

LOCAL LEAKAGE RESILIENCE OF MULTIPLICATION-FRIENDLY SHAMIR SECRET SHARING

JIHUN HWANG¹ AND HEMANTA K. MAJI¹ AND ILYA D. SHKREDOV² AND XIUYU YE¹

¹*Department of Computer Science, Purdue University, West Lafayette, IN, USA 47906*

²*Department of Mathematics, Purdue University, West Lafayette, IN, USA 47906*

ABSTRACT. Shamir secret sharing lets a quorum of $d + 1$ parties recover a secret, while smaller coalitions learn nothing. What if every party reveals one bit computed from its share? Our results reach the so-called multiplication-friendly regime, where d is below half the number of parties, down to any constant fraction of parties. Proving leakage resilience for Shamir sharing in this regime has been challenging. We prove resilience to arbitrary single-bit-per-share leakage throughout this regime. The party count may grow polynomially in the field's bit length, with the permitted growth exponent depending on the quorum fraction.

This guarantee holds for nearly every choice of public evaluation points over sufficiently large prime fields. For each such choice, $(2/\pi)^{(1-o(1))d}$ bounds the distinguishing advantage between any two secrets, even under the most damaging choice of local bit functions. This bound is essentially tight: a known leakage strategy achieves distinguishing advantage $(2/\pi)^d$.

The technical core is a new pseudorandomness theorem for Reed–Solomon codes, proved by combining higher-order Fourier analysis with tools for describing these codes through local constraints.

1. INTRODUCTION

Shamir secret sharing gives a simple way to split a secret among several parties [Sha79]. A quorum can recover the secret, while any smaller coalition learns nothing, even from its full shares. Yet could a little leakage from every share reveal something about the secret?

The scheme is based on polynomial interpolation. Let p be an odd prime, write $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, and choose m distinct nonzero public evaluation places $a_1, \dots, a_m \in \mathbb{F}_p$. Fix an integer $1 \leq d < m$. To share a secret $s \in \mathbb{F}_p$, choose independent uniform $R_1, \dots, R_d \in \mathbb{F}_p$ and give party i its *share* $P_s(a_i)$, where

$$P_s(X) = s + \sum_{j=1}^d R_j X^j.$$

Any $d + 1$ shares recover the secret by polynomial interpolation. Any d shares are jointly uniform and independent of the secret, by the usual Vandermonde argument [Sha79, pp. 612–613].

Leakage asks for only a little from each party, but it asks everyone. Suppose party i reveals $f_i(P_s(a_i))$, where $f_i: \mathbb{F}_p \rightarrow \{0, 1\}$ is arbitrary. Together, these bits form the *leakage profile*

$$\left(f_i(P_s(a_i)) \right)_{i=1}^m.$$

Of course, any d of these bits still reveal nothing about the secret. But $d + 1$ bits may already reveal something about the secret. The observer gets all m .

Benhamouda, Degwekar, Ishai, and Rabin proposed *local leakage resilience* to ask this question precisely [BDIR18, BDIR21], a notion also implicit in the work of Goyal and Kumar [GK18]. The

Email address: {hwang102, hmaji, ishkredo, ye151}@purdue.edu.

Date: October 28, 2026.

requirement is that the leakage profiles for any two secrets have small statistical distance,¹ uniformly over the allowed local functions. Since then, arbitrary one-bit leakage from Shamir sharing has received sustained attention, but progress has been hard to come by [MNP⁺21, MNP⁺22, KK23, Ngu24, Kas24].

We focus on the regime $0 < c < 1/2$, with sharing degree $d = \lfloor cm \rfloor$. Here, Shamir sharing has the algebraic room needed to multiply shared secrets, a basic ingredient in many cryptographic applications [BGW88].² Yet guarantees against arbitrary one-bit leakage long remained confined to $c > 1/2$, with a sequence of works bringing the sufficient relative threshold down to 0.668 under their respective field-size assumptions [BDIR18, BDIR21, KK23, Kas24]. Below one half, guarantees largely covered physical-bit probing [MNP⁺21], highly biased leakage [KK23], or most leakage functions [Ngu24]. But the adversary can choose any leakage functions it likes, including the exceptions.

Even very simple leakage functions can reveal information about the secret. To see this, take the first $d+1$ shares and write the secret as their weighted sum using public reconstruction coefficients. Each of these parties leaks the least significant bit (LSB) of its weighted share; all remaining parties leak constants. Here LSB means the parity of the representative in $\{0, \dots, p-1\}$. This gives a distinguishing advantage of at least $(2/\pi)^d$ (Appendix A; c.f., [MNP⁺22, Theorem 6, p. 16:7]). These non-constant maps are highly structured, nearly balanced, and need not probe physical bits of the original shares. Could this be just the tip of the iceberg, with other local bit functions revealing much more about the secret?

Excitingly, very recent independent work by Sun and Wootters [SW26] and Nguyen [Ngu26] broke the one-half barrier for arbitrary one-bit leakage. Each proves resilience for every fixed relative degree $c > 1/2 - \delta$, for some absolute constant $\delta > 0$, and for every valid choice of evaluation places.³ Compared with these works, we reach every fixed $0 < c < 1/2$, with a different tradeoff: nearly all valid evaluation tuples are secure, with the party count growing as a small power of the field's bit length. Our guarantee uses the standard worst-case notion of security: the adversary may choose the leakage functions after all evaluation places are public.

Summary: our result. For public evaluation places $a = (a_1, \dots, a_m)$ and local single-bit leakage functions $f = (f_1, \dots, f_m)$, let $\mathcal{L}_s^{a,f}$ denote the leakage probability law for secret s . Following Benhamouda, Degwekar, Ishai, and Rabin [BDIR18, BDIR21], we measure insecurity by the worst-case statistical distinguishing advantage

$$\text{Insec}(a; d) = \sup_{f_1, \dots, f_m: \mathbb{F}_p \rightarrow \{0,1\}} \max_{s, s' \in \mathbb{F}_p} \text{SD}(\mathcal{L}_s^{a,f}, \mathcal{L}_{s'}^{a,f}).$$

Small insecurity means that the leakage profiles of any two secrets are statistically indistinguishable, regardless of the chosen local leakage functions. The worst-case leakage is taken separately for each public tuple a , allowing the leakage functions to be chosen after a is known.

Write n for the field's bit length, so $2^{n-1} < p < 2^n$. Averages over a use uniform ordered tuples of m distinct nonzero field elements. For every fixed $0 < c < 1/2$, there is a positive exponent $u_*(c)$ such that the following holds. Fix $0 < u < u_*(c)$, and set $m = \lfloor n^u \rfloor$ and $d = \lfloor cm \rfloor$. As $n \rightarrow \infty$, we obtain the nearly optimal guarantee

$$\mathbb{E}_a [\text{Insec}(a; d)] \leq \left(\frac{2}{\pi} \right)^{(1-o(1))d}$$

¹The statistical distance between probability laws μ and ν on a finite set Ω is $\text{SD}(\mu, \nu) = \frac{1}{2} \sum_{y \in \Omega} |\mu(y) - \nu(y)|$. Equivalently, it is the largest difference in the probabilities they assign to the same event.

²Multiplying two degree- d sharings gives evaluations of a polynomial of degree at most $2d$. Thus any $2d+1$ shares suffice to reconstruct the product. Since $d = \lfloor cm \rfloor$ and $c < 1/2$, we have $2d < m$, so there are enough shares.

³Nguyen gives the explicit, unoptimized choice $\delta = 10^{-14}$ [Ngu26, Corollary 6]. Sun and Wootters leave their constant unspecified; see [SW26, Theorem 1.3 and Appendix A].

For $1/3 < c < 1/2$, we may take $u_*(c) = 1/2$. For smaller c , choose an integer $\ell > 1/c$; we give a finite exponent $\Theta_\ell > 0$, allowing $u_*(c) = \Theta_\ell^{-1}$. Incurring a further $o(d)$ loss in the exponent, Markov's inequality shows that nearly every evaluation tuple also satisfies $\text{Insec}(a; d) \leq (2/\pi)^{(1-o(1))d}$. The reconstruction-based LSB attack above achieves advantage at least $(2/\pi)^d$, so our upper bound is optimal up to $o(d)$ in the exponent. We state the precise bounds below.

Our approach. Our proof uses higher-order Fourier analysis, particularly Gowers uniformity norms and nilsequences, together with tools for local constraints on Reed–Solomon codes. We build on Leng's PhD thesis [Len25] and papers by Levi, Mosheiff, and Shagrithaya [LMS25a] and Krick, Pardo, and Sombra [KPS01]. The technical overview explains the specific ideas we draw from these works and how they come together in our argument.

1.1. Our Contributions. We prove our results through a stronger comparison: local bits look nearly the same on a random Reed–Solomon codeword as on independent uniform field elements.

For an integer $1 \leq r \leq m$ and $x = (x_0, \dots, x_{r-1}) \in \mathbb{F}_p^r$, define the *moment linear maps*, which compute the shares from the polynomial coefficient vector x :

$$L_i: \mathbb{F}_p^r \rightarrow \mathbb{F}_p, \quad L_i(x) = \sum_{h=0}^{r-1} x_h a_i^h \quad (1 \leq i \leq m).$$

These evaluate the polynomial with coefficient vector x . The corresponding r -dimensional Reed–Solomon code is

$$C_r(a) = \{(L_i(x))_{i=1}^m : x \in \mathbb{F}_p^r\} \subseteq \mathbb{F}_p^m.$$

If x is uniform in $\mathbb{F}_p^r$, then $(L_i(x))_i$ is a uniform codeword from $C_r(a)$. Let $U_1, \dots, U_m$ instead be independent uniform elements of $\mathbb{F}_p$. Define

$$\begin{array}{l} \text{codeword versus} \\ \text{independent-input:} \\ \text{comparison} \end{array} \quad \Delta(a; r) = \sup_{f_1, \dots, f_m: \mathbb{F}_p \rightarrow \{0,1\}} \text{SD} \left((f_i(L_i(x)))_{i=1}^m, (f_i(U_i))_{i=1}^m \right) \quad (1)$$

This is the pseudorandomness property we need: the leaked bits should reveal very little about the dependence among the codeword coordinates. For Shamir sharing, this measures how different the leakage profile of shares of the 0 secret can look from that of independent uniform field elements.

Why this comparison suffices. Write $P_s(X) = s + XQ(X)$, where Q is uniform among polynomials of degree less than d . For fixed s and leakage maps f_i , put $\tilde{f}_i(z) = f_i(s + a_i z)$. The Shamir leakage probability law is therefore the law of $\tilde{f}$ on a uniform codeword of $C_d(a)$. Since every a_i is nonzero, the independent-input comparison law is $(f_i(U_i))_{i=1}^m$ for every secret s . The triangle inequality gives, for every fixed a ,

$$\text{Insec}(a; d) \leq 2\Delta(a; d). \quad (2)$$

Thus, a bound on Reed–Solomon pseudorandomness immediately gives a leakage-resilience bound for Shamir sharing. See also [DLM⁺26, Lemma 6.8, p. 48].

Theorem 1 (Relative degrees above one third). *Fix $1/3 < c < 1/2$ and $0 < \varepsilon < 1$. There are constants $C_{c,\varepsilon}, m_{c,\varepsilon} > 0$ such that, for $m \geq m_{c,\varepsilon}$, $d = \lfloor cm \rfloor$, and every odd prime satisfying*

$$\log p \geq C_{c,\varepsilon} \cdot m^2 \log(2m),$$

we have

$$\mathbb{E}_a \Delta(a; d) \leq \left(\frac{2}{\pi}\right)^{(1-\varepsilon)d}, \quad \mathbb{E}_a \text{Insec}(a; d) \leq 2 \left(\frac{2}{\pi}\right)^{(1-\varepsilon)d}. \quad (3)$$

Consequently, for every fixed $0 < u < 1/2$, if $m = \lfloor n^u \rfloor$ and $2^{n-1} < p < 2^n$, then, uniformly over these primes,

$$\mathbb{E}_a \Delta(a; d) \leq \left(\frac{2}{\pi}\right)^{(1-o(1))d}, \quad \mathbb{E}_a \text{Insec}(a; d) \leq \left(\frac{2}{\pi}\right)^{(1-o(1))d}, \quad (4)$$

where $o(1) \rightarrow 0$ as $n \rightarrow \infty$ for fixed c, u . More generally, the same asymptotic conclusion holds when $m \rightarrow \infty$ and $\log p / (m^2 \log(2m)) \rightarrow \infty$.

The restriction $c > 1/3$ comes from using quadratic structure. Using higher-order structure lets us reach $c > 1/\ell$ for any fixed $\ell \geq 4$. The exponential base stays $2/\pi$; the required field grows.

Theorem 2 (Arbitrary positive relative degrees). *For every fixed integer $\ell \geq 4$, there is a finite exponent $\Theta_\ell > 0$ with the following property. Fix $1/\ell < c < 1/2$ and $0 < \varepsilon < 1$, and put $d = \lfloor cm \rfloor$. There are constants $C_{c,\ell,\varepsilon}, m_{c,\ell,\varepsilon} > 0$ such that, for every $m \geq m_{c,\ell,\varepsilon}$ and every prime p satisfying*

$$\log p \geq C_{c,\ell,\varepsilon} \cdot m^{\Theta_\ell} \log(2m),$$

we have

$$\mathbb{E}_a \Delta(a; d) \leq \left(\frac{2}{\pi}\right)^{(1-\varepsilon)d}, \quad \mathbb{E}_a \text{Insec}(a; d) \leq 2 \left(\frac{2}{\pi}\right)^{(1-\varepsilon)d}. \quad (5)$$

If $m \rightarrow \infty$ and $\log p / (m^{\Theta_\ell} \log(2m)) \rightarrow \infty$, both quantities are at most $(2/\pi)^{(1-o(1))d}$. For n -bit primes, this holds uniformly with $m = \lfloor n^u \rfloor$ and any fixed $0 < u < u_0(\ell) = 1/\Theta_\ell$; see Equation 77 in Section 10 for a permissible Θ_ℓ .

Nearly optimal security for almost every evaluation tuple. Put $\tau = 2/\pi$ and write our asymptotic mean bound as $\mathbb{E}_a \text{Insec}(a; d) \leq \tau^{(1-\varepsilon_m)d}$, where $\varepsilon_m \rightarrow 0$. Markov's inequality gives

$$\Pr_a[\text{Insec}(a; d) > \exp(\sqrt{m}) \cdot \tau^{(1-\varepsilon_m)d}] \leq \exp(-\sqrt{m}).$$

Since $d = \lfloor cm \rfloor$ and $m = \lfloor n^u \rfloor$ for fixed $c, u > 0$, the threshold remains $\tau^{(1-o(1))d}$ and the exceptional fraction is negligible in n . Thus nearly every public evaluation tuple gives the nearly optimal security bound simultaneously for every secret pair and every family of local leakage functions. The LSB attack gives $\text{Insec}(a; d) \geq \tau^d$ for every valid evaluation tuple (Appendix A), so our bound is optimal up to $o(d)$ in the exponent.

The parameter tradeoff. The table compares arbitrary one-bit leakage under each result's respective field-size hypotheses. Our quadratic result requires $\log p \geq C_{c,\varepsilon} m^2 \log(2m)$. For smaller fixed c , choose $\ell \geq 4$ with $1/\ell < c$; the requirement becomes $\log p \geq C_{c,\ell,\varepsilon} m^{\Theta_\ell} \log(2m)$. For n -bit primes and fixed c , the number of parties may grow as a small positive power of n .

Result	Relative degree	Evaluation places	Field/party regime
Sun–Wootters [SW26]; Nguyen [Ngu26]	$c > 1/2 - \delta$, absolute $\delta > 0$	Every valid tuple	Any prime $p > m$
Our quadratic result	Fixed $1/3 < c < 1/2$	Nearly all valid evaluation tuples	$m = \lfloor n^u \rfloor$, $0 < u < 1/2$
Our higher-order result	Fixed $0 < c < 1/2$	Nearly all valid evaluation tuples	$m = \lfloor n^u \rfloor$, $1/\ell < c$, $0 < u < \Theta_\ell^{-1}$

For example, with $c = 0.4$, we may take $d = \lfloor cm \rfloor$ and $m = \lfloor n^{0.49} \rfloor$. For $c = 0.2$, choose $\ell = 6$ and take $d = \lfloor cm \rfloor$ and $m = \lfloor n^u \rfloor$ for any fixed $0 < u < \Theta_6^{-1}$. Both choices give security for all sufficiently large n , uniformly over n -bit primes.

Related work. Our question concerns a scheme that is already in use: how much protection does Shamir sharing itself provide? There is also a substantial literature on designing new secret-sharing schemes specifically for leakage resilience [SV19, ADN⁺19, KMS19, MPSW21, CKOS21, CKOS22]. In this direction, Doron, Leonov, Mosheiff, Navas, Resch, and Ribeiro [DLM⁺26, Theorem 6.5, pp. 46–47] use discrepancy bounds for random linear codes to construct locally leakage-resilient linear ramp schemes with arbitrarily small fixed positive relative reconstruction thresholds. The correlations among polynomial evaluations make Shamir sharing difficult to analyze. For additive sharing, where the shares are linked by a single sum constraint, the corresponding question is much better understood, with tight threshold estimates and refined leakage bounds in several models [BDIR21, MNP⁺21, MNP⁺22, FMM⁺24, KA25, BHM⁺26].

1.2. Technical Overview. Our proof follows a simple chain. We first control structured leakage. An approximation argument then gives Reed–Solomon pseudorandomness, and an affine change of variables gives Shamir leakage resilience:

$$\begin{array}{ccccc}
 \text{Technical} & & \text{Reed–Solomon} & & \text{Shamir} \\
 \text{Core Lemma} & \implies & \text{Pseudorandomness} & \implies & \text{Local Leakage Resilience} \\
 \text{Lemma 7.1 (Equation 6 below)} & & \text{Theorem 3} & & \text{Theorem 4}
 \end{array}$$

To explain the main ideas, we first prove the following bounds with a relaxed exponential rate. Sections 2–8 establish these statements. Section 9 sharpens the rate to prove Theorem 1. And Section 10 extends the argument to prove Theorem 2.

Working bounds for the overview

Theorem 3 (Reed–Solomon pseudorandomness). *Fix $1/3 < c < 1/2$ and put $d = \lfloor cm \rfloor$. There is a constant $C_c > 0$ such that for all sufficiently large m ,*

$$\log p \geq C_c \cdot m^2 \log(2m) \implies \mathbb{E}_a [\Delta(a; d)] \leq 2^{-d/4},$$

where a is uniform over the ordered distinct nonzero evaluation places.

Theorem 4 (Local leakage resilience of Shamir sharing). *Under the hypotheses of Theorem 3,*

$$\mathbb{E}_a \text{Insec}(a; d) \leq 2 \cdot 2^{-d/4}.$$

In particular, this holds for n -bit primes and $m = \lfloor n^u \rfloor$ whenever $0 < u < 1/2$ is fixed and n is sufficiently large.

Equation 2, proved in Section 1.1, gives the per- a insecurity bound from pseudorandomness. Taking expectations gives Theorem 4 from Theorem 3.

The technical core. For the approximation argument, we allow a leakage function to take values in $[0, 1]$: on input z , a probability map f outputs one with probability $f(z)$. The outputs at different shares are sampled independently conditional on their inputs. The original deterministic leakage functions are exactly the maps taking values in $\{0, 1\}$.

Let $\mathcal{H}$ be any nonempty family of such maps $\mathbb{F}_p \rightarrow [0, 1]$. For $f = (f_1, \dots, f_m) \in \mathcal{H}^m$, let $P_{a,f}$ be the probability law of these bits on a uniform codeword of $C_r(a)$, and let Q_f be their probability law on independent uniform inputs. Define

$$\Delta_{\mathcal{H}}(a; r) = \sup_{f_1, \dots, f_m \in \mathcal{H}} \text{SD}(P_{a,f}, Q_f).$$

Taking $\mathcal{H}$ to be all deterministic functions recovers $\Delta(a; r)$. We will construct a fixed structured family $\mathcal{G}_{p,3}$ for which the following estimate holds.

Technical core: Lemma 7.1. Fix $1/3 < c < 1/2$ and $b \in (0, 1)$, and let $\mathcal{G}_{p,3}$ be the fixed family described in Step 1 below. There are constants $A, C > 0$, depending only on b and the family data fixed by c , such that, for all sufficiently large m ,

$$\log p \geq C \cdot m^2 \log(2m), \quad bm \leq r < m \implies \mathbb{E}_a \Delta_{\mathcal{G}_{p,3}}(a; r) \leq A 2^{-r/2}. \quad (6)$$

The supremum over the maps is inside the expectation over a .

From the technical core to pseudorandomness. For $k \geq 2$ and $f: \mathbb{F}_p \rightarrow \mathbb{C}$, the Gowers uniformity norm is defined by

$$\|f\|_{U^k}^{2^k} = \mathbb{E}_{z, h_1, \dots, h_k \in \mathbb{F}_p} \prod_{\omega \in \{0,1\}^k} \mathcal{C}^{|\omega|} f\left(z + \sum_{j=1}^k \omega_j h_j\right),$$

where $\mathcal{C}$ denotes complex conjugation, $|\omega| = \sum_j \omega_j$, and all variables are independent and uniform. Small U^3 norm is the notion of quadratic uniformity used in our approximation.

The approximation described in Step 1 below lets us write each leakage function as $f_i = g_i + h_i$, where $g_i \in \mathcal{G}_{p,3}$ is the *structured part* and h_i is a *perturbation* that is pseudorandom in the U^3 sense: $\|h_i\|_{U^3} \leq \eta$. Expanding the leakage probabilities gives products of structured factors and perturbations. We split these terms according to how many perturbation factors they contain, using a cutoff t .

For a suitably chosen threshold t , the terms with at least t perturbation factors are the easy part to bound. Simultaneous Cauchy–Schwarz gains a factor η^t from t selected perturbations, making their total contribution small when η is sufficiently small. This step requires $m - t \leq 3(d - t)$.

As is typical in such estimates, the main difficulty lies in bounding the terms with few perturbations. For a term with $j < t$ perturbation factors, we condition on the field inputs at their locations and delete those coordinates; we call this operation *shortening*. The remaining inputs form a Reed–Solomon code of length $m - j$ and dimension $d - j$, up to invertible affine changes in each coordinate. The structured family is closed under these changes, so the remaining factors are still structured maps. The technical core lemma then bounds the comparison in Equation 1 for the shortened code, with the supremum restricted to structured leakage maps and the bound averaged over the evaluation places.

Since $c > 1/3$, we may take $t = \lfloor \alpha m \rfloor$ for a sufficiently small fixed $\alpha > 0$. Choosing α small ensures the Cauchy–Schwarz condition and controls the number of terms with few perturbations. Once α is fixed, choosing η small suppresses the terms with many perturbations. Together these estimates give $\mathbb{E}_a \Delta(a; d) \leq 2^{-d/4}$, proving Theorem 3. The reduction is Lemma 2.1 in Section 2.

Four steps for the core estimate. The work in Sections 3–8 establishes the core estimate. First, we fix the approximation family. Next, we see how the argument would work if its maps already followed ordinary quadratic torus orbits. We then explain the two structural lemmas that bring us to this model. Finally, we track the bounds to see how large the field must be.

Step 1. Fixing the approximation family. Nilsequences provide the structured models used in higher-order Fourier analysis. For the background and quantitative conventions used here, see Leng [Len25, Sections 2.2–2.3, pp. 17–36] and Green–Tao [GT12, Section 2 and Appendix A]. Ordinary quadratic phases are the simplest examples in the U^3 setting; nilsequences also capture bracket-type correlations. We write a nilsequence as $F(g(z)\Gamma)$: the orbit point $g(z)\Gamma$ moves on a compact space G/Γ , and a smooth readout F turns its position into an output probability.

Fix $c \in (1/3, 1/2)$ and the required accuracy $\eta = \eta(c)$. Choose a finite collection $\mathcal{N}$ of nilmanifolds of degree at most two, a dimension bound D , and a smoothness bound Q . With $R = 2D + 4$,

our family consists of maps

$$f_{\text{nil}}(z) = F(g(z)\Gamma), \quad F: G/\Gamma \rightarrow [0, 1], \quad \|F\|_{CR} \leq Q.$$

Here $G/\Gamma \in \mathcal{N}$, and $g(z)\Gamma$ is a p -periodic polynomial orbit of degree at most two. Periodicity makes the map depend only on $z \bmod p$. The norm controls derivatives through order R in fixed local coordinates. The data $\mathcal{N}, D, Q$ are fixed before choosing m, p, a or the leakage maps.

Periodic arithmetic regularity, followed by positive smoothing, gives the desired approximations [CS14, Theorem 3.4, p. 10]. Positive smoothing averages against a nonnegative smooth bump function of total mass one, preserving the range $[0, 1]$ while controlling the derivatives. The family is closed under complements and invertible affine changes of the input, as needed for the outer reduction. Section 3 and Section 4 give the definitions and the construction. In the sharp-rate argument of Section 9, the chosen accuracy and hence the family data may also depend on ε .

One good set for all leakage choices. We will work on a **Good** set of evaluation tuples on which the required certificate conclusions hold simultaneously for every allowed coefficient array, every solution, and every required shortening. Its defining constraints are independent of the leakage maps. Step 4 constructs this set and shows that its complement has exponentially small probability.

Step 2. Wishful Thinking. Let us dream for a moment.

