

Rare-Event Decorrelation for Linear Tests of Low-Complexity Pseudorandom Functions

Nikolas Melissaris

IRIF, CNRS & Université Paris Cité, nikolas@irif.fr

Abstract

Linear tests are a basic obstacle to pseudorandom functions built from low-depth circuits and secret linear or affine maps. Ball, Ducros, Erabelli, Kohl, Resch, and Scholl left open whether their bounded-query strong-PRF candidate in $\text{AC}^0[2]$ resists general non-adaptive linear tests with polynomially many chosen queries. Boyle, Couteau, Gilboa, Ishai, Kohl, and Scholl left linear-attack resistance of their Sipser-based weak-PRF candidate open.

We address both questions through a common correlation argument based on rare local events. Repeated Cauchy–Schwarz reduces the original correlation to a parity correlation of indicators of substantially rarer events while preserving the independence supplied by the underlying linear forms. Pairwise independence suffices to bound the resulting parity correlation by a constant below 1, while higher-order independence gives stronger cancellation.

For the candidate of Ball et al., this resolves the open problem for arbitrary polynomially many non-adaptive queries with unrestricted linear or affine dependencies. More generally, for every fixed $a \in (0, 1)$ and $q \leq 2^{a(\log_2 \lambda)^2}$, every such linear test has bias $\exp(-\lambda^{1-a-o(1)})$.

For the Sipser-based family of Boyle et al., for every fixed $0 < \varepsilon < 2$, with overwhelming probability over $Q \leq 2^{(2-\varepsilon)(\log_2 \lambda)^2}$ random samples, every linear attack has bias $\exp(-\lambda^{\varepsilon-o(1)})$. This establishes linear-attack resistance throughout this quasipolynomial-sample regime. Extending the guarantee to the conjectured subexponential weak-PRF regime remains open.

Keywords: Pseudorandom functions, linear tests, low-depth circuits, bounded-query security, small-bias distributions.

1 Introduction

Pseudorandom functions with very small evaluation circuits are of interest both in complexity theory and in cryptographic applications. From a complexity perspective, they ask how little computation can support pseudorandomness. From a cryptographic perspective, low-depth evaluation is valuable when a PRF is used inside secure computation or another protocol whose cost reflects the circuit complexity of the primitive [BDE⁺25, BCG⁺21]. Recent constructions in this line use simple Boolean predicates and ask which attacks can still be ruled out.

A recurring source of difficulty is the reuse of a secret linear or affine map. Across multiple queries, the same map preserves relations among the inputs. Linear maps preserve linear relations, while affine maps preserve affine relations. A linear test takes a parity of output bits and asks whether these preserved relations induce a detectable bias. For low-complexity candidates, controlling such correlations is a basic requirement for pseudorandomness under repeated use of the same key.

We study this problem for two recent candidates with different security notions and query models. In the bounded-query strong-PRF candidate of Ball, Ducros, Erabelli, Kohl, Resch, and Scholl [BDE⁺25], the

adversary chooses the queries and can therefore impose arbitrary affine dependencies. In the weak-PRF construction of Boyle, Couteau, Gilboa, Ishai, Kohl, and Scholl [BCG⁺21], the inputs are uniform random samples, so the adversary does not control the dependency structure. Arbitrary distinct binary linear forms guarantee only pairwise independence, whereas random samples typically have no short linear dependencies and therefore retain much higher-order independence.

The two candidates also have the same basic shape. The candidate of Ball et al. uses TRIBES; the candidate of Boyle et al. uses a depth-three Sipser predicate. In both cases, the Boolean function is a balanced OR of many independent rare local events. Our main observation is that repeated Cauchy–Schwarz can be organized around this rare-event decomposition. Each round introduces an independent copy of one group’s randomness, and after expanding the square the condition that no local event occurs in that group is replaced by the event that the corresponding indicators in the two copies disagree. This makes the event that remains in the final parity substantially rarer while preserving the independence supplied by the underlying linear forms. The number of Cauchy–Schwarz rounds is governed by the rarity of the local event rather than by the full dependency structure of the query system.

1.1 Two linear-test problems

We first recall the chosen-query problem of Ball et al. For $z = (z_{j,r}) \in \mathbb{F}_2^{mn}$, write $\text{TRIBES}_{m,n}(z) = \bigvee_{j=1}^m \bigwedge_{r=1}^n z_{j,r}$. Their candidate applies one secret affine map and then XORs ℓ TRIBES blocks on disjoint coordinate sets:

$$F_{A,b}(x) = \bigoplus_{w=1}^{\ell} \text{TRIBES}_{m,n}\left((Ax \oplus b)^{(w)}\right), \quad (1)$$

where $A \leftarrow \text{GL}_N(\mathbb{F}_2)$, $b \leftarrow \mathbb{F}_2^N$, and $N = \ell mn$. A non-adaptive linear test chooses distinct inputs $x_1, \dots, x_q$ and a nonzero $u \in \mathbb{F}_2^q$ and considers $\bigoplus_{i=1}^q u_i F_{A,b}(x_i)$. For example, suppose four distinct queries satisfy $x_1 + x_2 + x_3 + x_4 = 0$. Their images under $x \mapsto Ax \oplus b$ satisfy the same relation, since the four copies of b cancel. A linear test can then ask whether $F_{A,b}(x_1) \oplus F_{A,b}(x_2) \oplus F_{A,b}(x_3) \oplus F_{A,b}(x_4)$ has a detectable bias. For a uniformly random function this parity is unbiased. Since coordinates with $u_i = 0$ can be discarded, the problem is to control the all-ones parity on an arbitrary set of distinct chosen queries. We write $\text{Bias}(x_1, \dots, x_q)$ for the absolute expectation of this parity over the secret key.

Ball et al. prove a strong bound for every fixed $q = O(1)$, even when the queries have arbitrary dependencies [BDE⁺25, Theorem 5.11]. They also obtain a large-query result for adaptively chosen linearly independent queries [BDE⁺25, Theorem 5.13]. The remaining case is a growing chosen-query set with arbitrary rank and affine dependencies. Their constrained Fourier-sum argument deteriorates rapidly with the query count, and they explicitly ask whether the constant-query result can be extended to an arbitrary polynomial number of queries [BDE⁺25, Section 1.3].

The weak-PRF problem of Boyle et al. has a different form. They use a depth-three Sipser predicate G_λ as deterministic noise in the learning-parity-with-simple-deterministic-noise family

$$f_{s,K}(x) = \langle s, x \rangle \oplus G_\lambda(Kx). \quad (2)$$

Here the public inputs $x_1, \dots, x_Q$ are uniform random samples, and a linear attack may choose its coefficient vector after seeing them. If $\sum_i u_i x_i \neq 0$, the uniform secret s makes the corresponding output parity exactly unbiased. The only potentially useful attacks are therefore supported on linear relations among the public samples. Boyle et al. isolate this linear-attack problem but do not prove resistance for their Sipser-noise construction [BCG⁺21].

Ball et al. must handle a worst-case system of relations selected by the adversary. Boyle et al. need a stronger quantitative bound on a random relation, but random samples give the additional fact that every short subfamily is linearly independent. The same rare-event argument exploits these two levels of independence.

1.2 Our results

For the balanced parameters of Ball et al., $\ell = \lambda$, $mn = \lambda$, $N = \lambda^2$, and $2^{-n} = \Theta\left(\frac{\log \lambda}{\lambda}\right)$, and the key length is $\kappa = N^2 + N = \Theta(\lambda^4)$.

Our first result resolves the open polynomial-query problem of Ball et al. for general non-adaptive linear tests. For every fixed $C > 0$, uniformly over every set of $q \leq \lambda^C$ distinct chosen queries, $\text{Bias}(x_1, \dots, x_q) \leq \exp(-\Omega_C(\lambda))$, with no restriction on the rank or affine-dependence structure of the queries. The same argument extends substantially beyond every fixed polynomial. For every fixed $a \in (0, 1)$,

$$q \leq 2^{a(\log_2 \lambda)^2} \implies \text{Bias}(x_1, \dots, x_q) \leq \exp(-\lambda^{1-a-o(1)}). \quad (3)$$

In particular, writing the bound in terms of the key length, for every fixed $0 < c < 1/4$, $q \leq \kappa^{\log^c \kappa}$ implies $\text{Bias}(x_1, \dots, x_q) \leq 2^{-\kappa^c}$ for all sufficiently large κ . Thus, within non-adaptive linear tests, the result matches the query/advantage form of the bounded-query conjecture of Ball et al. throughout this range. The theorem with explicit parameter bounds and the full tradeoff are given in [Sections 5](#) and [5.2](#).

The second result uses the stronger form of the same decorrelation theorem. Suppose every subfamily of at most k of the underlying linear forms is linearly independent. The Cauchy–Schwarz transformation then preserves k -wise independence, and the final parity estimate is exponentially small in k before taking the 2^t th root introduced by the Cauchy–Schwarz iterations. Random samples have no short linear dependencies with overwhelming probability, which makes this stronger bound available in the weak-PRF setting.

For the Sipser-based family of Boyle et al., for every fixed $0 < \varepsilon < 2$ and $Q \leq 2^{(2-\varepsilon)(\log_2 \lambda)^2}$, with probability $1 - 2^{-\Omega(n_\lambda)}$ over the Q public samples, *simultaneously for every* nonzero linear attack u ,

$$\left| \mathbb{E}_{s,K} \left[(-1)^{\sum_i u_i f_{s,K}(x_i)} \right] \right| \leq \exp(-\lambda^{\varepsilon-o(1)}). \quad (4)$$

The attack vector may be chosen after seeing the entire sample tuple. This establishes linear-attack resistance in this quasipolynomial-sample regime. Extending the guarantee to the conjectured subexponential weak-PRF regime remains open.

The correlation bound for the candidate of Ball et al. has consequences beyond parity tests. We show that at the quasipolynomial query scale the vector of responses on any fixed query set fools subexponential-size constant-depth AC^0 distinguishers. We also obtain unrestricted statistical security for a smaller adaptive query range, and a stronger adaptive theorem when every execution path contains only $O(\lambda/\log \lambda)$ independent affine relations. These consequences are developed in [Sections 7](#) and [8](#).

1.3 Relation to prior work

The two candidates belong to a broader line of work on shallow weak pseudorandom functions; see Boura, Couteau, Perrin, and Rotella [[BCPR25](#)] for a recent systematization. Akavia, Bogdanov, Guo, Kamath, and Rosen [[ABG⁺14](#)] combined a secret linear transformation with TRIBES. Bogdanov and Rosen later observed that this construction has rational degree $O(\log \lambda)$, yielding a quasipolynomial-time distinguisher [[BR17](#)]. Boyle et al. replace the Boolean predicate by a depth-three Sipser function and prove large rational degree for the resulting candidate [[BCG⁺21](#)]. In the chosen-query setting, Ball et al. identify a different weakness of the earlier TRIBES architecture and introduce the affine-map/XOR-of-TRIBES construction studied here [[BDE⁺25](#)]. Their Theorem 5.11 handles arbitrary dependencies for a constant number of non-adaptive queries, while Theorem 5.13 allows many adaptive queries under a linear-independence condition. The missing regime is a growing adversarial query set with arbitrary affine relations, and Section 1.3 explicitly asks whether the constant-query result can be extended to an arbitrary polynomial number of queries [[BDE⁺25](#)].

The bounded-query formulation is also related to standard limitations on low-depth circuits. Razborov and Smolensky showed that bounded-depth circuits with modular gates admit low-degree polynomial approximations in the regimes relevant to their lower bounds [Raz87, Smo87], and Krause and Lucks made the corresponding cryptographic implication explicit, showing that sufficiently effective approximation or learning procedures yield distinguishers against PRFs in the class [KL01]. These results help explain why quasipolynomially many evaluations are already meaningful, but they do not control the joint distribution when the same secret linear or affine map is reused across many queries.

The Boyle et al. problem is similar, but the structure comes from the random samples rather than chosen queries. If the attack coefficients do not define a linear relation among the samples, the parity term makes the attack exactly unbiased, leaving the deterministic Sipser noise to be analyzed only on linear dependencies among random samples [BCG⁺21]. Other structured LPN-based constructions obtain pseudorandom correlation through different algebraic or coding mechanisms [BCG⁺20, CD23, BCG⁺22]. Our result addresses how much cancellation this constant-depth deterministic predicate exhibits on linear relations among the random weak-PRF samples.

On the analytic side, repeated Cauchy–Schwarz for functions evaluated on systems of linear forms is classical. Gowers introduced the uniformity norms that underlie this theory [Gow01]; Green and Tao formalized Cauchy–Schwarz complexity [GT10], and Gowers–Wolf and Hatami–Hatami–Lovett developed the corresponding true-complexity theory [GW10, HHL16]. Austrin and Mossel study a related setting. They consider correlations over pairwise-independent distributions and record an observation of Hamed Hatami that a related linear-form correlation can be treated using Gowers–Cauchy–Schwarz and Fourier truncation [AM13]. Manners [Man21] and Candela, González-Sánchez, and Szegedy [CGSS22] show that the sequence of Cauchy–Schwarz steps can sometimes be organized far more efficiently than a naive complexity bound suggests.

The saving here comes from a different source. Those works exploit structure in the *system of forms*; we exploit structure in the *function*. A balanced OR is built from events of probability $p \ll 1$, and each Cauchy–Schwarz round can be arranged so that the condition that no local event occurs in one group is replaced by the rarer event that two indicators disagree. This makes $O\left(1 + \frac{\log q}{\log(1/p)}\right)$ rounds sufficient uniformly over an arbitrary set of distinct binary forms. The higher-independence extension connects this argument with the literature on limited independence. Benjamini, Gurel-Gurevich, and Peled study Boolean functions, including TRIBES, under k -wise independent inputs [BGGP12], while Gopalan and Yehudayoff obtain general product bounds under limited independence [GY20]. These results do not apply directly to the correlation considered here. Before Cauchy–Schwarz, the evaluations across forms may have arbitrary higher-order dependencies, and in the worst-case application we have only pairwise independence. The useful k -wise independent variables arise *after* the Cauchy–Schwarz transformation. At that point, it remains only to bound the parity of sparse Bernoulli indicators with total mean $\Theta(k)$. The resulting parity bound follows from the k -wise independence of the transformed indicators.

The nonlinear consequences then follow from standard pseudorandomness results applied to our small-bias bound. We use Alon–Goldreich–Mansour [AGM03] to pass from almost bounded independence to exact bounded independence and then apply Bazzi [Baz07] and Tal [Tal17]. The relevant small-bias literature begins with Naor and Naor [NN90]; local and efficiently computable constructions and linear-attack results appear in [MST03, GV04, ABR16, Bog18, AL16, GMOR15].

1.4 Scope and organization

Our focus is linear-test security and the consequences of the resulting correlation bounds. For the candidate of Ball et al., algebraic degree gives a barrier at $2^{\lambda+1}$ queries. The function has degree at most $mn = \lambda$, and an affine cube of this size gives a perfect linear relation. Our positive result covers the quasipolynomial regime in (3), leaving an open range between this guarantee and the affine-cube attack. For the Sipser

application, the limitation comes from the 2^t th root introduced by Cauchy–Schwarz. This eventually dominates the minimum dependency size available among random samples. Extending the Sipser bound further would require avoiding this loss. We return to these limitations in [Section 9](#).

