a PRF, then the distribution $\{f'_n(k_2, z^{(i)})\}_{i \in [t(n)], k_2 \in \{0,1\}^{n^{b\epsilon}}}$ is indistinguishable from $\mathcal{U}_{t(n)}$ (with some negligible advantage δ) which means g_n is a PRF by a hybrid argument.

We upper bound $\Pr\left[z^{(a)} = z^{(b)}\right]$ for distinct $a, b \in [t(n)]$. We know that $q^{(a)}$ and $q^{(b)}$ must differ in at least one block i^*, so the distribution $y^{(a)}_{i^*} || y^{(b)}_{i^*}$ (the outputs from step 3) is indistinguishable (with some negligible advantage δ') from $\mathcal{U}_{2n^\epsilon}$ since f'_n is a PRF. Thus,

$$\Pr\left[z^{(a)} = z^{(b)}\right] < \left(\frac{1}{2}\right)^{n^\epsilon} + \delta'$$

By the union bound, we have

$$\Pr\left[z^{(i)} \text{ are not distinct}\right] < t(n)^2 \left(\frac{1}{2}\right)^{n^\epsilon} + t(n)^2 \delta'.$$

Finally, we obtain the advantage of the PRF game for g_n is at most

$$\delta + t(n)^2 \left(\frac{1}{2}\right)^{n^\epsilon} + t(n)^2 \delta',$$

which is negligible if δ and δ' are negligible and $t(n)$ is polynomial. □

We now give the explicit construction of UG.

Construction 6.3 (Universal candidate UG) *Fix integers $b, d \geq 1$. We define*

$$UG_n^{(b,d)} : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$$

by the following two steps.

1. Construct the intermediate family UF. *Take any enumeration of all Turing Machines M_i. We enumerate all possible function families that take $\{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ in $\mathsf{AC}^0[2]$ as follows:*

   (a) *For each $i \in \mathbb{N}$, run the Turing Machine M_i for n^2 steps.*

   (b) *Interpret the resulting output string from M_i as an $\mathsf{AC}^0[2]$ circuit of depth $\leq d$ (as all circuits can be represented by binary strings) and call this circuit F_i. If the depth is greater than d, skip this iteration.*

   If a polynomial-time constructible family of PRFs exists in $\mathsf{AC}^0[2]$, combining Lemma 6.1 and Lemma 6.2, shows that there must be a PRF family computable in time $O(n^2)$. Its index in the enumeration is therefore constant, so it appears among the first $\sqrt{n}$ machines for all sufficiently large n. Note that $\sqrt{n}$ is chosen arbitrarily, we could also choose any $\mathsf{poly}(n)$. We define

$$UF(k_1, \ldots, k_{\sqrt{n}}, x) = \bigoplus_{i=1}^{\sqrt{n}} F_i(k_i, x).$$

2. Define UG by rescaling and domain extension. *First, by applying the domain shrinkage technique from Lemma 6.1, we obtain the PRF $UF' : \{0,1\}^{n^{3b\epsilon/2}} \times \{0,1\}^{n^\epsilon} \to \{0,1\}$ of depth $d+1$ for some $\epsilon \in (0,1)$. Then, using the domain extension technique from Lemma 6.2, we obtain the PRF $UG : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ of depth $2d+3$.*

Theorem 6.4 (Universal PRF) *Suppose $f_n : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$ is a PRF secure against polynomial-time and polynomial-query adversaries, and that f_n can be represented by an $\mathsf{AC}^0[2]$ circuit of depth d constructible in polynomial time and size for some $b \geq 1$. Then*

$$UG : \{0,1\}^{n^b} \times \{0,1\}^n \to \{0,1\}$$

is a PRF which remains secure against polynomial-time and polynomial-query adversaries and can be represented as an $\mathsf{AC}^0[2]$ circuit constructible in polynomial time and size and depth $2d+3$.

Proof. We show that UF in Construction 6.3 is a PRF secure against $\mathsf{poly}(n)$ time and query adversaries. Assume UF is not secure against $\mathsf{poly}(n)$ time or query adversaries for the sake of contradiction. Then there is a polynomial-time or query distinguisher $\mathcal{D}$ such that

$$\left| \Pr_{\{k_1,\ldots,k_{\sqrt{n}}\} \leftarrow \{0,1\}^{n \times \sqrt{n}}} \left[\mathcal{D}^{UF_k} = 1\right] - \Pr_{R \leftarrow RF_n} \left[\mathcal{D}^R = 1\right] \right| > \frac{1}{n^c}$$

where RF_n is the set of all functions that map $\{0,1\}^n \to \{0,1\}$ and for some constant c. Since we know there exists a PRF in the truncated list, let j be the first index $1 \leq j \leq \sqrt{n}$ such that F_j is a PRF. We construct a distinguisher $\mathcal{A}$ for F_j as follows:

- Sample $k_i \leftarrow \{0,1\}^n$ randomly for all $1 \leq i \neq j \leq \sqrt{n}$.
- Let $\mathcal{O}$ be the oracle representing F_j or function sampled randomly from RF_n.
- We construct the oracle $\mathcal{U}$ from $\mathcal{O}$ as follows:

$$\mathcal{U}(x) = \bigoplus_{i=1}^{j-1} F_i(k_i, x) \bigoplus \mathcal{O}(x) \bigoplus_{i=j+1}^{\sqrt{n}} F_i(k_i, x).$$

- Run the distinguisher $\mathcal{D}$ with $\mathcal{U}$ as the oracle and output the outcome.

If $\mathcal{O}$ represents F_j, then $\mathcal{U}$ has the same distribution as UF_k where $\{k_1, \ldots, k_{\sqrt{n}}\} \leftarrow \{0,1\}^{n \times \sqrt{n}}$. Otherwise, if $\mathcal{O}$ is a function sampled randomly from RF^n, then $\mathcal{U}$ also has the same distribution as a function sampled randomly from RF^n. Thus,

$$\left|\Pr_{k_j \leftarrow \{0,1\}^n}\left[\mathcal{A}^{F_j} = 1\right] - \Pr_{R \leftarrow RF_n}\left[\mathcal{A}^{R} = 1\right]\right| > \frac{1}{n^c}.$$

If $\mathcal{D}$ uses $\mathsf{poly}(n)$ time or queries, then $\mathcal{A}$ also uses $\mathsf{poly}(n)$ time or queries. This is a contradiction that F_j is a PRF secure against $\mathsf{poly}(n)$ time and query adversaries, so we have shown UF is a PRF secure against $\mathsf{poly}(n)$ time and query adversaries. The size of UF is at most $\sqrt{n} \cdot \mathsf{poly}(n)$ and the depth is at most $d+1$, where d is the maximum depth of the enumerated functions F_i. This ensures UF stays in the class $\mathsf{AC}^0[2]$.

Finally, UG is a PRF from Lemma 6.1 and Lemma 6.2. □

7 Limitations of Natural Property-Based Learners

In this section, we provide a more detailed discussion on the rationales for the hardness of learning in $\mathsf{AC}^0[p]$ (Conjecture 1.4). The CIKK learning algorithm [CIKK16] for $\mathsf{AC}^0[p]$ uses a natural property, which we define below.

7.1 Natural Properties

Let F_n be the collection of all Boolean functions on n variables. Λ and Γ denote complexity classes. A combinatorial property is a sequence of subsets of F_n for each n.