Quadratic torus orbits. Keep the moment maps L_i defined in Section 1.1. Write $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, so $\mathbb{T}^D = \mathbb{R}^D/\mathbb{Z}^D$ is the D -dimensional torus: its coordinates are real numbers taken modulo one. Suppose our leakage maps (magically) already had the form⁴

$$f_i(z) = A_i\left(\frac{\beta_{i,1}z + \beta_{i,2}z^2}{p}\right), \quad A_i: \mathbb{T}^D \rightarrow [0, 1], \quad \beta_{i,1}, \beta_{i,2} \in \mathbb{F}_p^D.$$

For $w \in \mathbb{F}_p$, the notation $w/p \in \mathbb{T}$ uses any integer representative of w ; changing representatives changes the quotient by an integer. Here A_i is the probability-valued readout on the torus, and $\beta_{i,1}, \beta_{i,2}$ are the coefficient vectors of the linear and quadratic terms. Thus the input z determines a point on the torus, and A_i reads that point to give the probability of leaking one.

We dream of quadratic orbits because we work with U^3 ; for U^ℓ , the analogous wishful model uses polynomial orbits of degree at most $(\ell - 1)$. Suppose also that each A_i is a trigonometric polynomial of cutoff T , meaning that

$$A_i(\theta) = \sum_{k \in [-T, T]^D \cap \mathbb{Z}^D} \hat{A}_i(k) \exp(2\pi i k \cdot \theta), \quad \theta \in \mathbb{T}^D,$$

for an integer $T \geq 1$ with $\log T = \mathcal{O}(m \log(2m))$. In other words, the high frequencies have already been truncated, while the readout still takes values in $[0, 1]$. By translating A_i , we may take the torus orbit to have zero constant term. The original nilmanifolds are drawn from the fixed catalog $\mathcal{N}$, chosen using only c and the approximation accuracy. The readouts and orbit coefficients may depend on a and the leakage, within uniform dimension and cutoff bounds. Steps 3 and 4 will justify this picture, up to a small perturbation of the leakage probability laws.

Local Fourier weights. The torus map $A_i: \mathbb{T}^D \rightarrow [0, 1]$ and coefficient vectors $\beta_{i,1}, \beta_{i,2} \in \mathbb{F}_p^D$ define a function $B_i: \mathbb{F}_p^2 \rightarrow [0, 1]$ by

$$B_i(s, t) = A_i\left(\frac{\beta_{i,1}s + \beta_{i,2}t}{p}\right).$$

⁴We use the same dimension D at every share: if a map originally uses fewer coordinates, pad its coefficient vectors with zeros and let A_i ignore the extra coordinates.

Here we let the linear and quadratic inputs vary independently; the leakage map is recovered by setting $(s, t) = (z, z^2)$: $f_i(z) = B_i(z, z^2)$. We use the normalized Fourier transform on $\mathbb{F}_p^2$:

$$\widehat{B}_i(u, v) = \mathbb{E}_{s, t \in \mathbb{F}_p} B_i(s, t) e_p(-us - vt), \quad e_p(w) = \exp(2\pi i w/p).$$

We now expand B_i in these characters. Writing $\lambda_i = (u_i, v_i) \in \mathbb{F}_p^2$, we get

$$f_i(z) = B_i(z, z^2) = \sum_{\lambda_i \in \mathbb{F}_p^2} \widehat{B}_i(\lambda_i) e_p(u_i z + v_i z^2).$$

To compute $\widehat{B}_i(u_i, v_i)$, we collect all integer frequencies of A_i that produce the pair (u_i, v_i) :

$$\widehat{B}_i(u_i, v_i) = \sum_{\substack{k \in \mathbb{Z}^D, \|k\|_\infty \leq T \\ k \cdot \beta_{i,1} = u_i, k \cdot \beta_{i,2} = v_i}} \widehat{A}_i(k),$$

where the equalities in the summation condition are taken in $\mathbb{F}_p$. If $\widehat{B}_i(u_i, v_i) \neq 0$, this sum must contain at least one nonzero term. Thus (u_i, v_i) has an integer representative k with $\|k\|_\infty \leq T$. Here $\|k\|_\infty$ is the largest absolute coordinate of k ; such a vector k is a *short integer lift* of the frequency pair. These short lifts will let us apply the moment certificate, [Lemma 6.1](#), to bound the span of the supported Fourier relations below. The probability maps for leaking one and zero are represented by B_i and $1 - B_i$, respectively, restricted to $(s, t) = (z, z^2)$. To account for both possible outputs of the leaked bit, define

$$\mu_i(\lambda_i) := \left| \widehat{B}_i(\lambda_i) \right| + \left| \widehat{(1 - B_i)}(\lambda_i) \right| = \begin{cases} 1, & \lambda_i = 0, \\ 2|\widehat{B}_i(\lambda_i)|, & \lambda_i \neq 0. \end{cases}$$

At zero frequency, the two averages sum to one; at nonzero frequencies, the coefficients have equal magnitudes. The set of frequencies with positive weight will be important. Define the *local Fourier support* by

$$S_i = \text{supp } \mu_i := \{\lambda_i \in \mathbb{F}_p^2 : \mu_i(\lambda_i) > 0\}.$$

Thus S_i contains the zero frequency, since $\mu_i(0) = 1$, and exactly those nonzero frequencies that occur in the local Fourier expansion. Parseval, positivity, and conjugate symmetry give

$$\sum_{\lambda_i} \mu_i(\lambda_i)^2 \leq 2, \quad \mu_i(\lambda_i) \leq \sigma := 2^{-1/2} \quad (\lambda_i \neq 0), \quad |S_i| \leq N := (2T + 1)^D.$$

Why the full dual space is too large. For $x \in \mathbb{F}_p^r$, a frequency tuple $\lambda = ((u_i, v_i))_{i=1}^m \in (\mathbb{F}_p^2)^m$ produces the phase

$$Q_\lambda(x) = \sum_i (u_i L_i(x) + v_i L_i(x)^2), \quad \mathcal{R} = \{\lambda : Q_\lambda \equiv 0\}.$$

Here $Q_\lambda \equiv 0$ means that this polynomial vanishes identically. Writing $u = (u_i)_i$ and $v = (v_i)_i$, these are dual Reed–Solomon relations: u annihilates evaluations of polynomials of degree less than r , and v annihilates those of degree less than $2r - 1$. A relation uses share i when $(u_i, v_i) \neq (0, 0)$. Every nonzero relation uses at least $r + 1$ shares, by the Vandermonde property. Each such share contributes a factor at most σ . The dual code $C_r(a)^\perp$ consists of the vectors $u \in \mathbb{F}_p^m$ with $\sum_i u_i c_i = 0$ for every $c = (c_i)_i \in C_r(a)$. Even if we restrict to relations with $v = 0$, every dual codeword $u \in C_r(a)^\perp$ gives such a relation. There are already $p^{m-r} - 1$ nonzero relations of this form. Bounding each contribution by σ^{r+1} therefore gives

$$(p^{m-r} - 1)\sigma^{r+1},$$

which grows with p rather than giving the decay we need. A large support makes each relation's contribution small; we must also control the total Fourier weight of all contributing relations.

Only supported relations contribute. This brings us to the local-constraints viewpoint of Levi–Mosheiff–Shagrirhaya [LMS25a]: study Reed–Solomon relations subject to restrictions at each share. Here the restriction is $\lambda_i \in S_i$: a relation contributes only if each coordinate occurs in the corresponding local Fourier expansion. These supported relations need not form a vector space. To use linear-algebraic tools, we take their $\mathbb{F}_p$ -linear span:

$$V_0 = \text{Span}_{\mathbb{F}_p} \left(\mathcal{R} \cap \prod_i S_i \right), \quad s_0 = \dim_{\mathbb{F}_p} V_0.$$

Taking the span may introduce unsupported relations, but their Fourier weights are zero, so the sum we must bound is unchanged. The space $\mathcal{R}$ is determined by the code, while S_i and V_0 also depend on the chosen leakage maps from the family fixed by c . Although these objects vary with a and $(f_i)_i$, we will prove

$$\dim_{\mathbb{F}_p} V_0 = \mathcal{O}_{c,b}(1)$$

simultaneously for every $a \in \text{Good}$ and every choice of leakage maps, including choices made after seeing a .

Example: A relation that contributes nothing. Take $r = d = 2$ and any prime $p \geq 5$, and consider $L_i(x, y) = x + iy$ for $i = 1, 2, 3$, corresponding to evaluations of linear polynomials at $1, 2, 3$. Let each leakage map be $f_i(z) = (1 + \cos(2\pi z/p))/2$, with $B_i(s, t) = f_i(s)$. Then $S_i = \{(0, 0), (\pm 1, 0)\}$. The identity $L_1 - 2L_2 + L_3 = 0$ gives

$$\lambda^* = ((1, 0), (-2, 0), (1, 0)) \in \mathcal{R}, \quad \text{but } \lambda^* \notin S_1 \times S_2 \times S_3.$$

The middle map has no frequency -2 , so this relation has zero Fourier weight. In fact, $\mathcal{R}$ is the line spanned by λ^* , and none of its nonzero vectors is supported. Thus $\dim \mathcal{R} = 1$, but $s_0 = 0$.

Now change only the middle map to $f_2(z) = (1 + \cos(4\pi z/p))/2$. Its nonzero frequencies are now ± 2 , so $\pm \lambda^*$ are supported and $s_0 = 1$. The code and its full relation space stay the same, while the local Fourier supports determine which relations contribute. This is why we study Reed–Solomon relations subject to local constraints.

Bounding the supported span. Our moment certificate, Lemma 6.1, uses the short integer lifts to prove $s_0 = \mathcal{O}(1)$ on **Good**, uniformly over all leakage choices. Step 3 explains the certificate mechanism, and Step 4 tracks its field-size requirement.

Two small sets of shares determine every relation in V_0 . An information set for V_0 is a set of shares whose coordinates determine every vector in V_0 . Since s_0 is bounded and every nonzero relation uses at least $r + 1$ shares, for sufficiently large m we can choose two disjoint information sets I_1, I_2 , each of size at most s_0 . Outside their union, every nonzero relation has at least $r + 1 - 2s_0$ nonzero coordinates, each of weight at most σ . Cauchy–Schwarz and Parseval on the two information sets give

$$\sum_{\lambda \in V_0 \setminus \{0\}} \prod_i \mu_i(\lambda_i) \leq 2^{s_0} \sigma^{r+1-2s_0} = \mathcal{O}_{c,b}(2^{-r/2}).$$

This is the standard two-information-set argument [BDIR19, Lemma 4.20 and its proof, pp. 31–33]. The bounded dimension controls the summation cost; the remaining coordinates supply the exponential decay.

The remaining quadratic phases. Write $P = P_{a,f}$ and $Q = Q_f$ for the two leakage probability laws. Their all-zero Fourier terms agree. On the code, every tuple outside $\mathcal{R}$ produces a nonconstant phase of degree at most two. Its average is zero if it is linear and has magnitude at most $p^{-1/2}$ otherwise, by the quadratic Gauss-sum estimate [TV06, Lemma 4.14 and its proof, p. 162]. On independent inputs, every nonzero tuple also produces a nonconstant phase of degree at most two

and satisfies the same bound. Since every contributing relation lies in V_0 , we get

$$\begin{aligned} \text{SD}(\mathbf{P}, \mathbf{Q}) &\leq \frac{1}{2} \sum_{\lambda \in \mathcal{R} \setminus \{0\}} \prod_i \mu_i(\lambda_i) + p^{-1/2} \prod_i \sum_{\lambda_i} \mu_i(\lambda_i) \\ &= \frac{1}{2} \sum_{\lambda \in V_0 \setminus \{0\}} \prod_i \mu_i(\lambda_i) + p^{-1/2} \prod_i \sum_{\lambda_i} \mu_i(\lambda_i) \\ &\leq \mathcal{O}\left(2^{-r/2}\right) + p^{-1/2}(2N)^{m/2}. \end{aligned}$$

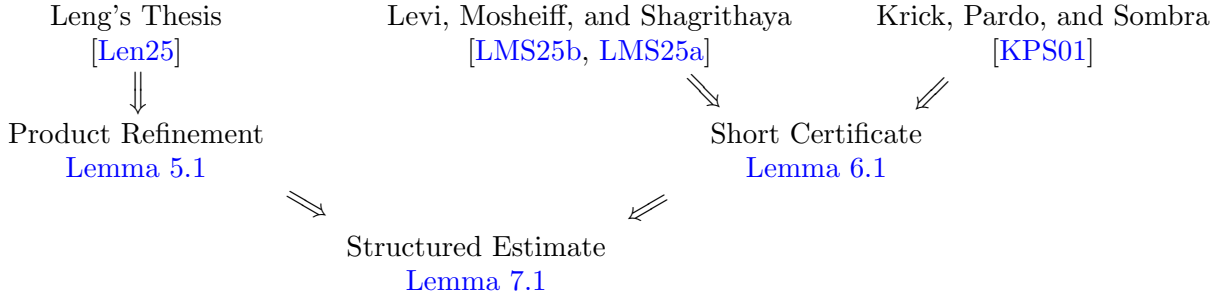
For the last term, we used $\sum_{\lambda_i} \mu_i(\lambda_i) \leq \sqrt{2N}$, where $N = (2T + 1)^D$ is the upper bound on each local Fourier support size $|S_i|$ established above. The desired bound follows once p is large enough to absorb this Gauss-sum contribution.

Recovering the sharp rate. The sharp coefficient bound from [Lemma 9.1](#) is

$$\mu_i(\lambda_i) \leq M_p := \frac{1}{p \sin(\pi/(2p))} = \frac{2}{\pi} + \mathcal{O}(p^{-2}) \quad (\lambda_i \neq 0).$$

Replacing σ by M_p gives the structured rate $\mathcal{O}((2/\pi)^r)$. Choosing a smaller fixed α in the cutoff $t = \lfloor \alpha m \rfloor$ of the outer reduction then gives the nearly matching bound for arbitrary leakage in [Theorem 1. Section 9](#) proves the coefficient bound and explains its interval extremizers; see also [\[Lev01\]](#) and [Appendix A](#).

Step 3. Reaching the wishful model. General degree-two nilsequences need not follow ordinary quadratic torus orbits. Bringing their leakage analysis to the preceding model is the main structural work. Our two principal lemmas and their sources fit together as follows.



Only the LMS branch uses randomness in the evaluation places: on **Good**, the argument holds for each fixed a uniformly over the leakage maps, and averaging over a gives the structured estimate.

Which products matter. To analyze the leakage laws $\mathbf{P}_{a,f}$ and $\mathbf{Q}_f$, we expand both output readouts, for leaking zero or one, into Fourier components in their central degree-two directions and normalize each nonzero component by its smoothness norm. Choose one normalized Fourier component at each share, from either output readout, and write $\varphi_i(z)$ for its value at input $z \in \mathbb{F}_p$. We call the resulting product *significant* if

$$\max \left\{ \underbrace{\left| \mathbb{E}_{x \in \mathbb{F}_p^r} \prod_{i=1}^m \varphi_i(L_i(x)) \right|}_{\text{codeword inputs}}, \underbrace{\left| \mathbb{E}_{U_1, \dots, U_m} \prod_{i=1}^m \varphi_i(U_i) \right|}_{\text{independent inputs}} \right\} \geq \delta,$$

where x is uniform in $\mathbb{F}_p^r$ and $U_1, \dots, U_m$ are independent uniform field elements. The second average factors as $\prod_i \mathbb{E}_{z \in \mathbb{F}_p} \varphi_i(z)$. Step 4 chooses δ exponentially small in m .

Finding global equations ([Lemma 5.1](#)). Our product refinement lemma is obtained by adapting Leng's degree-two argument and sharp bracket lemma, while keeping the product structure of the

readout [Len25, Theorem 3.3.1 and proof, pp. 51–54; Lemma 4.2.1, p. 123]. The statement holds for arbitrary linear forms, so it applies to any linear code, including Reed–Solomon codes. Each local orbit has first-level torus coordinates, obtained by forgetting its degree-two directions. These coordinates evolve affinely in the input. Writing these coordinates as $\theta \in \mathbb{T}^{d_1}$, a first-level character has the form

$$\xi_k(\theta) = k \cdot \theta \pmod{1}, \quad k \in \mathbb{Z}^{d_1}.$$

A product character adds one such character from each share. We call it *short* if its integer coefficients have magnitude at most a common cutoff B , whose size is quantified in Step 4. For each significant product, our lemma supplies short product characters that are constant modulo one on the orbit. The lemma further guarantees that, for any two infinitesimal directions preserving these equations, the sum of their vertical commutator contributions is zero. These equations therefore control the non-commutative interactions relevant to that product.

A toy bracket calculation makes the short equation visible. Write $e(t) = \exp(2\pi it)$, and let $\{t\}_c$ denote the representative of t modulo one in $(-1/2, 1/2]$. For uniform $h \in \mathbb{F}_p$, put $t_h = \{h/p\}_c$. Consider

$$F_1(h) = e(8t_h^2), \quad F_2(h) = e(-4t_h\{2t_h\}_c).$$

Factoring the combined phase gives

$$F_1(h)F_2(h) = e(4t_h \cdot (2t_h - \{2t_h\}_c)), \quad \mathbb{E}_h F_1(h)F_2(h) = \frac{1}{2} + \mathcal{O}(p^{-1}).$$

The bracket difference vanishes as a real number on the central arc $|t_h| < 1/4$, while the short torus equation $2t_h - \{2t_h\}_c = 0 \pmod{1}$ holds everywhere. We chose the bracket arguments so that this short relation is built in. The calculation illustrates the kind of relation that the general argument must discover. The general argument works for any significant nilsequence product correlation satisfying the hypotheses of Lemma 5.1: it first produces frequent bracket identities, from which Leng’s sharp bracket lemma extracts short integer constraints.

Making global equations local (Lemma 6.1). For nearly every public Reed–Solomon evaluation tuple, our short moment certificate lemma guarantees that every sufficiently large collection of short global relations has a short integer combination, with coefficients not all zero, that vanishes separately at every share. The local-profile theorem of Levi–Mosheiff–Shagrirhaya supplies locality [LMS25a, Proposition 6.1, pp. 28–29], while the arithmetic height bound of Krick–Pardo–Sombra keeps the certificate short [KPS01, Corollary 2.11, author version p. 25].

To see where the moment equations come from, write the first-level orbit at share i as $\theta_i + \beta_i z/p$, where $\beta_i \in \mathbb{F}_p^{h_i}$ and h_i is the number of first-level coordinates. Padding the local vectors with zeros when necessary, take $\beta_i \in \mathbb{F}_p^D$ and let $k_i \in \mathbb{Z}^D$ be the integer coefficient row at share i . The product character is constant on the code orbit precisely when

$$\sum_i a_i^h (k_i \cdot \beta_i) = 0 \quad (0 \leq h < r),$$

where each k_i is reduced coordinate-wise modulo p before taking the dot product.

For a toy version, work over $\mathbb{F}_p$ with $p > 5$ and take five evaluation places $a_i = i$. Define the local spaces and choose vectors in them:

$$V_i = \begin{cases} \mathbb{F}_p(2, 3)^\top, & i \leq 3, \\ \mathbb{F}_p(1, 0)^\top, & i = 4, 5, \end{cases} \quad y_i = \begin{cases} t_i(2, 3)^\top, & i \leq 3, \\ s_i(1, 0)^\top, & i = 4, 5, \end{cases}$$

where $t_1, t_2, t_3, s_4, s_5 \in \mathbb{F}_p$. Thus $y_i \in V_i \subseteq \mathbb{F}_p^2$. Suppose the global Reed–Solomon moment equations hold:

$$\sum_{i=1}^5 y_i = 0, \quad \sum_{i=1}^5 a_i y_i = 0.$$

Choose the short integer certificate $u = (3, -2)^\top$. Its transpose annihilates the common direction of the first three spaces, since $u^\top(2, 3)^\top = 0$. Applying this row to the two global equations gives

$$0 = u^\top \left(\sum_{i=1}^5 y_i \right) = 3(s_4 + s_5), \quad 0 = u^\top \left(\sum_{i=1}^5 a_i y_i \right) = 3(4s_4 + 5s_5).$$

Dividing by 3 gives $s_4 + s_5 = 0$ and $4s_4 + 5s_5 = 0$, hence $s_4 = s_5 = 0$. Thus the same certificate vanishes locally: $u^\top y_i = 0$ at every share. The tuple $(y_1, \dots, y_5)$ need not be zero: taking $(t_1, t_2, t_3, s_4, s_5) = (1, -2, 1, 0, 0)$ gives a nonzero solution of both moment equations.

In general, KPS controls the certificate's height, and LMS supplies locality at typical evaluation places. Our [Lemma 6.1](#) makes this uniform over the allowed coefficient arrays and solutions, including those chosen after the public places are known.

First impose all sufficiently short equations that hold separately at each share. The certificate shows that only a constant number of additional short global equations remain independent. The product lemma and a block-diagonal rank argument then identify a common bounded set J of exceptional shares: outside J , the components in every significant product are invariant under local commutator translations. The same J works for every output profile. We can therefore average the readouts over these translations outside J , changing both leakage laws only slightly. We then shorten the code by revealing the field inputs in J and deleting those coordinates, reducing its dimension by only a constant. The geometric reduction expresses the remaining leakage maps using ordinary quadratic torus orbits. Finally, a second certificate application bounds $\dim_{\mathbb{F}_p} V_0$, allowing the Fourier argument from Step 2 to finish the proof.

Step 4. Tracking the field size. We now track why $\log p \gg m^2 \log(2m)$ suffices. For $\alpha, \beta, h \in \mathbb{F}_p$, differencing turns a quadratic phase into an affine-linear one:

$$D_h e_p(\alpha z^2 + \beta z) = e_p(2\alpha h z + \alpha h^2 + \beta h), \quad D_h f(z) = f(z + h) \overline{f(z)}.$$

After separating the readouts into degree-two Fourier components, the corresponding derivative readouts live on ordinary tori of bounded local dimension. Fourier orthogonality then detects exact slope relations.

Differencing also preserves products. For each normalized local derivative readout $\Psi_{i,h}$, uniform smoothness gives a nonnegative Fourier envelope w_i with $|\widehat{\Psi}_{i,h}(k)| \leq w_i(k)$ for every h , and

$$\sum_{k_1, \dots, k_m} \prod_i w_i(k_i) = \prod_i \left(\sum_{k_i} w_i(k_i) \right) \leq C_0^m,$$

where C_0 depends only on the fixed family. At significance threshold $\delta = \exp(-\kappa m)$, with κ a sufficiently large fixed constant, this gives an exact bracket identity on at least an $\exp(-\mathcal{O}(m))$ fraction of shifts. Leng's sharp bracket lemma gives an integer coefficient bound B with $\log B = \mathcal{O}(m \log(2m))$ [[Len25](#), Lemma 4.2.1, p. 123]. There are only $(2B + 1)^{\mathcal{O}(m)} = \exp(\mathcal{O}(m^2 \log(2m)))$ possible short characters; our field-size bound lets us choose a line on which none of the non-constant ones becomes constant. Thus the short equations found on that line hold on the full orbit.

From short equations to Fourier cutoffs. The moment certificate and Leng's geometric estimates make the local coordinate and regularity costs polynomial in B [[Len25](#), Lemmas 2.3.11, 7.2.11, and 7.2.15]. Positive Fourier approximation therefore gives a cutoff $\log T = \mathcal{O}(m \log(2m))$, preserving the range $[0, 1]$ and making the total approximation error exponentially small in m .

Keeping the certificates valid modulo p . The certificate proof turns qualitative confinement into integer polynomial identities. Its explicit characteristic bound comes from a nonzero determinant polynomial Δ_C for each integer coefficient array C of height at most T , with

$$\|\Delta_C\|_{\text{coeff}} \leq (mD)! T^{mD}.$$

Here the norm is the largest absolute coefficient. Taking $p > (mD)!T^{mD}$ keeps Δ_C nonzero modulo p ; requiring $\Delta_C(a) \neq 0$ then preserves the needed rank at a . The logarithm of the characteristic bound is $\mathcal{O}(m^2 \log(2m))$.

One good set for every leakage choice. The defining polynomials have degree $\mathcal{O}(m^2)$, and their family is fixed by the global parameters and cutoffs. A union bound over all coefficient arrays and required shortenings gives

$$\Pr_a[a \notin \text{Good}] \leq p^{-1} \exp(\mathcal{O}(m \log(2T) + \log(2m))) \leq \exp(-2m).$$

Increasing the constant in the field-size hypothesis also absorbs the Gauss-sum error from Step 2. The good set therefore works simultaneously for every leakage choice, including choices made after a is known.

Higher-order structure. In our wishful model, when using U^3 , we had (z, z^2) terms. For U^ℓ , the torus model uses $(z, z^2, \dots, z^{\ell-1})$. The certificate and two-information-set argument still control supported relations, while repeated differencing controls the remaining polynomial phases. The larger field-size exponent comes from reaching this model: one difference need not produce linear torus motion, and Leng’s general estimate ([Imported Theorem 10.2](#)) depends on the growing total dimension. Successive restrictions and quotients also enlarge complexity. [Section 10](#) tracks these losses and proves the extension.

1.3. Organization of the Paper. We begin in [Section 2](#) by reducing arbitrary one-bit leakage to structured leakage and explaining where the threshold $c > 1/3$ enters. [Section 3](#) introduces the nilmanifold background, organized by its role in the proof, and [Section 4](#) constructs the fixed family of smooth approximations.