[Section 2](#) gives a high-level account of the proof. [Section 3](#) fixes notation and the basic linear-algebraic facts. In [Section 4](#) we prove the rare-event decorrelation theorem, first for arbitrary systems of forms and then under the stronger assumption that every small subfamily is linearly independent. [Section 5](#) applies [Theorem 4.6](#) to the candidate of Ball et al., and [Section 5.2](#) derives the polynomial and quasipolynomial query ranges. [Section 6](#) applies the higher-independence theorem to the Sipser-based family of Boyle et al. The remaining sections develop further consequences for the candidate of Ball et al. and give the information-theoretic barriers.

2 Technical overview

We abbreviate Ball, Ducros, Erabelli, Kohl, Resch, and Scholl [[BDE⁺25](#)] as BDEKRS and Boyle, Couteau, Gilboa, Ishai, Kohl, and Scholl [[BCG⁺21](#)] as BCGIKS throughout the technical sections.

The proof rests on a correlation bound for balanced ORs of rare events evaluated on systems of binary linear forms. We first explain this bound and then its use in the two applications. The formal statements and constants begin in [Sections 3](#) and [4](#).

2.1 Rare-event decorrelation

Let $\varphi : \mathbb{F}_2^n \rightarrow \{0, 1\}$ be a local predicate with $p := \mathbb{P}[\varphi(U) = 1] \ll 1$, and consider the balanced composition

$$H(z^{(1)}, \dots, z^{(m)}) = \bigvee_{j=1}^m \varphi(z^{(j)}), \quad (5)$$

where the blocks are independent and m is chosen so that $\beta := (1 - p)^m$ is bounded away from both 0 and 1. TRIBES corresponds to $\varphi = \text{AND}_n$; the Sipser predicate used by BCGIKS has the same form after grouping its middle gates.

Suppose H is evaluated on distinct nonzero binary linear forms $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$. For uniform $Z \leftarrow V$, any two of these forms evaluate to independent unbiased bits, but higher-order independence can fail arbitrarily badly. If Q_i denotes the event that none of the m local events occurs for form i , then the output correlation, in absolute value, is the parity bias of $Q_1, \dots, Q_q$. Pairwise independence alone gives no useful direct control over this parity.

The key step is to apply Cauchy–Schwarz to groups of the local events rather than to the form system itself. Partition some of the m blocks into t groups. At each round, Cauchy–Schwarz introduces an independent copy of one group’s randomness. After expanding the square, the condition that no event occurs in that group is replaced by the event that the corresponding indicators in the two copies disagree. After t rounds,

$$|\mathcal{R}|^{2^t} \leq \mathbb{E}[(-1)^T], \quad T = \sum_{i=1}^q C_i, \quad (6)$$

where $\mathcal{R}$ is the original correlation and C_i records a disagreement in every duplicated group. If the duplicated groups have sizes $r_1, \dots, r_t$, then

$$\mathbb{P}[C_i = 1] = \beta 2^t \prod_{g=1}^t (1 - (1 - p)^{r_g}). \quad (7)$$

For a singleton group the new factor is essentially $2p$, so each Cauchy–Schwarz round makes the remaining event roughly a factor p rarer.

We take the first $t - 1$ groups to be single blocks and use the size of the last group only to tune the final mean. Then $t = O\left(1 + \frac{\log q}{\log(1/p)}\right)$ rounds suffice, with fewer rounds when the local events are rarer.

The transformed indicators preserve the same independence. For arbitrary distinct forms, the C_i remain pairwise independent. We tune $\mathbb{E}[T]$ to a small constant; then $\mathbb{P}[T = 1] \geq \mathbb{E}[T] - \mathbb{E}[T(T - 1)] = \Omega(1)$, so $\mathbb{E}[(-1)^T] \leq 1 - \Omega(1)$. Taking the 2^t th root in (6) gives a nontrivial one-block correlation bound.

If every subfamily of at most k forms is linearly independent, the C_i are instead k -wise independent. We tune $\mathbb{E}[T] = \Theta(k)$. For every $r < k$, the value of $\mathbb{E}[\binom{T}{r}]$ is then the same as in the independent case, and a truncated expansion of $(-1)^T$ gives $|\mathbb{E}[(-1)^T]| \leq \exp(-\Omega(k))$. Consequently,

$$|\mathcal{R}| \leq \exp\left(-\Omega\left(\frac{k}{2^t}\right)\right). \quad (8)$$

This is the stronger form used in the random-sample application.

2.2 The BDEKRS candidate

For one row a of a uniform matrix and one shift bit c , a query x_i to the BDEKRS candidate sees the linear form $L_i(a, c) = \langle a, x_i \rangle + c$. Its coefficient vector is $(x_i, 1)$. Distinct queries therefore give distinct nonzero forms, regardless of how many affine relations the query set contains. [Theorem 4.6](#) applies to a single TRIBES block using only the resulting pairwise independence.

There are two additional steps before this becomes a bound for the full candidate. First, the secret matrix is required to be invertible. We cannot simply multiply signed one-block correlations and then condition on invertibility, because conditioning can destroy cancellation. We instead average a one-block correlation over the shift, call the resulting quantity $z(R)$, and bound $\mathbb{E}_R[|z(R)|]$ by a second moment. Squaring introduces two independent shifts and hence the $2q$ forms $(x_i, 1, 0)$ and $(x_i, 0, 1)$, which are again all distinct. The one-block theorem controls this second moment.

Second, under a uniform matrix the ℓ row blocks are independent. Independence across these blocks raises the one-block bound on $\mathbb{E}_R[|z(R)|]$ to the ℓ th power. Only after this multiplication do we condition the matrix on invertibility, which costs a constant factor because a uniform binary matrix is invertible with constant probability. If t_{2q} denotes the number of Cauchy–Schwarz rounds for the resulting $2q$ forms, the final bound has the shape

$$\text{Bias}(x_1, \dots, x_q) \leq \exp\left(-\Omega\left(\frac{\lambda}{2^{t_{2q}}}\right)\right).$$

For the BDEKRS parameters, $p = \Theta(\log \lambda / \lambda)$, so $t_{2q} = O(1 + \log q / \log \lambda)$. This immediately gives exponential bias for every fixed polynomial number of queries and yields (3) throughout the larger quasipolynomial range.

2.3 The BCGIKS Sipser application

The stronger theorem becomes useful when the forms have no short dependency. For Q independent uniform samples in $\mathbb{F}_2^{n_\lambda}$, a fixed nonempty subset sums to zero with probability 2^{-n_λ} . A union bound therefore shows that, with overwhelming probability, every nonzero dependency has support at least $k_{\text{dep}} = \Omega\left(\frac{n_\lambda}{\log Q}\right)$. In the family (2), if the attack coefficients do not define a linear relation among the samples, the uniform secret s makes the attack exactly unbiased. If the attack is supported on a relation, then its support has size larger than k_{dep} , and every subfamily of at most k_{dep} corresponding forms is linearly independent.

The Sipser predicate is again a balanced OR of rare independent events, now with local acceptance probability $p_\lambda = \Theta(1/\lambda)$. Applying (8) with $k = k_{\text{dep}}$ gives $\exp\left(-\Omega\left(\frac{k_{\text{dep}}}{2^{t_{Q,k_{\text{dep}}}}}\right)\right)$. Here $t_{Q,k_{\text{dep}}} = O\left(1 + \frac{\log(Q/k_{\text{dep}})}{\log \lambda}\right)$. Since $n_\lambda = \Theta(\lambda^2 \log \lambda)$, the choice $Q \leq 2^{(2-\varepsilon)(\log_2 \lambda)^2}$ yields

$$\frac{k_{\text{dep}}}{2^{t_{Q,k_{\text{dep}}}}} \geq \lambda^{\varepsilon-o(1)},$$

which gives (4) simultaneously for every linear attack on the sampled points.

2.4 Further consequences

The correlation theorem for the BDEKRS candidate controls every parity of the responses on a fixed query set and can therefore be used as a small-bias statement. Two consequences follow from standard results. First, sufficiently accurate small bias gives almost bounded independence; Alon–Goldreich–Mansour [AGM03] approximate such a distribution by an exactly bounded-independent one, after which Tal’s theorem [Tal17] fools AC^0 .

The adaptive result with few affine relations uses a different idea. For a Fourier subset of the response coordinates, consider the affine relations whose supports are contained in that subset. Coordinates that occur in no such relation are independent of the remaining coordinates and contribute the very small one-query bias as separate factors. The number of possible subspaces of relations is controlled by the dimension of the full affine relation space, which allows us to sum the complete Fourier spectrum when that dimension is $O(\lambda/\log \lambda)$. A coupling between a uniform affine map and one with an invertible linear part then transfers the resulting pointwise estimate to adaptive queries.

3 Preliminaries

This section fixes the notation, security notions, and elementary facts used in the proofs. All vector spaces and linear maps are over $\mathbb{F}_2$. We write $[q] := \{1, \dots, q\}$, use $\log = \log_2$, and reserve $\ln$ for the natural logarithm. Unless otherwise specified, all expectations and probabilities are over all random choices appearing in their subscripts.

For a finite set Ω , the notation $X \leftarrow \Omega$ means that X is sampled uniformly from Ω . We write $\text{GL}_N(\mathbb{F}_2)$ for the group of invertible $N \times N$ matrices over $\mathbb{F}_2$. For a random variable X , write $\text{Law}(X)$ for its distribution. We identify a bit $b \in \{0, 1\}$ with its signed encoding $(-1)^b \in \{-1, 1\}$. We switch freely between the $\{0, 1\}$ and $\{-1, 1\}$ representations coordinatewise. This preserves total variation distance and changes an AC^0 circuit only by possible input negations.

3.1 Statistical distance and bounded-query pseudorandomness

For two distributions P, Q on a finite set Ω , their total variation distance is

$$\text{TV}(P, Q) := \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|.$$

Equivalently, $\text{TV}(P, Q) = \max_{T \subseteq \Omega} |P(T) - Q(T)|$. Thus $\text{TV}(P, Q)$ is precisely the maximum distinguishing advantage of an unbounded distinguisher receiving one sample from either P or Q .

We use the standard PRF experiment of Goldreich, Goldwasser, and Micali [GGM86], with the bounded-query restriction studied by BDEKRS [BDE⁺25]. Let $\mathcal{F} = \{F_k : \mathcal{X} \rightarrow \{0, 1\}\}_{k \in \mathcal{K}}$ be a keyed family. In the *real* experiment, a key $k \leftarrow \mathcal{K}$ is sampled and an adversary $\mathcal{A}$ receives oracle access to F_k . In the

random experiment, $\mathcal{A}$ instead receives oracle access to a uniformly random function $R : \mathcal{X} \rightarrow \{0, 1\}$. The distinguishing advantage of $\mathcal{A}$ is $\text{Adv}_{\mathcal{F}}(\mathcal{A}) := |\mathbb{P}[\mathcal{A}^{F_k} = 1] - \mathbb{P}[\mathcal{A}^R = 1]|$.

A q -query adversary makes at most q oracle queries in either experiment. It is *non-adaptive* if all queried inputs are chosen before any oracle answers are seen, and *adaptive* otherwise. Fixing its internal randomness turns an adaptive adversary into a deterministic binary decision tree of depth at most q . Its internal nodes are labeled by oracle queries, and its outgoing edges correspond to the two possible answers.

We call a statement computational when it restricts the adversary's running time or circuit size, and statistical when it allows computationally unbounded adversaries. Our main linear-test result is information-theoretic for the specified class of tests. Later, a standard Fourier argument also gives information-theoretic security against arbitrary adaptive adversaries for a smaller query range.

Repeated queries never provide additional information to a deterministic oracle algorithm. We may therefore assume, whenever convenient, that the queries appearing along any root-to-leaf path of an adaptive decision tree are distinct.

3.2 Linear tests and small bias

Following the linear-test model used by BDEKRS [BDE⁺25], let $F : \mathbb{F}_2^N \rightarrow \mathbb{F}_2$. A non-adaptive linear test chooses distinct queries $x_1, \dots, x_q \in \mathbb{F}_2^N$ and a nonzero attack vector $u = (u_1, \dots, u_q) \in \mathbb{F}_2^q$, and considers the parity $\bigoplus_{i=1}^q u_i F(x_i)$. For a uniformly random function, every such nontrivial parity is an unbiased bit.

For the keyed family considered in this paper, the bias of this test is

$$\left| \mathbb{E}_{A,b} \left[(-1)^{\sum_{i=1}^q u_i F_{A,b}(x_i)} \right] \right|$$

. Since the coordinates for which $u_i = 0$ do not contribute, every nonzero attack vector reduces to the all-ones attack on the set of queries indexed by $\text{supp}(u)$. Accordingly, for distinct $x_1, \dots, x_q \in \mathbb{F}_2^N$, we write

$$\text{Bias}(x_1, \dots, x_q) := \left| \mathbb{E}_{A,b} \left[\prod_{i=1}^q (-1)^{F_{A,b}(x_i)} \right] \right|. \quad (9)$$

A uniform bound on (9) therefore controls every non-adaptive linear test. The same statement holds for randomized non-adaptive tests by conditioning on their internal randomness.

The same notion can be stated in the standard language of small-bias distributions [NN90]. Let $Z = (Z_1, \dots, Z_q)$ be a random vector in $\{-1, 1\}^q$. We say that Z is ε -biased if, for every nonempty $S \subseteq [q]$, $|\mathbb{E} [\prod_{i \in S} Z_i]| \leq \varepsilon$. Thus our linear-test problem asks for a small-bias bound on the signed outputs of the candidate on an arbitrary chosen query set.

3.3 Fourier analysis on the Boolean cube

For every finite index set I , let U_I denote a uniform random vector in $\{-1, 1\}^I$, and abbreviate $U_{[q]}$ by U_q . We use the same symbol for its law when it appears as an argument of TV or χ^2 . For $S \subseteq [q]$, define the Fourier character $\chi_S(z) := \prod_{i \in S} z_i$, $z \in \{-1, 1\}^q$. The characters form an orthonormal basis under U_q . Hence every function $D : \{-1, 1\}^q \rightarrow \mathbb{R}$ has the expansion

$$D(z) = \sum_{S \subseteq [q]} \widehat{D}(S) \chi_S(z), \quad \widehat{D}(S) := \mathbb{E}_{U_q} [D(U_q) \chi_S(U_q)]. \quad (10)$$

Parseval's identity gives $\sum_{S \subseteq [q]} \widehat{D}(S)^2 = \mathbb{E}_{U_q} [D(U_q)^2]$.

The Fourier degree of D is $\deg_F(D) := \max\{|S| : \widehat{D}(S) \neq 0\}$. We call $\|\widehat{D}\|_1 := \sum_{S \subseteq [q]} |\widehat{D}(S)|$ the *Fourier L_1 norm* of D .¹

Let P be a distribution on $\{-1, 1\}^q$, and define its density relative to uniform by $f_P(z) := 2^q P(z)$. Then

$$\widehat{f_P}(S) = \mathbb{E}_{Z \leftarrow P} [\chi_S(Z)]. \quad (11)$$

Thus the nonconstant Fourier coefficients of f_P are the parity biases of P .