Definition 7.1 (Natural Property [Raz87]) *A combinatorial property $R_n \subseteq F_n$ is Γ-natural against Λ with density δ_n if it satisfies the following three conditions:*

1. ***Constructivity:*** *The predicate $f_n \stackrel{?}{\in} R_n$ is computable in Γ. We call D a* natural property distinguisher *if it computes the predicate $f_n \stackrel{?}{\in} R_n$.*
2. ***Largeness:*** *$|R_n| \geq \delta_n \cdot |F_n|$.*
3. ***Usefulness:*** *For any sequence of functions f_n, if $f_n \in \Lambda$, then $f_n \notin R_n$.*

Remark 7.2 *Note that this implies the natural property distinguisher D can distinguish any Boolean function $f_n \in \Lambda$ from a randomly chosen function in F_n with advantage δ_n.*

7.2 Limitations of Natural Properties Against $\mathsf{AC}^0[p]$

In order to invalidate the security of a PRF candidate, it suffices to distinguish it from random, and a natural property precisely promises such a distinguisher. For $\mathsf{AC}^0[2]$, Razborov [Raz87] and Smolensky [Smo87] argue that $\mathsf{AC}^0[2]$ can be approximated by polynomials of polylogarithmic-degree. Krause and Lucks [KL01] point out that so long as one can *tolerantly test* low-degree polynomials – namely, given oracle access to a function which is either promised to be close to a low-degree polynomial, or is far from any low-degree polynomial – then one has an effective distinguisher for any PRF candidate computable in $\mathsf{AC}^0[2]$. Indeed, by Razborov-Smolensky [Raz87, Smo87] if one has oracle access to the PRF candidate then it is close to a low-degree polynomial, while if the function is random then it is far from any low-degree polynomial (with very high probability).

It is reasonable to consider the sample complexity of such an attack on our PRF. We note that the depth D of our candidate is 3 (at least; it is even more if we view the key as an explicit input), and the size S is at least $N(\lambda) = \sqrt{\kappa(\lambda)}$, where we recall again that λ is the security parameter, N the input size, and κ the key-size. Hence, even if one is happy with a very large approximation error ε the degree of the Razborov-Smolensky approximating polynomial is at least $\log(S/\varepsilon)^D \geq \log^3(N) = \log^3(\sqrt{\kappa}) = \log^3(\kappa)/8$. We thus have a polylogarithmic degree d for the Razborov-Smolensky approximating polynomial.

Now, how many samples would one need to take to determine if the oracle is implementing a function which is close to a degree-d polynomial? As mentioned above, this is a *tolerant testing* problem, but we can consider the formally easier task where, in the yes case, the oracle is promised to be *exactly implementing* a degree-d polynomial. This problem has seen much attention [AKK+03, JPRZ09] and in

particular it is known that $\Omega(2^{d-1})$ queries are *necessary*, at least if one wishes to distinguish with constant advantage. This follows from the fact that the family of degree-d multivariate polynomials forms of $\Omega(2^{d-1})$-wise independent hash function family; thus, with less queries it is information-theoretically impossible to distinguish the case that the function is implementing a uniformly random degree-d polynomial from the case it is implementing a uniformly random function. This implies that if one hopes for a distinguisher with constant advantage, then $2^{\log^3(\kappa)/8}$ samples – that is, a quasi-polynomially large number – must be taken.

Lastly, we would like to point out that this also means that this recipe cannot yield a distinguisher with, say, $\mathsf{poly}(\kappa)$ query complexity. Indeed, by the reasoning above seeing only $\mathsf{poly}(\kappa)$ values one cannot learn anything in general about a degree $d = \Omega(\log^3 \kappa)$ polynomial.

Thus, to speed up this attack one would have find a better approximating polynomial for our PRF (or, more ambitiously, improve on Razborov-Smolensky, which hasn't been done in the last 35 years). In any case, in light of the current state-of-affairs, we do not see any reason to suspect that distinguishing attacks based on natural properties defined in terms of approximating polynomials could invalidate our candidate's conjectured security (Conjecture 1.2).

7.3 Limitations of Learning in $\mathsf{AC}^0[p]$

We recall the CIKK [CIKK16] learner for classes $\mathcal{C}$ that contain $\mathsf{AC}^0[p]$ and admit a natural property. Besides the natural property, the learner uses a family of low-intersection sets (a Nisan-Wigderson design).

7.4 Nisan–Wigderson designs and generator

Definition 7.3 (NW design [NW94]) *For $n, m, L \in \mathbb{N}$, a sequence of sets $S_1, \ldots, S_L \subseteq [m]$ forms an NW design if*

- *$|S_i| = n$ for all $1 \leq i \leq L$ and*
- *$|S_i \cap S_j| \leq \log L$ for all $1 \leq i \neq j \leq L$.*

Definition 7.4 (NW generator [NW94]) *Let $f : \{0,1\}^n \to \{0,1\}$, set $m = n^2$, and let $L(m) < 2^n$ be a stretch function. For an NW design $S_1, \ldots, S_L \subseteq [m]$, define the NW pseudorandom generator $G_f : \{0,1\}^m \to \{0,1\}^{L(m)}$ as:*

$$G_f(z) = f(z|_{S_1}) f(z|_{S_2}) \ldots f(z|_{S_{L(m)}}),$$

where $z|_S$ is the restriction of z to coordinates in S. We call g_z the NW function, where $g_z(i) = f(z|_{S_i})$ is the i^{th} truth table entry of G_f.

Theorem 7.5 (NW function circuit size [CIKK16]) *There exists an NW design such that the function $MX_{NW} : \{0,1\}^\ell \times \{0,1\}^m \to \{0,1\}^n$, defined by*

$$MX_{NW}(i, z) = z|_{S_i} \ ,$$

is computable by an $\mathsf{AC}^0[p]$ circuit of size $O(n^3 \log(n) \log(L))$.

On the other hand, providing a lower bound on the size of the NW function g_z remains an open question.

7.5 The CIKK Learner

We now formalize a CIKK-style learner. Our definition is deliberately liberal so that the CIKK query lower bound applies to a broad class of learning algorithms. In particular, we place no restrictions on how the hypothesis h is obtained; the only requirement is that the procedure eventually invokes a natural property distinguisher on an NW function.

Definition 7.6 (CIKK Learner) *Let $\mathcal{C}$ be a circuit class, and suppose there exists a* natural property distinguisher *D for $\mathcal{C}$. A* CIKK learner for $\mathcal{C}$ with respect to D *for a function f is a probabilistic algorithm $\mathcal{A}^D$ such that*

- *$\mathcal{A}^D$ runs D on the Nisan-Wigderson function $g_z(i)$ (whose truth table is the string $G_f(z)$) of size L for some stretch function $L(m) > m$.*

Algorithm 1 CIKK Learning Algorithm

Input: Oracle to function $f : \{0,1\}^n \to \{0,1\}$, natural property distinguisher D
Output: Circuit C approximating f

1. Randomly sample $i \leftarrow \{1, \ldots, L\}$.
2. For each $1 \le j \le n^2$ with $j \notin S_i$, randomly sample $z_j \leftarrow \{0,1\}$.
3. For each $k < i$, query f on all $2^{|S_i \cap S_k|} \le L$ strings that might be $z|_{S_k}$ consistent with $z'_j s$. Store these in table T.
4. Randomly sample $b_k \leftarrow \{0,1\}$ for all $i \le k \le L$.