The next two sections establish the main structural inputs. [Section 5](#) proves [Lemma 5.1](#), extracting short character equations from significant product averages while controlling the dependence on the number of shares. [Section 6](#) proves [Lemma 6.1](#), which supplies a short combination of global moment constraints that vanishes separately at every share. [Section 7](#) develops the remaining geometric and Fourier tools needed for [Lemma 7.1](#). [Section 8](#) then proves [Lemma 7.1](#) by combining these ingredients: it removes the relevant noncommutative contributions, passes to quadratic torus orbits, and completes the Fourier argument.

[Section 9](#) proves [Theorem 1](#), sharpening the exponential base to $2/\pi$. [Section 10](#) proves [Theorem 2](#), extending the argument to $c > 1/\ell$ using U^ℓ , with a stronger field-size requirement. Finally, [Appendix A](#) proves the LSB lower bound using public reconstruction coefficients and explains its connection to interval leakage.

1.4. AI Methodology. We used GPT-6 ASTRA (Ultra) to help prepare and revise the manuscript, find relevant literature, and develop proof ideas. In particular, it brought Leng’s thesis [[Len25](#)] to our attention and suggested several ideas for turning the strategy outlined in [Section 1.2](#) into the technical arguments developed in [Section 5](#), [Section 6](#), [Section 7](#), [Section 8](#), and [Section 10](#). We have read and checked the arguments and revised the proofs to clarify the reasoning and make the presentation accessible to a broader audience. The authors take full responsibility for the correctness, originality, and attribution of the paper’s content, including its references.

2. REDUCTION FROM ARBITRARY LEAKAGE TO STRUCTURED LEAKAGE

This first step only needs an approximation class that is preserved by affine changes of the input. This lets us isolate the reason for the threshold $1/3$ before developing the structured estimate.

Retain the codeword and independent-input experiments from [Section 1.2](#), with $1 \leq r < m$. All finite averages use the uniform probability measure. Write $[m] = \{1, \dots, m\}$, $J^c = [m] \setminus J$, and $a_J = (a_i)_{i \in J}$; empty products are 1. Here P is uniform among polynomials of degree less than r , and U has independent uniform coordinates.

For a probability map $f: \mathbb{F}_p \rightarrow [0, 1]$, write

$$K_f(1 | z) = f(z), \quad K_f(0 | z) = 1 - f(z).$$

The output bits are independent conditional on their inputs. Then

$$\Delta_{\mathcal{H}}(a; r) = \sup_{f_1, \dots, f_m \in \mathcal{H}} \frac{1}{2} \sum_{y \in \{0, 1\}^m} \left| \mathbb{E}_P \prod_i K_{f_i}(y_i | P(a_i)) - \mathbb{E}_U \prod_i K_{f_i}(y_i | U_i) \right|. \quad (7)$$

The quantity $\Delta(a; r)$ is given by the same formula with the supremum taken over all maps $f_i: \mathbb{F}_p \rightarrow \{0, 1\}$.

For a finite abelian group A and a function $h: A \rightarrow \mathbb{C}$, define the Gowers U^3 norm by

$$\|h\|_{U^3(A)}^8 = \mathbb{E}_{x, u_1, u_2, u_3 \in A} \prod_{\omega \in \{0, 1\}^3} \mathcal{C}^{|\omega|} h(x + \omega_1 u_1 + \omega_2 u_2 + \omega_3 u_3), \quad (8)$$

where $\mathcal{C}$ denotes complex conjugation and $|\omega| = \omega_1 + \omega_2 + \omega_3$. The notation $\|h\|_{\infty}$ denotes the supremum norm. This is the normalized convention in [GT10, Appendix B, p. 1816].

Lemma 2.1 (Simultaneous Cauchy–Schwarz). *Fix $\eta > 0$. Let $\mathcal{H}$ be a class of probability maps closed under bijective affine precomposition: if $g \in \mathcal{H}$ and $\theta(z) = \lambda z + \mu$ with $\lambda \in \mathbb{F}_p^*$ and $\mu \in \mathbb{F}_p$, then $g \circ \theta \in \mathcal{H}$. Assume that for every $f: \mathbb{F}_p \rightarrow \{0, 1\}$ there is $g \in \mathcal{H}$ such that*

$$\|f - g\|_{U^3(\mathbb{F}_p)} \leq \eta.$$

For every ordered tuple a of distinct nonzero evaluation places, if t is an integer with $1 \leq t < r$ and $m - t \leq 3(r - t)$, then

$$\Delta(a; r) \leq \sum_{\substack{J \subseteq [m] \\ |J| < t}} 2^{|J|} \Delta_{\mathcal{H}}(a_{J^c}; r - |J|) + 4^m \eta^t. \quad (9)$$

Section 2.2 proves Lemma 2.1. Let us use it to prove Theorem 3 first. The reduction holds for every fixed evaluation tuple a ; only the structured estimate requires averaging over the places.

Throughout this section, we abbreviate $\mathcal{G}_{p,3}$ to $\mathcal{G}$. The class $\mathcal{G}$ supplied by Definition 4.1 in Section 4.1 satisfies the approximation and affine-closure hypotheses above (Proposition 4.2 in Section 4.2). Its defining geometric and smoothness bounds depend only on the fixed accuracy η , hence only on c . The remaining input is the following bound from Lemma 7.1:

$$\mathbb{E}_a \Delta_{\mathcal{G}}(a; r') \leq A 2^{-r'/2} \quad (bm' \leq r' < m', \quad \log p \geq C(m')^2 \log(2m')). \quad (10)$$

Here $0 < b < 1$ is fixed, and the estimate holds for all sufficiently large m' . The constants A, C and the lower bound on m' depend only on b and the fixed class. The tuple a is uniform among ordered distinct nonzero m' -tuples. The codeword has length m' , which may be smaller than m . Throughout, $\log$ is natural and $\log_2$ is binary. Later, we set $\mathcal{H} = \mathcal{G}$ in Lemma 2.1.

2.1. Proof of Theorem 3 using Equation 10. Set $d = \lfloor cm \rfloor$ and choose the approximation parameters (all constants)

$$s = \left(c - \frac{1}{3}\right) \in \left(0, \frac{1}{6}\right), \quad \alpha = \frac{s^2}{8} \in \left(0, \frac{1}{288}\right), \quad \eta = 2^{-3/\alpha}, \quad b = \frac{c}{2}.$$

Choose $\mathcal{G}$ at accuracy η , fix the constants A, C and the length threshold in Equation 10, and put $t = \lfloor \alpha m \rfloor$. Choose the constant C_c in Theorem 3 at least as large as C . Since $0 < \alpha < c$, we have $1 \leq t < d$ for sufficiently large m . Moreover, every shortened length $m - j$ with $0 \leq j < t$ exceeds

the threshold in Equation 10, since $m - j > (1 - \alpha)m$. The remaining hypothesis of Lemma 2.1, with $r = d$, is $m - t \leq 3(d - t)$. Indeed,

$$3d - 2t - m \geq (3c - 2\alpha - 1)m - 3 = \underbrace{\left(3s - \frac{s^2}{4}\right)}_{>0} m - 3 > 0$$

for sufficiently large m . Thus Lemma 2.1, with $\mathcal{H} = \mathcal{G}$, gives

$$\Delta(a; d) \leq \sum_{|J| < t} 2^{|J|} \Delta_{\mathcal{G}}(a_{J^c}; d - |J|) + 4^m \eta^t.$$

Next fix J with $j = |J| < t$, and put $m' = m - j$ and $r' = d - j$. The actual lower-rank condition in Equation 10 for this shortened tuple is $r' \geq bm'$. The following stronger bound is convenient:

$$\begin{aligned} d - j - bm &> d - t - bm \\ &\geq \underbrace{\left(\frac{c}{2} - \frac{s^2}{8}\right)}_{=1/6+s(4-s)/8 > 1/6} m - 1 > 0. \end{aligned}$$

Hence $r' \geq bm \geq bm'$, and $r' < m'$ because $d < m$. Also, the original field bound implies $\log p \geq C(m')^2 \log(2m')$, and the marginal tuple a_{J^c} is uniform among ordered distinct nonzero m' -tuples. The constants A, C do not change because b and $\mathcal{G}$ are fixed. Applying Equation 10 after averaging over a therefore gives

$$\begin{aligned} \mathbb{E}_a \Delta(a; d) &\leq \sum_{j < t} \binom{m}{j} 2^j A 2^{-(d-j)/2} + 4^m \eta^t \\ &= A 2^{-d/2} \sum_{j < t} \binom{m}{j} 2^{3j/2} + 4^m \eta^t. \end{aligned} \tag{11}$$

Let $h_2(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$, with $0 \log_2 0 = 0$. The elementary binomial bound gives

$$\sum_{j < t} \binom{m}{j} 2^{3j/2} \leq 2^{m(h_2(\alpha) + 3\alpha/2)} \leq 2^{m/25}.$$

The last inequality follows from the monotonicity of $h_2(\alpha)$ and $3\alpha/2$ and evaluating their sum at $\alpha = 1/288$. Consequently,

$$\mathbb{E}_a \Delta(a; d) \leq A 2^{-d/2 + m/25} + \eta^{-1} 2^{-m} \leq 2^{-d/4}$$

for sufficiently large m .

2.2. Proof of Lemma 2.1.

Notation and roadmap. Identify a polynomial of degree less than r with its coefficient vector $x = (x_0, \dots, x_{r-1}) \in \mathbb{F}_p^r$, and write

$$P_x(X) = \sum_{k=0}^{r-1} x_k X^k, \quad L_i(x) = P_x(a_i), \quad L_J = (L_i)_{i \in J}. \tag{12}$$

For $|J| \leq r$, the map $L_J: \mathbb{F}_p^r \rightarrow \mathbb{F}_p^J$ is surjective by polynomial interpolation at distinct places. Expand each leakage map into its structured approximation and its perturbation. For fewer than t perturbations, condition on their input coordinates and interpolate the remaining polynomial. For at least t perturbations, use repeated applications of Cauchy–Schwarz on vector variables, so that all t selected perturbations contribute their small U^3 norms at once.

Imported Lemma 2.2 (Box Cauchy–Schwarz; [GT10, Lemma B.2, pp. 1812–1813; (B.7), p. 1814]). *The following form uses the normalized counting measure. Let T be a nonempty finite index set and put $A = \mathbb{F}_p^T$. For $H: A \rightarrow \mathbb{C}$ and $B_1, B_2, B_3: A^2 \rightarrow \mathbb{C}$ with $\|B_j\|_\infty \leq 1$,*

$$\left| \mathbb{E}_{z_1, z_2, z_3 \in A} H(z_1 + z_2 + z_3) B_1(z_2, z_3) B_2(z_1, z_3) B_3(z_1, z_2) \right| \leq \|H\|_{U^3(A)} \quad (13)$$

The additive identification is on p. 1816 of the cited paper. Indeed, three successive applications of Cauchy–Schwarz bound the eighth power of the left-hand side by

$$\mathbb{E}_{\substack{z_j^0, z_j^1 \in A \\ 1 \leq j \leq 3}} \prod_{\omega \in \{0,1\}^3} \mathcal{C}^{|\omega|} H(z_1^{\omega_1} + z_2^{\omega_2} + z_3^{\omega_3}) = \|H\|_{U^3(A)}^8$$

For the displayed equality, set $x = z_1^0 + z_2^0 + z_3^0$ and $u_j = z_j^1 - z_j^0$; these four variables are independent and uniform in A . In particular, taking $B_1 = B_2 = B_3 = 1$ gives $|\mathbb{E}_{z \in A} H(z)| \leq \|H\|_{U^3(A)}$.

Proof of Lemma 2.1. Step 1: Expand the leakage probabilities. Fix one-bit leakage functions $f_1, \dots, f_m$. Let us decompose each leakage function f_i : choose $g_i \in \mathcal{H}$ (called the *structured approximation* of f_i) and put $v_i = f_i - g_i$ (called its *perturbation*) such that

$$\|v_i\|_{U^3} \leq \eta, \quad \|v_i\|_\infty \leq 1.$$

For each output vector $y \in \{0,1\}^m$, expand the probabilities in the two experiments using

$$K_{f_i}(y_i|z) = K_{g_i}(y_i|z) + (2y_i - 1)v_i(z)$$

Thus $P_{a,f}(y) - Q_f(y)$ (the difference between probabilities that the leaked bits equal y on a uniform codeword and on independent uniform inputs) can be written as $P_{a,f}(y) - Q_f(y) = \sum_{J \subseteq [m]} D_J(y)$ where

$$D_J(y) = \mathbb{E}_{x \in \mathbb{F}_p^r} [T_J(L_1(x), \dots, L_m(x))] - \mathbb{E}_{U \in \mathbb{F}_p^m} [T_J(U_1, \dots, U_m)].$$

where $T_J(z) = \prod_{i \notin J} K_{g_i}(y_i | z_i) \prod_{i \in J} (2y_i - 1)v_i(z_i)$; note that averaging over $x \in \mathbb{F}_p^r$ is the same as averaging over polynomials P , since $x \mapsto P_x$ (see Equation 12) is a bijection. Essentially, $D_J(y)$ is the contribution to $P_{a,f}(y) - Q_f(y)$ obtained by choosing the perturbation term $(2y_i - 1)v_i$ precisely at the coordinates in $J \subseteq [m]$. The total variation distance (Equation 7) is then at most

$$\frac{1}{2} \sum_{J \subseteq [m]} \sum_{y \in \{0,1\}^m} |D_J(y)|.$$

Step 2: Few perturbations (condition and interpolate). Suppose first that $|J| < t$. In both experiments, the inputs at the coordinates in J have the uniform law on $\mathbb{F}_p^J$: for the codeword experiment this follows from the surjectivity of L_J , since $|J| < r$. Condition on these inputs having value $u = (u_j)_{j \in J}$. Every polynomial P of degree less than r with $P(a_j) = u_j$ for $j \in J$ has a unique representation

$$P(X) = P_u(X) + M_J(X) \cdot Q(X), \quad M_J(X) = \prod_{j \in J} (X - a_j), \quad \deg Q < r - |J|,$$

where P_u is the interpolating polynomial of degree less than $|J|$. When $J = \emptyset$, set $P_u = 0$ and $M_J = 1$. Conditional on u , the polynomial Q is uniform. For each $i \notin J$, the affine map

$$\theta_{i,u}(z) = P_u(a_i) + M_J(a_i)z$$

is bijective because $M_J(a_i) \neq 0$. For $i \notin J$, put $g_{i,u} = g_i \circ \theta_{i,u} \in \mathcal{H}$. Conditional on u , the codeword experiment uses these maps on a uniform codeword of $C_{r-|J|}(a_{J^c})$. The independent-input experiment has the same output law as these maps applied to independent uniform inputs, because each $\theta_{i,u}$ is a bijection. Thus the statistical distance between the two conditional structured output laws is at most $\Delta_{\mathcal{H}}(a_{J^c}; r - |J|)$.

Let $D_{J,u}(y_{J^c})$ be the difference between these conditional structured probabilities. Then,

$$D_J(y) = \mathbb{E}_{u \in \mathbb{F}_p^J} \left[\left(\prod_{j \in J} (2y_j - 1)v_j(u_j) \right) D_{J,u}(y_{J^c}) \right].$$

The product multiplying $D_{J,u}$ has magnitude at most one. Summing over the $2^{|J|}$ choices of y_J therefore gives

$$\frac{1}{2} \sum_y |D_J(y)| \leq 2^{|J|} \mathbb{E}_u \left[\frac{1}{2} \sum_{y_{J^c}} |D_{J,u}(y_{J^c})| \right] \leq 2^{|J|} \Delta_{\mathcal{H}}(a_{J^c}; r - |J|).$$

Step 3: Many perturbations (apply Cauchy–Schwarz simultaneously). Now suppose that $|J| \geq t$. Choose $T \subseteq J$ with $|T| = t$. The hypothesis $m - t \leq 3(r - t)$ allows a disjoint partition $T^c = I_1 \sqcup I_2 \sqcup I_3$ with $|I_j| \leq r - t$. Since $|T \cup I_j| \leq r$, the evaluation map $L_{T \cup I_j}$ is surjective. Choose a linear right inverse $R_j: \mathbb{F}_p^{T \cup I_j} \rightarrow \mathbb{F}_p^r$ and define

$$W_j \mathbf{z} = R_j(\mathbf{z}, 0_{I_j}).$$

Thus $W_j: \mathbb{F}_p^T \rightarrow \mathbb{F}_p^r$ satisfies

$$L_T W_j = \text{id}_{\mathbb{F}_p^T}, \quad L_{I_j} W_j = 0.$$

Choose independently and uniformly $x^{(0)} \in \mathbb{F}_p^r$ and three vectors

$$\mathbf{z}_j = (z_{j,i})_{i \in T} \in \mathbb{F}_p^T \quad (j = 1, 2, 3).$$

Here $z_{j,i} \in \mathbb{F}_p$ is the i th coordinate of the j th vector. In the codeword average, use the substitution

$$x = x^{(0)} + W_1 \mathbf{z}_1 + W_2 \mathbf{z}_2 + W_3 \mathbf{z}_3.$$

Averaging over $x^{(0)}$ preserves the uniform law of x . For each $i \in T$, the identities $L_T W_j = \text{id}_{\mathbb{F}_p^T}$ give

$$L_i(x) = L_i(x^{(0)}) + z_{1,i} + z_{2,i} + z_{3,i}.$$

For fixed $x^{(0)}$, define $H: \mathbb{F}_p^T \rightarrow [-1, 1]$ by

$$H(\mathbf{w}) = \prod_{i \in T} v_i(L_i(x^{(0)}) + w_i), \quad \mathbf{w} = (w_i)_{i \in T} \in \mathbb{F}_p^T.$$

Thus the product of the selected perturbations is

$$\prod_{i \in T} v_i(L_i(x)) = H(\mathbf{z}_1 + \mathbf{z}_2 + \mathbf{z}_3).$$

The corresponding factors in the expansion also have the constant sign $\prod_{i \in T} (2y_i - 1)$, which does not affect their magnitude. Group the remaining factors according to I_1, I_2, I_3 . The product associated with I_j is bounded in magnitude by one and is independent of the entire vector $\mathbf{z}_j$, because $L_{I_j} W_j = 0$. For every fixed $x^{(0)}$, Equation 13 bounds the absolute value of the inner average over $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ by

$$\|H\|_{U^3(\mathbb{F}_p^T)} = \prod_{i \in T} \|v_i\|_{U^3(\mathbb{F}_p)} \leq \eta^t.$$

The equality follows from translation invariance and the independence of the coordinates of each cube variable in Equation 8. This bound is uniform in $x^{(0)}$. Averaging over $x^{(0)}$ and applying the triangle inequality therefore gives the same bound for the codeword average. In the independent-input experiment, the average factors over the coordinates, and its magnitude is at most

$$\prod_{i \in T} \mathbb{E} |v_i| \leq \prod_{i \in T} \|v_i\|_{U^3(\mathbb{F}_p)} \leq \eta^t.$$

Thus $|D_J(y)| \leq 2\eta^t$ for every J with $|J| \geq t$ and every output vector y . There are 2^m output vectors and at most 2^m such subsets, so

$$\frac{1}{2} \sum_{\substack{J \subseteq [m] \\ |J| \geq t}} \sum_{y \in \{0,1\}^m} |D_J(y)| \leq \frac{1}{2} \cdot 2^m \cdot 2^m \cdot 2\eta^t = 4^m \eta^t.$$

Combining the two cases and taking the supremum over the original one-bit leakage functions proves [Equation 9](#). $\square$

2.3. Averages and conventions. All averages and integrals over compact groups use the Haar probability measure; on a finite group, it is the uniform measure, and on a compact torus, it is the unique translation-invariant regular Borel measure of total mass one (see [\[Rud62\]](#) for more details). In particular, for $f: A \rightarrow \mathbb{C}$ on a finite group or torus A , $\|f\|_q = (\mathbb{E}_{z \in A} |f(z)|^q)^{1/q}$ for $1 \leq q < \infty$, and we say $\|f\|_\infty = \sup_{z \in A} |f(z)|$ is the supremum norm.

Put $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and write $e(t) = \exp(2\pi it)$ for $t \in \mathbb{T}$ and, for $z \in \mathbb{F}_p$, $e_p(z) = e(\tilde{z}/p)$, where $\tilde{z} \in \mathbb{Z}$ is any integer satisfying $\tilde{z} \bmod p = z$. The same convention interprets $\beta/p \in (\mathbb{T})^k$ for $\beta \in \mathbb{F}_p^k$. For real or integer coordinate vectors, $\|v\|_\infty = \max_j |v_j|$ and $\|v\|_1 = \sum_j |v_j|$. An unsubscripted coordinate norm or coefficient size means the maximum norm by default; for example, $\|v\| = \|v\|_\infty$. The maximum norm of a matrix is the maximum absolute value of its entries. For $\varphi: (\mathbb{T})^s \rightarrow \mathbb{C}$, its torus Fourier coefficient is

$$\widehat{\varphi}(k) = \int_{(\mathbb{T})^s} \varphi(t) e(-k \cdot t) dt, \quad k \in \mathbb{Z}^s. \quad (14)$$

Here φ is integrable, and $\text{Supp}(\widehat{\varphi}) = \{k \in \mathbb{Z}^s : \widehat{\varphi}(k) \neq 0\}$ denotes its Fourier support. These conventions also apply after choosing coordinates on any compact torus. The notation span denotes linear span over the indicated field.

We write $\mathbb{F}_p^J$ for vectors indexed by J . The notations $X \ll_{\mathcal{P}} Y$ and $X = O_{\mathcal{P}}(Y)$ allow constants depending only on the displayed parameters $\mathcal{P}$. The notation $o_{p \rightarrow \infty}(1)$ denotes a quantity tending to zero as $p \rightarrow \infty$, with the other indicated data fixed. Constants denoted by C may change between inequalities. Unless stated otherwise, after the local smooth class is fixed, constants may depend on its data but not on m, p, a , or the chosen leakage maps.

3. THE DEGREE-TWO STRUCTURE BEHIND THE PROOF

This section collects the degree-two nilmanifold background used in the proof. Our coordinates, metrics, and rationality conventions follow Green–Tao [\[GT12, §2 and Appendix A\]](#), while our treatment of periodic nilsequences and first-level characters follows Candela–Sisask [\[CS14, §§2,4\]](#). The vertical Fourier expansion and derivative construction are degree-two specializations of the framework in [\[GT12, §§3,7\]](#). We spell out the coordinate identities needed later and organize the material by where it enters the proof.

The geometric preliminaries are grouped by where they first enter the proof. They can be read in the following order:

- Sections [3.1–3.2](#) give the nilmanifolds and periodic orbits needed to construct the smooth class in [Section 4](#).
- Sections [3.3–3.5](#) give the additional coordinates, characters, and derivative tori needed for the product lemma in [Section 5](#).
- [Section 3.6](#) gives the rational-subgroup terminology needed for the local reduction in [Section 7](#).

Thus the reader can proceed to [Section 4](#) after [Section 3.2](#), returning for the later blocks when needed. [Section 6](#) uses the character-to-equation translation from [Section 3.4](#); its certificate proof needs no additional nilmanifold geometry. [Section 10](#) uses higher-degree analogues of these notions.

3.1. For the smooth class: Nilmanifolds, Coordinates, and Complexity. A degree-two filtered nilmanifold consists of a connected, simply connected nilpotent Lie group G , a discrete cocompact subgroup Γ , and a filtration

$$G = G_0 = G_1 \supseteq G_2 \supseteq G_3 = \{\text{id}_G\}, \quad [G_i, G_j] \subseteq G_{i+j}.$$

Here $G_j = \{\text{id}_G\}$ for $j \geq 3$, each G_j is closed and connected, and $\Gamma_j = \Gamma \cap G_j$ is cocompact in G_j . The group commutator is $[u, v] = uvu^{-1}v^{-1}$; for subgroups, $[A, B]$ denotes the subgroup generated by these commutators. In particular, G_2 is central and contains $[G, G]$.

Write $\text{Lie}(G)$ for the Lie algebra, with Lie bracket $[U, V]$, and $\exp: \text{Lie}(G) \rightarrow G$ for the exponential map, whose inverse is $\log$.

Put $d_G = \dim G$, $d_2 = \dim G_2$, and $h = d_G - d_2$. An adapted Mal'cev basis $\mathcal{X} = (X_1, \dots, X_{d_G})$ is an ordered basis of $\text{Lie}(G)$ with the following properties. Every final span $\text{span}_{\mathbb{R}}\{X_j, \dots, X_{d_G}\}$ is an ideal; the final d_2 vectors span $\text{Lie}(G_2)$; and ordered exponentials give unique coordinates in which the lattice is precisely the set of integer points:

$$\psi(\exp(t_1 X_1) \cdots \exp(t_{d_G} X_{d_G})) = (t_1, \dots, t_{d_G}), \quad \psi(\Gamma) = \mathbb{Z}^{d_G}. \quad (15)$$

We call $\psi: G \rightarrow \mathbb{R}^{d_G}$ the *Mal'cev coordinate map associated with $\mathcal{X}$* , or simply the *coordinate map*. Its components are the Mal'cev coordinates of the second kind. The ideal condition means that bracketing a vector from a final span with any Lie-algebra vector stays in that span. The coordinates in Equation 15 describe an ordered product; the coefficients of $\log g$ instead describe a single exponential $g = \exp(\sum_j s_j X_j)$. These coordinate systems need not agree.⁵

Metrics and complexity. The metric $d_{\mathcal{X}}$ is the largest right-invariant metric on G satisfying $d_{\mathcal{X}}(g, \text{id}_G) \leq \|\psi(g)\|_{\infty}$. The quotient metric and Lipschitz norm are

$$d_{G/\Gamma}(x\Gamma, y\Gamma) = \inf_{\substack{x' \in x\Gamma, \\ y' \in y\Gamma}} d_{\mathcal{X}}(x', y'),$$

$$\|F\|_{\text{Lip}(\mathcal{X})} = \|F\|_{\infty} + \sup_{x\Gamma \neq y\Gamma} \frac{|F(x\Gamma) - F(y\Gamma)|}{d_{G/\Gamma}(x\Gamma, y\Gamma)}.$$

And we say a function $F: G/\Gamma \rightarrow \mathbb{C}$ is *Lipschitz* if it is bounded and there is a constant L such that $|F(x\Gamma) - F(y\Gamma)| \leq L \cdot d_{G/\Gamma}(x\Gamma, y\Gamma)$ for all $x, y \in G$ (i.e. $\|F\|_{\text{Lip}(\mathcal{X})} < \infty$). For a reduced rational number a/b , $b > 0$, its height is $\max\{|a|, b\}$. The structure constants are defined by $[X_i, X_j] = \sum_k c_{i,j,k} X_k$; the basis is M -rational when every $c_{i,j,k}$ has height at most M . Complexity at most M bounds d_G , the filtration degree, and the rationality of the basis; for a nilsequence, it also bounds the Lipschitz norm of its readout. The dimension bound D used below always means the ordinary dimension d_G . The total dimension of the filtration is $\dim G_0 + \dim G_1 + \dim G_2 = 2d_G + d_2$.