We will also use the standard χ^2 identity

$$\chi^2(P \| U_q) = \mathbb{E}_{U_q} [(f_P(U_q) - 1)^2] = \sum_{\emptyset \neq S \subseteq [q]} \widehat{f_P}(S)^2 \quad (12)$$

and the consequence

$$\text{TV}(P, U_q) \leq \frac{1}{2} \sqrt{\chi^2(P \| U_q)}. \quad (13)$$

In particular, if P is ε -biased, then $\text{TV}(P, U_q) \leq \frac{1}{2} \sqrt{2^q - 1} \varepsilon$. We will later use the same Fourier viewpoint to control distinguishers with small Fourier L_1 norm and, via decision trees, arbitrary bounded-query adaptive distinguishers.

3.4 TRIBES and the BDEKRS candidate

We next give the candidate and its parameters in the form used in the proofs. Put $M := mn$. For $z = (z_{j,k})_{j \in [m], k \in [n]} \in \mathbb{F}_2^M$, define the TRIBES function of Ben-Or and Linial [BL85] by $\text{TRIBES}_{m,n}(z) := \bigvee_{j=1}^m \bigwedge_{k=1}^n z_{j,k}$. We refer to the n coordinates $(z_{j,1}, \dots, z_{j,n})$ as the j th *tribe*.

If $Z \leftarrow \mathbb{F}_2^M$ is uniform, a fixed tribe consists entirely of ones with probability $p := 2^{-n}$. Since the m tribes depend on disjoint sets of coordinates, $\mathbb{P}[\text{TRIBES}_{m,n}(Z) = 0] = (1-p)^m$. We write $\beta := (1-p)^m$. Thus $\mathbb{P}[\text{TRIBES}_{m,n}(Z) = 1] = 1 - \beta$. The balanced regime chooses m and n so that β is close to $1/2$.

Let $N := \ell M = \ell mn$, and decompose $y \in \mathbb{F}_2^N$ as $y = (y^{(1)}, \dots, y^{(\ell)})$, where $y^{(w)} \in \mathbb{F}_2^M$. The candidate of BDEKRS [BDE⁺25] is defined in (1), with $A \leftarrow \text{GL}_N(\mathbb{F}_2)$ and $b \leftarrow \mathbb{F}_2^N$ independent. We write $\mathcal{F}_{\text{BDEKRS}} := \{F_{A,b} : A \in \text{GL}_N(\mathbb{F}_2), b \in \mathbb{F}_2^N\}$ for this keyed family.

For the correlation calculations, we pass to the $\{-1, 1\}$ representation. Define $h(z) := (-1)^{\text{TRIBES}_{m,n}(z)}$ and, for $y = (y^{(1)}, \dots, y^{(\ell)}) \in (\mathbb{F}_2^M)^\ell$, define $g(y) := \prod_{w=1}^\ell h(y^{(w)})$. Since XOR becomes multiplication after applying $t \mapsto (-1)^t$, $(-1)^{F_{A,b}(x)} = g(Ax \oplus b)$.

For the balanced parameters considered by BDEKRS [BDE⁺25] and throughout our main asymptotic consequences, $\ell = \lambda$, $M = mn = \lambda$, and $N = \lambda^2$, with $n = \log_2 \lambda - \log_2 \ln \lambda + O(1)$ and $m = \ln(2) 2^n + O(1)$. Consequently, $p = \Theta(\log \lambda / \lambda)$, $\beta = 1/2 + o(1)$, and $m = \Theta(\lambda / \log \lambda)$. The key is specified by A and b , so its description length is $\kappa = N^2 + N = \lambda^4 + \lambda^2 = \Theta(\lambda^4)$.

For $s \geq 1$, we write t_s for the least integer $t \geq 1$ such that

$$s\beta(2p)^t \leq \frac{1}{4}. \quad (14)$$

Thus t_s is the number of Cauchy–Schwarz rounds used below for a system of s forms.²

¹In Boolean-function analysis this quantity is also commonly called the *spectral norm*. We use “Fourier L_1 norm” to avoid confusion with matrix spectral norms.

²The threshold $1/4$ is not intrinsic. Any fixed $\tau \in (0, 1/3)$ could be used here, with corresponding changes to the absolute constants below; we choose $1/4$ for convenience in the tuning argument.

3.5 Linear forms over $\mathbb{F}_2$

The proof uses the following elementary property of linear forms over $\mathbb{F}_2$. It is the reason no higher-order independence assumption on the queries is needed.

Lemma 3.1 (Distinct linear forms are pairwise independent). *Let V be a finite-dimensional $\mathbb{F}_2$ -vector space, and let $L, L' : V \rightarrow \mathbb{F}_2$ be distinct nonzero linear forms. If $Z \leftarrow V$ is uniform, then $(L(Z), L'(Z))$ is uniform on $\mathbb{F}_2^2$. In particular, $L(Z)$ and $L'(Z)$ are independent unbiased bits.*

Proof. Over $\mathbb{F}_2$, distinct nonzero linear forms are linearly independent. Hence the map

$$z \mapsto (L(z), L'(z))$$

has rank 2 and is therefore surjective onto $\mathbb{F}_2^2$. Every point of $\mathbb{F}_2^2$ has the same number of preimages, so a uniform input induces the uniform distribution on $\mathbb{F}_2^2$. $\square$

More generally, if $L_1, \dots, L_r$ are linearly independent forms, then $(L_1(Z), \dots, L_r(Z))$ is uniform on $\mathbb{F}_2^r$. For an arbitrary collection of distinct forms, we only have pairwise linear independence in general. The family $L_1(Z), \dots, L_q(Z)$ need not be jointly independent when the forms satisfy higher-order linear relations.

We will repeatedly distinguish arbitrary systems of forms from systems with no short linear dependency. For nonzero linear forms $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$, we say that there is *no linear dependency of size at most k* if

$$\sum_{i \in I} L_i \neq 0 \quad \text{for every nonempty } I \subseteq [q] \text{ with } |I| \leq k.$$

Equivalently, every subfamily of at most k forms is linearly independent. Over $\mathbb{F}_2$, a relation is determined by its support, so no coefficients are lost in this formulation.³

We use the analogous condition for vectors. In particular, if $L_i(a) = \langle a, x_i \rangle$ for nonzero $x_i \in \mathbb{F}_2^n$, then the forms have no dependency of size at most k exactly when

$$\sum_{i \in I} x_i \neq 0 \quad \text{for every nonempty } I \subseteq [q] \text{ with } |I| \leq k.$$

Affine expressions arising from the secret shift can be incorporated by adding one coordinate. In particular, for fixed $x \in \mathbb{F}_2^N$, the expression $\langle a, x \rangle + c$, $(a, c) \in \mathbb{F}_2^{N+1}$, is the linear form whose coefficient vector is $(x, 1)$. Hence, for distinct $x_1, \dots, x_q$, the forms $(a, c) \mapsto \langle a, x_i \rangle + c$ are distinct and nonzero, even if one of the queries is $x_i = 0$. By [Lemma 3.1](#), their evaluations are therefore pairwise independent under a uniform choice of (a, c) .

For a nonempty tuple $X = (x_1, \dots, x_q)$, we write

$$d_{\text{aff}}(X) := \dim \text{span}\{(x_1, 1), \dots, (x_q, 1)\} - 1$$

for its affine dimension, and $\nu(X) := q - d_{\text{aff}}(X) - 1$ for the dimension of its affine relation space. Thus $\nu(X)$ is simply the number of independent affine relations among the augmented query vectors.

This elementary observation is the only independence property of the query set that our one-block argument requires. In particular, we impose no assumption on the rank of $\{x_1, \dots, x_q\}$ or on their higher-order affine dependencies.

³In matroid terminology, the minimum support size of a nonzero dependency is the *girth*. Equivalently, it is the minimum distance of the binary linear code consisting of coefficient vectors $c \in \mathbb{F}_2^q$ such that $\sum_i c_i L_i = 0$. We use the explicit “no short dependency” formulation in the paper.

4 Rare-event decorrelation for arbitrary linear forms

The proof below applies to balanced ORs of rare local events, rather than only to AND gates. We first introduce the general notation, then prove the concrete TRIBES statement needed for BDEKRS before giving the general theorem.

For a Boolean predicate $\varphi : \mathbb{F}_2^n \rightarrow \{0, 1\}$, let $U_n \leftarrow \mathbb{F}_2^n$ be uniform and put $p_\varphi := \mathbb{P}[\varphi(U_n) = 1]$ and $\beta_\varphi := (1 - p_\varphi)^m$, and define

$$H_{\varphi,m}(z^{(1)}, \dots, z^{(m)}) := \bigvee_{j=1}^m \varphi(z^{(j)}), \quad h_{\varphi,m} := (-1)^{H_{\varphi,m}}.$$

In this general setting, “balanced” means only that β_φ is bounded away from 0 and 1; the BDEKRS tuning has the stronger property $\beta = 1/2 + o(1)$. For $s \geq 1$, let $t_s(\varphi)$ be the least integer $t \geq 1$ satisfying $s\beta_\varphi(2p_\varphi)^t \leq \frac{1}{4}$. For $\varphi = \text{AND}_n$, this agrees with the BDEKRS parameter t_s defined in (14).

We first prove the concrete TRIBES statement needed for the BDEKRS application. The proof uses only the probability of one local event, independence across blocks, and pairwise independence across forms. We then give the corresponding statement for an arbitrary local predicate φ ; that general form is what we use for the Sipser application.

Thus we begin with a single TRIBES block evaluated on an arbitrary system of distinct linear forms. Here we use only the pairwise independence from Lemma 3.1. The outer XOR is handled later.

4.1 The one-block correlation problem

Let V be a finite-dimensional $\mathbb{F}_2$ -vector space, and let $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$ be distinct nonzero linear forms. For every tribe $j \in [m]$ and coordinate $k \in [n]$, sample an independent uniform vector $Z_{j,k} \leftarrow V$. For $i \in [q]$, define

$$Y_i := (L_i(Z_{j,k}))_{j \in [m], k \in [n]} \in \mathbb{F}_2^M$$

and $\mathcal{R}(L_1, \dots, L_q) := \mathbb{E}[\prod_{i=1}^q h(Y_i)]$.

We prove the following estimate. Here $t_q = t_q(\text{AND}_n)$ as above. Qualitatively, the argument only needs β bounded away from 0 and 1, p and tp sufficiently small, and t bounded away from m .

Proposition 4.1 (One-block decorrelation). *Assume $\frac{1}{3} \leq \beta \leq \frac{2}{3}$ and $p \leq \frac{1}{8}$. Let $t = t_q$ and suppose $tp \leq \frac{1}{10}$ and $t \leq \frac{m}{2}$. Then for arbitrary distinct nonzero linear forms $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$,*

$$|\mathcal{R}(L_1, \dots, L_q)| \leq \eta^{1/2^{t_q}}, \quad \eta := \frac{15}{16}.$$

The value $\eta = 15/16$ is not optimized; only an absolute gap below 1 is used in the subsequent asymptotic bounds.

The proof has two stages. We first use repeated Cauchy–Schwarz to reduce the original correlation to the parity of much rarer events while preserving the pairwise independence available from the linear forms. We then choose the number of Cauchy–Schwarz rounds and the tribe groups so that the expected number of those rare events is a small constant, at which point a second-moment argument bounds the resulting parity correlation by a constant below 1.

By Lemma 3.1, for $i \neq i'$ and uniform $Z \leftarrow V$, the pair $(L_i(Z), L_{i'}(Z))$ is uniform on $\mathbb{F}_2^2$.

For tribe j and form i , let $B_{j,i} := \prod_{k=1}^n L_i(Z_{j,k}) \in \{0, 1\}$. Then $\mathbb{P}[B_{j,i} = 1] = p$, and for $i \neq i'$, the variables $B_{j,i}$ and $B_{j,i'}$ are independent. Different tribes are mutually independent. Thus $B_{j,i}$ is the indicator that tribe j evaluates to 1 for form i .

Define $Q_i := \prod_{j=1}^m (1 - B_{j,i})$. Thus $Q_i = 1$ iff no tribe evaluates to 1 on Y_i , and $h(Y_i) = (-1)^{1-Q_i}$. Taking absolute values gives

$$|\mathcal{R}(L_1, \dots, L_q)| = \left| \mathbb{E} \left[(-1)^{\sum_{i=1}^q Q_i} \right] \right|. \quad (15)$$

4.2 Repeated Cauchy–Schwarz

For an arbitrary system of forms, the variables $Q_1, \dots, Q_q$ are not guaranteed to be jointly independent, so pairwise independence alone does not control the parity in (15). We instead expose the tribe randomness in groups and apply Cauchy–Schwarz successively to those groups. Each application introduces an independent copy of one group’s randomness, and after expanding the square the corresponding factor becomes the XOR of the two copies. After t rounds, the original correlation is bounded by the parity correlation of a family of much rarer events.

Lemma 4.2 (Iterated Cauchy–Schwarz). *Let $Y_0, Y_1, \dots, Y_t$ be mutually independent random variables. For each $i \in [q]$, let $R_i = R_i(Y_0) \in \{0, 1\}$ and $U_{g,i} = U_{g,i}(Y_g) \in \{0, 1\}$ for $g \in [t]$. Set*

$$\mathcal{R} := \mathbb{E} \left[(-1)^{\sum_{i=1}^q R_i \prod_{g=1}^t U_{g,i}} \right].$$

For every g , let $Y_g^{(0)}, Y_g^{(1)}$ be independent copies of Y_g and write $\Delta U_{g,i} := U_{g,i}(Y_g^{(0)}) \oplus U_{g,i}(Y_g^{(1)})$. Then

$$|\mathcal{R}|^{2^t} \leq \mathbb{E} \left[(-1)^{\sum_{i=1}^q R_i \prod_{g=1}^t \Delta U_{g,i}} \right].$$

The expectation on the right is nonnegative.

Proof. For $s \in \{0, \dots, t\}$ define

$$\Gamma_s := \mathbb{E} \left[(-1)^{\sum_i R_i (\prod_{g=1}^s \Delta U_{g,i}) (\prod_{g=s+1}^t U_{g,i})} \right].$$

Thus $\Gamma_0 = \mathcal{R}$.

We prove inductively that for every $s \geq 1$, $|\mathcal{R}|^{2^s} \leq \Gamma_s$ and $\Gamma_s \geq 0$. For $s = 1$, condition on all variables except Y_1 . If $\Phi_0(X, Y_1) = (-1)^{\sum_i R_i U_{1,i} \prod_{g=2}^t U_{g,i}}$, then

$$|\mathcal{R}|^2 \leq \mathbb{E}_X \left[\left| \mathbb{E}_{Y_1} [\Phi_0(X, Y_1)] \right|^2 \right] = \Gamma_1.$$

The middle quantity is an expectation of a squared absolute value, so $\Gamma_1 \geq 0$.