Output the following circuit C: on input x, set $z|_{S_i} = x$, while other bits of z are fixed to random.

- Set $b_k = f(z|_{S_k})$ for all $k < i$ using the table T.
- If $D(b_1, \ldots, b_L) = 1$, output b_i. Else, output $1 - b_i$.

- $\mathcal{A}^D$ *outputs a hypothesis* $h : \{0,1\}^n \to \{0,1\}$ *such that*

$$\Pr_{x \sim U_n}[h(x) = f(x)] \;\ge\; \tfrac{1}{2} + \tfrac{1}{\mathsf{poly}(n)}.$$

The expensive step is recovering the true values $f(z|_{S_k})$ for indices $k < i$. Because each S_k overlaps previously recovered sets in at most $\log L$ positions, the number of oracle calls needed to compute b_k is at most $2^{\log L} = L$ for each k. Summing over k gives $Q(n) = \Theta(L^2)$ queries in the standard implementation.

7.6 A query lower bound for CIKK

Oliveira and Santhanam [OS16] (Lemma 6) show two speedup implications for CIKK:

1. A $2^{O(n)}$-time natural property for $\mathcal{C}[\mathsf{poly}(n)]$ gives a learning algorithm for $\mathcal{C}[\mathsf{poly}(n)]$ running in time $O(2^{n^\varepsilon})$ for any $\varepsilon > 0$.
2. A $2^{O(n)}$-time natural property for $\mathcal{C}[2^{n^\varepsilon}]$ yields a learning algorithm for $\mathcal{C}[\mathsf{poly}(n)]$ running in time $2^{(\log n)^{O(1)}}$.

Analogous statements hold when one plugs CIKK into the standard reductions to improve the *query* complexity of learning.

However, these speedups disappear once circuit size is accounted for. The statements above parameterize running time solely by the distinguisher/learner input length n, suppressing the circuit-size parameter. Rewriting them in terms of the circuit size m makes this explicit:

1. If $m = n^c$ for some $c > 0$, then a $2^{O(m^{1/c})}$-time natural property for $\mathcal{C}[m]$ yields a learner for $\mathcal{C}[m]$ running in time $O\big(2^{m^{\varepsilon/c}}\big)$ for any $\varepsilon > 0$. In other words, we start and end with subexponential-time procedures giving no asymptotic improvement.
2. If $m = 2^{n^\varepsilon}$ (with $\varepsilon, c > 0$), then a $2^{O\left((\log m)^{1/\varepsilon}\right)}$-time natural property for $\mathcal{C}[m]$ gives a learner for $\mathcal{C}[s = n^c]$ running in time $2^{(\log s)^{O(1)}}$. Here both sides are quasipolynomial.

The apparent advantage comes from viewing the property as a function of n alone. Once the necessary size threshold is enforced, the overall learning complexity does not actually improve.

Moreover, for classes such as $AC^0[p]$ any meaningful natural property must depend on circuit size since every Boolean function has a (potentially exponential-size) AC^0 implementation (e.g., as a 2-DNF).

Lemma 7.7 (CIKK learners are limited for size-dependent natural properties) *Let $\mathcal{C}$ be a typical circuit class. Suppose $\mathcal{C}$ admits a CIKK learner A^D as in Definition 7.6, which uses a natural property distinguisher D making $g(s)$ queries, where s is the size of the target circuit. Let t be the circuit size of the NW function g_z. Then A^D requires at least $g(t)$ queries.*

Proof. By the definition of a CIKK learner (Definition 7.6), A^D must run D on the NW function g_z. Thus, D must run on circuits of size at least t, requiring at least $g(t)$ queries. From the definition of a CIKK learner (Definition 7.6) which runs D on G_f, since D requires at least $\Omega(t)$ queries, then A^D also requires at least $g(t)$ queries. □

Applying this lemma to the best known natural property (Razborov-Smolensky) for $\mathsf{AC}^0[p]$, which uses $\Omega(2^{\log^d(m)})$ queries, where m is the circuit size and d is the depth of the circuit and the smallest known NW circuit size $t = \Omega(n^3 \log n \log L)$ from Theorem 7.5 gives us the following corollary.

Corollary 7.8 *Suppose a CIKK learner uses the Razborov-Smolensky natural property for* $\mathsf{AC}^0[p]$ *and an NW circuit size of* $\Omega(n)$*. Then the learner requires at least quasipolynomially many queries.*

Remark 7.9 *Note that the above result* does not *extend to the regime where one measures the query complexity of the distinguisher D solely as a function of the input length n. The subtlety is that the NW function* g_z *is parameterized by an input length of* $\Theta(\log L)$*. Hence its* input length *can be as small as* $O(\log n)$ *even when the relevant* circuit size *parameters might remain as large as* $\Omega(n)$*.*

To obtain a polynomial-query learner, we would require a natural property that uses polynomially many queries in the circuit size or a smaller circuit size for the NW function, which is currently not known for $\mathsf{AC}^0[p]$.

8 Acknowledgements

The work of L. Kohl is funded by NWO Talent Programme Veni (VI.Veni.222.348) and by NWO Gravitation project QSC. Part of this work was conducted while the author was visiting the Simons Institute for the Theory of Computing. N. Resch is supported by an NWO (Dutch Research Council) grant with number C.2324.0590. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2039655 as well as the National Science Foundation CAREER Award #2443735. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.

Following the preparation of an earlier draft of this manuscript, the proofs of Theorems 5.11, 5.13, 5.19 and 5.23 were produced with assistance from Gemini Deep Think, strengthening our initial results. Concretely, we initially only had an incomplete proof attempt for Theorem 5.11, lacked the support for adaptive queries in Theorems 5.13 and 5.19, and had a weaker average-case bound on sparsity instead of the worst-case bound in Theorem 5.23. We manually edited the final writeups and take full responsibility for their content.

References

[ABG+14] Adi Akavia, Andrej Bogdanov, Siyao Guo, Akshay Kamath, and Alon Rosen. Candidate weak pseudorandom functions in AC^0 *o* MOD_2. In Moni Naor, editor, *ITCS 2014*, pages 251–260. ACM, January 2014.

[AKK+03] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing low-degree polynomials over GF(2(. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques, 6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2003, Princeton, NJ, USA, August 24-26, 2003, Proceedings*, volume 2764 of *Lecture Notes in Computer Science*, pages 188–199. Springer, 2003.

[AKPW13] Joël Alwen, Stephan Krenn, Krzysztof Pietrzak, and Daniel Wichs. Learning with rounding, revisited - new reduction, properties and applications. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 57–74. Springer, Berlin, Heidelberg, August 2013.

[AR16] Benny Applebaum and Pavel Raykov. Fast pseudorandom functions based on expander graphs. In Martin Hirt and Adam D. Smith, editors, *TCC 2016-B, Part I*, volume 9985 of *LNCS*, pages 27–56. Springer, Berlin, Heidelberg, October / November 2016.

[ARS+15] Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 430–454. Springer, Berlin, Heidelberg, April 2015.

[BCG+20] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Correlated pseudorandom functions from variable-density LPN. In *61st FOCS*, pages 1069–1080. IEEE Computer Society Press, November 2020.