3.2. For the smooth class: Quadratic Orbits and Periodicity. A polynomial sequence of degree at most two has the Taylor form

$$g(z) = g_0 g_1^z g_2^{\binom{z}{2}}, \quad g_0, g_1 \in G, \quad g_2 \in G_2. \quad (16)$$

We use $g \in \text{poly}(\mathbb{Z}, G_{\bullet})$ for precisely this form [CS14, Lemma 2.8, pp. 8–9].

⁵More precisely, these coordinate systems always share their first-level coordinates as $\psi_1(g) = (s_1, \dots, s_h)$, but their central coordinates can differ by commutator corrections. For example, in the Heisenberg group, take a basis (X, Y, Z) with $[X, Y] = Z$ and Z central. If $g = \exp(xX)\exp(yY)\exp(tZ)$, then $\psi(g) = (x, y, t)$, whereas the Baker–Campbell–Hausdorff formula gives

$$\log g = xX + yY + \left(t + \frac{xy}{2}\right)Z.$$

Thus the central coordinate in the ordered product differs from the corresponding coefficient of $\log g$ by a commutator correction.

The sequence is p -periodic modulo Γ when

$$g(z+p)\Gamma = g(z)\Gamma \quad (z \in \mathbb{Z}). \quad (17)$$

For odd p , this is equivalent to the exact coefficient conditions

$$g_1^p \in \Gamma, \quad g_2^p \in \Gamma_2. \quad (18)$$

Indeed, centrality gives

$$g(z)^{-1}g(z+p) = g_1^p g_2^{pz + \binom{p}{2}}.$$

If these elements lie in Γ , their values at $z = 0, 1$ first give $g_2^p \in \Gamma_2$, and then $g_1^p \in \Gamma$, since $\binom{p}{2} = p(p-1)/2$. The converse follows from the same identity. Since we use cosets of the form $g\Gamma$, periodicity is equivalent to $g(z)^{-1}g(z+p) \in \Gamma$. The expression $g(z+p)g(z)^{-1}$ is a conjugate of this element and need not belong to Γ , because Γ need not be normal.

A periodic nilsequence is the scalar function $z \mapsto F(g(z)\Gamma)$ where $F: G/\Gamma \rightarrow \mathbb{C}$ is a Lipschitz function, and we call this function F the *readout function*. We reserve g for the group-valued sequence and f_{nil} for a scalar approximation to a leakage map. The probability-valued readouts used below take values in $[0, 1]$.

Affine restriction. For integers ℓ_0, ℓ_1 ,

$$g(\ell_0 + \ell_1 z) = g(\ell_0) \left(g_1^{\ell_1} g_2^{\ell_0 \ell_1 + \binom{\ell_1}{2}} \right)^z \left(g_2^{\ell_1^2} \right)^{\binom{z}{2}}.$$

Thus affine restriction preserves degree two and p -periodicity. If $p \mid \ell_1$, then $\ell_0 + \ell_1 z \equiv \ell_0 \pmod{p}$ for every $z \in \mathbb{Z}$, so p -periodicity gives

$$g(\ell_0 + \ell_1 z)\Gamma = g(\ell_0)\Gamma \quad (z \in \mathbb{Z}).$$

This is all the orbit information needed for the class definition. The compact normalization used in [Section 5](#) is recorded in [Section 3.3](#) below.

3.3. For the product lemma: Coordinates and Normalization. Split $\psi = (\psi_1, \psi_{\text{cen}})$ into its first h and final d_2 coordinates. The first-level map ψ_1 is additive and identifies G/G_2 with $\mathbb{R}^h$. On G_2 , the map $\psi_2 = \psi_{\text{cen}}|_{G_2}$ is additive as well; it is the second-level coordinate map of Candela–Sisask [[CS14](#), p. 12, before Lemma 4.4]. Under multiplication, the first-level coordinates add, while the central coordinates acquire a bilinear correction. Thus the full coordinate map ψ need not be additive.⁶

Representatives and elementary identities. The *fractional part* of g is $\{g\} = g[g]^{-1}$, where its *integral part* $[g] \in \Gamma$ is uniquely determined by the requirement $\psi(\{g\}) \in [0, 1)^{d_G}$. For real numbers let $\{t\}_c$ denote the representative in $(-1/2, 1/2]$ modulo $\mathbb{Z}$, and use this notation coordinate-wise. There is also a centered decomposition

$$g = c_{\text{rep}} \gamma_{\text{rep}}, \quad \gamma_{\text{rep}} \in \Gamma, \quad \psi(c_{\text{rep}}) \in (-1/2, 1/2]^{d_G}.$$

Additivity at the first level gives

$$\psi_1(\gamma_{\text{rep}}) = \psi_1(g) - \{\psi_1(g)\}_c.$$

Neither $\psi(\{g\})$ nor $\psi(c_{\text{rep}})$ is in general obtained by taking coordinate-wise fractional parts of the full vector $\psi(g)$: the central coordinates include multiplication corrections.

⁶More precisely, there is a bilinear map $\beta: \mathbb{R}^h \times \mathbb{R}^h \rightarrow \mathbb{R}^{d_2}$, determined by the chosen basis, such that

$$\psi_{\text{cen}}(gg') = \psi_{\text{cen}}(g) + \psi_{\text{cen}}(g') + \beta(\psi_1(g), \psi_1(g')).$$

The correction records the central commutators produced when rearranging the product into ordered exponential form. For $g, g' \in G_2$, both first-level coordinates vanish, so the correction is zero. This explains why $\psi_2 = \psi_{\text{cen}}|_{G_2}$ is additive.

For the quadratic orbit in Equation 16, the real Mal'cev coordinates have the shape

$$\psi(g(n)) = \left(u_0 + nu_1, v_0 + nv_1 + \binom{n}{2} v_2 \right), \quad u_0, u_1 \in \mathbb{R}^h, \quad v_0, v_1, v_2 \in \mathbb{R}^{d_2}.$$

The vectors v_1, v_2 include central multiplication corrections. They need not be the central coordinates of g_1, g_2 . Coordinates on a fundamental domain are obtained only after replacing $g(n)$ by $\{g(n)\}$.

Commutators. The degree-two filtration gives

$$[\text{Lie}(G), \text{Lie}(G)] \subseteq \text{Lie}(G_2), \quad [\text{Lie}(G), \text{Lie}(G_2)] = \{0\}.$$

Thus the Lie bracket of two vectors may be nonzero, but it lies in the central subspace $\text{Lie}(G_2)$. Bracketing any vector in this subspace with another vector gives zero. In particular, all nested brackets involving three or more vectors vanish. Consequently, for $U, V \in \text{Lie}(G)$, the Baker–Campbell–Hausdorff commutator formula gives the exact identity [Len25, §2.2.1, equation (2.2.2), p. 20]

$$\exp(U) \exp(V) \exp(-U) \exp(-V) = \exp([U, V]). \quad (19)$$

We next record the group identities used to rearrange products and powers in the compact normalization below. Here $[u, v] = uvu^{-1}v^{-1}$ denotes the group commutator; its relation to the Lie bracket is given by Equation 19. Since group commutators lie in the central subgroup G_2 , for $u, v, w \in G$ and $j, k \in \mathbb{Z}$ we have

$$[uv, w] = [u, w][v, w], \quad [u^j, v^k] = [u, v]^{jk}, \quad (uv)^j = u^j [v, u]^{\binom{j}{2}} v^j.$$

Also $uvu^{-1} = [u, v]v$. These identities follow by commuting adjacent factors and collecting central commutators. In particular, conjugation acts trivially on G/G_2 , and conjugation by a lattice element preserves Γ . On products, order the Mal'cev bases by filtration level, with all G_2 -directions last.

Compact normalization. For $c_{\text{rep}} \in G$, write $F_{c_{\text{rep}}}(u\Gamma) = F(c_{\text{rep}}u\Gamma)$. Decomposing $g_0 = c_{\text{rep}}\gamma_{\text{rep}}$ absorbs the compact factor into this readout and the lattice factor by conjugating the remaining sequence. The group identities above also replace the linear Taylor coefficient by a compact representative, at the cost of changing the central quadratic coefficient. Both operations preserve the scalar nilsequence. The proof of Lemma 5.1 will spell out this normalization when it is used.

3.4. For the product lemma: First-level Equations and Vertical Frequencies. A *first-level character* is a continuous homomorphism $\xi: G \rightarrow \mathbb{R}$, meaning that $\xi(uv) = \xi(u) + \xi(v)$ for all $u, v \in G$, that vanishes on G_2 and is integer-valued on Γ . In first-level coordinates, it has the form

$$\xi(g) = k \cdot \psi_1(g), \quad k \in \mathbb{Z}^h.$$

We define its *complexity* by

$$\text{comp}(\xi) = \|k\|_1 = \sum_{j=1}^h |k_j|.$$

The associated torus is $G/(G_2\Gamma) \simeq \mathbb{T}^h$. It can be smaller than the usual horizontal torus $G/([G, G]\Gamma)$; every short equation below (in the degree-two argument⁷) uses a first-level character. In Candela–Sisask's notation, $G_1^\nabla = G_2$ and $G_2^\nabla = [G, G]$ [CS14, Definition 4.2, Lemma 4.4, and Definition 4.5, p. 12].

On a quadratic orbit, each first-level character evolves affinely:

$$\xi(g(z)) = \xi(g_0) + z\xi(g_1),$$

⁷Section 10 uses broader horizontal characters, which need not annihilate G_2

with initial value $\xi(g_0)$ and slope $\xi(g_1)$. It is therefore constant modulo 1 precisely when its slope is an integer.

For a p -periodic quadratic orbit, the first-level coordinates take the form

$$\psi_1(g(z)) = \theta + \frac{z\beta}{p} \pmod{\mathbb{Z}^h}, \quad \theta \in \mathbb{T}^h, \quad \beta \in \mathbb{F}_p^h.$$

The *initial position* is $\theta = \psi_1(g_0) \pmod{\mathbb{Z}^h}$. The *slope* β/p is the displacement at each step. Here β is the reduction modulo p of the integer vector $p\psi_1(g_1)$, whose integrality follows from [Equation 18](#).

For local groups $G^{(i)}$, set $h_i = \dim G^{(i)} - \dim G_2^{(i)}$. A product first-level character is

$$\xi((g_i)_i) = \sum_i k_i \cdot \psi_1^{(i)}(g_i), \quad k_i \in \mathbb{Z}^{h_i}, \quad \text{comp}(\xi) = \sum_i \|k_i\|_1.$$

We now specialize to the Reed–Solomon code orbit. Recall that $a_1, \dots, a_m$ are the fixed evaluation places and

$$L_i(x) = \sum_{j=0}^{r-1} \alpha_i^j x_j.$$

The i th local sequence is evaluated at $L_i(x)$. We next express constancy of a product character on this orbit as a system of moment equations. Writing θ_i and β_i/p for the local initial positions and first-level slopes, respectively, its value on the code orbit is

$$\gamma_\xi + \frac{\lambda_\xi \cdot x}{p} \pmod{1}, \quad \gamma_\xi = \sum_i k_i \cdot \theta_i, \quad (\lambda_\xi)_j = \sum_i \alpha_i^j (k_i \cdot \beta_i) \quad (0 \leq j < r).$$

Here $\gamma_\xi \in \mathbb{T}$ and $\lambda_\xi \in \mathbb{F}_p^r$. Thus constancy on the code orbit is exactly $\lambda_\xi = 0$. On the independent-input orbit, it is exactly $k_i \cdot \beta_i = 0$ for every i .

The real kernel of a family of equations. The differential of a character at the identity is

$$d\xi(v) = \left. \frac{d}{dt} \xi(\exp(tv)) \right|_{t=0}, \quad v \in \text{Lie}(G). \quad (20)$$

Let $v = \sum_j u_j X_j$ and $\xi = k \cdot \psi_1$. By the Baker–Campbell–Hausdorff formula, $\exp v$ differs from the ordered product $\exp(u_1 X_1) \cdots \exp(u_{d_G} X_{d_G})$ by a factor in G_2 . But since ψ_1 is additive and factors through G/G_2 (i.e. vanishes on G_2), these two elements have the same first-level coordinates which are, by [Equation 15](#), $\psi_1(\exp v) = (u_1, \dots, u_h)$. Hence, $\xi(\exp v) = \sum_{j \leq h} k_j u_j$. Replacing v by tv and differentiating at $t = 0$ gives the same value for $d\xi(v)$, and thus $\xi(\exp v) = d\xi(v)$ and $d\xi$ vanishes on $\text{Lie}(G_2)$. For product characters $\xi_1, \dots, \xi_t$, their common real kernel is

$$K_0 = \bigcap_{\ell=1}^t \ker(d\xi_\ell) = \left\{ v \in \bigoplus_i \text{Lie}(G^{(i)}) : d\xi_\ell(v) = 0 \text{ for every } 1 \leq \ell \leq t \right\}. \quad (21)$$

When $t = 0$, the common kernel is the entire space $\bigoplus_i \text{Lie}(G^{(i)})$. If $k_{\ell,i,j}$ are the character coefficients, these are the exact real equations

$$\sum_i \sum_{j=1}^{h_i} k_{\ell,i,j} u_{i,j} = 0 \quad (1 \leq \ell \leq t).$$

Constancy modulo 1 means that each character ξ_ℓ has the same value modulo $\mathbb{Z}$ at every point of the orbit; its real values may differ by integers. Membership in K_0 means that a Lie algebra vector v satisfies the exact real equalities $d\xi_\ell(v) = 0$ for every ℓ . Later, our product lemma produces characters that are constant modulo 1 on the orbit and establishes a commutator constraint on their common real kernel K_0 .

Fourier expansion in the central layer. The *vertical torus*

$$T_2 = G_2/\Gamma_2 \simeq \mathbb{T}^{d_2}$$

parametrizes central motion on G/Γ : moving a point by an element of T_2 leaves its first-level position unchanged. We apply ordinary Fourier analysis along these central directions.

A *vertical frequency* is a continuous homomorphism $\nu: T_2 \rightarrow \mathbb{T}$. In central Mal'cev coordinates, it has the form

$$\nu(u\Gamma_2) = k_\nu \cdot \psi_2(u) \pmod{1}, \quad k_\nu \in \mathbb{Z}^{d_2}.$$

For continuous $F: G/\Gamma \rightarrow \mathbb{C}$, its component at frequency ν is

$$F_\nu(x) = \int_{T_2} F(ux) e(-\nu(u)) du, \quad x \in G/\Gamma,$$

where du is normalized Haar measure on T_2 . Thus, at each point x , we take the usual Fourier coefficient of the function $u \mapsto F(ux)$. The resulting function has a single central frequency:

$$F_\nu(ux) = e(\nu(u)) F_\nu(x), \quad u \in T_2.$$

In particular, central translation changes its phase but preserves its magnitude.

For differentiation, we also use ν to denote the corresponding real homomorphism on G_2 :

$$\nu(u) = k_\nu \cdot \psi_2(u), \quad u \in G_2.$$

This is the unique real lift of the frequency pulled back to G_2 , and it is integer-valued on Γ_2 . For a central direction $z = \sum_{j=1}^{d_2} z_j X_{h+j}$, its differential is

$$d\nu(z) = \left. \frac{d}{dt} \nu(\exp(tz)) \right|_{t=0} = \sum_{j=1}^{d_2} (k_\nu)_j z_j.$$

The same frequency also measures the central displacement produced by commutators. Since $[v, w] \in \text{Lie}(G_2)$, we obtain the alternating bilinear form

$$(v, w) \mapsto d\nu([v, w]), \quad v, w \in \text{Lie}(G).$$

A subspace $W \subseteq \text{Lie}(G)$ is *isotropic* for this form if

$$d\nu([v, w]) = 0 \quad \text{for every } v, w \in W.$$

This means that the frequency detects no commutator displacement between directions in W ; the commutators themselves need not vanish.

Our vertical frequencies need not be second-level characters in the sense of Candela–Sisask [CS14, Definition 4.2, p. 12]. Those characters must vanish on $G_2^\nabla = [G, G]$, which would make the commutator form identically zero. Here we allow a nonzero form and study the subspaces on which it vanishes.

3.5. For the product lemma: the derivative torus and bracket relations. To compare group elements with the same first-level coordinates, define the *fiber product* (or *relative square* as [GT12, Lemma 7.4] refers it to) and its lattice by

$$G^\square = \{(u, v) \in G \times G : uv^{-1} \in G_2\}, \quad \Gamma^\square = G^\square \cap (\Gamma \times \Gamma).$$

The condition $uv^{-1} \in G_2$ says that u and v differ only by a central displacement. Thus every pair has a unique expression

$$(u, v) = (zv, v), \quad z = uv^{-1} \in G_2.$$

Here z records their relative central displacement, while v specifies their common first-level position and a central position.

First, we identify the quotient that records the common first-level position modulo the lattice. The first-level coordinate map $\psi_1: G \rightarrow \mathbb{R}^h$ is a continuous surjective homomorphism with kernel

$\ker(\psi_1) = G_2$ and $\psi_1(\Gamma) = \mathbb{Z}^h$. Consequently, $G_2\Gamma = \psi_1^{-1}(\mathbb{Z}^h)$. Indeed, if $\psi_1(g) \in \mathbb{Z}^h$, choose $\gamma \in \Gamma$ with $\psi_1(\gamma) = \psi_1(g)$. Then $g\gamma^{-1} \in G_2$, so $g \in G_2\Gamma$. The reverse inclusion is immediate. Because $\mathbb{Z}^h$ is a closed normal subgroup of the additive group $\mathbb{R}^h$, its preimage is closed and normal. Hence,

$$G/(G_2\Gamma) \simeq \mathbb{R}^h/\mathbb{Z}^h = \mathbb{T}^h.$$

We now remove simultaneous central motion of both entries. Define the diagonal central subgroup

$$\text{diag } G_2 = \{(z, z) : z \in G_2\}.$$

Multiplying (u, v) by (z, z) leaves both the relative displacement uv^{-1} and the common first-level position unchanged. After also quotienting by the lattice, we obtain the *derivative torus*

$$T^\square = G^\square/(\text{diag } G_2 \cdot \Gamma^\square) \simeq (G_2/\Gamma_2) \times G/(G_2\Gamma).$$

Explicitly, the above isomorphism is induced by the continuous surjective homomorphism $\Pi: G^\square \rightarrow (G_2/\Gamma_2) \times G/(G_2\Gamma)$,

$$(u, v) \mapsto (uv^{-1}\Gamma_2, v(G_2\Gamma)).$$

The first factor records the relative central displacement modulo the central lattice; the second records the shared first-level position modulo its lattice. Centrality of G_2 makes the map Π a homomorphism, and its kernel is exactly $\text{diag } G_2 \cdot \Gamma^\square$. Consequently,

$$\dim T^\square = d_2 + (d_G - d_2) = d_G.$$

Let $\Phi: G/\Gamma \rightarrow \mathbb{C}$ be a continuous readout with vertical frequency ν , meaning that

$$\Phi(zx) = e(\nu(z))\Phi(x) \quad (z \in G_2, x \in G/\Gamma).$$

For example, Φ may be the Fourier component F_ν defined above. Fix $c_{\text{rep}} \in G$. We construct a new readout on the derivative torus by multiplying one evaluation of Φ by the complex conjugate of another. For the fixed translation c_{rep} , define

$$\Psi_{c_{\text{rep}}}(u, v) = \Phi(c_{\text{rep}}u\Gamma)\overline{\Phi(v\Gamma)}, \quad (u, v) \in G^\square.$$

We next check that this function is well-defined on $T^\square$. This function compares the values of Φ at two points using a complex conjugate, so that their common central phase cancels.

Indeed, right multiplication by $\Gamma^\square$ preserves both cosets. Also, since G_2 is central, simultaneous translation by $z \in G_2$ gives

$$\Psi_{c_{\text{rep}}}(zu, zv) = e(\nu(z))e(-\nu(z))\Psi_{c_{\text{rep}}}(u, v) = \Psi_{c_{\text{rep}}}(u, v).$$

Thus $\Psi_{c_{\text{rep}}}$ is unchanged by the identifications defining $T^\square$, and hence defines a function on this torus.

Translating only the first entry retains one vertical phase:

$$\Psi_{c_{\text{rep}}}(zu, v) = e(\nu(z))\Psi_{c_{\text{rep}}}(u, v).$$

This translation changes the relative central displacement uv^{-1} while leaving the shared first-level position fixed. Consequently, as a function on

$$T^\square \simeq (G_2/\Gamma_2) \times G/(G_2\Gamma),$$

the function $\Psi_{c_{\text{rep}}}$ has frequency ν in the first factor.

More explicitly, every Fourier character of $T^\square$ has the form

$$(u, v) \mapsto \chi(v) + \mu(uv^{-1}) \pmod{1},$$

where χ is a first-level character and μ is a vertical frequency. The transformation rule above and Fourier orthogonality on G_2/Γ_2 imply that its Fourier coefficient vanishes unless $\mu = \nu$. Thus the only characters that can occur are

$$(u, v) \mapsto \chi(v) + \nu(uv^{-1}) \pmod{1}.$$

We have therefore passed to ordinary Fourier analysis on a torus while retaining the original vertical frequency ν . This is the frequency used to track the commutator pairing $d\nu([\cdot, \cdot])$ in the product lemma.

Integer equations from bracket relations. We first introduce the terminology needed for Leng’s sharp bracket lemma [Len25, Lemma 4.2.1, p. 123], whose exact periodic form we state as [Imported Lemma 5.2](#). This result converts a bracket relation holding on many inputs into linear relations with bounded integer coefficients. For a real subspace $\mathcal{W} \subseteq \mathbb{R}^{N_{\text{lin}}}$, $\mathcal{W}^\perp$ is its orthogonal complement for the usual dot product. Two integer-vector families are *complementary orthogonal families* when each is linearly independent over $\mathbb{R}$, every vector in one is orthogonal to every vector in the other, and their sizes add to N_{lin} . Their real spans are then orthogonal complements; no assertion about bases of integer lattices is intended. An integer vector has height at most P_{br} when every coordinate has magnitude at most P_{br} ; its ℓ_1 norm is then at most $N_{\text{lin}}P_{\text{br}}$.

An exact bracket relation is a congruence

$$\gamma_{\text{br}} + \mathbf{a}_{\text{br}} \cdot \{z\vartheta\}_{\text{c}} = 0 \pmod{1}, \quad \gamma_{\text{br}} \in \mathbb{R}, \quad \mathbf{a}_{\text{br}} \in \mathbb{R}^{N_{\text{lin}}}, \quad \vartheta \in p^{-1}\mathbb{Z}^{N_{\text{lin}}}.$$

Its density is the proportion of $z \in \mathbb{F}_p$ on which it holds. Here the word “bracket” refers to the fractional-part expression.

3.6. For the structured estimate: rational subgroups. The local reduction in [Section 7](#) restricts an orbit to a subgroup defined by local character equations. We need this subgroup to give a compact nilmanifold, and we need quantitative control of its description in the ambient coordinates.

A connected subgroup $H \leq G$ is *rational* if

$$H/(H \cap \Gamma)$$

is compact; equivalently, $H \cap \Gamma$ is a cocompact lattice in H . In coordinates, this means that $\text{Lie}(H)$ has a basis $Y_1, \dots, Y_{\dim H}$ that are rational linear combinations of the adapted Mal’cev basis $\mathcal{X} = (X_1, \dots, X_{d_G})$ of $\text{Lie}(G)$,

$$Y_\ell = \sum_{j=1}^{d_G} b_{\ell j} X_j, \quad b_{\ell j} \in \mathbb{Q}.$$

Thus rationality requires rational coefficients for a basis of directions spanning $\text{Lie}(H)$ over $\mathbb{R}$. For example, in $G = \mathbb{R}^2$ with $\Gamma = \mathbb{Z}^2$, the line $H = \mathbb{R}(1, 2)$ is rational: its lattice is $\mathbb{Z}(1, 2)$, and the quotient is a circle.

For $M \geq 1$, we call H *M-rational relative to $\mathcal{X}$* if $\text{Lie}(H)$ has a basis whose vectors, when expanded in the ambient basis $\mathcal{X}$, use only rational coefficients of height at most M . Here the height of a reduced fraction r/q , with $q > 0$, is $\max\{|r|, q\}$. Thus M bounds both the numerators and denominators needed to describe these directions [GT12, Definition 1.10, p. 471; Definitions 2.3–2.5, pp. 479–480].