Now suppose the claim holds for $s < t$. Condition in Γ_s on all variables except the single uncopied variable Y_{s+1} , and call the resulting sign $\Phi_s(X, Y_{s+1})$. By the induction hypothesis and $\Gamma_s \geq 0$,

$$|\mathcal{R}|^{2^{s+1}} \leq \Gamma_s^2.$$

Applying Cauchy–Schwarz to the conditional expectation defining Γ_s gives

$$\Gamma_s^2 \leq \mathbb{E}_X \left[\left| \mathbb{E}_{Y_{s+1}} [\Phi_s(X, Y_{s+1})] \right|^2 \right].$$

Expanding the square with two independent copies $Y_{s+1}^{(0)}, Y_{s+1}^{(1)}$ replaces $U_{s+1,i}(Y_{s+1}^{(0)}) + U_{s+1,i}(Y_{s+1}^{(1)})$ modulo 2 by $\Delta U_{s+1,i}$, while all previously duplicated groups remain outside the inner conditional expectation. The resulting expression is Γ_{s+1} , and it is nonnegative because it is an expectation of a square. Taking $s = t$ proves the lemma. $\square$

Partition some of the m tribes into disjoint nonempty groups $G_1, \dots, G_t$, and let the remaining tribes form G_0 , which may be empty. For form i , define

$$U_{g,i} := \prod_{j \in G_g} (1 - B_{j,i}), \quad R_i := \prod_{j \in G_0} (1 - B_{j,i}).$$

Then $Q_i = R_i \prod_{g=1}^t U_{g,i}$. Applying [Lemma 4.2](#),

$$|\mathcal{R}(L_1, \dots, L_q)|^{2^t} \leq \mathbb{E}[(-1)^T], \quad T := \sum_{i=1}^q C_i, \quad (16)$$

where $C_i := R_i \prod_{g=1}^t (U_{g,i}^{(0)} \oplus U_{g,i}^{(1)})$. Thus $C_i = 1$ precisely when no tribe in G_0 evaluates to 1 for form i , and in every duplicated group G_g , one copy has no tribe evaluating to 1 while the other does. The next lemma computes their probability and shows that they retain the pairwise independence needed for the second-moment step.

Lemma 4.3 (Pairwise-independent rare events). *Let $r_g := |G_g|$ and $r_0 := |G_0|$. Then every C_i is Bernoulli with*

$$\alpha := \mathbb{P}[C_i = 1] = \beta 2^t \prod_{g=1}^t (1 - (1-p)^{r_g}). \quad (17)$$

For every $i \neq i'$, the variables C_i and $C_{i'}$ are independent.

Proof. Put $a_0 := (1-p)^{r_0}$ and $a_g := (1-p)^{r_g}$. For fixed i , the group G_0 and the duplicated groups depend on disjoint sets of tribe samples. Hence

$$\mathbb{P}[R_i = 1] = a_0, \quad \mathbb{P}[U_{g,i}^{(0)} \oplus U_{g,i}^{(1)} = 1] = 2a_g(1 - a_g).$$

Therefore $\mathbb{P}[C_i = 1] = a_0 \prod_{g=1}^t 2a_g(1 - a_g)$. Since $a_0 \prod_{g=1}^t a_g = (1-p)^m = \beta$, this is (17).

Fix $i \neq i'$. Let $\mathcal{Z}$ be the complete list of independent vector samples present after the Cauchy-Schwarz duplication: one sample for every coordinate in G_0 and two independent samples for every coordinate in a duplicated group. Define

$$V_i := (L_i(Z) : Z \in \mathcal{Z}), \quad V_{i'} := (L_{i'}(Z) : Z \in \mathcal{Z}).$$

For each individual $Z \in \mathcal{Z}$, the pair $(L_i(Z), L_{i'}(Z))$ is uniform on $\mathbb{F}_2^2$. Since the samples in $\mathcal{Z}$ are mutually independent, the joint law factors coordinatewise, so V_i and $V_{i'}$ are independent. The variable C_i is a deterministic function only of V_i , and $C_{i'}$ only of $V_{i'}$. Hence C_i and $C_{i'}$ are independent. $\square$

Lemma 4.4 (Constant mean bounds the parity correlation away from 1). *Suppose $\frac{1}{16} \leq \mu := q\alpha \leq \frac{1}{4}$. Then*

$$|\mathcal{R}(L_1, \dots, L_q)|^{2^t} \leq \eta, \quad \eta := \frac{15}{16}.$$

Consequently, $|\mathcal{R}(L_1, \dots, L_q)| \leq \eta^{1/2^t}$.

Proof. By [Lemma 4.3](#),

$$\mathbb{E}[T] = \mu, \quad \mathbb{E}[T(T-1)] = q(q-1)\alpha^2 \leq \mu^2.$$

Since $r \leq r(r-1)$ for every integer $r \geq 2$,

$$\mathbb{P}[T = 1] = \mathbb{E}[T] - \mathbb{E}[T\mathbf{1}_{\{T \geq 2\}}] \geq \mu - \mathbb{E}[T(T-1)] \geq \mu - \mu^2.$$

On the interval $\mu \in [1/16, 1/4]$, the function $\mu - \mu^2$ is increasing, and therefore

$$\mathbb{P}[T = 1] \geq \frac{1}{16} - \frac{1}{256} = \frac{15}{256} > \frac{1}{32}.$$

Hence $\mathbb{P}[T \text{ odd}] > 1/32$ and

$$\mathbb{E}[(-1)^T] = 1 - 2\mathbb{P}[T \text{ odd}] < 1 - \frac{1}{16} = \eta.$$

The expectation is nonnegative by [Lemma 4.2](#) and [Eq. \(16\)](#). $\square$

4.3 Tuning the tribe groups

It remains to choose the duplicated tribe groups so that the rare-event count lies in the constant range required by [Lemma 4.4](#). We set the number of Cauchy–Schwarz rounds to t_q and choose the size of the final tribe group so that the expected number of rare events lies in a fixed constant interval.

Lemma 4.5 (Group-size tuning). *Assume $\frac{1}{3} \leq \beta \leq \frac{2}{3}$ and $p \leq \frac{1}{8}$. Let $t = t_q$ be defined by [\(14\)](#), and suppose $tp \leq \frac{1}{10}$ and $t \leq \frac{m}{2}$. Then one can choose disjoint nonempty groups $G_1, \dots, G_t$ so that the resulting mean $\mu = q\alpha$ satisfies $\frac{1}{16} \leq \mu \leq \frac{1}{4}$.*

Proof. We keep the first $t - 1$ groups as single tribes and vary only the size r of the final group.

Take $r_1 = \dots = r_{t-1} = 1$ and let $r_t = r$ vary from 1 to $r_{\max} := m - t + 1$. The remaining tribes form G_0 . By [\(17\)](#),

$$\mu(r) = q\beta 2^t p^{t-1} (1 - (1 - p)^r).$$

At $r = 1$, $\mu(1) = q\beta(2p)^t \leq 1/4$.

Minimality of t gives $q\beta(2p)^{t-1} > 1/4$. For $t = 1$, the same inequality follows from $q\beta \geq 1/3 > 1/4$. We also have $1 - (1 - p)^{r_{\max}} = 1 - \beta(1 - p)^{-(t-1)}$. Since $p \leq 1/8$, $-\ln(1 - p) \leq p/(1 - p) < 2p$, and therefore, using $tp \leq 1/10$,

$$(1 - p)^{-(t-1)} \leq e^{2pt} \leq e^{1/5} < \frac{5}{4}.$$

Together with $\beta \leq 2/3$, this gives $1 - (1 - p)^{r_{\max}} > 1/6$. Hence

$$\mu(r_{\max}) > \frac{1}{3} q\beta(2p)^{t-1} > \frac{1}{12} > \frac{1}{16}.$$

If $\mu(1) \geq 1/16$, we are done. Otherwise, let r be the least integer with $\mu(r) \geq 1/16$. Put $s(r) := 1 - (1 - p)^r$. Since $s(r) \geq p$,

$$s(r + 1) = s(r) + p(1 - p)^r \leq s(r) + p \leq 2s(r).$$

Thus $\mu(r) \leq 2\mu(r - 1) < 1/8 < 1/4$. $\square$

Proof of [Proposition 4.1](#). Choose the groups supplied by [Lemma 4.5](#) with $t = t_q$, and apply [Lemma 4.4](#). $\square$

The preceding proof uses φ only through p_φ . We now give the general form needed later.

Theorem 4.6 (Rare-event decorrelation). *Assume $\frac{1}{3} \leq \beta_\varphi \leq \frac{2}{3}$ and $p_\varphi \leq \frac{1}{8}$. Let $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$ be distinct nonzero linear forms, let $t = t_q(\varphi)$, and suppose $tp_\varphi \leq \frac{1}{10}$ and $t \leq \frac{m}{2}$. For every $j \in [m]$ and $k \in [n]$, sample an independent uniform $Z_{j,k} \leftarrow V$, and for $i \in [q]$ write*

$$Y_i := ((L_i(Z_{j,1}), \dots, L_i(Z_{j,n})))_{j \in [m]}.$$

Then

$$\left| \mathbb{E} \left[\prod_{i=1}^q h_{\varphi,m}(Y_i) \right] \right| \leq \eta^{1/2^t}, \quad \eta := \frac{15}{16}.$$

Proof of Theorem 4.6. For block j and form i , define

$$B_{j,i} := \varphi(L_i(Z_{j,1}), \dots, L_i(Z_{j,n})),$$

the indicator that the j th local event occurs for form i . For $i \neq i'$, the two n -bit vectors

$$(L_i(Z_{j,1}), \dots, L_i(Z_{j,n})), \quad (L_{i'}(Z_{j,1}), \dots, L_{i'}(Z_{j,n}))$$

are independent and uniform on $\mathbb{F}_2^n$. This follows coordinatewise from Lemma 3.1 and from independence of the $Z_{j,k}$. Hence $B_{j,i}$ and $B_{j,i'}$ are independent Bernoulli- p_φ variables. Different values of j use independent randomness.

The proof of Proposition 4.1 uses no further property of AND_n . After the indicators $B_{j,i}$ are defined, it uses only their Bernoulli parameter, independence across blocks, and pairwise independence across distinct forms. Substituting $p_\varphi, \beta_\varphi, t_q(\varphi)$ for p, β, t_q therefore gives the stated bound. $\square$

4.4 Higher-order independence gives stronger cancellation

For worst-case query sets we only have pairwise independence. If more independence is available, the Cauchy–Schwarz step is unchanged; only the final parity estimate improves. Instead of tuning the transformed rare-event count to have constant mean, we tune its mean to be proportional to the number of independent forms.

We first show that the transformed rare events inherit the independence order of the original forms.

Lemma 4.7 (Rare events inherit k -wise independence). *Consider the repeated Cauchy–Schwarz construction in the proof of Theorem 4.6, for an arbitrary local predicate φ . Suppose that every subfamily of at most k of $L_1, \dots, L_q$ is linearly independent. Then the transformed indicators $C_1, \dots, C_q$ are k -wise independent. Each has the Bernoulli distribution with parameter*

$$\alpha = \beta_\varphi 2^t \prod_{g=1}^t (1 - (1 - p_\varphi)^{r_g}).$$

Proof. Fix $I \subseteq [q]$ with $|I| \leq k$. The forms $(L_i)_{i \in I}$ are linearly independent. Hence, for every uniform vector Z used anywhere in the construction, $(L_i(Z))_{i \in I}$ is uniform on $\mathbb{F}_2^I$. The vectors Z used for different tribe coordinates and for the independent copies created by Cauchy–Schwarz are mutually independent. It follows that the complete evaluation vectors

$$V_i := (L_i(Z) : Z \in \mathcal{Z}), \quad i \in I,$$

are mutually independent. Each C_i is a deterministic function of V_i . Thus $(C_i)_{i \in I}$ are mutually independent. Their common marginal is the Bernoulli- α distribution computed in Lemma 4.3. $\square$

The next lemma gives the needed parity bound.

Lemma 4.8 (Parity of a k -wise independent Bernoulli sum). *Let $X_1, \dots, X_q$ be k -wise independent Bernoulli- α random variables, let $T := \sum_{i=1}^q X_i$ and $\mu := q\alpha$, and assume $4\mu \leq k$. Then*

$$|\mathbb{E}[(-1)^T] - (1 - 2\alpha)^q| \leq 4 \left(\frac{3e\mu}{k} \right)^k. \quad (18)$$

Consequently, there are absolute constants $c > 0$ and $k_0 \in \mathbb{N}$ such that, whenever $k_0 \leq k < q$ and

$$\frac{k}{36e} \leq \mu \leq \frac{k}{12e}, \quad (19)$$

we have

$$|\mathbb{E}[(-1)^T]| \leq \exp(-ck).$$

Proof. Define the degree- $(k-1)$ polynomial on nonnegative integers $P_{k-1}(x) := \sum_{r=0}^{k-1} (-2)^r \binom{x}{r}$. For every integer $0 \leq x < k$, the binomial theorem gives $P_{k-1}(x) = (1-2)^x = (-1)^x$.

We first bound the error on the event $T \geq k$. For integers $x \geq k$ and $r < k$,

$$\binom{x}{r} \leq \binom{k}{r} \binom{x}{k}.$$

Indeed, $\binom{x}{k} \binom{k}{r} = \binom{x}{r} \binom{x-r}{k-r}$. Therefore

$$|(-1)^x - P_{k-1}(x)| \leq 2 \cdot 3^k \binom{x}{k}.$$

Taking expectations and using k -wise independence, $\mathbb{E} \left[\binom{T}{k} \right] = \binom{q}{k} \alpha^k$, so

$$|\mathbb{E}[(-1)^T] - \mathbb{E}[P_{k-1}(T)]| \leq 2 \cdot 3^k \binom{q}{k} \alpha^k \leq 2 \left(\frac{3e\mu}{k} \right)^k. \quad (20)$$

For every $r < k$, k -wise independence also gives $\mathbb{E} \left[\binom{T}{r} \right] = \binom{q}{r} \alpha^r$. Hence

$$\mathbb{E}[P_{k-1}(T)] = \sum_{r=0}^{k-1} (-2)^r \binom{q}{r} \alpha^r,$$

which is the first k terms in the expansion of $(1-2\alpha)^q$. If $4\mu \leq k$, then for every $r \geq k$ the absolute values of consecutive terms in the omitted tail have ratio at most

$$\frac{q-r}{r+1} 2\alpha \leq \frac{2\mu}{k} \leq \frac{1}{2}.$$

Thus

$$|(1-2\alpha)^q - \mathbb{E}[P_{k-1}(T)]| \leq 2 \binom{q}{k} (2\alpha)^k \leq 2 \left(\frac{2e\mu}{k} \right)^k. \quad (21)$$

Combining (20) and (21) proves (18).

Now assume (19). Since $k < q$ and $\mu \leq k/(12e)$, we have $\alpha = \mu/q < 1/(12e) < 1/2$. Hence

$$|(1-2\alpha)^q| \leq e^{-2\alpha q} = e^{-2\mu} \leq e^{-k/(18e)}.$$

Moreover,

$$4 \left(\frac{3e\mu}{k} \right)^k \leq 4^{1-k}.$$

For all $k \geq k_0$ these two terms are at most e^{-ck} for a suitable absolute $c > 0$, after decreasing c if necessary. $\square$

We now tune the duplicated groups to the scale k . Put $c_0 := \frac{1}{12e}$. We choose this constant so that $3e\mu/k \leq 1/4$ in Lemma 4.8, while leaving slack for the group-tuning step. Any sufficiently small positive constant would work. For $1 \leq k < q$, define $t_{q,k}(\varphi)$ to be the least integer $t \geq 1$ such that

$$q\beta_\varphi(2p_\varphi)^t \leq c_0 k. \quad (22)$$

When φ is clear from context we write $t_{q,k}$.