[BCG+21] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Low-complexity weak pseudorandom functions in AC0[MOD2]. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part IV*, volume 12828 of *LNCS*, pages 487–516, Virtual Event, August 2021. Springer, Cham.

[BCH+25] Fabrice Benhamouda, Caicai Chen, Shai Halevi, Yuval Ishai, Hugo Krawczyk, Tamer Mour, Tal Rabin, and Alon Rosen. Encrypted matrix-vector products from secret dual codes. In Chun-Ying Huang, Jyh-Cheng Chen, Shiuh-Pyng Shieh, David Lie, and Véronique Cortier, editors, *ACM CCS 2025*, pages 394–408. ACM Press, October 2025.

[BHI+20] Marshall Ball, Justin Holmgren, Yuval Ishai, Tianren Liu, and Tal Malkin. On the complexity of decomposable randomized encodings, or: How friendly can a garbling-friendly PRF be? In Thomas Vidick, editor, *ITCS 2020*, volume 151, pages 86:1–86:22. LIPIcs, January 2020.

[BIP+18] Dan Boneh, Yuval Ishai, Alain Passelègue, Amit Sahai, and David J. Wu. Exploring crypto dark matter: New simple PRF candidates and their applications. In Amos Beimel and Stefan Dziembowski, editors, *TCC 2018, Part II*, volume 11240 of *LNCS*, pages 699–729. Springer, Cham, November 2018.

[BJMM12] Anja Becker, Antoine Joux, Alexander May, and Alexander Meurer. Decoding random binary linear codes in $2^{n/20}$: How $1 + 1 = 0$ improves information set decoding. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 520–536. Springer, Berlin, Heidelberg, April 2012.

[BKS+10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. In *51st FOCS*, pages 488–497. IEEE Computer Society Press, October 2010.

[BLP11] Daniel J. Bernstein, Tanja Lange, and Christiane Peters. Smaller decoding exponents: Ball-collision decoding. In Phillip Rogaway, editor, *CRYPTO 2011*, volume 6841 of *LNCS*, pages 743–760. Springer, Berlin, Heidelberg, August 2011.

[BM18] Leif Both and Alexander May. Decoding linear codes with high error rate and its impact for LPN security. In Tanja Lange and Rainer Steinwandt, editors, *PQCrypto 2018*, pages 25–46. Springer, Cham, 2018.

[BN25] Mark Braverman and Stephen Newman. Practical secure delegated linear algebra with trapdoored matrices. In Benny Applebaum and Huijia (Rachel) Lin, editors, *TCC 2025, Part I*, volume 16268 of *LNCS*, pages 97–118. Springer, Cham, December 2025.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 719–737. Springer, Berlin, Heidelberg, April 2012.

[BR17] Andrej Bogdanov and Alon Rosen. Pseudorandom functions: Three decades later. Cryptology ePrint Archive, Report 2017/652, 2017.

[CCF+16] Anne Canteaut, Sergiu Carpov, Caroline Fontaine, Tancrède Lepoint, María Naya-Plasencia, Pascal Paillier, and Renaud Sirdey. Stream ciphers: A practical solution for efficient homomorphic-ciphertext compression. In Thomas Peyrin, editor, *Fast Software Encryption - 23rd International Conference, FSE 2016, Bochum, Germany, March 20-23, 2016, Revised Selected Papers*, volume 9783 of *Lecture Notes in Computer Science*, pages 313–333. Springer, 2016.

[CG20] Alessio Caminata and Elisa Gorla. Solving multivariate polynomial systems and an invariant from commutative algebra. In Jean-Claude Bajard and Alev Topuzoglu, editors, *Arithmetic of Finite Fields - 8th International Workshop, WAIFI 2020, Rennes, France, July 6-8, 2020, Revised Selected and Invited Papers*, volume 12542 of *Lecture Notes in Computer Science*, pages 3–36. Springer, 2020.

[CIKK16] Marco L. Carmosino, Russell Impagliazzo, Valentine Kabanets, and Antonina Kolokolova. Learning algorithms from natural proofs. In Ran Raz, editor, *31st Conference on Computational Complexity, CCC 2016, May 29 to June 1, 2016, Tokyo, Japan*, volume 50 of *LIPIcs*, pages 10:1–10:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[CJL25a] Yilei Chen, Liheng Ji, and Wenjie Li. Learning with alternating moduli, arora-ge over composite moduli, and weak PRFs. Cryptology ePrint Archive, Report 2025/968, 2025.

[CJL25b] Yilei Chen, Liheng Ji, and Wenjie Li. Learning with alternating moduli, arora-ge over composite moduli, and weak prfs. *IACR Cryptol. ePrint Arch.*, page 968, 2025.

[CKPS00] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir. Efficient algorithms for solving overdefined systems of multivariate polynomial equations. In Bart Preneel, editor, *EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 392–407. Springer, Berlin, Heidelberg, May 2000.

[CRR21] Geoffroy Couteau, Peter Rindal, and Srinivasan Raghuraman. Silver: Silent VOLE and oblivious transfer from hardness of decoding structured LDPC codes. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part III*, volume 12827 of *LNCS*, pages 502–534, Virtual Event, August 2021. Springer, Cham.

[DI14] Erez Druk and Yuval Ishai. Linear-time encodable codes meeting the gilbert-varshamov bound and their cryptographic applications. In Moni Naor, editor, *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014*, pages 169–182. ACM, 2014.

[DJK25] Youlong Ding, Aayush Jain, and Ilan Komargodski. A new approach for LPN-based pseudorandom functions: Low-depth and key-homomorphic. In Michal Koucký and Nikhil Bansal, editors, *57th ACM STOC*, pages 1898–1909. ACM Press, June 2025.

[DLM+07] Ilias Diakonikolas, Homin K. Lee, Kevin Matulef, Krzysztof Onak, Ronitt Rubinfeld, Rocco A. Servedio, and Andrew Wan. Testing for concise representations. In *48th FOCS*, pages 549–558. IEEE Computer Society Press, October 2007.

[EKM17] Andre Esser, Robert Kübler, and Alexander May. LPN decoded. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part II*, volume 10402 of *LNCS*, pages 486–514. Springer, Cham, August 2017.

[FS09] Matthieu Finiasz and Nicolas Sendrier. Security bounds for the design of code-based cryptosystems. In Mitsuru Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 88–105. Springer, Berlin, Heidelberg, December 2009.

[Gol00] Oded Goldreich. Candidate one-way functions based on expander graphs. Cryptology ePrint Archive, Report 2000/063, 2000.

[JPRZ09] Charanjit S. Jutla, Anindya C. Patthak, Atri Rudra, and David Zuckerman. Testing low-degree polynomials over prime fields. *Random Struct. Algorithms*, 35(2):163–193, 2009.

[Kha93] Michael Kharitonov. Cryptographic hardness of distribution-specific learning. In *25th ACM STOC*, pages 372–381. ACM Press, May 1993.

[KL01] Matthias Krause and Stefan Lucks. On the minimal hardware complexity of pseudorandom function generators. In Afonso Ferreira and Horst Reichel, editors, *STACS 2001, 18th Annual Symposium on Theoretical Aspects of Computer Science, Dresden, Germany, February 15-17, 2001, Proceedings*, volume 2010 of *Lecture Notes in Computer Science*, pages 419–430. Springer, 2001.

[KM91] Eyal Kushilevitz and Yishay Mansour. Learning decision trees using the Fourier sprectrum (extended abstract). In *23rd ACM STOC*, pages 455–464. ACM Press, May 1991.