These coefficients describe how $\text{Lie}(H)$ sits inside $\text{Lie}(G)$. The rationality bound for the ambient basis $\mathcal{X}$, defined in [Section 3.1](#), instead bounds the structure constants in the brackets $[X_i, X_j]$. When the coefficients $b_{\ell j}$ are integers, the height bound is simply $|b_{\ell j}| \leq M$, as in the conventions of [Section 3.5](#).

4. A FIXED CLASS OF SMOOTH APPROXIMATIONS

We now construct the fixed approximation family described in the introduction. Periodic arithmetic regularity gives the initial approximations [CS14, Theorem 3.4, p. 10]; positive smoothing preserves their range $[0, 1]$ and gives uniform derivative bounds. All geometric and smoothness data are fixed before m, p , and the leakage maps are chosen.

The Fourier and central-averaging arguments follow the vertical Fourier framework of [GT12, Definitions 3.4–3.5 and the proof of Lemma 3.7, pp. 485–486]. The subsections have the following roles:

- Section 4.1 defines the class from a fixed finite collection of nilmanifolds and a fixed smoothness bound.
- Section 4.2 chooses these data and proves the approximation and closure properties required by the reduction in Section 2.
- Section 4.3 gives the Fourier envelopes and shift identities used by the product lemma in Section 5.
- Section 4.4 gives the vertical decomposition and central averaging used in the structured estimate of Section 7.

4.1. The smooth class. Fix a nonempty finite collection $\mathcal{N}$ of degree-two filtered nilmanifolds $(G/\Gamma, G_\bullet, \mathcal{X})$ with adapted Mal'cev bases. Let $D \geq 1$ be an integer bounding their ordinary dimensions, and fix $Q \geq 1$. For differentiable F and $X \in \text{Lie}(G)$, put

$$D_X F(x\Gamma) = \left. \frac{d}{dt} F(\exp(tX)x\Gamma) \right|_{t=0}.$$

Think of X as specifying a direction of motion. Then $D_X F(x\Gamma)$ measures the slope of F at $x\Gamma$ along the curve $t \mapsto \exp(tX)x\Gamma$, just as a directional derivative measures slope in Euclidean space. For a nonnegative integer R , let $C^R(G/\Gamma; \mathbb{C})$ denote the space of complex-valued functions on G/Γ that are R times continuously differentiable. We write $C^R(G/\Gamma; [0, 1])$ for those functions whose values lie in $[0, 1]$. For $F \in C^R(G/\Gamma; \mathbb{C})$, define

$$\|F\|_{C^R(\mathcal{X})} = \max_{\substack{0 \leq j \leq R \\ X_{i_1}, \dots, X_{i_j} \in \mathcal{X}}} \|D_{X_{i_1}} \cdots D_{X_{i_j}} F\|_\infty,$$

where the order-zero derivative is F itself. Thus $\|F\|_{C^R(\mathcal{X})} \leq Q$ means that F and all its iterated directional derivatives of order at most R are uniformly bounded in absolute value by Q . Informally, this controls both the size of the readout and how rapidly it varies along the basis directions. Set

$$K = D + 2, \quad R = 2D + 4.$$

We use K derivatives for Fourier analysis on the derivative tori, and the additional $R - K = D + 2$ derivatives to make the vertical components summable in the C^K norm.

For the fixed finite collection $\mathcal{N}$, there exists a constant $C_{\text{Lip}} \geq 1$ such that

$$\|F\|_{\text{Lip}(\mathcal{X})} \leq C_{\text{Lip}} \|F\|_{C^1(\mathcal{X})} \leq C_{\text{Lip}} \|F\|_{C^R(\mathcal{X})}.$$

By [GT12, Lemma A.2, Lemma A.4], whenever $d_{\mathcal{X}}(x, y)$ is sufficiently small, the vector $Z = \log(xy^{-1})$ satisfies $x = \exp(Z)y$ and

$$\|Z\|_1 \ll \|\psi(xy^{-1})\|_\infty \ll d_{\mathcal{X}}(xy^{-1}, \text{id}_G) = d_{\mathcal{X}}(x, y),$$

where $\|Z\|_1$ is the sum of the absolute values of its coefficients in $\mathcal{X}$. Consequently,

$$\begin{aligned} |F(x\Gamma) - F(y\Gamma)| &\leq \int_0^1 |D_Z F(\exp(sZ)y\Gamma)| ds \\ &\leq \|F\|_{C^1(\mathcal{X})} \|Z\|_1 \\ &\ll \|F\|_{C^1(\mathcal{X})} d_{\mathcal{X}}(x, y). \end{aligned}$$

For distances above the fixed local threshold, the same estimate follows from $|F(x\Gamma) - F(y\Gamma)| \leq 2\|F\|_\infty$. Taking the infimum over representatives gives the Lipschitz bound for the quotient metric. Including the supremum norm proves the first inequality and the second follows from $R \geq 1$. Since $\mathcal{N}$ is finite, the constant C_{Lip} can be chosen uniformly. Section 4.2 will choose $\mathcal{N}, D, Q$ depending only on c so that the resulting class approximates every one-bit leakage function.

For $z \in \mathbb{Z}$, write $\bar{z} = z + p\mathbb{Z} \in \mathbb{F}_p$ for its residue class modulo p .

Definition 4.1 (Class $\mathcal{G}_{p,3}$). For an odd prime p , with $\mathcal{N}, R, Q$ fixed, let $\mathcal{G} = \mathcal{G}_{p,3}$ consist of all probability maps $f_{\text{nil}}: \mathbb{F}_p \rightarrow [0, 1]$ of the form

$$f_{\text{nil}}(\bar{z}) = F\left(g_0 g_1^z g_2^{\binom{z}{2}} \Gamma\right),$$

where $(G/\Gamma, G_\bullet, \mathcal{X}) \in \mathcal{N}$ and

$$g_0, g_1 \in G, \quad g_2 \in G_2, \quad g_1^p \in \Gamma, \quad g_2^p \in \Gamma_2,$$

while $F \in C^R(G/\Gamma; [0, 1])$ and $\|F\|_{C^R(\mathcal{X})} \leq Q$. The two power conditions make this independent of the integer representative z of $\bar{z}$. The subscript 3 records the U^3 norm; the corresponding structured approximations use nilmanifolds of filtration degree at most $3 - 1 = 2$.

4.2. For the reduction: approximation and closure. We apply Candela–Sisask’s periodic regularity lemma [CS14, Theorem 3.4, p. 10], then smooth the readouts and verify the required closure properties, with all bounds depending only on c . Recall the choices

$$\alpha = \frac{(c - 1/3)^2}{8}, \quad \eta = 2^{-3/\alpha} = 2^{-24/(c-1/3)^2}.$$

Proposition 4.2 (Approximation and closure). *There is a choice of $\mathcal{N}, D, Q$, depending only on c , such that for every odd prime p and every one-bit leakage function $f: \mathbb{F}_p \rightarrow \{0, 1\}$, there is $f_{\text{nil}} \in \mathcal{G}_{p,3}$ with*

$$\|f - f_{\text{nil}}\|_{U^3} < \eta.$$

The class is closed under complements and bijective affine precompositions. These closure properties hold for every choice of the class data with $Q \geq 1$.

Proof roadmap. Periodic arithmetic regularity first gives a Lipschitz approximation on a nilmanifold of bounded complexity. There are only finitely many possible based nilmanifolds at that complexity. Positive smoothing then gives a uniform C^R bound while preserving the range $[0, 1]$.

Imported Theorem 4.3 (Periodic arithmetic regularity; [CS14, Theorem 3.4, p. 10]). *Fix $\varepsilon > 0$ and a growth function $\mathcal{F}: \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$. There exists $M_0 = M_0(\varepsilon, \mathcal{F})$ such that, for every odd prime p and every function $f: \mathbb{F}_p \rightarrow [0, 1]$, there is a number $0 < M_f < M_0$ and a decomposition*

$$f = f_{\text{nil}} + f_{\text{sml}} + f_{\text{unf}}, \quad \|f_{\text{sml}}\|_2 \leq \varepsilon, \quad \|f_{\text{unf}}\|_{U^3} \leq 1/\mathcal{F}(M_f).$$

The function f_{nil} is a p -periodic polynomial nilsequence of degree at most two and complexity at most M_f , taking values in $[0, 1]$. The bound M_0 is independent of p and f .

This is the $s = 2$ case over the characteristic-zero family of prime moduli; see Definition 3.1 and Theorem 3.3 on p. 10 of the same source. Its complexity conventions are Definitions 2.5 and 2.9, pp. 7 and 9.

Imported Lemma 4.4 (Gowers norm versus L^2 ; [ET12, equations (5)–(6), p. 3]). *For a finite abelian group A with normalized counting measure, $\|h\|_{U^3(A)} \leq \|h\|_2$.*

Proof of Proposition 4.2. Step 1: Choose the geometric data. Apply periodic arithmetic regularity with

$$s = 2, \quad \varepsilon_{\text{reg}} = \eta/4, \quad \mathcal{F}_{\text{reg}}(t) = 4(1 + t)/\eta. \quad (22)$$

Let $M_0 = M_0(c)$ be the bound supplied by this application. The regularity lemma is deduced from the periodic inverse theorem, [CS14, Theorem 3.3 and the discussion preceding Theorem 3.4, p. 10]; in particular, M_0 is fixed independently of m, p and the leakage maps.

The conclusion we use is the following. For every odd prime p and every one-bit leakage function $f: \mathbb{F}_p \rightarrow \{0, 1\}$ there is a decomposition

$$f = f_{\text{nil}} + f_{\text{sml}} + f_{\text{unf}}, \quad \|f_{\text{sml}}\|_2 \leq \frac{\eta}{4}, \quad \|f_{\text{unf}}\|_{U^3} \leq \frac{\eta}{4(1 + M_f)}.$$

in which $f_{\text{nil}}(\bar{z}) = F_0(g(z)\Gamma)$ is a p -periodic nilsequence of degree at most two and complexity at most M_0 . Here $F_0: G/\Gamma \rightarrow \mathbb{C}$ is the Lipschitz readout supplied by the regularity lemma, with $\|F_0\|_{\text{Lip}(\mathcal{X})} \leq M_0$. The subscript distinguishes it from the smooth readout F constructed in Step 2. Thus, the dimension and the rationality of the adapted Mal'cev basis, as well as $\|F_0\|_{\text{Lip}(\mathcal{X})}$, are at most M_0 ; see [CS14, Definitions 2.5 and 2.9, pp. 7, 9]. By [CS14, Theorem 3.4(iv), p. 10], $F_0(g(z)\Gamma) \in [0, 1]$ for every z .

The range condition holds along the orbit, but need not hold elsewhere on G/Γ . We therefore *clip* the readout, replacing F_0 by

$$x\Gamma \mapsto \min\{1, \max\{0, \text{Re}(F_0(x\Gamma))\}\}, \quad x \in G.$$

Clipping fixes every value already in $[0, 1]$, so it leaves the orbit values unchanged. It also does not increase the supremum norm or the Lipschitz constant. Keeping the notation F_0 for the clipped readout, we now have $F_0: G/\Gamma \rightarrow [0, 1]$ on the whole nilmanifold. Clipping need not preserve differentiability; the required smoothness will be obtained in Step 2.

Consequently,

$$\|f - f_{\text{nil}}\|_{U^3} \leq \|f_{\text{sml}}\|_2 + \|f_{\text{unf}}\|_{U^3} \leq \frac{\eta}{4} + \frac{\eta}{4(1 + M_f)} < \frac{\eta}{2}.$$

There are only finitely many filtered nilmanifolds of complexity at most M_0 , up to isomorphisms preserving the adapted Mal'cev basis $\mathcal{X} = (X_1, \dots, X_{d_G})$. Indeed, the dimension satisfies $d_G \leq M_0$, and each structure constant in

$$[X_i, X_j] = \sum_k c_{i,j,k} X_k$$

is a rational number whose numerator and denominator have absolute value at most M_0 . Thus there are only finitely many possible bracket tables. Each valid table determines the connected, simply connected Lie group G together with its specified basis $\mathcal{X}$, up to a basis-preserving isomorphism. Here, valid table means a table that is antisymmetric, satisfies the Jacobi identity, defines a Lie algebra and hence determines the connected, simply connected Lie group G together with its basis $\mathcal{X}$, up to basis-preserving isomorphism.

There are also only finitely many filtration choices: in degree two, $\text{Lie}(G_2)$ is the span of the final d_2 basis vectors of $\mathcal{X}$ for some $0 \leq d_2 \leq d_G$. The lattice introduces no further choice. If $\mathcal{X}$ is adapted to Γ , then by the adapted-basis convention that the lattice consists exactly of the points with integer Mal'cev coordinates, by Equation 15,

$$\Gamma = \{\exp(n_1 X_1) \cdots \exp(n_{d_G} X_{d_G}) : n_1, \dots, n_{d_G} \in \mathbb{Z}\},$$

so Γ is determined by $\mathcal{X}$, by the adapted-basis convention that the lattice consists exactly of the points with integer Mal'cev coordinates [CS14, Definition 2.5, p. 7; Definition A.1, p. 27].

Fix a finite collection $\mathcal{N}$ containing one representative of each such isomorphism class of degree at most two, and put

$$D = \max\{1, \lceil M_0 \rceil\}, \quad L = \max\{1, M_0\}. \quad (23)$$

Then D bounds the dimensions of these nilmanifolds, and L bounds the Lipschitz norms of the clipped readouts F_0 . This bound controls differences between function values and does not require differentiability. All these choices depend only on c . The finiteness of $\mathcal{N}$ will let us choose the smoothing constants over $\mathcal{N}$ uniformly in Step 2. Only the smoothness bound Q remains to be chosen.

Step 2: Choose the smoothness bound. We smooth F_0 by taking weighted averages of nearby values. For each member of $\mathcal{N}$, choose a nonnegative smooth bump on G supported near the identity. For $0 < \tau \leq 1$, shrink its support by a factor of τ in Mal'cev coordinates and divide by its Haar integral. The resulting smooth kernel κ_τ satisfies

$$\kappa_\tau \geq 0, \quad \int_G \kappa_\tau(u) du = 1,$$

where du denotes left Haar measure on G . These conditions ensure that averaging values in $[0, 1]$ keeps the result in $[0, 1]$. Define

$$(S_\tau F_0)(x\Gamma) = \int_G \kappa_\tau(u) F_0(u^{-1}x\Gamma) du.$$

Then $S_\tau 1 = 1$ and $0 \leq S_\tau F_0 \leq 1$. For $u \in \text{Supp}(\kappa_\tau)$, right invariance and symmetry of $d\mathcal{X}$ give

$$\begin{aligned} d_{G/\Gamma}(u^{-1}x\Gamma, x\Gamma) &\leq d\mathcal{X}(u^{-1}x, x) \\ &= d\mathcal{X}(\text{id}_G, u) = d\mathcal{X}(u, \text{id}_G) \\ &\leq \|\psi(u)\|_\infty = O(\tau). \end{aligned}$$

The Lipschitz bound therefore gives $\|S_\tau F_0 - F_0\|_\infty = O(L\tau)$. For differentiation, the substitution $u = \exp(tX)v$, justified by left invariance of Haar measure, gives

$$(S_\tau F_0)(\exp(tX)x\Gamma) = \int_G \kappa_\tau(\exp(tX)v) F_0(v^{-1}x\Gamma) dv.$$

Thus, differentiation falls entirely on the smooth kernel. Repeating this argument transfers all iterated derivatives to the kernel, whose derivatives of order j have $L^1(G)$ norm $O_j(\tau^{-j})$ by coordinate rescaling.

Hence, there are constants $c_{\text{sm}}, C_{\text{sm},R} \geq 1$, depending on the fixed geometric data and chosen bump functions (and also on R for $C_{\text{sm},R}$), independent of τ and uniform over the finite collection $\mathcal{N}$, such that

$$\|S_\tau F_0 - F_0\|_\infty \leq c_{\text{sm}} L \tau, \quad \|S_\tau F_0\|_{C^R(\mathcal{X})} \leq C_{\text{sm},R} \tau^{-R}.$$

Take

$$\tau = \frac{\eta}{4c_{\text{sm}}L}, \quad Q = \left\lceil C_{\text{sm},R} \left(\frac{4c_{\text{sm}}L}{\eta} \right)^R \right\rceil, \quad F = S_\tau F_0.$$

The quadratic Taylor expansion and the periodicity criterion from [Section 3.2](#) show that g has the form and periodicity required in the definition of $\mathcal{G}$. The nilmanifold belongs to $\mathcal{N}$, and the smoothed readout F takes values in $[0, 1]$ with $\|F\|_{C^R(\mathcal{X})} \leq Q$. Thus the probability map $\bar{z} \mapsto F(g(z)\Gamma)$ belongs to $\mathcal{G}$. Moreover,

$$\|f - F(g(\cdot)\Gamma)\|_{U^3} \leq \|f - f_{\text{nil}}\|_{U^3} + \|F - F_0\|_\infty < \frac{\eta}{2} + \frac{\eta}{4} < \eta.$$

Thus every one-bit leakage function has the required approximation in $\mathcal{G}$. All data defining the class depend only on c .

Step 3: Check the closure properties. For complements, replace F by $1 - F$. Its order-zero norm is at most $1 \leq Q$, and every positive-order derivative changes only its sign. For a bijective affine precomposition, replace $g(z)$ by $g(az + b)$, where $a, b \in \mathbb{Z}$ and $p \nmid a$. By [Section 3.2](#), this orbit is still quadratic and p -periodic. The readout F , its norm, and the underlying nilmanifold are unchanged. $\square$

From now on, fix the class data supplied by this proposition. All implicit constants below may depend on these data, but not on m, p , the evaluation places, or the chosen leakage maps.

4.3. For the product lemma: Fourier envelopes and shift correlations. The product lemma turns a large average into shift correlations, then applies ordinary Fourier analysis to each local factor. The derivative tori have dimension at most D . We record the Fourier bounds needed for this argument; the choice $K = D + 2$ will suffice.

From averages to shift correlations. For a test function $F_{\text{test}}: \mathbb{F}_p \rightarrow \mathbb{C}$, define its *multiplicative derivative* at shift $h \in \mathbb{F}_p$ by

$$\mathcal{D}_h F_{\text{test}}(z) = F_{\text{test}}(z+h) \overline{F_{\text{test}}(z)}.$$

This compares values at two inputs separated by h . The operator $\mathcal{D}_h$ is distinct from the differential operator D_X . Since $(z, z+h)$ is a uniformly distributed pair when z, h are independent and uniform in $\mathbb{F}_p$, we have

$$\left| \mathbb{E}_{z \in \mathbb{F}_p} F_{\text{test}}(z) \right|^2 = \mathbb{E}_{h \in \mathbb{F}_p} \mathbb{E}_{z \in \mathbb{F}_p} \mathcal{D}_h F_{\text{test}}(z).$$

Thus, a large average gives a large average of the shift correlations. For the nilsequence factors in the product lemma, the construction in [Section 3.5](#) expresses these correlations using readouts on derivative tori.

A Fourier bound uniform over the shifts. For each $h \in \mathbb{F}_p$, let $\varphi_h: \mathbb{T}^{d_{\text{tor}}} \rightarrow \mathbb{C}$, where $d_{\text{tor}} \leq D$. In the application, these are the local derivative readouts, one for each shift h . Assume that their coordinate derivatives through order K , including the functions themselves, are bounded by a common constant independent of h .

Their Fourier coefficients are

$$\widehat{\varphi}_h(k) = \int_{\mathbb{T}^{d_{\text{tor}}}} \varphi_h(t) e(-k \cdot t) dt, \quad k \in \mathbb{Z}^{d_{\text{tor}}},$$

where dt is the Haar probability measure. A *Fourier envelope* is a single nonnegative function $\mathfrak{w}: \mathbb{Z}^{d_{\text{tor}}} \rightarrow [0, \infty)$ that bounds these coefficients for every shift:

$$|\widehat{\varphi}_h(k)| \leq \mathfrak{w}(k) \quad \text{for every } h \in \mathbb{F}_p, k \in \mathbb{Z}^{d_{\text{tor}}}.$$

Smoothness supplies such an envelope. For $k \neq 0$, choose a coordinate j with $|k_j| = \|k\|_\infty$. Integrating by parts K times in that coordinate gives

$$|\widehat{\varphi}_h(k)| \leq \frac{\left\| \partial_j^K \varphi_h \right\|_\infty}{(2\pi \|k\|_\infty)^K}.$$

The zero coefficient is bounded by $\|\varphi_h\|_\infty$. Consequently, for a suitable constant C independent of h , we may take

$$\mathfrak{w}(k) = C(1 + \|k\|_\infty)^{-K}.$$

Since there are $\mathcal{O}(r^{d_{\text{tor}}-1})$ integer frequencies with $\|k\|_\infty = r$, summing over these shells gives

$$\sum_k \mathfrak{w}(k) \ll 1, \quad \sum_{\|k\|_\infty > T_{\text{der}}} \mathfrak{w}(k) \ll T_{\text{der}}^{d_{\text{tor}}-K} \leq T_{\text{der}}^{-2} \quad (T_{\text{der}} \geq 1).$$

The last inequality uses $d_{\text{tor}} \leq D$ and $K = D + 2$. Thus the total Fourier mass is uniformly bounded, and discarding frequencies above T_{der} costs at most $\mathcal{O}(T_{\text{der}}^{-2})$ in the supremum norm.

The constants depend only on the common derivative bound and the fixed torus geometry. The Fourier series therefore converge absolutely and uniformly, so finite products may be expanded and averaged term by term.

To apply these estimates to the derivative readouts from [Section 3.5](#), we verify the required coordinate-derivative bound explicitly. For a readout $\Phi \in C^K(G/\Gamma)$ with vertical frequency ν , put

$$a(t) = \exp(t_1 X_1) \cdots \exp(t_h X_h), \quad z(s) = \exp\left(\sum_{k=1}^{d_2} s_k X_{h+k}\right),$$

where $t \in \mathbb{R}^h$ and $s \in \mathbb{R}^{d_2}$. Under the product-torus identification of [Section 3.5](#), the pair

$$(z(s)a(t), a(t))$$

represents the coordinates $(s \bmod \mathbb{Z}^{d_2}, t \bmod \mathbb{Z}^h)$. Consequently, the derivative readout has the $\mathbb{Z}^{d_G}$ -periodic lift

$$\tilde{\Psi}_{c_{\text{rep}}}(s, t) = \Phi(c_{\text{rep}} z(s) a(t) \Gamma) \overline{\Phi(a(t) \Gamma)}.$$

Let $\mathcal{C} \subseteq G$ be a fixed compact set containing the chosen shift representatives c_{rep} . For $c_{\text{rep}} \in \mathcal{C}$ and $(s, t) \in [0, 1]^{d_G}$, the maps appearing in the expression above have uniformly bounded derivatives with respect to (s, t) . Expressing their derivatives in the smooth frames $D_{X_1}, \dots, D_{X_{d_G}}$ gives coefficients whose required derivatives are uniformly bounded on this compact domain. Repeated application of the product and chain rules yields

$$\sup_{c_{\text{rep}} \in \mathcal{C}} \max_{|\alpha| \leq K} \left\| \partial^\alpha \tilde{\Psi}_{c_{\text{rep}}} \right\|_\infty \leq C \|\Phi\|_{C^K(\mathcal{X})}^2.$$

The constant depends only on the fixed geometric data, K and $\mathcal{C}$, and can be chosen uniformly over the finite collection $\mathcal{N}$. Thus, for readouts with uniformly bounded $C^K(\mathcal{X})$ norm, the required coordinate-derivative bound is independent of the shift and the vertical frequency.

Averaging the Fourier modes. For a phase slope $\lambda_{\text{ph}} \in p^{-1}\mathbb{Z}$, the function $z \mapsto e(z\lambda_{\text{ph}})$ is p -periodic. Character orthogonality gives

$$\mathbb{E}_{z \in \mathbb{F}_p} e(z\lambda_{\text{ph}}) = \begin{cases} 1, & \lambda_{\text{ph}} \in \mathbb{Z}, \\ 0, & \lambda_{\text{ph}} \notin \mathbb{Z}. \end{cases}$$

Along a p -periodic affine-linear torus orbit, each Fourier mode has a fixed initial phase and a slope of this form. Averaging retains the initial phase factor when the slope is integral, and gives zero otherwise. This converts the Fourier expansion of the shift correlations into exact conditions on the phase slopes.

4.4. For the structured estimate: Vertical Decomposition and Central Averaging. [Section 7](#) expands each readout into vertical components, applies the product lemma to normalized components, and later averages over central subtori. The fixed C^R bound makes the vertical components summable in C^K and is preserved under central averaging.

Summable vertical components. Use the vertical components F_ν from [Section 3.4](#), with ν identified with its integer coefficient vector in the central Mal'cev coordinates. Then

$$F = \sum_{\nu} F_\nu, \quad \sum_{\nu} \|F_\nu\|_{C^K(\mathcal{X})} \leq C_{\mathcal{N}, R} \cdot \|F\|_{C^R(\mathcal{X})}, \quad (24)$$

with convergence in C^K . In particular, the sum is bounded uniformly for all readouts F in the class.