Lemma 4.9 (Group tuning at scale k). Assume $\frac{1}{3} \leq \beta_\varphi \leq \frac{2}{3}$ and $p_\varphi \leq \frac{1}{8}$. Let $1 \leq k < q$, let $t = t_{q,k}(\varphi)$, and suppose $tp_\varphi \leq \frac{1}{10}$ and $t \leq \frac{m}{2}$. Then one can choose disjoint nonempty groups $G_1, \dots, G_t$ so that the mean $\mu = q\alpha$ of the transformed rare-event count satisfies

$$\frac{k}{36e} \leq \mu \leq \frac{k}{12e}.$$

Proof. Use the same one-parameter choice as in Lemma 4.5. The first $t - 1$ groups are single blocks and the last has size $r \in [1, m - t + 1]$. Then

$$\mu(r) = q\beta_\varphi 2^t p_\varphi^{t-1} (1 - (1 - p_\varphi)^r).$$

The definition of t gives $\mu(1) \leq c_0 k$, while the estimate at $r = m - t + 1$ from Lemma 4.5, together with minimality of t (or $q > k$ when $t = 1$), gives $\mu(m - t + 1) > c_0 k/3$. If $\mu(1) \geq c_0 k/3$, take $r = 1$. Otherwise take the first $r > 1$ with $\mu(r) \geq c_0 k/3$. Consecutive values differ by at most a factor of two, by the same estimate $1 - (1 - p_\varphi)^{r+1} \leq 2(1 - (1 - p_\varphi)^r)$, so in the second case $\mu(r) < 2c_0 k/3$. In either case

$$\frac{c_0 k}{3} \leq \mu(r) \leq c_0 k.$$

Since $c_0 = 1/(12e)$, this is the required interval. $\square$

Theorem 4.10 (Rare-event decorrelation with no short dependencies). There are absolute constants $c > 0$ and $k_0 \in \mathbb{N}$ with the following property. Assume $\frac{1}{3} \leq \beta_\varphi \leq \frac{2}{3}$ and $p_\varphi \leq \frac{1}{8}$. Let $L_1, \dots, L_q : V \rightarrow \mathbb{F}_2$ be distinct nonzero linear forms such that every subfamily of at most k forms is linearly independent, where $k_0 \leq k < q$. Let $t = t_{q,k}(\varphi)$ be defined by (22), and suppose $tp_\varphi \leq \frac{1}{10}$ and $t \leq \frac{m}{2}$. With Y_i and $h_{\varphi,m}$ as in Theorem 4.6,

$$\left| \mathbb{E} \left[\prod_{i=1}^q h_{\varphi,m}(Y_i) \right] \right| \leq \exp \left(-\frac{ck}{2^t} \right). \quad (23)$$

Proof. Choose groups using Lemma 4.9 and perform the same Cauchy–Schwarz iteration as in Lemma 4.2. Writing $\mathcal{R} := \mathbb{E} [\prod_{i=1}^q h_{\varphi,m}(Y_i)]$, we obtain

$$|\mathcal{R}|^{2^t} \leq \mathbb{E}[(-1)^T], \quad T = \sum_{i=1}^q C_i.$$

By Lemma 4.7, the C_i are k -wise independent Bernoulli- α variables. Their mean lies in the interval (19), so Lemma 4.8 gives

$$|\mathbb{E}[(-1)^T]| \leq e^{-ck}.$$

The expectation on the right of the Cauchy–Schwarz inequality is nonnegative, by Lemma 4.2. Taking the 2^t th root proves (23). $\square$

From (22),

$$t_{q,k}(\varphi) \leq 1 + \frac{\log(q/k) + \log(1/c_0) + O(1)}{\log(1/(2p_\varphi))}. \quad (24)$$

We will use this estimate with $p_\varphi = \Theta(1/\lambda)$ and with k coming from the minimum size of a nonzero dependency among the random samples.

5 Application to the BDEKRS candidate

We now apply [Theorem 4.6](#) to the candidate $\mathcal{F}_{\text{BDEKRS}}$. The exact correlation bound is the following.

Theorem 5.1 (General non-adaptive linear tests). *Assume $\frac{1}{3} \leq \beta \leq \frac{2}{3}$ and $p \leq \frac{1}{8}$. Let $x_1, \dots, x_q \in \mathbb{F}_2^N$ be distinct and suppose $t_{2q}p \leq \frac{1}{10}$ and $t_{2q} \leq \frac{m}{2}$. Then*

$$\left| \mathbb{E}_{\substack{A \leftarrow \text{GL}_N(\mathbb{F}_2) \\ b \leftarrow \mathbb{F}_2^N}} \left[(-1)^{\sum_{i=1}^q F_{A,b}(x_i)} \right] \right| \leq 4\eta^{\ell/2^{t_{2q}+1}}, \quad \eta := \frac{15}{16}. \quad (25)$$

In particular,

$$\text{Bias}(x_1, \dots, x_q) \leq 4 \exp\left(-\frac{\ell}{2^{t_{2q}+5}}\right).$$

5.1 From one block to the full candidate

The one-block estimate of [Proposition 4.1](#) assumes independent uniform rows. To recover the full candidate we must both combine ℓ blocks and restore the requirement that the secret matrix be invertible. Under a uniform matrix the row blocks are independent, so their bounds multiply, but conditioning a signed expectation on invertibility can destroy cancellation. We therefore first use a second moment to bound $\mathbb{E}_R[|z(R)|]$, multiply this bound across independent blocks, and only then condition on invertibility.

Fix distinct PRF queries $x_1, \dots, x_q \in \mathbb{F}_2^N$. For $R \in \mathbb{F}_2^{M \times N}$ define

$$z(R) := \mathbb{E}_{c \leftarrow \mathbb{F}_2^M} \left[\prod_{i=1}^q h(Rx_i \oplus c) \right].$$

Lemma 5.2 (Second moment). *Under the assumptions of [Theorem 5.1](#), for uniform $R \leftarrow \mathbb{F}_2^{M \times N}$,*

$$\mathbb{E}_R[z(R)^2] \leq \eta^{1/2^{t_{2q}}}.$$

Consequently,

$$\mathbb{E}_R[|z(R)|] \leq \eta^{1/2^{t_{2q}+1}}.$$

Proof. Let $c, c' \leftarrow \mathbb{F}_2^M$ be independent. Expanding the square,

$$\mathbb{E}_R[z(R)^2] = \mathbb{E}_{R, c, c'} \left[\prod_{i=1}^q h(Rx_i \oplus c) h(Rx_i \oplus c') \right].$$

For one row $a \in \mathbb{F}_2^N$ of R and corresponding shift coordinates $\sigma, \sigma' \in \mathbb{F}_2$, define the $2q$ forms on $\mathbb{F}_2^{N+2}$ by

$$L_i^{(0)}(a, \sigma, \sigma') = \langle a, x_i \rangle + \sigma, \quad L_i^{(1)}(a, \sigma, \sigma') = \langle a, x_i \rangle + \sigma'.$$

Their coefficient vectors are $(x_i, 1, 0)$ and $(x_i, 0, 1)$. They are all nonzero, including when $x_i = 0$, and they are pairwise distinct. Equal superscripts use the distinctness of the x_i , while different superscripts differ in the last two coordinates.

To identify the preceding expectation with the one-block problem, index the $M = mn$ rows of R by pairs $(j, k) \in [m] \times [n]$. If $a_{j,k}$ is the corresponding row and $c_{j,k}, c'_{j,k}$ are the corresponding shift coordinates, then $Z_{j,k} := (a_{j,k}, c_{j,k}, c'_{j,k})$ is uniform in $\mathbb{F}_2^{N+2}$, and the $Z_{j,k}$ are independent across (j, k) .

Evaluating the $2q$ forms $L_i^{(0)}, L_i^{(1)}$ on these vectors produces the coordinate bits appearing in $h(Rx_i \oplus c)$ and $h(Rx_i \oplus c')$. Thus the second-moment expression is precisely the one-block correlation of these $2q$ forms.

Applying [Proposition 4.1](#) therefore gives $\mathbb{E}_R[z(R)^2] \leq \eta^{1/2^{t_{2q}}}$. Finally,

$$\mathbb{E}_R[|z(R)|] \leq \sqrt{\mathbb{E}_R[z(R)^2]} \leq \eta^{1/2^{t_{2q}+1}}.$$

□

Lemma 5.3 (Invertibility has constant probability). *If $\tilde{A} \leftarrow \mathbb{F}_2^{N \times N}$ is uniform, then*

$$\mathbb{P}[\tilde{A} \in \text{GL}_N(\mathbb{F}_2)] = \prod_{j=1}^N (1 - 2^{-j}) \geq \frac{1}{4}.$$

Proof. Exposing the columns successively gives the product formula. Moreover,

$$\prod_{j=2}^N (1 - 2^{-j}) \geq 1 - \sum_{j=2}^N 2^{-j} \geq \frac{1}{2}.$$

The $j = 1$ factor is $1/2$, so the full product is at least $1/4$.

□

Proof of Theorem 5.1. Write a matrix $A \in \mathbb{F}_2^{N \times N}$ in ℓ row blocks

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_\ell \end{pmatrix}, \quad A_w \in \mathbb{F}_2^{M \times N}.$$

For fixed A , averaging over the independent shift blocks $b_1, \dots, b_\ell$ gives

$$\mathbb{E}_b \left[\prod_{i=1}^q (-1)^{F_{A,b}(x_i)} \right] = \prod_{w=1}^\ell z(A_w) =: W(A).$$

Let $\tilde{A} \leftarrow \mathbb{F}_2^{N \times N}$ be uniform and let $E := \{\tilde{A} \in \text{GL}_N(\mathbb{F}_2)\}$. Conditioned on E , the matrix $\tilde{A}$ is uniform in $\text{GL}_N(\mathbb{F}_2)$. Therefore

$$\text{Bias}(x_1, \dots, x_q) = |\mathbb{E}[W(\tilde{A}) \mid E]| \leq \mathbb{E}[|W(\tilde{A})| \mid E] \leq \frac{\mathbb{E}[|W(\tilde{A})|]}{\mathbb{P}[E]}.$$

Under the unconditioned distribution, the row blocks $\tilde{A}_1, \dots, \tilde{A}_\ell$ are independent uniform $M \times N$ matrices. Hence, by [Lemma 5.2](#),

$$\mathbb{E}[|W(\tilde{A})|] = \left(\mathbb{E}_R[|z(R)|] \right)^\ell \leq \eta^{\ell/2^{t_{2q}+1}}.$$

Combining with [Lemma 5.3](#),

$$\text{Bias}(x_1, \dots, x_q) \leq 4\eta^{\ell/2^{t_{2q}+1}}.$$

Since $\eta = 15/16 \leq e^{-1/16}$,

$$\eta^{\ell/2^{t_{2q}+1}} \leq \exp\left(-\frac{\ell}{2^{t_{2q}+5}}\right).$$

□

Corollary 5.4 (Parity bound before conditioning on invertibility). *Under the hypotheses of Theorem 5.1, if $\tilde{A} \leftarrow \mathbb{F}_2^{N \times N}$ is uniform, then*

$$\left| \mathbb{E}_{\tilde{A}, b} \left[\prod_{i=1}^q (-1)^{F_{\tilde{A}, b}(x_i)} \right] \right| \leq \eta^{\ell/2^{t_{2q}+1}}.$$

Here $F_{\tilde{A}, b}$ denotes the same evaluation formula as in (1), without conditioning $\tilde{A}$ on invertibility.

Proof. In the proof of Theorem 5.1, the unconditioned quantity is $\mathbb{E}[W(\tilde{A})]$. Before conditioning on invertibility we have

$$|\mathbb{E}[W(\tilde{A})]| \leq \mathbb{E}[|W(\tilde{A})|] \leq \eta^{\ell/2^{t_{2q}+1}}.$$

□

5.2 Parameter consequences

We next derive the polynomial and quasipolynomial parameter ranges stated in the introduction. The calculation uses the slow growth of the number of Cauchy–Schwarz rounds with the number of forms.

Lemma 5.5 (Growth of t_s). *Assume $p \leq 1/8$. For every $s \geq 1$, $t_s \leq 1 + \frac{\log s + 2}{\log(1/(2p))}$. In particular,*

$$t_s = O\left(1 + \frac{\log s}{\log(1/(2p))}\right), \quad (26)$$

with an absolute constant in the $O(\cdot)$ notation.

Proof. Put $D := \log(1/(2p))$. Since $p \leq 1/8$, we have $D \geq 2$. Let $t := \left\lceil \frac{\log s + 2}{D} \right\rceil$. Using $\beta \leq 1$ and the definition of D ,

$$s\beta(2p)^t \leq s(2p)^t = 2^{\log s - tD} \leq \frac{1}{4}.$$

Hence $t_s \leq t$, and $t_s \leq 1 + (\log s + 2)/D$. The final estimate follows from $D \geq 2$. □

Corollary 5.6 (Every fixed polynomial number of queries). *For every fixed $C > 0$, uniformly over every $q \leq \lambda^C$ and every set of q distinct chosen queries,*

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\Omega_C(\lambda)).$$

Equivalently, the candidate resists every non-adaptive linear test making an arbitrary fixed polynomial number of queries in the key length κ .

Proof of Corollary 5.6. For the balanced parameters, $p = \Theta(\log \lambda / \lambda)$, $\beta = 1/2 + o(1)$, $\ell = \lambda$, and $m = \Theta(\lambda / \log \lambda)$. If $q \leq \lambda^C$ for fixed C , then $\log(1/(2p)) = (1 + o(1)) \log \lambda$, and (26) gives $t_{2q} = O_C(1)$. Hence $t_{2q}p = o(1)$ and $t_{2q}/m = o(1)$, so the hypotheses of Theorem 5.1 hold for all sufficiently large λ . Therefore

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\Omega_C(\lambda)).$$

Finally, $\kappa = \lambda^4 + \lambda^2$, so every fixed polynomial in κ is a fixed polynomial in λ . □

Corollary 5.7 (Quasipolynomial linear-test regime). *Fix $\theta \in (0, 1)$ and put $Q := \kappa^{\log^\theta \kappa}$. Uniformly over every $q \leq Q$ and every set of q distinct chosen queries,*

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\lambda^{1-o(1)}).$$

Consequently, for every fixed $0 < \rho < 1/4$ and all sufficiently large κ ,

$$q \leq Q \quad \implies \quad \text{Bias}(x_1, \dots, x_q) \leq 2^{-\kappa^\rho}.$$

Proof of Corollary 5.7. Let $Q = \kappa^{\log^\theta \kappa}$. Since $\kappa = \Theta(\lambda^4)$,

$$\log Q = (\log \kappa)^{1+\theta} = \Theta((\log \lambda)^{1+\theta}).$$

Hence $t_{2Q} = O((\log \lambda)^\theta)$. For every fixed $0 < \theta < 1$, $t_{2Q}p = o(1)$ and $t_{2Q}/m = o(1)$, so Theorem 5.1 applies uniformly for every $q \leq Q$. Moreover,

$$2^{t_{2Q}} = \exp(O((\log \lambda)^\theta)) = \lambda^{o(1)}.$$

Therefore

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\lambda^{1-o(1)}).$$

Finally fix $0 < \rho < 1/4$. Since $\kappa^\rho = \Theta(\lambda^{4\rho}) = o(\lambda^{1-o(1)})$, the same estimate is eventually at most $2^{-\kappa^\rho}$. $\square$

Corollary 5.8 (Explicit quasipolynomial query/advantage tradeoff). *Fix $a \in (0, 1)$ and put $Q_a := 2^{a(\log_2 \lambda)^2}$. Uniformly over every $q \leq Q_a$ and every set of q distinct chosen queries,*