[LMN89] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, Fourier transform, and learnability. In *30th FOCS*, pages 574–579. IEEE Computer Society Press, October / November 1989.

[LMN93] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *Journal of the ACM (JACM)*, 40(3):607–620, 1993.

[Lyu05] Vadim Lyubashevsky. The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem. In Chandra Chekuri, Klaus Jansen, José D. P. Rolim, and Luca Trevisan, editors, *Approximation, Randomization and Combinatorial Optimization, Algorithms and Techniques, 8th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005 and 9th InternationalWorkshop on Randomization and Computation, RANDOM 2005, Berkeley, CA, USA, August 22-24, 2005, Proceedings*, volume 3624 of *Lecture Notes in Computer Science*, pages 378–389. Springer, 2005.

[MJSC16] Pierrick Méaux, Anthony Journault, François-Xavier Standaert, and Claude Carlet. Towards stream ciphers for efficient FHE with low-noise ciphertexts. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part I*, volume 9665 of *LNCS*, pages 311–343. Springer, Berlin, Heidelberg, May 2016.

[MMT11] Alexander May, Alexander Meurer, and Enrico Thomae. Decoding random linear codes in $\tilde{\mathcal{O}}(2^{0.054n})$. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 107–124. Springer, Berlin, Heidelberg, December 2011.

[MO15] Alexander May and Ilya Ozerov. On computing nearest neighbors with applications to decoding of binary linear codes. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 203–228. Springer, Berlin, Heidelberg, April 2015.

[MV12] Eric Miles and Emanuele Viola. Substitution-permutation networks, pseudorandom functions, and natural proofs. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 68–85. Springer, Berlin, Heidelberg, August 2012.

[NR97] Moni Naor and Omer Reingold. Number-theoretic constructions of efficient pseudo-random functions. In *38th FOCS*, pages 458–467. IEEE Computer Society Press, October 1997.

[NRR00] Moni Naor, Omer Reingold, and Alon Rosen. Pseudo-random functions and factoring (extended abstract). In *32nd ACM STOC*, pages 11–20. ACM Press, May 2000.

[NW88] Noam Nisan and Avi Wigderson. Hardness vs. randomness (extended abstract). In *29th FOCS*, pages 2–11. IEEE Computer Society Press, October 1988.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of computer and System Sciences*, 49(2):149–167, 1994.

[O'D14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014.

[OS16] Igor C. Oliveira and Rahul Santhanam. Conspiracies between learning algorithms, circuit lower bounds and pseudorandomness. *CoRR*, abs/1611.01190, 2016.

[Pra62] Eugene Prange. The use of information sets in decoding cylcic codes. *IRE Transactpons on Information Theory*, 8(5):5–9, 1962.

[Raz87] Alexander A Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mat. Zametki*, 41(4):598–607, 1987.

[RR94] Alexander A. Razborov and Steven Rudich. Natural proofs. In *26th ACM STOC*, pages 204–213. ACM Press, May 1994.

[RST15] Benjamin Rossman, Rocco A. Servedio, and Li-Yang Tan. An average-case depth hierarchy theorem for boolean circuits. *CoRR*, abs/1504.03398, 2015.

[Smo87] Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In Alfred Aho, editor, *19th ACM STOC*, pages 77–82. ACM Press, May 1987.

[Ste89] Jacques Stern. A method for finding codewords of small weight. *Coding Theory and Applications*, 388(10):106–113, 1989.

[Vio13] Emanuele Viola. The communication complexity of addition. In Sanjeev Khanna, editor, *24th SODA*, pages 632–651. ACM-SIAM, January 2013.

[VZ25] Vinod Vaikuntanathan and Or Zamir. Improving algorithmic efficiency using cryptography. *CoRR*, abs/2502.13065, 2025.

[Wil13] Ryan Williams. Natural proofs versus derandomization. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th ACM STOC*, pages 21–30. ACM Press, June 2013.

[YS16] Yu Yu and John P. Steinberger. Pseudorandom functions in almost constant depth from low-noise LPN. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part II*, volume 9666 of *LNCS*, pages 154–183. Springer, Berlin, Heidelberg, May 2016.

A Proof of Theorem 5.11

We prove that our candidate resists every fixed non-adaptive linear test making $Q = O(1)$ distinct queries. Recall that

$$g(Y) = (-1)^{G(Y)} = \prod_{i=1}^{\ell} h(Y_i), \qquad h(z) = (-1)^{\mathsf{TRIBES}_{m,n}(z)},$$

where the Y_i's are disjoint $M = mn$-bit blocks.

We first analyze a slightly relaxed test in which A is sampled uniformly from all matrices in $\mathbb{F}_2^{N\times N}$, rather than from $\mathcal{GL}_N(\mathbb{F}_2)$. We later transfer the bound back to the actual distribution $A \leftarrow \mathcal{GL}_N(\mathbb{F}_2)$.

Fix distinct query points $x^{(1)}, \ldots, x^{(Q)} \in \mathbb{F}_2^N$. For a function $f : \mathbb{F}_2^N \to \{-1, 1\}$, define

$$\Xi_f(x^{(1)}, \ldots, x^{(Q)}) := \mathop{\mathbb{E}}_{A,b}\left[\prod_{r=1}^{Q} f(Ax^{(r)} + b)\right],$$

where, for now, A is uniform over $\mathbb{F}_2^{N\times N}$ and b is uniform over $\mathbb{F}_2^N$.

Expanding each $f(Ax^{(r)} + b)$ in the Fourier basis gives

$$\Xi_f = \sum_{\xi_1,\ldots,\xi_Q \in \mathbb{F}_2^N} \left(\prod_{r=1}^{Q} \widehat{f}(\xi_r)\right) \mathop{\mathbb{E}}_{A,b}\left[(-1)^{\sum_{r=1}^{Q}\langle \xi_r, Ax^{(r)}+b\rangle}\right].$$

Using the matrix trace inner product $\langle U, V\rangle := \mathrm{tr}(U^\top V)$, we have

$$\sum_{r=1}^{Q}\langle \xi_r, Ax^{(r)}\rangle = \left\langle \sum_{r=1}^{Q} \xi_r (x^{(r)})^\top, A\right\rangle, \qquad \sum_{r=1}^{Q}\langle \xi_r, b\rangle = \left\langle \sum_{r=1}^{Q} \xi_r, b\right\rangle.$$

By character orthogonality, the expectation over b evaluates to 1 if $\sum_r \xi_r = 0$ and 0 otherwise. Similarly, the expectation over A is 1 if $\sum_r \xi_r (x^{(r)})^\top = 0$ and 0 otherwise. Since a matrix is zero if and only if its transpose is zero, the condition $\sum_r \xi_r (x^{(r)})^\top = 0$ is equivalent to $\sum_r x^{(r)} \xi_r^\top = 0$. Hence,

$$\Xi_f = \sum_{\substack{\xi_1,\ldots,\xi_Q \in \mathbb{F}_2^N \\ \sum_r \xi_r = 0 \\ \sum_r x^{(r)}\xi_r^\top = 0}} \prod_{r=1}^{Q} \widehat{f}(\xi_r). \tag{8}$$

Applying the triangle inequality, we have:

$$|\Xi_f| \le \sum_{\substack{\xi_1,\dots,\xi_Q \in \mathbb{F}_2^N \\ \sum_r \xi_r = 0 \\ \sum_r x^{(r)} \xi_r^\top = 0}} \prod_{r=1}^{Q} |\widehat{f}(\xi_r)|. \tag{9}$$

Let $X \in \mathbb{F}_2^{N \times Q}$ be the query matrix whose r-th column is $x^{(r)}$, and let $y_1, \dots, y_N \in \mathbb{F}_2^Q$ be the rows of X. Define

$$V := \mathsf{Span}\{y_1, \dots, y_N, \mathbf{1}\} \le \mathbb{F}_2^Q,$$

where $\mathbf{1} = (1, \dots, 1)$.