To prove this, central derivatives commute with every D_X . For any derivative of order at most K and any nonzero ν , integrate by parts $R - K = D + 2$ times in a central coordinate attaining $\|\nu\|_\infty$. The resulting derivative of F has order at most R , so

$$\|F_\nu\|_{C^K(\mathcal{X})} \ll_{\mathcal{N}, R} \|F\|_{C^R(\mathcal{X})} (1 + \|\nu\|_\infty)^{-(D+2)}.$$

For $\nu = 0$ the same estimate follows directly from the integral. Since $\dim T_2 \leq D$, these bounds are summable. The Fourier series on each central fiber recovers F , and the derivative bounds give convergence in C^K . If T_2 is trivial, the expansion has just its zero-frequency term.

Positive central averaging. Let $Z \leq T_2$ be a closed subgroup, equipped with Haar probability measure, and define

$$(A_Z F)(x\Gamma) = \int_Z F(zx\Gamma) dz.$$

The action is well defined because G_2 is central. Positivity and normalization give

$$0 \leq F \leq 1 \implies 0 \leq A_Z F \leq 1, \quad A_Z(1 - F) = 1 - A_Z F.$$

Commuting each D_X with the central action also gives $\|A_Z F\|_{C^R(\mathcal{X})} \leq \|F\|_{C^R(\mathcal{X})}$. Finally, character orthogonality shows that A_Z retains exactly the vertical components whose characters are trivial on Z , meaning $e(\nu(z)) = 1$ for every $z \in Z$:

$$A_Z F = \sum_{\nu: \nu|_Z=0} F_\nu.$$

These are the range, smoothness, and Fourier-support properties used when [Section 7](#) averages over commutator tori.

5. FROM A LARGE PRODUCT AVERAGE TO SHORT EQUATIONS

This section proves the product refinement lemma ([Lemma 5.1](#)) for arbitrary linear forms $L_i: \mathbb{F}_p^r \rightarrow \mathbb{F}_p$. A large average of smooth readouts, each with a single vertical frequency, forces short first-level character equations on the whole orbit. Their common real kernel is isotropic for the combined commutator form from [Section 3.4](#). This constrains the combined vertical phase; the individual commutators need not vanish. For Reed–Solomon evaluation forms, [Lemma 6.1](#) will turn suitable combinations of these global equations into local equations.

Technical approach summary. We adapt Leng’s degree-two argument [[Len25](#), proof of Theorem 3.3.1, pp. 52–54; Corollary 3.2.9 and its proof, pp. 49–50], using his sharp bracket lemma [[Len25](#), Lemma 4.2.1, p. 123]. Our refinement is quantitative. Applying the general dimension-dependent bounds directly would give only an unspecified positive range $0 < u < u_0$. By exploiting the product structure, we obtain the range $0 < u < 1/2$.

We take one multiplicative derivative and apply ordinary Fourier estimates separately to the derivative readout formed from each local factor Φ_i . Each such readout lives on a torus of dimension at most D , independent of m . This gives a fixed Fourier cost per factor, which we track across the m factors.

The argument takes place on a carefully chosen line: the product average remains large there, and every relevant short character that is non-constant on the full orbit remains non-constant on the line. Consequently, any short character equation obtained on this line already holds on the entire orbit. For correlation thresholds satisfying $\log(1/\delta) = \mathcal{O}(m)$, these bounds give the field-size requirement

$$\log p \gg m^2 \log(2m).$$

Lemma statement. Fix the finite collection $\mathcal{N}$ from [Section 4](#). Each local group has dimension at most D , and $K = D + 2$. We state the lemma for arbitrary linear forms. Its applications will use the code forms $L_i(x) = \sum_{j < r} a_i^j x_j$ and the coordinate projections on $\mathbb{F}_p^m$ for independent inputs. The code application requires no randomness in the evaluation places.

Lemma 5.1 (Product refinement). *Let $r \geq 1$ and let $L_1, \dots, L_m: \mathbb{F}_p^r \rightarrow \mathbb{F}_p$ be arbitrary linear forms. For each $i \in [m]$, choose $(G^{(i)}/\Gamma^{(i)}, G_\bullet^{(i)}, \mathcal{X}^{(i)}) \in \mathcal{N}$. Let $\Phi_i: G^{(i)}/\Gamma^{(i)} \rightarrow \mathbb{C}$ have vertical frequency ν_i and satisfy $\|\Phi_i\|_{C^K(\mathcal{X}^{(i)})} \leq 1$. Let $g^{(i)} \in \text{poly}(\mathbb{Z}, G_\bullet^{(i)})$ be p -periodic modulo $\Gamma^{(i)}$, and*

let $0 < \delta \leq 1/10$. There is a constant C , depending only on the fixed local data and independent of r and the forms L_i , such that the following holds. Suppose

$$\left| \mathbb{E}_{x \in \mathbb{F}_p^r} \prod_{i=1}^m \Phi_i \left(g^{(i)}(L_i(x)) \Gamma^{(i)} \right) \right| \geq \delta \quad (25)$$

and

$$\log p \geq C(m^2 \log(2m) + m \log(1/\delta)). \quad (26)$$

Then there are product first-level characters $\xi_1, \dots, \xi_t$, with $0 \leq t \leq Dm$, each constant modulo 1 on the orbit

$$\left(g^{(i)}(L_i(x)) \Gamma^{(i)} \right)_{i \in [m]}, \quad x \in \mathbb{F}_p^r,$$

and each of complexity at most

$$B = \exp(C(m \log(2m) + \log(1/\delta))), \quad (27)$$

such that their common real kernel

$$K_0 = \bigcap_{j=1}^t \ker(d\xi_j) \subseteq \bigoplus_{i=1}^m \text{Lie}(G^{(i)})$$

satisfies

$$\sum_{i=1}^m d\nu_i([v_i, w_i]) = 0 \quad \text{for every } v = (v_i)_i, \ w = (w_i)_i \in K_0. \quad (28)$$

For $t = 0$, the intersection means the whole direct sum. Integer representatives may be chosen arbitrarily when evaluating the group-valued sequences at field elements: periodicity makes the cosets well-defined.

The external bracket lemma. Our proof uses an exact periodic specialization of Leng's sharp bracket lemma [Len25, Lemma 4.2.1, p. 123] in Step 5 below. It converts a fractional-part relation holding for many shifts into two complementary orthogonal families of integer vectors. The key quantitative feature is the bound on their heights, whose exponent is independent of the dimension. We fixed the terminology of orthogonal families and height at the end of Section 3.

Imported Lemma 5.2 (Exact periodic sharp bracket lemma). *Let $N_{\text{lin}} \geq 1$, $M_{\text{br}} \geq 1$, and $0 < \kappa < 1/10$. Suppose*

$$\vartheta \in p^{-1}\mathbb{Z}^{N_{\text{lin}}}, \quad \mathbf{a}_{\text{br}} \in \mathbb{R}^{N_{\text{lin}}}, \quad \|\mathbf{a}_{\text{br}}\|_{\infty} \leq M_{\text{br}}, \quad \gamma_{\text{br}} \in \mathbb{R}$$

satisfy

$$\gamma_{\text{br}} + \mathbf{a}_{\text{br}} \cdot \{h\vartheta\}_{\text{c}} = 0 \pmod{1}$$

for at least κp residues $h \in \mathbb{F}_p$. Then, for some $0 \leq \ell \leq N_{\text{lin}}$, there are integer-vector families

$$w_1, \dots, w_{\ell} \in \mathbb{Z}^{N_{\text{lin}}}, \quad e_1, \dots, e_{N_{\text{lin}}-\ell} \in \mathbb{Z}^{N_{\text{lin}}},$$

each linearly independent over $\mathbb{R}$, such that

$$w_s \cdot \mathbf{a}_{\text{br}} = 0, \quad e_j \cdot \vartheta \in \mathbb{Z}, \quad w_s \cdot e_j = 0$$

for every $1 \leq s \leq \ell$ and $1 \leq j \leq N_{\text{lin}} - \ell$. Thus the real spans of the two families are orthogonal complements. Every vector has height at most

$$P_{\text{br}} = (\kappa^{-1}(2N_{\text{lin}})^{N_{\text{lin}}} M_{\text{br}})^{C_{\text{br}}},$$

where $C_{\text{br}} > 0$ is an absolute constant.

Why the specialization is exact. Leng's finite-interval lemma gives approximate relations, with vector heights bounded independently of the interval length. Periodicity lets us repeat our exact input relation on arbitrarily long intervals. The approximation bounds then tend to zero, while only finitely many output families are possible.

More precisely, fix all the data in the statement and put $N = qp$, where q is a positive integer. Since $p\vartheta \in \mathbb{Z}^{N_{\text{lin}}}$, the expression

$$\gamma_{\text{br}} + \mathbf{a}_{\text{br}} \cdot \{h\vartheta\}_{\text{c}}$$

is p -periodic in h . Its exact congruence therefore holds for at least κN integers in $[N]$. In particular, at each of these integers,

$$\|\gamma_{\text{br}} + \mathbf{a}_{\text{br}} \cdot \{h\vartheta\}_{\text{c}}\|_{\mathbb{R}/\mathbb{Z}} = 0 < \frac{2}{N}.$$

Apply Leng's lemma with its perturbation parameter set to 2. For sufficiently large q , neither the small-interval alternative nor the alternative $2/N > 1/100$ can occur. The lemma therefore supplies two linearly independent integer-vector families of complementary sizes, orthogonal to one another, with

$$|w_s \cdot \mathbf{a}_{\text{br}}| \leq \frac{A}{N}, \quad \|e_j \cdot \vartheta\|_{\mathbb{R}/\mathbb{Z}} \leq \frac{A}{N}.$$

Here A depends only on N_{lin} , M_{br} , and κ , and is independent of q . The source's height bounds are

$$\left(\kappa^{-1} N_{\text{lin}}^{N_{\text{lin}}} M_{\text{br}}\right)^{\mathcal{O}(1)},$$

so every output vector has height at most P_{br} after choosing the absolute constant C_{br} sufficiently large.

The vectors are integral and their height bound is independent of q , so only finitely many different pairs of families can occur as q varies. Since there are infinitely many sufficiently large integers q , one pair must recur infinitely often; restricting to those values gives a sequence $q_k \rightarrow \infty$ along which ℓ and all the vectors in both families remain fixed. Along this sequence, the vectors are fixed while A/N tends to zero. Hence

$$w_s \cdot \mathbf{a}_{\text{br}} = 0, \quad e_j \cdot \vartheta \in \mathbb{Z}.$$

Linear independence and orthogonality already hold exactly for this fixed pair. This proves the specialization, without requiring $\mathbf{a}_{\text{br}}$ to be rational or imposing any further lower bound on p .

Proof of Lemma 5.1.

(Step 1) *Choose a line that remembers all short characters.* We reduce the r -variable average to a one-variable average along an affine line. The line must retain a large product average, and any short character equation that holds on it must already hold on the whole orbit. We choose the direction first: the field-size hypothesis lets us avoid every direction that would make a relevant nonconstant short character constant. Once this direction is fixed, averaging over the parallel lines gives a starting point for which the product average still has magnitude at least δ .

Fix a sufficiently large constant C_{line} , depending only on the local data, and put

$$B_{\text{line}} = \exp(C_{\text{line}}(m \log(2m) + \log(1/\delta))).$$

The height calculation in Step 6 will fit every character we construct inside this bound. Write $L_i(x) = \sum_{j=0}^{r-1} b_{ij} x_j$, with $b_{ij} \in \mathbb{F}_p$. For a product character ξ , let k_i be its local coefficient vectors, and let $\theta_i, \beta_i/p$ be the local initial positions and slope vectors from Section 3.4. Its composition with the orbit is

$$\gamma_\xi + \frac{\lambda_\xi \cdot x}{p} \pmod{1}, \quad \gamma_\xi = \sum_i k_i \cdot \theta_i, \quad (\lambda_\xi)_j = \sum_i b_{ij}(k_i \cdot \beta_i) \quad (0 \leq j < r).$$

Here $\lambda_\xi \in \mathbb{F}_p^r$, and the character is constant modulo 1 on the whole orbit exactly when $\lambda_\xi = 0$. There are at most $(2B_{\text{line}} + 1)^{Dm}$ candidate characters. By Equation 26, after increasing C ,

$$(2B_{\text{line}} + 1)^{Dm} < p.$$

For each candidate with $\lambda_\xi \neq 0$, the directions on which its restriction becomes constant form the hyperplane $\{\mathbf{v}: \lambda_\xi \cdot \mathbf{v} = 0\}$, of size p^{r-1} . Fewer than p such hyperplanes cannot cover $\mathbb{F}_p^r$. Each excludes a fraction $1/p$ of all directions, regardless of r ; the candidate count depends only on the total local dimension, at most Dm . Choose a nonzero direction outside their union; if there are no nonconstant candidates, choose any nonzero direction. For any function $F_{\text{test}}: \mathbb{F}_p^r \rightarrow \mathbb{C}$,

$$\left| \mathbb{E}_x F_{\text{test}}(x) \right| \leq \mathbb{E}_x \left| \mathbb{E}_{z \in \mathbb{F}_p} F_{\text{test}}(x + z\mathbf{v}) \right|,$$

by translation invariance and the triangle inequality. Apply this to the product in Equation 25. Some $x^{(0)} \in \mathbb{F}_p^r$ therefore satisfies

$$\left| \mathbb{E}_{z \in \mathbb{F}_p} \prod_i \Phi_i \left(g^{(i)}(L_i(x^{(0)}) + zL_i(\mathbf{v})) \Gamma^{(i)} \right) \right| \geq \delta.$$

By Section 3.2, the restricted sequences remain quadratic and p -periodic, including when $L_i(\mathbf{v}) = 0$. For Steps 2–6, write $g^{(i)}$ for these restricted sequences.

(Step 2) *Put the local linear coefficients in a compact set.* We use compact normalization to bound the first-level slopes uniformly before forming the bracket relation.

Using compact normalization from Section 3.3, write

$$g^{(i)}(z) = u_i^z b_i^{\binom{z}{2}}, \quad b_i \in G_2^{(i)},$$

where u_i lies in a fixed compact fundamental set. To remove the initial coefficient, translate the readout by its compact representative and conjugate the remaining coefficients by its lattice representative. For a linear coefficient $c\gamma$, the identity

$$(c\gamma)^z = c^z [\gamma, c]^{\binom{z}{2}} \gamma^z,$$

removes the right lattice factor after changing the central coefficient. These operations preserve the scalar sequence, vertical frequency, and first-level slopes modulo the integer lattice. The readout's C^K norm grows by at most a fixed local factor. Conjugation also preserves first-level coordinates and central commutators.

We henceforth denote the normalized readout and group-valued sequence by Φ_i and $g^{(i)}$, respectively. The normalized coset orbit is a fixed left translate of the previous coset orbit, so it remains p -periodic and Equation 18 applies. Put $h_i = \dim G^{(i)} - \dim G_2^{(i)}$ and $\vartheta_i = \psi_1^{(i)}(u_i) \in \mathbb{R}^{h_i}$. By Equation 18,

$$u_i^p \in \Gamma^{(i)}, \quad b_i^p \in \Gamma_2^{(i)}, \quad \vartheta_i \in p^{-1}\mathbb{Z}^{h_i}, \quad \|\vartheta_i\|_\infty \ll 1.$$

Here and below, the suppressed constants depend only on the fixed local data. We absorb the fixed C^K -norm increase into the Fourier-envelope constant C_0 below.

(Step 3) *Differentiate and pass to ordinary tori.* We start with the normalized quadratic nilsequences and want to bring them within reach of ordinary Fourier analysis. The idea is to compare each local sequence at $z + h$ and z , multiplying one value by the conjugate of the other. Just as differencing turns a quadratic polynomial into a linear one, this comparison removes the common quadratic motion. What remains is a smooth function evaluated along an orbit on an ordinary torus that is affine-linear in z for each fixed h . The next step will use this description to extract exact frequency relations from the large product average.

For a shift $h \in \mathbb{F}_p$, choose an integer representative and write

$$u_i^h = c_{i,h} \gamma_{i,h}, \quad \psi^{(i)}(c_{i,h}) \in (-1/2, 1/2]^{\dim G^{(i)}}, \quad \gamma_{i,h} \in \Gamma^{(i)}.$$

The derivative readout

$$\Psi_{i,h}(v, w) = \Phi_i(c_{i,h} v \Gamma^{(i)}) \overline{\Phi_i(w \Gamma^{(i)})}$$

descends to the derivative torus $T_i^\square$ of [Section 3.5](#). The multiplicative derivative $\Phi_i(g^{(i)}(z + h)\Gamma^{(i)}) \overline{\Phi_i(g^{(i)}(z)\Gamma^{(i)})}$ is obtained by evaluating $\Psi_{i,h}$ at

$$\left(\gamma_{i,h} g^{(i)}(z) \gamma_{i,h}^{-1} b_i^{zh + \binom{h}{2}}, g^{(i)}(z) \right). \quad (29)$$

Indeed, multiplying the first entry by $c_{i,h}$ gives the coset of $g^{(i)}(z + h)$, by the quadratic Taylor formula. The relative central coordinate in [Equation 29](#) is $[\gamma_{i,h}, u_i] b_i^{zh + \binom{h}{2}}$; the second entry has first-level coordinate $z\vartheta_i$. Both coordinates on $T_i^\square$ are therefore affine-linear in z . This is the degree-two differentiation used in [\[Len25, proof of Theorem 3.3.1, pp. 52–54\]](#).

In the product coordinates on $T_i^\square$, a supported Fourier character has the form

$$(v, w) \mapsto \chi_i(w) + \nu_i(vw^{-1}) \pmod{1},$$

where χ_i is a first-level character. Its slope on [Equation 29](#) is

$$\lambda_i(h) = \chi_i(u_i) + h\nu_i(b_i) + \nu_i([\gamma_{i,h}, u_i]). \quad (30)$$

Every term lies in $p^{-1}\mathbb{Z}$. For the commutator term, this follows from $[\gamma_{i,h}, u_i]^p = [\gamma_{i,h}, u_i^p] \in \Gamma_2^{(i)}$. Thus the average vanishes unless its slope is integral. When the slope is integral, the average equals its initial phase and has modulus one.

(*Step 4*) *Extract one exact relation from the Fourier expansion.* We start with the large product average and the description of each local multiplicative derivative on an ordinary torus. The output is one fixed choice of Fourier modes $k_1, \dots, k_m$, with $\|k_i\|_\infty \leq T_{\text{der}}$, whose slopes satisfy

$$\sum_i \lambda_i(h) \in \mathbb{Z}$$

for at least κp shifts $h \in \mathbb{F}_p$. For each such shift, the selected modes have a product that is independent of z .

Write $d_i = \dim G^{(i)} \leq D$. Compactness of the $c_{i,h}$ and the Fourier-envelope estimate of [Section 4](#) give envelopes $\mathfrak{w}_i: \mathbb{Z}^{d_i} \rightarrow [0, \infty)$ and a fixed $C_0 \geq 1$ such that

$$\left| \widehat{\Psi}_{i,h}(k) \right| \leq \mathfrak{w}_i(k), \quad \sum_k \mathfrak{w}_i(k) \leq C_0, \quad \sum_{\|k\|_\infty > T_{\text{der}}} \mathfrak{w}_i(k) \leq C_0 T_{\text{der}}^{-2}.$$

Set the envelope to zero off the vertical-frequency support just described. For $F_{\text{test}}(z) = \prod_i \Phi_i(g^{(i)}(z)\Gamma^{(i)})$, the shift-correlation identity in [Section 4](#), followed by Fourier expansion and orthogonality, gives

$$\delta^2 \leq \sum_{k_1, \dots, k_m} \prod_i \mathfrak{w}_i(k_i) \Pr_{h \in \mathbb{F}_p} \left[\sum_i \lambda_i(h) \in \mathbb{Z} \right].$$

Each Fourier tuple in this sum has its own collection of slopes. Choose $T_{\text{der}} \geq 2$ with $mC_0^m T_{\text{der}}^{-2} \leq \delta^2/2$ and $\log T_{\text{der}} \ll m + \log(1/\delta)$. The tuples with some $\|k_i\|_\infty > T_{\text{der}}$ have total envelope mass at most $mC_0^m T_{\text{der}}^{-2}$. Since the remaining total mass is at most C_0^m , one remaining tuple satisfies

$$\Pr_h \left[\sum_i \lambda_i(h) \in \mathbb{Z} \right] \geq \kappa, \quad \kappa = \frac{\delta^2}{2C_0^m} < \frac{1}{10}. \quad (31)$$

Fix such a tuple. Its central Fourier coordinates are the integer coefficients of ν_i , so $\|\nu_i\|_\infty \ll T_{\text{der}}$; here the norm is the maximum magnitude of those coefficients.

(Step 5) *Write the relation in coordinates.* We start with the fixed Fourier modes whose slopes satisfy

$$\sum_i \lambda_i(h) \in \mathbb{Z}$$

for at least κp shifts. The output is two complementary orthogonal families of integer vectors of controlled height. One family gives exact integer relations among the orbit coefficients, with one auxiliary coordinate; the other gives exact real relations for the coefficients describing the combined commutator phase. These are the families from which the next step will construct the required first-level character equations.

Put $N_{\text{hor}} = \sum_i h_i \leq Dm$ and concatenate the ϑ_i into $\vartheta \in p^{-1}\mathbb{Z}^{N_{\text{hor}}}$. Let $\pi_1: \bigoplus_i \text{Lie}(G^{(i)}) \rightarrow \mathbb{R}^{N_{\text{hor}}}$ retain the first-level Lie-algebra coordinates. For each i , define the linear lift

$$\iota_i: \mathbb{R}^{h_i} \longrightarrow \text{Lie}(G^{(i)}), \quad \iota_i(t) = \sum_{j=1}^{h_i} t_j X_j^{(i)}.$$

This lift inserts t in the first-level positions of the Lie algebra and puts zero in its central positions. For $y = (y_i)_i$ and $z = (z_i)_i$ in $\mathbb{R}^{N_{\text{hor}}}$, put

$$\Omega(y, z) = \sum_i d\nu_i([\iota_i(y_i), \iota_i(z_i)]).$$

The matrix of Ω is integral. Indeed, the basis elements $\exp X_j^{(i)}$ lie in the lattice; their commutators lie in $\Gamma_2^{(i)}$, and Equation 19 identifies their values under ν_i with the matrix entries. Moreover, $\|\Omega\|_\infty \ll T_{\text{der}}$.

The first-level coordinates of $\gamma_{i,h}$ are $h\vartheta_i - \{h\vartheta_i\}_c$. Their contribution to Equation 30 is $\Omega(h\vartheta - \{h\vartheta\}_c, \vartheta) = \Omega(\vartheta, \{h\vartheta\}_c)$. Thus Equation 31 gives

$$\gamma_{\text{br}} + \beta_{\text{br}}h + \Omega(\vartheta, \{h\vartheta\}_c) = 0 \pmod{1} \quad (32)$$

on at least κp shifts, where

$$\gamma_{\text{br}} = \sum_i \chi_i(u_i), \quad \beta_{\text{br}} = \sum_i \nu_i(b_i) \in p^{-1}\mathbb{Z}.$$

If $N_{\text{hor}} = 0$, every group is central and the conclusion holds with no characters. Assume henceforth that $N_{\text{hor}} > 0$. Define $\mathbf{a}_{\text{br}} \in \mathbb{R}^{N_{\text{hor}}}$ by $\mathbf{a}_{\text{br}} \cdot z = \Omega(\vartheta, z)$ and put

$$\tilde{\mathbf{a}} = (\mathbf{a}_{\text{br}}, 1), \quad \tilde{\vartheta} = (\vartheta, \beta_{\text{br}}).$$

Because $\beta_{\text{br}}h \equiv \{h\beta_{\text{br}}\}_c \pmod{1}$, Equation 32 is exactly the hypothesis of Imported Lemma 5.2, applied to $\tilde{\mathbf{a}}$ and $\tilde{\vartheta}$ in $N_{\text{hor}} + 1$ coordinates. It gives complementary orthogonal integer families

$$w_s = (\mathbf{q}_s, v_s) \quad (1 \leq s \leq \ell), \quad e_j = (s_j, t_j) \quad (1 \leq j \leq N_{\text{hor}} + 1 - \ell),$$

where $\mathbf{q}_s, s_j \in \mathbb{Z}^{N_{\text{hor}}}$ and $v_s, t_j \in \mathbb{Z}$, with

$$w_s \cdot \tilde{\mathbf{a}} = 0, \quad e_j \cdot \tilde{\vartheta} \in \mathbb{Z}.$$

Their heights are at most P_{br} , where

$$\log P_{\text{br}} \ll N_{\text{hor}} \log(2N_{\text{hor}}) + \log(1/\kappa) + \log(2 + \|\mathbf{a}_{\text{br}}\|_\infty). \quad (33)$$

The absolute exponent in the imported height bound is crucial here.

(Step 6) *Eliminate the auxiliary coordinate and obtain an isotropic kernel.* We have the two complementary orthogonal integer families from the previous step. We need to identify at most $N_{\text{hor}} \leq Dm$ first-level characters ξ_j , each of complexity at most B_{line} and constant on the selected line. Their common real kernel $K_0 = \bigcap_j \ker(d\xi_j)$ satisfies

$$\sum_i d\nu_i([v_i, w_i]) = 0 \quad \text{for all } v, w \in K_0.$$

Thus, motions preserving all the character equations produce no combined vertical phase through their commutators.