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\lambda^{1-a-o(1)}).$$

In key-length notation, if $\alpha_q, \rho > 0$ are fixed and $16\alpha_q + 4\rho < 1$, then for all sufficiently large κ ,

$$q \leq 2^{\alpha_q(\log_2 \kappa)^2} \implies \text{Bias}(x_1, \dots, x_q) \leq 2^{-\kappa^\rho}.$$

Proof of Corollary 5.8. Fix $a \in (0, 1)$ and let $Q_a = 2^{a(\log \lambda)^2}$. For the balanced parameters, $D_\lambda := \log \frac{1}{2p} = (1 + o(1)) \log \lambda$. By Lemma 5.5, uniformly for $q \leq Q_a$,

$$t_{2q} \leq 1 + \frac{\log(2Q_a) + 2}{D_\lambda} = (a + o(1)) \log \lambda.$$

Hence $2^{t_{2q}} \leq \lambda^{a+o(1)}$. Also

$$t_{2q}p = O\left(\frac{(\log \lambda)^2}{\lambda}\right) = o(1), \quad \frac{t_{2q}}{m} = O\left(\frac{(\log \lambda)^2}{\lambda}\right) = o(1),$$

so the hypotheses of Theorem 5.1 hold for all sufficiently large λ . Since $\ell = \lambda$, Theorem 5.1 gives

$$\text{Bias}(x_1, \dots, x_q) \leq 4 \exp\left(-\frac{\lambda}{2^{t_{2q}+5}}\right) = \exp(-\lambda^{1-a-o(1)}).$$

For the key-length formulation, fix $\alpha_q, \rho > 0$ with $16\alpha_q + 4\rho < 1$. Since $\log \kappa = (4 + o(1)) \log \lambda$, choose a fixed a satisfying $16\alpha_q < a < 1 - 4\rho$. For all sufficiently large λ , $2^{\alpha_q(\log \kappa)^2} \leq 2^{a(\log \lambda)^2}$. The first part therefore gives

$$\text{Bias}(x_1, \dots, x_q) \leq \exp(-\lambda^{1-a-o(1)}).$$

Because $1 - a > 4\rho$ and $\kappa^\rho = \Theta(\lambda^{4\rho})$, this is eventually at most $2^{-\kappa^\rho}$. $\square$

6 Application to the BCGIKS Sipser-noise family

We now use the stronger rare-event theorem in a setting where the queries are random rather than adversarial. BCGIKS [BCG⁺21] propose a weak-PRF candidate in $\text{AC}^0 \circ \text{MOD}_2$ that uses a biased depth-three Sipser predicate as deterministic noise after a secret linear map. They analyze its algebraic structure and leave resistance of this candidate to linear attacks open. We show that random samples supply the additional independence needed by Theorem 4.10.

6.1 The Sipser noise family and linear attacks

Let $m_\lambda, r_\lambda = \Theta(\lambda)$, $w_\lambda = \Theta(\log \lambda)$, and $M_\lambda := m_\lambda r_\lambda w_\lambda$. For $z \in \mathbb{F}_2^{M_\lambda}$, indexed as $(z_{a,b,c})_{a \in [m_\lambda], b \in [r_\lambda], c \in [w_\lambda]}$, define

$$G_\lambda(z) := \bigvee_{a=1}^{m_\lambda} \bigwedge_{b=1}^{r_\lambda} \bigvee_{c=1}^{w_\lambda} z_{a,b,c}. \quad (27)$$

BCGIKS [BCG⁺21] use this depth-three Sipser shape with $w_\lambda \simeq \log \lambda - \log \log \lambda$ and explain how the integer fan-ins can be tuned so that the resulting predicate has the required balance. We use this tuning only through the following parameter bounds. Put

$$\varphi_\lambda(z) := \bigwedge_{b=1}^{r_\lambda} \bigvee_{c=1}^{w_\lambda} z_{b,c}, \quad p_\lambda := \mathbb{P}_{U \leftarrow \mathbb{F}_2^{r_\lambda w_\lambda}} [\varphi_\lambda(U) = 1].$$

We assume the standard balanced tuning, for which $p_\lambda = \Theta(1/\lambda)$ and $\frac{1}{3} \leq \beta_\lambda := (1 - p_\lambda)^{m_\lambda} \leq \frac{2}{3}$ for all sufficiently large λ . In particular, $M_\lambda = \Theta(\lambda^2 \log \lambda)$. No later argument uses the exact values of the three fan-ins or requires β_λ to converge to $1/2$.

Following BCGIKS [BCG⁺21], consider the learning-parity-with-simple-deterministic-noise family

$$f_{s,K}(x) := \langle s, x \rangle \oplus G_\lambda(Kx), \quad (28)$$

where $n_\lambda := M_\lambda + 1$, $s \leftarrow \mathbb{F}_2^{n_\lambda}$, and $K \leftarrow \mathbb{F}_2^{M_\lambda \times n_\lambda}$ are independent and uniform. The public inputs are uniform random samples, and an attack may choose a linear combination of the output bits after seeing those inputs.

For a fixed sample tuple $X = (x_1, \dots, x_Q)$ and $u \in \mathbb{F}_2^Q \setminus \{0\}$, write

$$\text{Bias}_{\text{Sip}}(X; u) := \left| \mathbb{E}_{s,K} \left[(-1)^{\sum_{i=1}^Q u_i f_{s,K}(x_i)} \right] \right|.$$

If $\sum_i u_i x_i \neq 0$, then

$$\text{Bias}_{\text{Sip}}(X; u) = 0, \quad (29)$$

because the expectation over the uniform s vanishes. Thus only coefficient vectors in the linear relation space of the public samples can contribute.

Our main result for this family is the following.

Theorem 6.1 (Quasipolynomial linear tests for the Sipser-based family). *Fix $0 < \varepsilon < 2$ and let $Q \leq 2^{(2-\varepsilon)(\log_2 \lambda)^2}$. Let $x_1, \dots, x_Q \leftarrow \mathbb{F}_2^{n_\lambda}$ be independent and uniform. With probability $1 - 2^{-\Omega(n_\lambda)}$ over these public samples, simultaneously for every nonzero $u \in \mathbb{F}_2^Q$,*

$$\left| \mathbb{E}_{s,K} \left[(-1)^{\sum_{i=1}^Q u_i f_{s,K}(x_i)} \right] \right| \leq \exp(-\lambda^{\varepsilon-o(1)}).$$

6.2 Random samples have no short linear dependencies

We use a simple union bound, but it is important that the no-short-dependency property holds for the entire sample tuple. Once it does, the bound applies simultaneously to every attack vector chosen after seeing that tuple.

Lemma 6.2 (No short dependencies among random binary samples). *Let $x_1, \dots, x_Q \leftarrow \mathbb{F}_2^{n_0}$ be independent and uniform. For every integer $k \geq 1$,*

$$\mathbb{P} \left[\exists \emptyset \neq I \subseteq [Q], |I| \leq k : \sum_{i \in I} x_i = 0 \right] \leq 2^{-n_0} \sum_{r=1}^k \binom{Q}{r}. \quad (30)$$

In particular, put $k := \left\lfloor \frac{n_0}{8 \log_2(eQ)} \right\rfloor$. Whenever $k \geq 1$, the probability on the left-hand side of (30) is at most $2^{-7n_0/8+1}$.

The factor $1/8$ is only convenient slack in the union bound. Any fixed sufficiently small positive constant in its place would give the same $2^{-\Omega(n_0)}$ conclusion, with a different absolute constant in the exponent.

Proof. For a fixed nonempty $I \subseteq [Q]$, the sum $\sum_{i \in I} x_i$ is uniform on $\mathbb{F}_2^{n_0}$, and hence vanishes with probability 2^{-n_0} . Summing over all I of size at most k gives (30).

For the stated choice of k ,

$$\sum_{r=1}^k \binom{Q}{r} \leq \sum_{r=1}^k (eQ)^r \leq 2(eQ)^k \leq 2^{n_0/8+1}.$$

Substitution in (30) proves the claim. $\square$

6.3 Quasipolynomial linear-test security

We now prove the theorem. The proof is uniform over the attack vector, so an adversary may choose its linear test as an arbitrary function of the entire public sample tuple.

Proof of Theorem 6.1. Put $k_{\text{dep}} := \left\lfloor \frac{n_\lambda}{8 \log_2(eQ)} \right\rfloor$. By Lemma 6.2, except with probability $2^{-\Omega(n_\lambda)}$, the sample tuple has no nonzero linear dependency of size at most k_{dep} . Fix a tuple in this event and a nonzero attack vector u .

If $\sum_i u_i x_i \neq 0$, (29) gives zero bias. Suppose therefore that $\sum_{i=1}^Q u_i x_i = 0$. Let $S := \text{supp}(u)$ and $q := |S|$. The no-short-dependency event implies $q > k_{\text{dep}}$. For each row $a \in \mathbb{F}_2^{n_\lambda}$ of K , the internal bit seen at query x_i is the linear form $L_i(a) := \langle a, x_i \rangle$. On the support S , every subfamily of at most k_{dep} of these forms is linearly independent.

The Sipser predicate in (27) is a balanced OR of $m = m_\lambda = \Theta(\lambda)$ independent copies of φ_λ . Apply Theorem 4.10 with $k = k_{\text{dep}}$, $p_\varphi = p_\lambda = \Theta(1/\lambda)$, and $\beta_\varphi = \beta_\lambda \in [1/3, 2/3]$. Let $t = t_{q, k_{\text{dep}}}$. Since $q \leq Q$, monotonicity in the number of forms gives $t_{q, k_{\text{dep}}} \leq t_{Q, k_{\text{dep}}}$. The growth estimate (24) yields

$$t_{Q, k_{\text{dep}}} \leq 1 + \frac{\log(Q/k_{\text{dep}}) + O(1)}{\log(1/(2p_\lambda))}. \quad (31)$$

Because $n_\lambda = \Theta(\lambda^2 \log \lambda)$ and $\log Q \leq (2 - \varepsilon)(\log \lambda)^2$, we have

$$k_{\text{dep}} = \Omega\left(\frac{\lambda^2}{\log \lambda}\right). \quad (32)$$

Also $\log(1/(2p_\lambda)) = (1 + o(1)) \log \lambda$, so

$$t_{Q, k_{\text{dep}}} \leq (2 - \varepsilon + o(1)) \log_2 \lambda \quad (33)$$

and therefore

$$2^{t_{Q,k_{\text{dep}}}} \leq \lambda^{2-\varepsilon+o(1)}.$$

In particular, $t_{Q,k_{\text{dep}}} p_\lambda = o(1)$ and, since $m_\lambda = \Theta(\lambda)$, $t_{Q,k_{\text{dep}}} = o(m_\lambda)$. Thus the hypotheses of [Theorem 4.10](#) hold for all sufficiently large λ .

Combining (32) with the preceding bound gives

$$\frac{k_{\text{dep}}}{2^{t_{Q,k_{\text{dep}}}}} \geq \frac{k_{\text{dep}}}{2^{t_{Q,k_{\text{dep}}}}} \geq \lambda^{\varepsilon-o(1)}.$$

Applying [Theorem 4.10](#) therefore gives

$$\left| \mathbb{E}_K \left[(-1)^{\sum_{i \in S} G_\lambda(Kx_i)} \right] \right| \leq \exp(-\lambda^{\varepsilon-o(1)}).$$

The linear secret has already canceled because S is a relation, so this is the claimed bound. $\square$

7 Further consequences for the BDEKRS candidate

We collect here consequences of the small-bias theorem for nonlinear distinguishers. The first two are standard Fourier arguments. The final subsection combines the quasipolynomial-regime bias with standard bounded-independence results for AC^0 .

7.1 Fixed non-adaptive query sets

Fix distinct inputs $X = (x_1, \dots, x_q)$ and define

$$Z_X := \left((-1)^{F_{A,b}(x_1)}, \dots, (-1)^{F_{A,b}(x_q)} \right) \in \{-1, 1\}^q.$$

Put $P_X := \text{Law}(Z_X)$. For $S \subseteq [q]$, write $\mu_X(S) := \mathbb{E} \left[\prod_{i \in S} (-1)^{F_{A,b}(x_i)} \right]$. Thus $\mu_X(\emptyset) = 1$, and the nonempty $\mu_X(S)$ are precisely the Fourier coefficients of the density of the response distribution relative to uniform.

Lemma 7.1 (Parity bias and Fourier L_1 norm). *Let $D : \{-1, 1\}^q \rightarrow [-1, 1]$. If $|\mu_X(S)| \leq \varepsilon$ for every nonempty $S \subseteq [q]$, then*

$$|\mathbb{E}[D(Z_X)] - \mathbb{E}[D(U_q)]| \leq \varepsilon \sum_{\emptyset \neq S \subseteq [q]} |\widehat{D}(S)|.$$

In particular,

$$\text{TV}(P_X, U_q) \leq \frac{1}{2} \sqrt{2^q - 1} \varepsilon.$$

Proof. By (10),

$$\mathbb{E}[D(Z_X)] - \mathbb{E}[D(U_q)] = \sum_{\emptyset \neq S} \widehat{D}(S) \mu_X(S),$$

which proves the first claim. For the second, recall that $P_X = \text{Law}(Z_X)$. By (11) and (12),

$$\chi^2(P_X \| U_q) = \sum_{\emptyset \neq S} \mu_X(S)^2 \leq (2^q - 1) \varepsilon^2.$$

Apply (13). $\square$

Whenever the hypotheses of [Theorem 5.1](#) hold for q , put $\varepsilon_q := 4\eta^{\ell/2^{t_{2q}+1}}$. Since t_s is nondecreasing, [Theorem 5.1](#) gives $|\mu_X(S)| \leq \varepsilon_q$ for every nonempty $S \subseteq [q]$. Thus every fixed set of q distinct chosen queries satisfies

$$\text{TV}(P_X, U_q) \leq \frac{1}{2}\sqrt{2^q - 1} \varepsilon_q,$$

so the same quantity bounds the advantage of every non-adaptive distinguisher on that query set.

The same bound also controls low-Fourier-degree postprocessing. Fix $0 < \theta < 1$ and $0 < \rho < 1/4$, let $q \leq \kappa^{\log^\theta \kappa}$, and suppose $D : \{-1, 1\}^q \rightarrow [-1, 1]$ has Fourier degree at most d . By Cauchy–Schwarz, Parseval, and [Lemma 7.1](#) and [Corollary 5.7](#),

$$|\mathbb{E}[D(Z_X)] - \mathbb{E}[D(U_q)]| \leq 2^{-\kappa^\rho} \sqrt{\sum_{j=1}^d \binom{q}{j}}.$$

In particular, if $d \log(eq/d) = o(\kappa^\rho)$, then the distinguishing advantage is $2^{-\kappa^\rho + o(\kappa^\rho)}$. We use a stronger bounded-independence argument below for constant-depth circuits.

7.2 Adaptive adversaries

Fixing its internal randomness turns an adaptive q -query adversary into a binary decision tree of depth at most q . Repeated queries may be deleted, so queries along a root-to-leaf path are distinct.

Lemma 7.2 (Fourier L_1 norm of a query tree). *Let D be the $\{0, 1\}$ -valued acceptance function of a deterministic binary decision tree of depth at most q . Then*

$$\sum_{\emptyset \neq S} |\widehat{D}(S)| \leq 2^{q-1}.$$

Here the Fourier expansion is taken as a function of the truth-table coordinates queried by the tree.