We now define the corresponding constrained absolute Fourier sum for a local L-bit predicate $F : \mathbb{F}_2^L \to \mathbb{R}$. This local definition will be applied first with $L = n$, then with $L = M = mn$, and finally blockwise to the full $N = \ell M$-bit function. For $\xi_1, \dots, \xi_Q \in \mathbb{F}_2^L$, write

$$\xi_r = (\xi_{r,1}, \dots, \xi_{r,L}),$$

and for each coordinate $t \in [L]$, form the column vector

$$c_t := (\xi_{1,t}, \dots, \xi_{Q,t})^\top \in \mathbb{F}_2^Q.$$

The global constraints in (9), restricted to any fixed set of coordinates, require precisely that each $c_t \in V^\perp$. Indeed, $\sum_r \xi_r = 0$ imposes $\langle \mathbf{1}, c_t \rangle = 0$, and $\sum_r x^{(r)} \xi_r^\top = 0$ imposes $\langle y_j, c_t \rangle = 0$ for each row y_j of X.

Thus, for a local predicate $F : \mathbb{F}_2^L \to \mathbb{R}$, define

$$S_F(V) := \sum_{\substack{\xi_1,\dots,\xi_Q \in \mathbb{F}_2^L \\ c_t = (\xi_{1,t},\dots,\xi_{Q,t})^\top \in V^\perp \text{ for all } t \in [L]}} \prod_{r=1}^{Q} |\widehat{F}(\xi_r)|. \tag{10}$$

Equivalently, one may arrange the coordinate-columns $c_1, \dots, c_L$ as the columns of a $Q \times L$ matrix C. The rows of C are precisely $\xi_1, \dots, \xi_Q$, and the condition is that every column of C lies in $V^\perp$.

We begin with a structural lemma about $V^\perp$.

Lemma A.1 *Assume the queries $x^{(1)}, \dots, x^{(Q)} \in \mathbb{F}_2^N$ are distinct. Then every nonzero vector in $V^\perp$ has Hamming weight at least 4. Equivalently, either $V^\perp = \{0\}$, or $V^\perp$ is a binary linear code of minimum distance at least 4.*

Proof. By definition,

$$V = \mathsf{Span}\{y_1, \dots, y_N, \mathbf{1}\} \subseteq \mathbb{F}_2^Q.$$

If $c \in V^\perp$, then $\langle c, \mathbf{1} \rangle = 0$, so c has even Hamming weight. Thus weights 1 and 3 are impossible.

If $c \in V^\perp$ has weight 2, then $c = e_r + e_s$ for some distinct r, s. Then

$$\langle c, y_j \rangle = y_{j,r} + y_{j,s} = 0$$

for every $j \in [N]$. Hence the r-th and s-th columns of X are identical, meaning $x^{(r)} = x^{(s)}$, contradicting the assumption that the queries are distinct. Therefore weight 2 is also impossible. □

We next analyze the constrained sum for one n-bit AND gate. In this part we temporarily use a $\{0,1\}$-valued function as an intermediate object. Let

$$u = 1 - \mathsf{AND}_n.$$

Proposition A.2 *For $u = 1 - \mathsf{AND}_n$, the Fourier coefficients are*

$$\widehat{u}(0) = 1 - 2^{-n},$$

and for every $\alpha \neq 0$,

$$\widehat{u}(\alpha) = -2^{-n}(-1)^{|\alpha|}.$$

In particular,

$$|\widehat{u}(\alpha)| = \begin{cases} 1 - 2^{-n}, & \alpha = 0, \\ 2^{-n}, & \alpha \neq 0. \end{cases}$$

Proof. Since $u(x) = 1$ except at $x = \mathbf{1}$, where $u(\mathbf{1}) = 0$, we have

$$u(x) = 1 - \mathbf{1}_{\{x=\mathbf{1}\}}.$$

Therefore

$$\widehat{u}(0) = 1 - 2^{-n},$$

and for $\alpha \neq 0$,

$$\widehat{u}(\alpha) = -2^{-n}(-1)^{\langle \alpha, \mathbf{1} \rangle} = -2^{-n}(-1)^{|\alpha|}.$$

□

Let

$$a := |\widehat{u}(0)| = 1 - 2^{-n}.$$

Proposition A.3 *For $u = 1 - \mathsf{AND}_n$,*

$$S_u(V) = a^Q + O_Q(2^{-3n}).$$

Proof. Using the matrix viewpoint from (10), write the constrained sum as a sum over $Q \times n$ matrices C whose columns lie in $V^\perp$. The rows of C are $\xi_1, \ldots, \xi_Q \in \mathbb{F}_2^n$. Thus

$$S_u(V) = \sum_{C \in (V^\perp)^n} \prod_{r=1}^{Q} |\widehat{u}(\xi_r)|.$$

The zero matrix $C = 0$ contributes exactly

$$|\widehat{u}(0)|^Q = a^Q.$$

Now fix a nonzero matrix $C \neq 0$. Let

$$U := \mathsf{Span}\{c_1, \ldots, c_n\} \subseteq V^\perp,$$

where $c_1, \ldots, c_n$ are the columns of C, and write $j := \dim(U) \geq 1$. Let k be the number of nonzero rows of C. Equivalently, if $\mathsf{supp}(U) := \{r \in [Q] : \exists u \in U \text{ with } u_r = 1\}$, then $k = |\mathsf{supp}(U)|$. Indeed, row r of C is nonzero exactly when some column c_t, and hence some vector in U, has a nonzero r-th coordinate.

Because $U \subseteq V^\perp$, every nonzero codeword of U has weight at least 4 by Lemma A.1. Puncturing U to its support gives a binary linear code of length k, dimension j, and minimum distance at least 4. By the Singleton bound,

$$k \geq j + 4 - 1 = j + 3.$$

Exactly k of the rows ξ_r are nonzero. Hence, by Proposition A.2,

$$\prod_{r=1}^{Q} |\widehat{u}(\xi_r)| = a^{Q-k}(2^{-n})^k \leq 2^{-nk} \leq 2^{-n(j+3)}.$$

For a fixed subspace U, the number of matrices C whose columns all lie in U is at most

$$|U|^n = 2^{jn}.$$

Thus the total contribution of all matrices whose column span is contained in a fixed j-dimensional subspace $U \subseteq V^\perp$ is at most

$$2^{jn} \cdot 2^{-n(j+3)} = 2^{-3n}.$$

Since Q is fixed, the number of subspaces of $\mathbb{F}_2^Q$ is a constant depending only on Q. Here, the $O(1)$ can be replaced by 2^{Q^2}. Summing over all nonzero subspaces $U \subseteq V^\perp$ gives total nonzero contribution $O_Q(2^{-3n})$. Therefore

$$S_u(V) = a^Q + O_Q(2^{-3n}).$$

□

We now tensorize from a single AND gate to a full TRIBES block. Let

$$q(z) := 1 - \mathsf{TRIBES}_{m,n}(z) = \prod_{j=1}^{m} u(z^{(j)}),$$

where the m copies of u depend on disjoint n-bit blocks.