We use the auxiliary-coordinate elimination from [Len25, Corollary 3.2.9 and its proof, pp. 49–50]. At least one t_j is nonzero. Otherwise $(0, 1)$ is orthogonal to every e_j , hence lies in the span of the w_s , contradicting $w_s \cdot \tilde{\mathbf{a}} = 0$ because $(0, 1) \cdot \tilde{\mathbf{a}} = 1$. Relabel so $t_1 \neq 0$, and define integer linear forms on $\mathbb{R}^{N_{\text{hor}}}$ by

$$\zeta_j = t_j s_1 - t_1 s_j \quad (j > 1), \quad \rho_s(z) = \Omega(\mathbf{q}_s, z) \quad (1 \leq s \leq \ell).$$

Composing these forms with the product first-level coordinate map gives first-level characters. They are constant modulo 1 on the selected line:

$$\zeta_j \cdot \vartheta \in \mathbb{Z}, \quad \rho_s(\vartheta) = -\mathbf{q}_s \cdot \mathbf{a}_{\text{br}} = v_s \in \mathbb{Z}.$$

The first assertion eliminates β_{br} from the two integer relations $e_1 \cdot \tilde{\vartheta}$ and $e_j \cdot \tilde{\vartheta}$.

It remains to check what these equations say about their real kernel. The $\mathbf{q}_s$ are independent: a relation among them gives a vector $(0, v)$ in the span of the w_s , and orthogonality to $\tilde{\mathbf{a}}$ forces $v = 0$; independence of the w_s then forces all coefficients to vanish. The ζ_j are also independent, since a relation among them gives

$$\left(\sum_{j>1} c_j t_j \right) e_1 - t_1 \sum_{j>1} c_j e_j = 0.$$

Furthermore $\zeta_j \cdot \mathbf{q}_s = 0$, by orthogonality of the two input families. Their numbers add to N_{hor} , so

$$\bigcap_{j>1} \ker \zeta_j = \text{span}_{\mathbb{R}}\{\mathbf{q}_s : 1 \leq s \leq \ell\}.$$

Consequently,

$$\overline{K}_0 = \bigcap_{j>1} \ker \zeta_j \cap \bigcap_s \ker \rho_s$$

is isotropic for Ω : its vectors are linear combinations of the $\mathbf{q}_s$ and annihilate every $\Omega(\mathbf{q}_s, \cdot)$. The corresponding common kernel in the full Lie algebra is $K_0 = \pi_1^{-1}(\overline{K}_0)$. For $v = (v_i)_i$ in the full Lie algebra, $v_i - \iota_i((\pi_1 v)_i) \in \text{Lie}(G_2^{(i)})$ is central. Hence, for any v, w in the full Lie algebra,

$$\sum_i d\nu_i([v_i, w_i]) = \Omega(\pi_1 v, \pi_1 w).$$

Hence the isotropy of $\overline{K}_0$ is exactly the required identity Equation 28 for the full common kernel.

There are at most N_{hor} resulting characters. Their coefficient heights are bounded by $2P_{\text{br}}^2 + N_{\text{hor}} \|\Omega\|_{\infty} P_{\text{br}}$; multiplying by N_{hor} bounds their complexities. Since $\|\vartheta\|_{\infty} \ll 1$, $\|\mathbf{a}_{\text{br}}\|_{\infty} \ll N_{\text{hor}} T_{\text{der}}$. Together with Equation 31 and Equation 33, this gives

$$\text{comp}(\xi_j) \leq \exp(C_*(m \log(2m) + \log(1/\delta)))$$

for a constant C_* depending only on the fixed local data. Choose $C_{\text{line}} \geq C_*$ in Step 1.

(Step 7) *Recover equations on the entire orbit.* Every character just constructed has complexity at most B_{line} and is constant on the chosen line. Step 1 ensured that a character in this range could become constant on the line only if it was already constant on the full orbit. The normalization preserved first-level slopes modulo the integer lattice and preserved central commutators. Thus the same character conditions and the same common real kernel apply to the original data. Increasing C gives both Equation 26 and Equation 27.

6. FROM MOMENT EQUATIONS TO LOCAL EQUATIONS

This section proves an algebraic certificate lemma for Reed–Solomon moment systems (Lemma 6.1). For a system with bounded integer coefficient matrices, we seek a nonzero integer combination, with controlled coefficients, that vanishes at each share separately for every solution of the moment equations. The conclusion holds with high probability over the public evaluation places; this is where their randomness enters the proof. Our certificate lemma is independent of Lemma 5.1, and the two results will be combined in Section 7 and Section 8.

The leakage maps may be chosen after the evaluation places are revealed, so the equations may also depend on those places. We therefore need one good global event on which the conclusion holds simultaneously for every bounded coefficient array and every solution of its moment equations. The certificate associated with each array is chosen using that array alone.

The main coding-theoretic input comes from the local-constraints analysis of Levi–Mosheiff–Shagrithaya, whose potential-maximization and quotient strategy we also follow [LMS25a, Proposition 6.1 and the quotient argument, pp. 28–29]. Their local-profile theorem applies because the Vandermonde moment matrix (a_i^h) has a generalized Reed–Solomon code as its kernel. This is the Reed–Solomon-specific probabilistic input, stated below as an imported vanishing theorem. We supplement this framework with arithmetic height bounds to keep the integer certificates short and a polynomial-identity argument to obtain an explicit field-size condition.

- Section 6.1 introduces the moment equations, explains their connection to the product lemma, and states our certificate lemma (Lemma 6.1).
- Section 6.2 presents the geometric idea and the external inputs needed for the proof.
- Section 6.3 proves the certificate lemma.

6.1. Preliminaries: The Equations and the Desired Certificate. Fix a positive real number b and a positive integer D . Here D bounds the number of local coordinates, and b denotes a fixed lower bound on the ratio r/m of the number of moments to the number of shares. Let m be the number of shares and r the number of moments, with $bm \leq r < m$. Set

$$q = \lceil 4D/b \rceil.$$

Thus q is a constant once b, D are fixed. It will be the number of global equations that we combine.

An array $C = (C_i)_{i=1}^m$ consists of integer matrices $C_i \in \mathbb{Z}^{q \times D}$. Its *height* is the largest absolute value of an entry. For column vectors $x_i \in \mathbb{F}_p^D$, put $y_i = C_i x_i \in \mathbb{F}_p^q$, reducing the integer entries modulo p . The moment equations are

$$\sum_{i=1}^m a_i^h y_i = 0 \quad (0 \leq h < r). \quad (34)$$

Each of these is a vector equation with q scalar coordinates. We want a nonzero integer column $u_C \in \mathbb{Z}^q$, the *certificate*, for which the same combination of those coordinates vanishes locally:

$$u_C^\top y_i = 0 \quad (1 \leq i \leq m).$$

The superscript $\top$ denotes transpose. The conclusion gives one common local equation; it need not force the vectors y_i themselves to vanish.

From character equations to moment equations. Roughly speaking, for q dual codewords represented locally by integer matrices with bounded entries, we give a short integer certificate of their linear dependence. To explain how character equations give a moment system, suppose $\xi_1, \dots, \xi_q$ are product first-level characters of complexity at most B that are constant on the moment orbit.⁸ At coordinate i , place their integer coefficient rows in C_i , adding zero columns if the local first-level dimension is smaller than D . By the affine-phase calculation in [Section 3.4](#), constancy on the moment orbit is exactly [Equation 34](#), with x_i equal to the local first-level slope numerator β_i , padded with zeros to length D . The row $u_C^\top C_i$ then describes one integer combination of those characters at coordinate i . If C has height at most B , its local height is bounded by

$$\|u_C^\top C_i\|_\infty \leq qB\|u_C\|_\infty. \quad (35)$$

This is the quantity the next lemma controls.

Only the evaluation tuple is random. As before, write

$$\mathcal{A}_{p,m} = \{(a_1, \dots, a_m) \in (\mathbb{F}_p^*)^m : a_i \neq a_j \text{ whenever } i \neq j\}.$$

It is equipped with the uniform distribution.

Lemma 6.1 (Uniform short moment certificate). *Fix $b > 0$ and $D \in \mathbb{N}$, and put $q = \lceil 4D/b \rceil$. Let $bm \leq r < m$, with m sufficiently large in terms of b, D . For every $B \geq 2$, there is a bound*

$$H = (2B)^{C_{b,D}},$$

where $C_{b,D} > 0$ depends only on b, D , with the following property. Suppose

$$p > \max\{H, (mD)!B^{mD}, m(m+1)\}. \quad (36)$$

For each integer array $C = (C_i)_{i=1}^m$ of height at most B , one can choose a nonzero $u_C \in \mathbb{Z}^q$, using C alone, such that

$$qB\|u_C\|_\infty \leq H. \quad (37)$$

Except on an event of probability at most

$$\frac{2Dm(r-1)(2B+1)^{qmD}}{p}, \quad (38)$$

over $a \in \mathcal{A}_{p,m}$, the implication

$$\left[\sum_{i=1}^m a_i^h C_i x_i = 0 \quad (0 \leq h < r) \right] \implies u_C^\top C_i x_i = 0 \quad (i \in [m]) \quad (39)$$

holds simultaneously for every such array and every choice of $x_i \in \mathbb{F}_p^D$.

Fix the assignment $C \mapsto u_C$ before sampling a . On the good global event, the array and its solutions may then be chosen after seeing a . The same assigned certificate works for every solution of that array; later, we use it for the two coefficient vectors of a quadratic orbit.

⁸The moment equations express membership in the dual of the Reed–Solomon code $\mathcal{C}_a = \{(f(a_1), \dots, f(a_m)) : f \in \mathbb{F}_p[X], \deg f < r\}$: a vector $z \in \mathbb{F}_p^m$ belongs to $\mathcal{C}_a^\perp$ exactly when $\sum_i a_i^h z_i = 0$ for every $0 \leq h < r$. Here the dual-code vectors are formed from the character coefficients and the local orbit slopes; the share vectors themselves lie in $\mathcal{C}_a$.

6.2. The geometric idea and the facts it needs. The two imported lemmas supply complementary ingredients: [Imported Lemma 6.2](#) controls the certificate's height independently of m , and [Imported Lemma 6.3](#) supplies locality after passing to a suitable quotient.

Work first over $\overline{\mathbb{Q}}$, and let V_i be the column space of C_i . Cancellation between shares can concentrate in a subspace containing many local directions. To identify such a subspace, put $\lambda = bm/2$ and define the potential

$$\mathcal{P}_C(U) = \sum_{i=1}^m \dim(V_i \cap U) - \lambda \dim U \quad (U \leq \overline{\mathbb{Q}}^q). \quad (40)$$

The first term counts local directions inside U ; the second charges λ for each dimension of U .

The proof chooses the largest subspace U_* maximizing this quantity. Two facts make that choice useful. First, U_* is proper and has a short integer annihilating row. Second, after quotienting by U_* , no subspace contains enough local directions to support a nonzero moment solution at typical evaluation places. We record precisely the two results used to justify these claims.

Why a rational subspace has a short description here. Choose s coordinate positions on which projection from an s -dimensional subspace is invertible. There is then a unique $q \times s$ matrix X spanning that subspace whose selected rows form the identity matrix. Such a representation is called a *graph chart*. The other $s(q-s)$ entries are its coordinates. Conditions on $\dim(V_i \cap U)$ become rank conditions on $[C_i \mid X]$, hence polynomial equations in those coordinates. The following height bound will turn uniqueness of this description into a short certificate.

Imported Lemma 6.2 (Height of a unique rational zero [[KPS01](#)]). *Fix $q \geq 1$ and $N \leq q^2$. Let a finite family of integer polynomials in N variables have degrees at most q and coefficients of magnitude at most $A \geq 2$. Suppose its common zero set over $\overline{\mathbb{Q}}$ consists of one point $z \in \mathbb{Q}^N$. Write $[1 : z]$ as a primitive integer vector $(Z_0, \dots, Z_N)$ with $Z_0 > 0$. Then*

$$\log \max_{0 \leq j \leq N} |Z_j| \leq C_q \log A,$$

where C_q depends only on q , independently of the number of polynomials.

This is the singleton case of Krick–Pardo–Sombra [[KPS01](#), Corollary 2.11 and the following affine-space specialization, printed p. 25]. In their author version [maino.pdf](#), equation (1.4), printed p. 10, defines the height of a polynomial family; for these integer coefficients it is at most $\log A$. Section 1.2.3, printed p. 14, gives the height of a singleton, which dominates the logarithmic projective maximum height above. Thus the number of rank conditions can grow with m without worsening the exponent in the certificate height.

Why uncrowded local spaces have no moment solution. The next statement is a form of the local-profile theorem of Levi–Mosheiff–Shagrithaya. Here it is stated for moments, as required by the proof.

Imported Lemma 6.3 (Qualitative vanishing for moment systems [[LMS25a](#), [LMS25b](#)]). *Fix positive integers q, m, r and a positive real number $b < 1$, with*

$$bm \leq r < m, \quad bm/2 \geq 2q(q+1).$$

For each prime $p > m$, let $V_1, \dots, V_m \leq \mathbb{F}_p^q$ satisfy

$$\sum_i \dim(V_i \cap W) \leq (bm/2) \dim W \quad (W \leq \mathbb{F}_p^q).$$

For uniform $a \in \mathcal{A}_{p,m}$, with probability $1 - o_{p \rightarrow \infty}(1)$ the only solution of

$$\sum_i a_i^h y_i = 0 \quad (0 \leq h < r), \quad y_i \in V_i,$$

is $y_1 = \dots = y_m = 0$. The parameters q, m, r, b are fixed in this limit. The probability bound is uniform in the spaces V_i , which may depend on p .

The source is [LMS25a, Proposition 6.1, equations (23)–(24), printed p. 28]; see also the conference version [LMS25b]. The proposition’s proof uses Lemma 7.7, equation (30), printed p. 35. Our size condition checks the hypothesis $\lambda \geq 2q(q+1)$ of that lemma. The related quotient-potential calculation appears on printed pp. 28–29. All these locators refer to arXiv:2406.02238v7.

Converting the cited theorem to moments. For distinct evaluation places, set $P(X) = \prod_{j=1}^m (X - a_j)$ and $w_i = P'(a_i)^{-1}$. Comparing the coefficient of X^{m-1} in the Lagrange interpolation formula for X^t gives

$$\sum_i w_i a_i^t = 0 \quad (0 \leq t \leq m-2).$$

It follows that

$$\ker(a_i^h)_{\substack{0 \leq h < r \\ 1 \leq i \leq m}} = \{(w_i f(a_i))_{i=1}^m : \deg f < m-r\}.$$

The displayed identity proves inclusion; both sides have dimension $m-r$, so they are equal. Apply it to each coordinate of a vector-valued polynomial $f \in \mathbb{F}_p[X]^q$. The nonzero scalar w_i preserves V_i , so it does not change the local constraint.

We apply [LMS25a, Proposition 6.1] with code length m , ambient dimension q , field size p , code dimension $m-r$, and slack parameter $b/2$. Its potential on a nonzero subspace W is

$$\sum_i \dim(V_i \cap W) - r \dim W \leq -(bm/2) \dim W.$$

The proposition therefore bounds failure for independent uniform places by

$$(2^q - 1) \left(\frac{2(4q)^{4q}(m-r)}{bp} \right)^{bm/(4q)}.$$

Conditioning on distinct nonzero places divides this by $\prod_{j=1}^m (1 - j/p) = 1 - O_m(p^{-1})$. This proves the stated qualitative form and its uniformity in the local spaces.

Two elementary tools for changing the field. (1) Let $f_1, \dots, f_t \in \mathbb{Q}[X]$ have no common zero over $\overline{\mathbb{Q}}$. Then, for all sufficiently large primes p , their reductions have no common zero over $\overline{\mathbb{F}}_p$. To see this, first clear denominators in each equation, so that $f_j \in \mathbb{Z}[X]$. The Nullstellensatz supplies integer polynomials $G_j \in \mathbb{Z}[X]$ whose polynomial combination is a nonzero integer:

$$N_0 = \sum_j G_j f_j, \quad N_0 \in \mathbb{Z} \setminus \{0\}.$$

One may use [KPS01, Theorem 1, printed p. 2] for this integer identity; its quantitative bounds are not needed here. For any prime $p \nmid N_0$, a common zero of the reduced polynomials over $\overline{\mathbb{F}}_p$ would make the right-hand side vanish, while the left-hand side remains nonzero. Thus no such common zero exists.

(2) We will also use the Schwartz–Zippel bound: for every nonzero multivariate polynomial $P \in \mathbb{F}_p[X_1, \dots, X_n]$ of total degree at most T ,

$$\Pr_{(x_1, \dots, x_n) \in \mathbb{F}_p^n} [P(x_1, \dots, x_n) = 0] \leq \frac{T}{p}.$$

6.3. Proof of the certificate lemma.

Proof of Lemma 6.1. The proof first constructs a short vector u_C from the coefficient array C alone. It then proves that, on one good global event, every moment solution satisfies

$$u_C^\top C_i x_i = 0 \quad (i \in [m]).$$

(Step 1) Choose the proper subspace that will contain all solutions. Fix C and use the spaces V_i and potential Equation 40. The potential takes only finitely many values, since it depends only on dimensions. Moreover,

$$\mathcal{P}_C(U) + \mathcal{P}_C(U') \leq \mathcal{P}_C(U + U') + \mathcal{P}_C(U \cap U').$$

To see this, apply the dimension formula to U, U' and use $(V_i \cap U) + (V_i \cap U') \subseteq V_i \cap (U + U')$ for each i . The sum of two maximizers is therefore a maximizer. A maximizer of largest dimension contains every other maximizer; call it U_* .

This subspace is proper. Indeed, $\dim V_i \leq D$ and $\lambda q \geq 2mD$, so

$$\mathcal{P}_C(0) = 0, \quad \mathcal{P}_C(\overline{\mathbb{Q}}^q) \leq mD - \lambda q \leq -mD < 0.$$

It is also defined over $\mathbb{Q}$. Every automorphism of $\overline{\mathbb{Q}}$ fixing $\mathbb{Q}$ preserves the matrices C_i and the potential, hence the unique largest maximizer U_* . Choose a graph chart for U_* . Its representing matrix is unique, so each entry is fixed by every such automorphism and is rational.

(Step 2) Find one short integer row annihilating U_* . Put $s = \dim U_*$ and $t_i = \dim(V_i \cap U_*)$. If $s = 0$, let X_* be the $q \times 0$ matrix and take $u_C = e_1 = (1, 0, \dots, 0)^\top \in \mathbb{Z}^q$. Suppose $s > 0$, and let X_* represent U_* in a graph chart. An s -dimensional subspace U satisfying $\dim(V_i \cap U) \geq t_i$ for every i has potential at least $\mathcal{P}_C(U_*)$. Hence it is a maximizer, lies in U_* , and equals U_* . Therefore the equations

$$\text{rank}[C_i \mid X] \leq \dim V_i + s - t_i \quad (i \in [m]) \tag{41}$$

have exactly one solution in this graph chart over $\overline{\mathbb{Q}}$, namely X_* .⁹

Write these rank conditions as vanishing minors,¹⁰ discarding vacuous conditions and zero polynomials. The number of polynomial equations may grow with m . However, adding more shares does not add variables or worsen the degree and coefficient bounds, which depend only on q and B . Since Imported Lemma 6.2 is independent of the number of equations, it gives a height exponent independent of m .

There are at most q^2 variables; each polynomial has degree at most q and coefficients of magnitude at most $q!B^q$. Imported Lemma 6.2 gives a common positive denominator d_* for X_* with

$$d_* \leq (2B)^{\mathcal{O}_q(1)}, \quad \|d_* X_*\|_\infty \leq (2B)^{\mathcal{O}_q(1)}.$$

Here the matrix norm is the maximum absolute value of an entry. Choose a row outside the prescribed identity rows; such a row exists because $s < q$, because U_* is proper. Expressing it as a linear combination of the identity rows, then multiplying by d_* , gives a nonzero integer column u_C with

$$u_C^\top z = 0 \quad (z \in U_*), \quad \|u_C\|_\infty \leq (2B)^{\mathcal{O}_q(1)}.$$

Increasing $C_{b,D}$ gives Equation 37, also when $s = 0$. The choice uses C alone.

⁹Note the equivalence:

$$\text{rank}[C_i \mid X] = \dim(V_i + U) = \dim V_i + s - \dim(V_i \cap U).$$

Therefore

$$\text{rank}[C_i \mid X] \leq \dim V_i + s - t_i \iff \dim(V_i \cap U) \geq t_i.$$

¹⁰Rank $\leq \rho$ is equivalent to the determinants of all $(\rho + 1) \times (\rho + 1)$ minors vanishing, which is a polynomial constraint.

(Step 3) *Show that the quotient has too few local directions.* We will prove that the preconditions of [Imported Lemma 6.3](#) are met. Let $\pi: \overline{\mathbb{Q}}^q \rightarrow \overline{\mathbb{Q}}^q/U_*$ be the quotient map. For a subspace W of the quotient,

$$\dim(\pi(V_i) \cap W) = \dim(V_i \cap \pi^{-1}(W)) - \dim(V_i \cap U_*).$$

Since U_* maximizes the potential, enlarging it to $\pi^{-1}(W)$ cannot increase the potential. Using the preceding identity and $\dim(\pi^{-1}(W)) - \dim U_* = \dim W$, we obtain

$$0 \geq \mathcal{P}_C(\pi^{-1}(W)) - \mathcal{P}_C(U_*) = \sum_i \dim(\pi(V_i) \cap W) - \lambda \dim W.$$

Thus the additional local directions captured by W cannot outweigh their dimension cost:

$$\sum_i \dim(\pi(V_i) \cap W) \leq \lambda \dim W. \quad (42)$$

We now transfer the quotient inequality to finite fields. For primes p not dividing the denominators of X_* , define

$$U_{*,p} = \text{im}_{\mathbb{F}_p}(X_* \bmod p), \quad \pi_p: \mathbb{F}_p^q \longrightarrow \mathbb{F}_p^q/U_{*,p}.$$

The prescribed identity rows of X_* remain unchanged modulo p , so $\dim U_{*,p} = s$. Write $V_{i,p} = \text{im}(C_i \bmod p)$. We claim that, for every sufficiently large prime,

$$\sum_i \dim(\pi_p(V_{i,p}) \cap W) \leq \lambda \dim W \quad \text{for every } W \leq \mathbb{F}_p^q/U_{*,p}.$$

To justify the claim simultaneously for all W , choose rational coordinates on the quotient. Excluding finitely many primes dividing the relevant denominators or nonzero minors ensures that these coordinates reduce correctly and that the projected local spaces retain their dimensions.

A violation specifies the dimension of W and lower bounds on the intersection dimensions whose sum exceeds $\lambda \dim W$. There are only finitely many such choices. In each graph chart, these intersection requirements become polynomial rank conditions, as in [Equation 41](#). After clearing denominators, they form a system of integer polynomial equations. Each corresponding system has no solution over $\overline{\mathbb{Q}}$, by [Equation 42](#). The Nullstellensatz reduction therefore excludes solutions modulo every sufficiently large prime. Since there are finitely many dimension profiles and charts, one prime threshold excludes every possible violation. This proves the claim, without yet making that threshold explicit.

(Step 4) *Apply the moment theorem, provisionally in large characteristic.* Take m large enough that $\lambda = bm/2 \geq 2q(q+1)$. The quotient has dimension $q - s \leq q$, so [Imported Lemma 6.3](#) applies to its local spaces $\pi_p(V_{i,p})$, using the finite-field version of [Equation 42](#) established in Step 3. With probability $1 - o_{p \rightarrow \infty}(1)$, every moment solution in the quotient is zero. Thus every solution of [Equation 34](#) with $y_i = C_i x_i$ satisfies

$$C_i x_i \in U_{*,p}, \quad u_C^T C_i x_i = 0 \quad (i \in [m]).$$

Here the integer array C and the parameters m, r are fixed, while p tends to infinity. This is the only application of [Imported Lemma 6.3](#). The argument so far applies only to sufficiently large primes, without an explicit lower bound on p . The next two steps turn this conclusion into integer polynomial identities and use them to obtain the explicit quantitative field-size condition in [Lemma 6.1](#).

(Step 5) *Replace the large-characteristic conclusion by integer identities.* We encode the moment equations by an integer polynomial matrix and choose a nonzero maximal minor Δ_C , taking $\Delta_C = 1$ when the matrix has rank zero. We will show that outside its zero set the certificate conclusion holds in every characteristic.

For this, introduce independent indeterminates $Y_1, \dots, Y_m$. The moment system is encoded by the $rq \times mD$ integer polynomial matrix

$$A_C(Y)_{(h,t),(i,j)} = (C_i)_{t,j} Y_i^h, \quad \begin{matrix} 0 \leq h < r, & 1 \leq t \leq q, \\ 1 \leq i \leq m, & 1 \leq j \leq D. \end{matrix} \quad (43)$$

For $x = (x_1^\top, \dots, x_m^\top)^\top$, the moments say $A_C(a)x = 0$. Let k_C be the rank of this matrix over $\mathbb{Q}(Y_1, \dots, Y_m)$. If $k_C = 0$, the rows with $h = 0$ show that every $C_i = 0$. The conclusion is then automatic; set $\Delta_C = 1$. Otherwise choose a $k_C \times k_C$ submatrix $S_C(Y)$ with nonzero determinant $\Delta_C(Y)$. Since $k_C \leq mD$, the formula for a determinant gives

$$\deg \Delta_C \leq mD(r-1), \quad \max_{\alpha} |[Y^\alpha] \Delta_C| \leq (mD)! B^{mD}. \quad (44)$$

The notation $[Y^\alpha] \Delta_C$ means the coefficient of the monomial Y^α .