Proof. Each leaf is represented by the indicator of its root-to-leaf path, a product of at most q literals. Its Fourier L_1 norm is exactly 1. Let a and r be the numbers of accepting and rejecting leaves. Writing D as the sum of accepting-leaf indicators gives $\sum_{S \neq \emptyset} |\widehat{D}(S)| \leq a$; applying the same argument to $1 - D$ gives the bound by r . Hence

$$\sum_{S \neq \emptyset} |\widehat{D}(S)| \leq \min\{a, r\} \leq \frac{a + r}{2} \leq 2^{q-1}.$$

□

Proposition 7.3 (From parity bias to adaptive distinguishers; cf. [\[BR17\]](#)). *Suppose that for every set of at most q distinct domain points and every nonempty subset of those points, the corresponding parity bias is at most ε_q . Then every randomized adaptive q -query adversary has advantage at most $2^{q-1}\varepsilon_q$.*

Proof. Fix the adversary's internal randomness. The resulting deterministic algorithm is a depth- q query tree. Every nonconstant Fourier monomial of a leaf indicator is supported on a subset of the fixed domain points queried along that path, so its expectation changes by at most ε_q . Summing over the nonconstant Fourier coefficients and applying [Lemma 7.2](#) gives the claim. Average over the adversary's coins. □

Theorem 7.4 (Unrestricted adaptive statistical security). *For the balanced BDEKRS parameters, there is an absolute $\delta > 0$ such that every randomized adaptive adversary making at most $q \leq \delta\lambda$ queries has advantage $\exp(-\Omega(\lambda))$. Since $\kappa = \Theta(\lambda^4)$, this gives information-theoretic security for $\Theta(\kappa^{1/4})$ unrestricted adaptive queries.*

Proof of Theorem 7.4. For $q \leq \delta\lambda$, we have $t_{2q} \leq 2$ for all sufficiently large λ , since

$$2q\beta(2p)^2 = O\left(\delta \frac{\log^2 \lambda}{\lambda}\right) = o(1).$$

Hence $\varepsilon_q \leq 4\eta^{\lambda/8}$. By Proposition 7.3,

$$\text{Adv} \leq 2^{q-1}\varepsilon_q \leq 2 \exp\left(q \ln 2 + \frac{\lambda}{8} \ln \eta\right).$$

Choosing $\delta > 0$ sufficiently small makes the exponent $-\Omega(\lambda)$. Since $\kappa = \Theta(\lambda^4)$, this is the $\Theta(\kappa^{1/4})$ query scale. $\square$

7.3 AC^0 distinguishers on the response vector

Theorem 7.5 (AC^0 distinguishers on the response vector). *Fix $\theta \in (0, 1)$, an integer $d \geq 1$, and $0 < \alpha < \frac{1}{12d+12}$. For the balanced BDEKRS parameters, let $Q = \kappa^{\log^\theta \kappa}$. Uniformly over every fixed set of $q \leq Q$ distinct non-adaptive queries and every depth- d AC^0 circuit D of size at most 2^{κ^α} applied to the q response bits,*

$$|\mathbb{E}[D(Z_X)] - \mathbb{E}[D(U_q)]| \leq 2^{-\Omega(\kappa^\alpha)}.$$

At quasipolynomially many queries the generic $2^{q/2}$ loss in the statistical-distance bound is useless, but the individual parity biases are much smaller than what is needed for bounded independence. We use two standard results.

We say that a distribution P on q bits is (δ, k) -wise independent if every marginal on at most k coordinates is within total variation δ of uniform.

Lemma 7.6 (Almost k -wise to exact k -wise; [AGM03]). *There is an absolute constant $C > 0$ such that every (δ, k) -wise independent distribution P on q bits is within total variation $q^{Ck}\delta$ of some exactly k -wise independent distribution.*

Theorem 7.7 (Bounded independence fools AC^0 ; [Tal17]). *For every fixed depth d there is a constant $C_d > 0$ such that the following holds. Every*

$$k \geq C_d (\log(s/\varepsilon))^{3d+3}$$

-wise independent distribution ε -fools every depth- d , size- s AC^0 circuit.

Lemma 7.8 (Small bias gives accurate marginals). *If every nonempty parity of a distribution P on q bits has bias at most ξ , then P is $(2^{k/2-1}\xi, k)$ -wise independent for every $1 \leq k \leq q$.*

Proof. Fix a set I of at most k coordinates. Every nonconstant Fourier coefficient of the marginal P_I has magnitude at most ξ . Therefore

$$\chi^2(P_I \| U_I) \leq (2^{|I|} - 1)\xi^2,$$

and (13) gives the result. $\square$

Proof of Theorem 7.5. Choose ρ so that $\alpha(3d+3) < \rho < \frac{1}{4}$. By Corollary 5.7, uniformly over the entire query range in the theorem, every nonempty response parity has bias at most $\xi := 2^{-\kappa^\rho}$ for all sufficiently large κ .

Let $s \leq 2^{\kappa^\alpha}$ be the circuit size and set $\varepsilon_0 := 2^{-2\kappa^\alpha}$. Choose

$$k_0 := \left\lceil C_d (\log(s/\varepsilon_0))^{3d+3} \right\rceil = O\left(\kappa^{\alpha(3d+3)}\right),$$

where C_d is from [Theorem 7.7](#). We separate the case in which the number of response bits is smaller than the bounded-independence parameter.

If $q < k_0$, the direct Fourier bound following (13) gives

$$\text{TV}(P_X, U_q) \leq \frac{1}{2} \sqrt{2^q - 1} \xi \leq 2^{q/2-1-\kappa^\rho}.$$

Since $q < k_0 = O(\kappa^{\alpha(3d+3)}) = o(\kappa^\rho)$, this is $2^{-\Omega(\kappa^\rho)}$. The claimed bound follows for every distinguisher, without using bounded independence.

Suppose now that $q \geq k_0$. By [Lemma 7.8](#), the response distribution is (δ_{k_0}, k_0) -wise independent with $\delta_{k_0} \leq 2^{k_0/2}\xi$. Apply [Lemma 7.6](#). There is an exactly k_0 -wise independent Q_X satisfying

$$\text{TV}(\text{Law}(Z_X), Q_X) \leq q^{O(k_0)} 2^{k_0/2} \xi.$$

Since $\log q = O((\log \kappa)^{1+\theta})$ and $\alpha(3d+3) < \rho$,

$$O(k_0 \log q) + \frac{k_0}{2} - \kappa^\rho = -\Omega(\kappa^\rho).$$

Thus

$$\text{TV}(\text{Law}(Z_X), Q_X) = 2^{-\Omega(\kappa^\rho)}.$$

By the definition of k_0 , the distribution Q_X fools the circuit to error ε_0 by [Theorem 7.7](#). The triangle inequality gives

$$|\mathbb{E}[D(Z_X)] - \mathbb{E}[D(U_q)]| \leq 2^{-2\kappa^\alpha} + 2^{-\Omega(\kappa^\rho)} = 2^{-\Omega(\kappa^\alpha)}.$$

□

Remark 7.9 (DNF/CNF distinguishers). Bazzi's depth-two theorem gives a slightly better numerical range for DNF and CNF formulas. Under the same quasipolynomial query bound as in [Theorem 7.5](#), formulas with at most 2^{κ^α} terms or clauses have distinguishing advantage $2^{-\Omega(\kappa^\alpha)}$ for every fixed $0 < \alpha < 1/8$. Indeed, choose ρ with $2\alpha < \rho < 1/4$ and apply the same approximation argument using Bazzi's $O(\log^2(s/\varepsilon))$ bounded-independence bound [[Baz07](#)].

8 Query sets with few affine relations

The generic Fourier bound treats all 2^q Fourier characters alike. We can do better when each transcript has only a few independent affine relations. For example, suppose an adversary starts with affinely independent queries and then adds only a few queries from their affine span. Each such query adds at most one independent affine relation.

For a fixed Fourier subset S , coordinates that do not participate in any affine relation supported on S are independent of the remaining coordinates. Their contributions to the Fourier coefficient factor into separate one-coordinate expectations. This lets us sum the Fourier L_1 mass according to the subspaces of affine relations supported on S , rather than according to all subsets of queries separately.

Theorem 8.1 (Adaptive statistical security with few affine relations). *Fix $\gamma \in (0, 1)$. There is an absolute constant $c_* > 0$ such that, for the balanced BDEKRS parameters and all sufficiently large λ , the following holds. Let $\mathcal{A}$ be an arbitrary computationally unbounded adaptive adversary making at most $q \leq (1-\gamma)N = (1-\gamma)\lambda^2$ distinct queries. Suppose that on every root-to-leaf path of its query tree the affine relation space of the queried points has dimension at most $\nu \leq c_*\lambda/\log \lambda$. Then*

$$\text{Adv}_{\mathcal{F}_{\text{BDEKRS}}}(\mathcal{A}) \leq \exp(-\Omega(\lambda)).$$

8.1 Relation spaces and their supports

Fix distinct $X = (x_1, \dots, x_q)$ and put $v_i := (x_i, 1) \in \mathbb{F}_2^{N+1}$. Define the affine relation space

$$\mathcal{K}_X := \left\{ c \in \mathbb{F}_2^q : \sum_{i=1}^q c_i v_i = 0 \right\}.$$

Recall that

$$d_{\text{aff}}(X) = \dim \text{span}\{v_1, \dots, v_q\} - 1, \quad \nu(X) = \dim \mathcal{K}_X = q - \dim \text{span}\{v_1, \dots, v_q\}.$$

For $S \subseteq [q]$, let $W_X(S) := \{c \in \mathcal{K}_X : \text{supp}(c) \subseteq S\}$, the space of affine relations whose support is contained in S . For any subspace $W \subseteq \mathcal{K}_X$, define its support by $\text{supp}(W) := \bigcup_{c \in W} \text{supp}(c)$, and abbreviate $C_X(S) := \text{supp}(W_X(S))$. Thus $C_X(S) = \emptyset$ when $W_X(S) = 0$.

For intuition, take five queries $x_1 = 0, x_2 = e_1, x_3 = e_2, x_4 = e_1 + e_2$, and $x_5 = e_3$, where e_1, e_2, e_3 are distinct standard basis vectors. For $S = \{1, 2, 3, 4, 5\}$, the only nonzero affine relation is $v_1 + v_2 + v_3 + v_4 = 0$, while $v_5 \notin \text{span}\{v_1, v_2, v_3, v_4\}$. Thus $W_X(S) = \text{span}\{(1, 1, 1, 1, 0)\}$ and $C_X(S) = \{1, 2, 3, 4\}$. The fifth coordinate occurs in no relation supported on S , so its contribution to the corresponding Fourier coefficient factors as a separate one-coordinate expectation.

Let $\tilde{A} \leftarrow \mathbb{F}_2^{N \times N}$ be uniform and $b \leftarrow \mathbb{F}_2^N$. Write $\tilde{P}_X$ for the law of the signed response vector produced by the same evaluation formula as the candidate with $\tilde{A}$ in place of A , and write

$$\tilde{\mu}_X(S) := \mathbb{E}_{\tilde{A}, b} \left[\prod_{i \in S} (-1)^{F_{\tilde{A}, b}(x_i)} \right].$$

Finally put $\zeta := \mathbb{E}_{Y \leftarrow \mathbb{F}_2^N} [g(Y)] = (2\beta - 1)^\ell$. For the balanced parameters, $|2\beta - 1| = O(p)$ and hence

$$|\zeta| = 2^{-\Omega(\lambda \log \lambda)}. \quad (34)$$

Lemma 8.2 (Factorization outside the relation support). *For every $S \subseteq [q]$, put $W = W_X(S)$ and $C = C_X(S)$. There is a quantity b_W , depending only on the queries indexed by C , such that*

$$\tilde{\mu}_X(S) = b_W \zeta^{|S \setminus C|}.$$

We have $b_0 = 1$. If $W \neq 0$, then

$$|b_W| \leq \varepsilon_q^{\text{unif}} := \eta^{\ell/2^{t_{2q}+1}}$$

whenever the hypotheses of [Theorem 5.1](#) hold for q .

Proof. The affine relation space among the augmented vectors indexed by S is $W_X(S)$. An index $i \in S \setminus C$ occurs in no relation supported in S . Equivalently, $v_i \notin \text{span}\{v_j : j \in S \setminus \{i\}\}$. We may therefore choose a basis of $\text{span}\{v_j : j \in S\}$ that contains every vector indexed by $S \setminus C$. The evaluations of a uniform linear functional on these basis vectors are independent unbiased bits and are independent of its evaluations on the remaining basis vectors. Applying this independently to each of the N rows of the random affine map shows that the vectors corresponding to indices in $S \setminus C$ are independent uniform elements of $\mathbb{F}_2^N$ and are independent of those corresponding to indices in C . Each therefore contributes the factor ζ .

The remaining correlation is b_W . If $W \neq 0$, then C is a nonempty set of distinct queries and $|C| \leq q$. Monotonicity of t_s and [Corollary 5.4](#) give $|b_W| \leq \varepsilon_q^{\text{unif}}$. $\square$

8.2 Counting the subspaces $W_X(S)$

Although S ranges over 2^q subsets, the subspaces $W_X(S)$ take far fewer distinct values when the affine relation space has small dimension.

Lemma 8.3 (Number of distinct subspaces $W_X(S)$). *Let $\nu = \dim \mathcal{K}_X$. The number $\mathcal{N}_X$ of distinct subspaces $W_X(S)$, as S ranges over all subsets of $[q]$, satisfies*

$$\mathcal{N}_X \leq \sum_{j=0}^{\nu} \binom{q}{j}.$$

In particular, if $1 \leq \nu \leq q/2$, then

$$\mathcal{N}_X \leq \left(\frac{eq}{\nu}\right)^{\nu}.$$

For $\nu = 0$, $\mathcal{N}_X = 1$.