Corollary A.4 *For $q = 1 - \mathsf{TRIBES}_{m,n}$,*

$$S_q(V) = a^{mQ} + O_Q(m2^{-3n}).$$

Proof. Write each $\xi_r \in \mathbb{F}_2^{mn}$ as

$$\xi_r = (\xi_r^{(1)}, \dots, \xi_r^{(m)}), \qquad \xi_r^{(j)} \in \mathbb{F}_2^n.$$

Since $q = \prod_{j=1}^m u_j$ is a product of functions on disjoint input blocks, its Fourier coefficients factor:

$$\widehat{q}(\xi_r) = \prod_{j=1}^m \widehat{u}(\xi_r^{(j)}).$$

Moreover, the condition that every coordinate-column $(\xi_{1,t}, \dots, \xi_{Q,t})^\top$ lies in $V^\perp$ is coordinatewise, and therefore separates across the m blocks. Hence

$$S_q(V) = \prod_{j=1}^m \sum_{\substack{\xi_1^{(j)}, \dots, \xi_Q^{(j)} \in \mathbb{F}_2^n \\ \text{all coordinate-columns lie in } V^\perp}} \prod_{r=1}^Q |\widehat{u}(\xi_r^{(j)})| = (S_u(V))^m.$$

By Proposition A.3,

$$S_u(V) = a^Q + O_Q(2^{-3n}).$$

Since $a = 1 - 2^{-n}$ and Q is fixed, we have $a^Q + O_Q(2^{-3n}) \le 1$ for all sufficiently large n. Hence

$$S_q(V) = \left(a^Q + O_Q(2^{-3n})\right)^m = a^{mQ} + O_Q(m2^{-3n}).$$

□

We next transfer this estimate to the signed TRIBES block

$$h(z) := (-1)^{\mathsf{TRIBES}_{m,n}(z)} = 2q(z) - 1.$$

Let

$$\beta := |2a^m - 1| = \left|2(1 - 2^{-n})^m - 1\right|.$$

Proposition A.5 *Assume $Q = O(1)$ and a^m is bounded below by an absolute positive constant. Then*

$$S_h(V) \le \beta^Q + O_Q(m2^{-3n}).$$

In particular, under the balanced-TRIBES parameter choice satisfying

$$\left|2(1 - 2^{-n})^m - 1\right| = O\left(\frac{\log \lambda}{\lambda}\right),$$

we have

$$S_h(V) \le O_Q\left(\left(\frac{\log \lambda}{\lambda}\right)^Q\right) + O_Q(m2^{-3n}).$$

Proof. For $\xi \neq 0$,

$$\widehat{h}(\xi) = 2\widehat{q}(\xi), \qquad \text{so} \qquad |\widehat{h}(\xi)| = 2|\widehat{q}(\xi)|.$$

At the origin,

$$|\widehat{h}(0)| = |2\widehat{q}(0) - 1| = |2a^m - 1| = \beta.$$

Write the constrained sum over $Q \times mn$ matrices C whose columns lie in $V^\perp$. The rows of C are $\xi_1, \dots, \xi_Q \in \mathbb{F}_2^{mn}$. Thus

$$S_h(V) = \sum_{C \in (V^\perp)^{mn}} \prod_{r=1}^Q |\widehat{h}(\xi_r)|.$$

The zero matrix contributes exactly β^Q.

For every nonzero matrix $C \neq 0$, at least one row ξ_r is nonzero. Since Q is fixed and a^m is bounded below by a positive constant, there is a constant C_Q such that

$$\prod_{r=1}^{Q} |\widehat{h}(\xi_r)| \leq C_Q \prod_{r=1}^{Q} |\widehat{q}(\xi_r)|$$

for all nonzero C. Therefore

$$S_h(V) \leq \beta^Q + C_Q \sum_{C \neq 0} \prod_{r=1}^{Q} |\widehat{q}(\xi_r)| = \beta^Q + C_Q\big(S_q(V) - a^{mQ}\big).$$

By Corollary A.4,

$$S_q(V) - a^{mQ} = O_Q(m2^{-3n}).$$

The balanced-TRIBES specialization follows from the assumed bound on β. □

Finally, we lift the estimate from one $M = mn$-bit TRIBES block to the full predicate

$$g(Y) = \prod_{i=1}^{\ell} h(Y_i),$$

where the Y_i's are disjoint M-bit blocks.

Corollary A.6 *Assume $Q = O(1)$, and suppose the TRIBES parameters are balanced so that*

$$\beta = \left|2(1-2^{-n})^m - 1\right| = O\left(\frac{\log \lambda}{\lambda}\right).$$

Then

$$S_g(V) \leq \left(O_Q\left(\left(\frac{\log \lambda}{\lambda}\right)^Q\right) + O_Q(m2^{-3n})\right)^{\ell}.$$

Proof. Since g is the product of ℓ disjoint copies of h, the constrained sum factors:

$$S_g(V) = \big(S_h(V)\big)^{\ell}.$$

The claim follows immediately from Proposition A.5. □

We now specialize to the parameter regime of our construction.

Corollary A.7 *Suppose*

$$\ell = \Theta(\lambda), \qquad m = \Theta(\lambda/\log\lambda), \qquad n = \Theta(\log\lambda - \log\log\lambda),$$

and the TRIBES parameters are chosen so that

$$2^{-n} = \Theta\left(\frac{\log\lambda}{\lambda}\right) \qquad \textit{and} \qquad \left|2(1-2^{-n})^m - 1\right| = O\left(\frac{\log\lambda}{\lambda}\right).$$

If $Q = O(1)$, then

$$S_g(V) \leq 2^{-\Omega(\lambda \log \lambda)}.$$

Proof. The balanced condition gives

$$\beta = O\left(\frac{\log\lambda}{\lambda}\right).$$

Moreover,

$$m2^{-3n} = \Theta\left(\frac{\lambda}{\log\lambda}\right) \cdot \Theta\left(\frac{\log^3\lambda}{\lambda^3}\right) = \Theta\left(\frac{\log^2\lambda}{\lambda^2}\right).$$

Therefore, for $Q = 1$,

$$S_h(V) \leq O\left(\frac{\log\lambda}{\lambda}\right),$$

and for every fixed $Q \geq 2$,

$$S_h(V) \leq O_Q\left(\frac{\log^2 \lambda}{\lambda^2}\right).$$

In either case, since $\ell = \Theta(\lambda)$,

$$S_g(V) = \big(S_h(V)\big)^\ell \leq 2^{-\Omega(\lambda \log \lambda)}.$$

□

It remains to transfer the relaxed all-matrices analysis back to the actual candidate, where $A \leftarrow \mathcal{GL}_N(\mathbb{F}_2)$.