Proof. Choose a full-row-rank generator matrix $H \in \mathbb{F}_2^{\nu \times q}$ for $\mathcal{K}_X$ and let its columns be $h_1, \dots, h_q \in \mathbb{F}_2^{\nu}$. The map $u \mapsto u^{\top} H$ identifies $\mathbb{F}_2^{\nu}$ with $\mathcal{K}_X$. Under this identification, $W_X(S)$ corresponds to

$$\{u : u \cdot h_i = 0 \text{ for every } i \notin S\} = \text{span}\{h_i : i \notin S\}^{\perp}.$$

Thus $W_X(S)$ is determined by a subspace spanned by a subset of the q columns of H . Every such span has a basis consisting of at most ν of those columns, so the number of possible spans is at most $\sum_{j=0}^{\nu} \binom{q}{j}$. The standard binomial bound gives the second claim. $\square$

Theorem 8.4 (Fourier L_1 mass with few affine relations). *Under the hypotheses of Theorem 5.1, let X be a fixed set of q distinct queries whose affine relation space has dimension ν . When $\tilde{A}$ is uniform over all $N \times N$ matrices,*

$$1 + \sum_{\emptyset \neq S \subseteq [q]} |\tilde{\mu}_X(S)| \leq (1 + |\zeta|)^q \left(1 + (\mathcal{N}_X - 1) \varepsilon_q^{\text{unif}}\right),$$

where $\mathcal{N}_X$ is bounded as in Lemma 8.3. In particular, for $1 \leq \nu \leq q/2$,

$$1 + \sum_{S \neq \emptyset} |\tilde{\mu}_X(S)| \leq (1 + |\zeta|)^q \left[1 + \left(\frac{eq}{\nu}\right)^{\nu} \varepsilon_q^{\text{unif}}\right].$$

Proof. Partition the subsets $S \subseteq [q]$ according to $W = W_X(S)$. By Lemma 8.2, for fixed W , put $C = \text{supp}(W)$. Then

$$\sum_{S: W_X(S)=W} |\tilde{\mu}_X(S)| \leq |b_W| \sum_{T \subseteq [q] \setminus C} |\zeta|^{|T|} \leq |b_W| (1 + |\zeta|)^q.$$

For $W = 0$ we have $b_W = 1$, and for every nonzero W we have $|b_W| \leq \varepsilon_q^{\text{unif}}$. Summing over the $\mathcal{N}_X$ possible subspaces $W_X(S)$ proves the claim. $\square$

Corollary 8.5 (Pointwise closeness to uniform with few affine relations). *There exists an absolute $c_* > 0$ such that the following holds for the balanced parameters. If $q \leq N = \lambda^2$ and $\nu(X) \leq c_* \lambda / \log \lambda$, then, when $\tilde{A}$ is uniform over all $N \times N$ matrices,*

$$\sum_{S \neq \emptyset} |\tilde{\mu}_X(S)| \leq \exp(-\Omega(\lambda)).$$

Consequently, for every $z \in \{-1, 1\}^q$,

$$\tilde{P}_X(z) = 2^{-q} (1 \pm \exp(-\Omega(\lambda))).$$

Proof. For $q \leq \lambda^2$, (34) gives $q|\zeta| = 2^{-\Omega(\lambda \log \lambda)}$. Moreover, Corollary 5.4 and Lemma 5.5 with $q \leq \lambda^2$ gives an absolute constant $c_{\text{corr}} > 0$ such that $\varepsilon_q^{\text{unif}} \leq e^{-c_{\text{corr}}\lambda}$ for all sufficiently large λ . If $\nu \leq q/2$, then $\nu \log(eq/\nu) \leq O(c_*\lambda)$ under the stated bound on the relation dimension. Choosing c_* sufficiently small compared with c_{corr} makes the contribution from nonzero relation subspaces in Theorem 8.4 exponentially small. If instead $\nu > q/2$, then $q < 2\nu = O(c_*\lambda/\log \lambda)$, and the crude bound $\mathcal{N}_X \leq 2^q$ gives the same conclusion. In both cases the factor $(1 + |\zeta|)^q$ is $1 + 2^{-\Omega(\lambda \log \lambda)}$.

For the pointwise statement, Fourier inversion gives

$$2^q \tilde{P}_X(z) - 1 = \sum_{S \neq \emptyset} \tilde{\mu}_X(S) \chi_S(z),$$

whose absolute value is bounded by the displayed Fourier L_1 mass. $\square$

8.3 Adaptive queries and the invertible matrix

The pointwise estimate can be summed over the leaves of an adaptive query tree without a multiplicative factor for the number of leaves. We then couple the uniform affine map to a uniform invertible one along the adaptively chosen affine span.

Lemma 8.6 (Adaptive affine-map coupling). *Let a deterministic adaptive algorithm make distinct queries $x_1, x_2, \dots$ in $\mathbb{F}_2^N$. Couple two experiments: in the first, $\tilde{A} \leftarrow \mathbb{F}_2^{N \times N}$ and $b \leftarrow \mathbb{F}_2^N$ are uniform; in the second, $A \leftarrow \text{GL}_N(\mathbb{F}_2)$ and $b' \leftarrow \mathbb{F}_2^N$ are uniform. If the affine dimension of every query prefix is at most r , the query images $\tilde{A}x_i + b$ and $Ax_i + b'$ in the two experiments can be coupled to agree throughout the execution except with probability less than 2^{r-N} . The same bound holds after applying any deterministic function to these images, in particular the candidate output predicate.*

Proof. Couple the image of the first query to be the same uniform point in both experiments. Relative to that point, expose images of direction vectors as the affine span grows. In the first experiment, when a new direction is linearly independent of the previous directions, its image is uniform in $\mathbb{F}_2^N$. In the second experiment, its image is uniform outside the span of the images of the previously exposed directions. If j independent directions have already been exposed, these two choices can be coupled except when the image in the first experiment falls in their j -dimensional span, an event of probability 2^{j-N} . A query in the existing affine span introduces no new randomness: its affine image is determined by the previous images in both experiments. Summing the failure probabilities over $j = 0, \dots, r-1$ gives

$$\sum_{j=0}^{r-1} 2^{j-N} < 2^{r-N}.$$

As long as the coupling has not failed, the observed oracle answers agree, so the adaptive algorithm asks the same next query in both experiments. $\square$

Proof of Theorem 8.1. Fix the adversary's internal randomness. Each leaf ℓ of the resulting query tree determines a fixed sequence X_ℓ of distinct queries and a fixed response string z_ℓ . By hypothesis, $|X_\ell| \leq q \leq N$ and $\nu(X_\ell) \leq c_*\lambda/\log \lambda$. For the experiment with $\tilde{A}$ uniform over all matrices, Corollary 8.5 gives, uniformly over leaves,

$$\mathbb{P}[\ell] = 2^{-|X_\ell|} \left(1 \pm e^{-\Omega(\lambda)}\right).$$

Under a random function, the same leaf has probability exactly $2^{-|X_\ell|}$. Summing over accepting leaves, whose random-function probabilities sum to at most 1, shows that this experiment and the random-function experiment differ in acceptance probability by $e^{-\Omega(\lambda)}$.

Every path has affine dimension at most $q - 1 \leq (1 - \gamma)N - 1$. By [Lemma 8.6](#), replacing $\tilde{A}$ by a uniform invertible matrix changes the entire adaptive experiment by at most $2^{-\Omega_\gamma(N)}$. Averaging over the adversary's internal randomness proves

$$\text{Adv}_{\mathcal{F}_{\text{BDEKRS}}}(\mathcal{A}) \leq e^{-\Omega(\lambda)} + 2^{-\Omega_\gamma(N)} = e^{-\Omega(\lambda)}.$$

□

9 Limitations and information-theoretic barriers

Two simple barriers place the positive results in context. Linear tests can eventually exploit the algebraic degree of the candidate, while arbitrary statistical distinguishers face a support-size barrier on low-dimensional query sets.

9.1 A linear attack from algebraic degree

Proposition 9.1 (Affine-cube attack). *For every fixed key (A, b) , the Boolean function $F_{A,b} : \mathbb{F}_2^N \rightarrow \mathbb{F}_2$ has algebraic degree at most $M = mn$. Consequently, whenever $N \geq M + 1$, there are 2^{M+1} distinct non-adaptive queries whose output parity is identically zero for every key. The corresponding linear-test bias is 1.*

Proof. Each tribe AND has degree n . Over $\mathbb{F}_2$,

$$\bigvee_{j=1}^m a_j = 1 + \prod_{j=1}^m (1 + a_j),$$

so $\text{TRIBES}_{m,n}$ has algebraic degree at most $mn = M$. Composition with an affine map does not increase degree, and the outer XOR does not increase it either. Hence $\deg_{\mathbb{F}_2}(F_{A,b}) \leq M$.

Choose linearly independent directions $v_1, \dots, v_{M+1} \in \mathbb{F}_2^N$ and a base point x_0 . On the affine cube

$$\mathcal{Q} = \left\{ x_0 + \sum_{r=1}^{M+1} u_r v_r : u \in \mathbb{F}_2^{M+1} \right\},$$

the $(M + 1)$ st discrete derivative of every degree- M polynomial vanishes. Equivalently,

$$\bigoplus_{x \in \mathcal{Q}} F_{A,b}(x) = 0.$$

□

For the balanced parameters, $M = \lambda$ and $N = \lambda^2$. The attack therefore uses $2^{\lambda+1}$ queries. By contrast, [Corollary 5.8](#) proves negligible bias through $2^{a(\log_2 \lambda)^2}$ queries for every fixed $0 < a < 1$. This leaves an open intermediate range between the positive theorem and the affine-cube attack. The present Cauchy-Schwarz argument stops giving a growing exponent when $\log q$ reaches order $(\log \lambda)^2$, because the final bound requires taking a 2^t th root, with $t = \Theta(\log q / \log \lambda)$.

9.2 A support-size barrier for low-dimensional query sets

Proposition 9.2 (Support bound from affine dimension). *Let $X = (x_1, \dots, x_q)$ be a fixed set of distinct queries with $d_{\text{aff}}(X) = d$. Let P_X be the law, under a uniform key, of the signed response vector $((-1)^{F_{A,b}(x_1)}, \dots, (-1)^{F_{A,b}(x_q)})$. Then*

$$|\text{supp}(P_X)| \leq 2^{N(d+1)}$$

and hence

$$\text{TV}(P_X, U_q) \geq 1 - 2^{N(d+1)-q}.$$

Proof. Choose a base point x_0 in the affine span of X and direction vectors $v_1, \dots, v_d$ spanning its direction space. The entire tuple $(Ax_i + b)_{i \leq q}$ is determined by $Ax_0 + b$ and $Av_1, \dots, Av_d$. There are at most $2^{N(d+1)}$ possible values of these $d+1$ vectors, so the same upper bound holds for the support of any deterministic postprocessing, including the response vector. A distribution supported on at most K points of $\{-1, 1\}^q$ has total variation distance at least $1 - K/2^q$ from uniform. $\square$

A set of q distinct points can have affine dimension as small as $\lceil \log_2 q \rceil$. Thus there are query sets of size $\Theta(N \log N)$ for which the support bound is already nontrivial. By contrast, the trivial support bound from the 2^κ possible keys becomes informative only around $q = \kappa = N^2 + N$.

References

- [ABG⁺14] Adi Akavia, Andrej Bogdanov, Siyao Guo, Akshay Kamath, and Alon Rosen. Candidate weak pseudorandom functions in $\text{AC}^0 \circ \text{MOD}_2$. In Moni Naor, editor, *ITCS 2014*, pages 251–260. ACM, January 2014.
- [ABR16] Benny Applebaum, Andrej Bogdanov, and Alon Rosen. A dichotomy for local small-bias generators. *Journal of Cryptology*, 29(3):577–596, July 2016.
- [AGM03] Noga Alon, Oded Goldreich, and Yishay Mansour. Almost k-wise independence versus k-wise independence. *Information Processing Letters*, 88(3):107–110, 2003.
- [AL16] Benny Applebaum and Shachar Lovett. Algebraic attacks against random local functions and their countermeasures. In Daniel Wichs and Yishay Mansour, editors, *48th ACM STOC*, pages 1087–1100. ACM Press, June 2016.
- [AM13] Per Austrin and Elchanan Mossel. Noise correlation bounds for uniform low degree functions. *Arkiv för Matematik*, 51(1):29 – 52, 2013.
- [Baz07] Louay Bazzi. Polylogarithmic independence can fool DNF formulas. In *48th FOCS*, pages 63–73. IEEE Computer Society Press, October 2007.
- [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 $\text{AC}^0[\text{MOD}_2]$. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part IV*, volume 12828 of *LNCS*, pages 487–516, Virtual Event, August 2021. Springer, Cham.

- [BCG⁺22] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, Nicolas Resch, and Peter Scholl. Correlated pseudorandomness from expand-accumulate codes. In Yevgeniy Dodis and Thomas Shrimpton, editors, *CRYPTO 2022, Part II*, volume 13508 of *LNCS*, pages 603–633. Springer, Cham, August 2022.
- [BCPR25] Christina Boura, Geoffroy Couteau, Léo Perrin, and Yann Rotella. SoK: On shallow weak PRFs A common symmetric building block for MPC protocols. *IACR Trans. Symm. Cryptol.*, 2025(3):289–336, 2025.
- [BDE⁺25] Marshall Ball, Clément Ducros, Saroja Erabelli, Lisa Kohl, Nicolas Resch, and Peter Scholl. Strong pseudorandom functions in $AC^0[2]$ in the bounded-query setting. Cryptology ePrint Archive, Paper 2025/2085, 2025.
- [BGGP12] Itai Benjamini, Ori Gurel-Gurevich, and Ron Peled. On k-wise independent distributions and boolean functions, 2012.
- [BL85] Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of banzhaf values. In *26th FOCS*, pages 408–416. IEEE Computer Society Press, October 1985.
- [Bog18] Andrej Bogdanov. Small bias requires large formulas. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *ICALP 2018*, volume 107 of *LIPIcs*, pages 22:1–22:12. Schloss Dagstuhl, July 2018.
- [BR17] Andrej Bogdanov and Alon Rosen. Pseudorandom functions: Three decades later. Cryptology ePrint Archive, Report 2017/652, 2017.
- [CD23] Geoffroy Couteau and Clément Ducros. Pseudorandom correlation functions from variable-density LPN, revisited. In Alexandra Boldyreva and Vladimir Kolesnikov, editors, *PKC 2023, Part II*, volume 13941 of *LNCS*, pages 221–250. Springer, Cham, May 2023.
- [CGSS22] Pablo Candela, Diego González-Sánchez, and Balázs Szegedy. A refinement of cauchy–schwarz complexity. *European Journal of Combinatorics*, 106:103592, 2022.
- [GGM86] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *Journal of the ACM*, 33(4):792–807, October 1986.
- [GMOR15] Siyao Guo, Tal Malkin, Igor Carboni Oliveira, and Alon Rosen. The power of negations in cryptography. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *TCC 2015, Part I*, volume 9014 of *LNCS*, pages 36–65. Springer, Berlin, Heidelberg, March 2015.
- [Gow01] W. T. Gowers. A new proof of szemerédi’s theorem. *GAFA, Geom. funct. anal.*, 11(3):465–588, 2001.
- [GT10] Benjamin Green and Terence Tao. Linear equations in primes. *Ann. Math. (2)*, 171(3):1753–1850, 2010.
- [GV04] Dan Gutfreund and Emanuele Viola. Fooling parity tests with parity gates. In Klaus Jansen, Sanjeev Khanna, José D. P. Rolim, and Dana Ron, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 381–392, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [GW10] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proceedings of the London Mathematical Society*, 100(1):155–176, 2010.

- [GY20] Parikshit Gopalan and Amir Yehudayoff. Concentration for limited independence via inequalities for the elementary symmetric polynomials. *Theory of Computing*, 16(17):1–29, 2020.
- [HHL16] Hamed Hatami, Pooya Hatami, and Shachar Lovett. General systems of linear forms: Equidistribution and true complexity. *Advances in Mathematics*, 292:446–477, 2016.
- [KL01] Matthias Krause and Stefan Lucks. Pseudorandom functions in TC^0 and cryptographic limitations to proving lower bounds. *Comput. complex.*, 10(4):297–313, 2001.
- [Man21] Freddie Manners. True complexity and iterated cauchy–schwarz, 2021.
- [MST03] Elchanan Mossel, Amir Shpilka, and Luca Trevisan. On e-biased generators in NC^0 . In *44th FOCS*, pages 136–145. IEEE Computer Society Press, October 2003.
- [NN90] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. In *22nd ACM STOC*, pages 213–223. ACM Press, May 1990.
- [Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987.
- [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.
- [Tal17] Avishay Tal. Tight Bounds on the Fourier Spectrum of AC^0 . In Ryan O’Donnell, editor, *32nd Computational Complexity Conference (CCC 2017)*, volume 79 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 15:1–15:31, Dagstuhl, Germany, 2017. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.