Lemma A.8 *Let $x^{(1)}, \ldots, x^{(Q)} \in \mathbb{F}_2^N$ be fixed distinct queries with $Q = O(1)$. Let*

$$\Xi_g^{\mathrm{all}} := \mathop{\mathbb{E}}_{A \leftarrow \mathbb{F}_2^{N \times N},\, b} \left[\prod_{r=1}^{Q} g(Ax^{(r)} + b)\right],$$

and

$$\Xi_g^{\mathrm{inv}} := \mathop{\mathbb{E}}_{A \leftarrow \mathcal{GL}_N(\mathbb{F}_2),\, b} \left[\prod_{r=1}^{Q} g(Ax^{(r)} + b)\right].$$

Then

$$|\Xi_g^{\mathrm{inv}}| \leq O_Q\left(|\Xi_g^{\mathrm{all}}|\right) + O_Q(2^{-N}).$$

Proof. Let

$$W := \mathsf{Span}\{x^{(1)}, \ldots, x^{(Q)}\}, \qquad r := \dim(W) \leq Q.$$

For $A \leftarrow \mathbb{F}_2^{N \times N}$, let $\mathcal{E}$ be the event that the map $A|_W : W \to \mathbb{F}_2^N$ is injective. Then

$$\Pr[\mathcal{E}] = \prod_{i=0}^{r-1}(1 - 2^{i-N}) \geq 1 - 2^{r-N} = 1 - O_Q(2^{-N}).$$

Conditioned on $\mathcal{E}$, the restriction $A|_W$ is a uniformly random injective linear map from W into $\mathbb{F}_2^N$. This is exactly the distribution of $A|_W$ when $A \leftarrow \mathcal{GL}_N(\mathbb{F}_2)$. Hence

$$\Xi_g^{\mathrm{inv}} = \mathbb{E}\left[\prod_{r=1}^{Q} g(Ax^{(r)} + b) \;\middle|\; \mathcal{E}\right],$$

where the expectation on the right is over $A \leftarrow \mathbb{F}_2^{N \times N}$ and uniform b.

Since the product is always in $\{-1, 1\}$,

$$|\Xi_g^{\mathrm{inv}}| \leq \frac{|\Xi_g^{\mathrm{all}}| + \Pr[\neg\mathcal{E}]}{\Pr[\mathcal{E}]} \leq O_Q\left(|\Xi_g^{\mathrm{all}}|\right) + O_Q(2^{-N}).$$

□

Theorem A.9 *Under the parameter regime of Corollary A.7, for every fixed $Q = O(1)$ and every set of distinct non-adaptive queries $x^{(1)}, \ldots, x^{(Q)} \in \mathbb{F}_2^N$,*

$$\left|\mathop{\mathbb{E}}_{A \leftarrow \mathcal{GL}_N(\mathbb{F}_2),\, b} \left[\prod_{r=1}^{Q} g(Ax^{(r)} + b)\right]\right| \leq 2^{-\Omega(\lambda \log \lambda)}.$$

Equivalently, every fixed non-adaptive linear test making $Q = O(1)$ distinct queries has negligible parity bias against the candidate.

Proof. By the constrained Fourier expansion and the triangle bound (9), together with the blockwise definition of $S_g(V)$, we have

$$|\Xi_g^{\mathrm{all}}| \leq S_g(V).$$

Corollary A.7 gives

$$S_g(V) \leq 2^{-\Omega(\lambda \log \lambda)}.$$

Finally, Lemma A.8 gives

$$|\Xi_g^{\mathrm{inv}}| \leq 2^{-\Omega(\lambda \log \lambda)} + O_Q(2^{-N}).$$

Since $N = \Theta(\lambda^2)$, the $O_Q(2^{-N})$ term is dominated by $2^{-\Omega(\lambda \log \lambda)}$. This proves the theorem. □

B Attempt at Proving Resistance Against General Linear Tests for a Random Matrix A

We conjecture that our candidate is resistant against more general *non-adaptive linear tests* if the matrix key $\mathbf{A}$ is chosen randomly. Recall, we define $g : \{0,1\}^N \to \{-1,1\}$ to be the signed version of G: $g(x) \coloneqq (-1)^{G(x)}$. For integers $Q \geq 1$ and $k \geq 1$, fix a matrix $U = (u_{ij}) \in \mathbb{F}_2^{Q \times k}$ and consider

$$x^{(i)} \;=\; \sum_{j=1}^{k} u_{ij}\, v_j \;\in\; \mathbb{F}_2^N \qquad (i = 1, \ldots, Q),$$

where the $v_1, \ldots, v_k \in \mathbb{F}_2^N$ are drawn uniformly at random conditioned on being linearly independent. We require that the resulting queries $x^{(1)}, \ldots, x^{(Q)}$ are distinct and nonzero.[10] The test's bias, as described in Section5.2 is

$$\delta \;\coloneqq\; \mathbb{E}_{v_1,\ldots,v_k}\Big[\,(-1)^{\sum_{i=1}^{Q} G(x^{(i)})}\Big] = \mathbb{E}_{v_1,\ldots,v_k}\Big[\prod_{i=1}^{Q} g\big(x^{(i)}\big)\Big] \;. \tag{11}$$

Our goal is to show $|\delta| \leq 2^{-\Omega(\lambda)}$.

Lemma B.1 (Reduction to a spectral sum) *Let $U \in \mathbb{F}_2^{Q \times k}$ have rank r, and write $S \coloneqq (s_1, \ldots, s_Q) \in (\mathbb{F}_2^N)^Q$. Then*

$$\delta \;=\; \sum_{S \in (\mathbb{F}_2^N)^Q \colon\, S \cdot U = 0} \;\prod_{i=1}^{Q} \widehat{g}(s_i),$$

where S is written as an $N \times Q$ matrix.

Proof. Expanding each $g(x^{(i)})$ into its Fourier expansion and plugging it into 11 gives:

$$\delta = \sum_{s_1,\ldots,s_Q} \Big(\prod_{i=1}^{Q} \widehat{g}(s_i)\Big)\mathbb{E}_{v_1,\ldots,v_k}\Big[\,(-1)^{\sum_{i=1}^{Q}\langle s_i,\, x^{(i)}\rangle}\Big].$$

Using $x^{(i)} = \sum_{j=1}^{k} u_{ij} v_j$ and independence of the v_j,

$$\mathbb{E}_{v_1,\ldots,v_k}\Big[(-1)^{\sum_i \langle s_i,\, x^{(i)}\rangle}\Big] = \prod_{j=1}^{k} \mathbb{E}_{v_j}\Big[(-1)^{\langle \sum_i u_{ij} s_i,\, v_j\rangle}\Big] = \prod_{j=1}^{k} \mathbf{1}\Big\{\sum_i u_{ij} s_i = 0\Big\},$$

since $\mathbb{E}_v[(-1)^{\langle t, v\rangle}] = \mathbf{1}\{t = 0\}$ for v uniform. This yields the claim. □

Since the number of possible $(s_1, \ldots, s_Q)$ in the spectral sum from Lemma B.1 is bounded by $2^{N(Q-r)} \leq 2^{N(Q - \log Q)}$ and from Proposition 5.2, we have $|\widehat{g}(s)| \leq O(2^{-\ell})$, using the Triangle Inequality, this gives the naive bound

$$|\delta| \;\leq\; O(2^{Q(N-\ell) - N\log Q}) \;.$$

However, this bound is clearly too weak whenever $Q = \Omega(\lambda)$. We leave proving a better bound on the sum in Lemma B.1 as an open question.

[10] This excludes degenerate tests where repeated queries make the XOR trivially cancel.