The polynomial Δ_C now has the required degree and coefficient bounds. It remains to prove that $\Delta_C(a) \neq 0$ guarantees the certificate conclusion. We do this by constructing integer polynomial vectors that span $\ker A_C(a)$ wherever $\Delta_C(a) \neq 0$, and proving that the certificate vanishes identically on each of them. Call the columns of $A_C(Y)$ used to form $S_C(Y)$ the pivot columns; the remaining columns correspond to free coordinates. For each nonpivot column j , let $b_j(Y)$ be that column restricted to the rows used to form $S_C(Y)$.

Setting free coordinate j equal to 1 and all other free coordinates equal to zero forces the vector of pivot coordinates to be

$$-S_C(Y)^{-1} b_j(Y) = -\frac{\text{adj}(S_C(Y)) b_j(Y)}{\Delta_C(Y)},$$

where adj denotes the adjugate matrix. Multiplying the entire solution by $\Delta_C(Y)$ clears the denominator. We obtain a vector $v^{(j)}(Y) \in \mathbb{Z}[Y]^{mD}$ whose coordinate j is $\Delta_C(Y)$, whose other free coordinates are zero, and whose pivot coordinates are $-\text{adj}(S_C(Y)) b_j(Y)$. Write $v_i^{(j)}(Y) \in \mathbb{Z}[Y]^D$ for its i th block.

The chosen rows annihilate $v^{(j)}$ by construction. Since they span the row space of $A_C(Y)$ over $\mathbb{Q}(Y)$, every row annihilates $v^{(j)}$. Thus

$$A_C(Y) v^{(j)}(Y) = 0$$

as an integer polynomial identity. In particular, $v^{(j)}(a)$ is a moment solution at every tuple a , whether or not $\Delta_C(a)$ vanishes.

Whenever $\Delta_C(a) \neq 0$, these vectors describe the entire kernel. Indeed, the chosen minor is nonzero, while all minors larger than k_C vanish identically, so $A_C(a)$ has rank exactly k_C . Each vector $v^{(j)}(a)$ has one free coordinate equal to $\Delta_C(a)$ and all others zero. They therefore form a basis of $\ker A_C(a)$. If there are no free coordinates, this kernel is zero.

It remains to show that the certificate vanishes on these kernel vectors. For each i, j , define

$$P_{i,j}(Y) = u_C^\top C_i v_i^{(j)}(Y) \in \mathbb{Z}[Y].$$

Fix C, m, r and let p tend to infinity. Since $v^{(j)}(a)$ is always a moment solution, Step 4 gives $P_{i,j}(a) = 0$ with probability $1 - o_{p \rightarrow \infty}(1)$ over $\mathcal{A}_{p,m}$. If $P_{i,j}$ were a nonzero integer polynomial, it would remain nonzero modulo every sufficiently large prime. The polynomial zero bound, conditioned on distinct nonzero places, would then give

$$\Pr_{a \in \mathcal{A}_{p,m}} [P_{i,j}(a) = 0] \leq \frac{2 \deg(P_{i,j})}{p} \rightarrow 0.$$

This contradicts the preceding conclusion. Hence every $P_{i,j}$ is identically zero over $\mathbb{Z}$.

These identities remain valid in every characteristic. For any prime and any tuple with $\Delta_C(a) \neq 0$, the certificate therefore vanishes on a basis of $\ker A_C(a)$, and hence on every moment

solution:

$$A_C(a)x = 0 \implies u_C^\top C_i x_i = 0 \quad (i \in [m]).$$

This proves Equation 39 at every such tuple.

(Step 6) *Exclude the zeros of the minors for all arrays at once.* By Equation 36 and Equation 44, the nonzero polynomial Δ_C remains nonzero modulo p : at least one nonzero integer coefficient has absolute value smaller than p . The polynomial zero bound gives

$$\Pr[\Delta_C(a) = 0] \leq \frac{mD(r-1)}{p}$$

for independent uniform $a_i \in \mathbb{F}_p$. The probability that some entry is zero or two entries coincide is at most

$$\frac{m + \binom{m}{2}}{p} = \frac{m(m+1)}{2p} < \frac{1}{2}.$$

Conditioning on $a \in \mathcal{A}_{p,m}$ therefore costs at most a factor of two. There are at most $(2B+1)^{qmD}$ integer arrays of height at most B . Their union bound is exactly Equation 38.

Each vector u_C was fixed before the union bound, and the event $\Delta_C(a) \neq 0$ covers the entire kernel of $A_C(a)$. Thus the good event is simultaneous over all arrays and all their solutions, as required.

Where the field bound enters. The condition $p > H$ keeps u_C nonzero modulo p ; $p > (mD)!B^{mD}$ preserves a nonzero coefficient of the maximal minor; and $p > m(m+1)$ controls conditioning on distinct nonzero places. The large-characteristic threshold used in Steps 3–4 has disappeared: it was needed only to establish integer polynomial identities. This is what allows the final field bound to be explicit.

7. THE STRUCTURED ESTIMATE: STATEMENT AND PREPARATIONS

This section states the structured estimate and develops the geometric and Fourier tools used in its proof. In Section 8, we combine these tools with the product lemma (Lemma 5.1) from Section 5 and the moment certificate (Lemma 6.1) from Section 6. The goal is the structured estimate Equation 10 used in Section 2. Unlike the reduction from arbitrary leakage, this estimate works at every fixed positive code rate. The restriction $c > 1/3$ has already entered through the simultaneous Cauchy–Schwarz argument of Lemma 2.1.

Recall the class $\mathcal{G} = \mathcal{G}_{p,3}$ from Definition 4.1. Its defining data $\mathcal{N}, D, Q$ depend only on the fixed relative degree c , through the approximation accuracy chosen in Section 2. In particular, these data are independent of m, p , the evaluation places, and the leakage maps. Thus, each local nilmanifold has dimension at most D , and each readout has C^R norm at most Q , where

$$K = D + 2, \quad R = 2D + 4.$$

Lemma 7.1 (Technical Lemma: Structured estimate). *Fix the data $\mathcal{N}, D, Q$ defining $\mathcal{G}_{p,3}$ and a real number $0 < b < 1$. There are constants $A, C > 0$, depending only on these data and b , such that for all sufficiently large m , every integer $bm \leq r < m$, and every odd prime p satisfying*

$$\log p \geq Cm^2 \log(2m),$$

we have

$$\mathbb{E}_{a \in \mathcal{A}_{p,m}} \Delta_{\mathcal{G}}(a; r) \leq A2^{-r/2}. \quad (45)$$

Here $\mathcal{A}_{p,m}$ is the set of ordered distinct nonzero evaluation tuples from Section 6.1. The supremum over the leakage maps is inside the expectation.

Proof roadmap. The Fourier argument needs probability maps on quadratic tori and a bound on the dimension of their supported Fourier relations. We use [Lemma 6.1](#) twice to obtain these controls. Together with [Lemma 5.1](#), its first application confines the relevant commutator phases to a bounded set of shares; central averaging and conditioning then give the torus models. Its second application bounds the span of the supported relations. We prepare these geometric and Fourier steps here; [Section 8](#) combines them to prove [Lemma 7.1](#). The preparation here is also used in proving [Proposition 10.3](#) in [Section 10.4](#).

7.1. The two experiments and their vertical products. Fix $a \in \mathcal{A}_{p,m}$ and choose arbitrary $f_i \in \mathcal{G}$, with representations

$$f_i(z) = F_i(g^{(i)}(z)\Gamma^{(i)}), \quad i \in [m].$$

The choices may depend on the entire tuple a . Write $F_{i,1} = F_i$ and $F_{i,0} = 1 - F_i$. Both readouts take values in $[0, 1]$ and have C^R norm at most Q . For a leakage profile $y \in \{0, 1\}^m$, the probabilities in the codeword and independent-input experiments are

$$\begin{aligned} P_f(y) &= \mathbb{E}_{x \in \mathbb{F}_p^r} \prod_i F_{i,y_i}(g^{(i)}(L_i(x))\Gamma^{(i)}), & L_i(x) &= \sum_{h=0}^{r-1} a_i^h x_h, \\ Q_f(y) &= \mathbb{E}_{z \in \mathbb{F}_p^m} \prod_i F_{i,y_i}(g^{(i)}(z_i)\Gamma^{(i)}). \end{aligned}$$

Thus $\Delta_{\mathcal{G}}(a; r) = \sup_f \text{SD}(P_f, Q_f)$.

Let $F_{i,\varepsilon,\nu}$ be the vertical components of $F_{i,\varepsilon}$ from [Section 4.4](#). By [Equation 24](#), there is a constant $C_0 \geq 1$, depending only on the fixed class data, such that

$$F_{i,\varepsilon} = \sum_{\nu} F_{i,\varepsilon,\nu}, \quad \sum_{\nu} \|F_{i,\varepsilon,\nu}\|_{C^K(\mathcal{X}^{(i)})} \leq C_0. \quad (46)$$

These series converge in C^K . For each nonzero component, write

$$F_{i,\varepsilon,\nu} = c_{i,\varepsilon,\nu} \Phi_{i,\varepsilon,\nu}, \quad c_{i,\varepsilon,\nu} = \|F_{i,\varepsilon,\nu}\|_{C^K(\mathcal{X}^{(i)})} > 0.$$

Then $\Phi_{i,\varepsilon,\nu}$ has vertical frequency ν and C^K norm 1, exactly as required by [Lemma 5.1](#).

Fix a threshold $0 < \delta \leq 1/10$. For each $i \in [m]$, choose an output bit $\varepsilon_i \in \{0, 1\}$ and a vertical frequency ν_i with $F_{i,\varepsilon_i,\nu_i} \neq 0$, and write

$$\Phi_i = \Phi_{i,\varepsilon_i,\nu_i}.$$

We call this choice of normalized components *significant* if

$$\max \left\{ \left| \mathbb{E}_{x \in \mathbb{F}_p^r} \prod_i \Phi_i(g^{(i)}(L_i(x))\Gamma^{(i)}) \right|, \left| \mathbb{E}_{z \in \mathbb{F}_p^m} \prod_i \Phi_i(g^{(i)}(z_i)\Gamma^{(i)}) \right| \right\} \geq \delta. \quad (47)$$

Significance refers to either experiment. This is necessary because the eventual perturbation estimate must control both leakage laws. The weights $c_{i,\varepsilon,\nu}$ will let us sum the insignificant products: for each profile, their total weight is at most C_0^m .

7.2. Global Equations and Local Commutator Directions. This subsection translates local and global character equations into spaces of allowable directions. We define the local subgroups H_i , the space E preserving all local equations, and its subspace V preserving the global equations as well. For each vertical frequency tuple, we define an alternating form that sums the commutator phases from individual shares. We then prove that, if these forms vanish on $V \times V$ for all significant products, only a common set of at most $\dim E - \dim V$ shares can contribute nonzero commutator phases.

The character conventions are those of [Section 3.4](#). Put $h_i = \dim G^{(i)} - \dim G_2^{(i)}$ and write the first-level positions as

$$\psi_1^{(i)}(g^{(i)}(z)) = \theta_i + \beta_i z/p \pmod{\mathbb{Z}^{h_i}}, \quad \theta_i \in \mathbb{T}^{h_i}, \quad \beta_i \in \mathbb{F}_p^{h_i}. \quad (48)$$

A local character with row $k_i \in \mathbb{Z}^{h_i}$ is constant modulo 1 precisely when $k_i \cdot \beta_i = 0$ in $\mathbb{F}_p$. A product character with rows $(k_i)_i$ is constant on the code orbit precisely when

$$\sum_i a_i^h(k_i \cdot \beta_i) = 0 \quad (0 \leq h < r). \quad (49)$$

On the independent-input orbit, constancy instead means $k_i \cdot \beta_i = 0$ for every i .

Choose cutoffs $B, H_{\text{loc}} \geq 2$, to be fixed in the proof, with $H_{\text{loc}} \geq B$. The following spaces are over $\mathbb{R}$; the preceding moment equations are over $\mathbb{F}_p$. Let W be the real span of the differentials of all product first-level characters of complexity at most B that are constant on the code orbit. For each i , let U_i be the real span of the differentials of all local first-level characters of complexity at most H_{loc} that are constant on $g^{(i)}$. Set

$$H_i = \exp(U_i^\perp), \quad E = \bigoplus_i \text{Lie}(H_i), \quad V = E \cap W^\perp, \quad s = \dim E - \dim V. \quad (50)$$

Here

$$U_i^\perp = \{v_i \in \text{Lie}(G^{(i)}) : \lambda(v_i) = 0 \text{ for every } \lambda \in U_i\}.$$

Thus $U_i^\perp$ consists of the directions of motion that preserve all the chosen local character equations at share i . The notation denotes a common kernel; no inner product is needed. Thus E consists of the local directions preserving all local equations, and V also preserves the global equations. The codimension s counts the global constraints that remain after imposing the local ones. The first certificate application will show that s is bounded independently of m .

Each $H_i = \exp(U_i^\perp)$ is a connected rational subgroup containing $G_2^{(i)}$. To see this, first note that the local characters depend only on first-level coordinates. Their differentials therefore vanish on $\text{Lie}(G_2^{(i)})$, so

$$\text{Lie}(G_2^{(i)}) \subseteq U_i^\perp.$$

Moreover, every bracket lies in $\text{Lie}(G_2^{(i)})$. In particular,

$$v, w \in U_i^\perp \implies [v, w] \in \text{Lie}(G_2^{(i)}) \subseteq U_i^\perp.$$

Thus $U_i^\perp$ is a Lie sub-algebra. Exponentiating it gives a connected subgroup H_i containing $G_2^{(i)}$.

Finally, the equations defining $U_i^\perp$ have integer coefficients relative to the ambient Mal'cev basis. Gaussian elimination therefore gives a basis with rational coordinates. By the criterion in [Section 3.6](#), H_i is rational: $H_i \cap \Gamma^{(i)}$ is a cocompact lattice in H_i .

For a tuple of vertical frequencies $\nu = (\nu_i)_i$, define the real alternating form on E by

$$\omega_\nu(v, w) = \sum_i d\nu_i([v_i, w_i]). \quad (51)$$

The form is block diagonal, with one block per share. Write $\omega_{\nu,i}$ for its restriction to $\text{Lie}(H_i) \times \text{Lie}(H_i)$. The identity $\omega_{\nu,i} = 0$ is equivalent to $d\nu_i$ vanishing on $\text{Lie}([H_i, H_i])$. Since $[H_i, H_i]$ is connected, this is equivalent to ν_i being trivial on the commutator torus

$$Z_i = [H_i, H_i]/([H_i, H_i] \cap \Gamma^{(i)}).$$

The subgroup is rational and central; Step (3) in the proof of [Lemma 7.3](#) below includes a verification. By [Section 4.4](#), averaging over Z_i keeps exactly the vertical components with the vanishing property, preserves the range $[0, 1]$, commutes with complements, and does not increase the C^R norm.

Why one exceptional set will suffice. We seek one small set of shares J such that averaging over Z_i for every $i \notin J$ preserves every significant vertical product. For each significant ν , define $J_\nu = \{i: \omega_{\nu,i} \neq 0\}$ as the shares that must be left unaveraged to preserve the vertical product. The product survives averaging outside J exactly when $J_\nu \subseteq J$. It suffices to bound the union of the nonzero block supports J_ν of all significant forms.

Suppose all the significant forms Equation 51 vanish on the same subspace $V \times V$. Write $E = V \oplus L$ where $\dim(L) = s$. Choose a basis of V , followed by a basis of L . The matrix for ω_ν has the form

$$M_\nu = \begin{pmatrix} 0 & A \\ -A^\top & C \end{pmatrix}$$

where the zero block captures exactly that $\omega_\nu(v, w) = 0$ for $v, w \in V$ and $A \in \mathbb{R}^{(\dim V) \times s}$, $C \in \mathbb{R}^{s \times s}$. It maps V into the annihilator of V , which has dimension s , and its image on L has dimension at most s . Thus, an alternating form ω_ν with an isotropic subspace of codimension s has rank at most $2s$. Recall ω_ν is block diagonal and each diagonal block is itself alternating, since $d\nu_i([v_i, v_i]) = 0$ and $d\nu_i([w_i, v_i]) = -d\nu_i([v_i, w_i])$. Each nonzero alternating block contributes at least two to the rank. Then,

$$2|J_\nu| \leq \text{rank}(\omega_\nu) \leq 2s, \quad |J_\nu| \leq s,$$

each alternating form ω_ν has at most s nonzero blocks.

But individual size bounds alone do not bound their union. To preserve every significant product while averaging outside a single set J , we need more: the *same* at most s blocks must contain the support of every significant form. We will construct a linear combination whose block support is exactly

$$J = \bigcup_{\nu \text{ significant}} J_\nu.$$

Take the real span of the significant forms and choose a finite basis $\omega^{(1)}, \dots, \omega^{(\ell)}$ of this span. If $\ell = 0$, every significant form is zero, so we take $J = \emptyset$ and the conclusion is immediate. Henceforth assume $\ell \geq 1$. Every linear combination still vanishes on $V \times V$, so it still has at most s nonzero blocks. For coefficients $\alpha = (\alpha_1, \dots, \alpha_\ell)$, consider $\omega(\alpha) = \sum_{j=1}^\ell \alpha_j \omega^{(j)}$ and its i th block is $\omega_i(\alpha) = \sum_{j=1}^\ell \alpha_j \omega_i^{(j)}$. For each block $i \in J$, the coefficient choices canceling the block form a proper linear subspace ¹¹

$$N_i = \left\{ \alpha \in \mathbb{R}^\ell: \sum_{j=1}^\ell \alpha_j \omega_i^{(j)} = 0 \right\}.$$

We choose the combination coefficients outside every N_i . Start with $\Theta_1 = \omega^{(1)}$. For $j = 2, \dots, \ell$, choose a real number t and set $\Theta_j = \Theta_{j-1} + t\omega^{(j)}$. For each share i in the union of the supports of Θ_{j-1} and $\omega^{(j)}$, the new block is $A_i + tB_i$, where $A_i = (\Theta_{j-1})_i$ and $B_i = \omega_i^{(j)}$. These two blocks are not both zero.

- If $B_i = 0$, then $A_i \neq 0$ and $A_i + tB_i = A_i \neq 0$ for every real number t .
- If $B_i \neq 0$, at most one value of t can make the block zero; otherwise, if $A_i + tB_i = 0$ and $A_i + t'B_i = 0$, subtraction gives $(t - t')B_i = 0$ and therefore $t = t'$.

There are finitely many blocks and each excludes at most one $t \in \mathbb{R}$. Choose t avoiding all these values. Every block occurring in either form Θ_{j-1} or $\omega^{(j)}$ remains nonzero while blocks absent from both remain zero. Repeating the steps gives a linear combination Θ_ℓ supported on J which is the union of J_ν over all significant ν . It follows from the same rank argument on Θ_ℓ that the union has size at most s .

¹¹The set N_i is a linear subspace because the map $\alpha \mapsto \omega_i(\alpha)$ is linear. It is proper because for $i \in J$, there exists some j with $\omega_i^{(j)} \neq 0$ and the coefficient vector having $\alpha_j = 1$ and all other entries zero does not belong to N_i .

This gives one set J for all output profiles and all significant frequency tuples, once the local representations are fixed.

7.3. Local Reduction and Positive Torus Approximation. We first record a geometric input from Leng's thesis [Len25] to control coordinates and metrics on rational subgroups and their quotients. We then prove two local lemmas: Lemma 7.3 reduces a commutator-invariant readout to a probability map on a quadratic torus orbit, and Lemma 7.4 gives a finite Fourier approximation that preserves its range $[0, 1]$. All local dimensions are at most the fixed bound D .

Imported Lemma 7.2 (Rational Subgroup and Quotient Coordinates [Len25]). *Fix a dimension bound D . Let G/Γ be a degree-two filtered nilmanifold with an adapted M_{geom} -rational Mal'cev basis, where $M_{\text{geom}} \geq 2$, and $\dim G \leq D$. Use the rationality convention of Section 3.6. Let $H \leq G$ be a connected M_{geom} -rational subgroup and $\Gamma_H = H \cap \Gamma$.*

There is an adapted Mal'cev basis for H/Γ_H with rationality at most $M_{\text{geom}}^{C_D}$, whose vectors are $M_{\text{geom}}^{C_D}$ -rational combinations of the ambient basis.¹² The quotient metric on H/Γ_H is comparable, within a factor $M_{\text{geom}}^{C_D}$, to the restriction of the quotient metric on G/Γ , where C_D depends only on D .¹³ If $Z \leq H$ is connected, central, and M_{geom} -rational, then the quotient

$$(H/Z)/\pi_Z(\Gamma_H)$$

has an adapted basis with the same type of bound, where $\pi_Z: H \rightarrow H/Z$ is the quotient map.¹⁴ A Z -invariant Lipschitz function on H/Γ_H descends to this quotient, with Lipschitz norm increased by at most $M_{\text{geom}}^{C_D}$.¹⁵ Multiplication by an ambient element in the fixed Mal'cev fundamental box also incurs at most $M_{\text{geom}}^{C_D}$ in the metric comparison.¹⁶

Lemma 7.3 (Local Reduction to a Quadratic Torus). *Let $D \geq 1$ be the fixed integer dimension bound. Fix one of the local degree-two filtered nilmanifolds G/Γ in the class $\mathcal{G}$, of dimension at most D . Let $H_0 \geq 2$ and let p be a prime satisfying $p > D!H_0^D$. Consider a finite family of first level characters $\xi_j: G \rightarrow \mathbb{R}$ with integer coefficient rows of maximum norm at most H_0 . Suppose $\xi_j(g(z))$ is constant modulo 1 for every j , where $g: \mathbb{Z} \rightarrow G$ is a degree-two polynomial sequence for the given filtration and $g(z)\Gamma$ is p -periodic. Set*

$$H = \exp\left(\bigcap_j \ker d\xi_j\right), \quad \Gamma_H = H \cap \Gamma, \quad Z = [H, H].$$

Then, there exists an element c in the fixed Mal'cev fundamental box in G and a degree-two polynomial sequence $h: \mathbb{Z} \rightarrow H$, for its inherited filtration, such that

$$g(z)\Gamma = ch(z)\Gamma, \quad h(0) = \text{id}_G, \quad h(z)\Gamma_H \text{ is } p\text{-periodic}.$$

The groups H and Z are connected and rational, and Z is central in G . Let $d_H := \dim(H) - \dim(Z)$. If $F: G/\Gamma \rightarrow [0, 1]$ is Lipschitz and invariant under Z , then $h \mapsto F(ch\Gamma)$ descends to

$$H/(Z\Gamma_H) \simeq \mathbb{T}^{d_H}, \quad 0 \leq d_H \leq D.$$

In suitable group coordinates on the torus, the descended function A satisfies

$$\|A\|_{\text{Lip}} \leq H_0^{C_{\mathcal{G}}} \|F\|_{\text{Lip}},$$

¹²Lemma 7.2.15 [Len25], p. 317, constructs the subgroup basis by taking its subgroup $G' = H$.

¹³Lemma 7.2.11 [Len25], pp. 312–313, compares the quotient metrics on H/Γ_H and G/Γ .

¹⁴Lemma 2.3.11 [Len25], pp. 33–34, treats rational quotients and descent of central-invariant functions. Apply it with ambient group H and normal subgroup $Z \leq H$.

¹⁵Lemma 2.3.11 [Len25], pp. 33–34, final clause. Descent of a Z -invariant function, with controlled Lipschitz norm

¹⁶Lemma 7.2.6 [Len25], pp. 308–309, controls bounded left translation. Their bounds have the form $M_{\text{geom}}^{O_{\ell-1}(D^{O(1)})}$, where the subscript records dependence on the filtration degree $\ell - 1$. Here $\ell = 3$, so the filtration degree is 2, and these bounds are $M_{\text{geom}}^{C_D}$ for a constant C_D depending only on D .

where $C_G \geq 1$ is a constant depending only on the fixed geometric data of the class $\mathcal{G}$, the dimension bound D and the chosen Mal'cev bases. Its orbit is a p -periodic quadratic torus polynomial. For odd p , it therefore has the form

$$\vartheta + \frac{\beta_1 z + \beta_2 z^2}{p} \pmod{\mathbb{Z}^{d_H}}, \quad \vartheta \in \mathbb{T}^{d_H}, \quad \beta_1, \beta_2 \in \mathbb{F}_p^{d_H}.$$

Proof of the local reduction lemma, Lemma 7.3. *Proof roadmap and relation to earlier work.* First normalize the constant term. Character constancy modulo 1 then gives integer values for the defining linear forms on the first-level slope, while periodicity makes this slope p -rational. A determinant bound and Bézout's identity then supply a lattice correction that makes these values vanish over $\mathbb{R}$. After this correction, the orbit lies in the connected subgroup H without changing its cosets or its period. This subgroup reduction follows the periodic factorization strategy of Candela–Sisask [CS14, arXiv version, Lemma 5.5, pp. 15–16; Lemma B.1, pp. 28–30] and Leng [Len25, Lemmas 3.7.1 and 3.7.3, pp. 94–96]. We give a direct degree-two argument to retain the explicit condition $p > D!H_0^D$. Finally, invariance under $[H, H]$ lets the readout descend to an ordinary torus. Imported Lemma 7.2 controls its Lipschitz norm, and periodicity gives the denominator- p quadratic representation.

To start the proof, write $d_1 = \dim G - \dim G_2$ for the number of first-level coordinates. Choose $q_1 \leq d_1$ independent rows from the defining characters, and put them in a matrix $C \in \mathbb{Z}^{q_1 \times d_1}$. The rows of C have the same real common kernel as the coefficient rows of all the characters ξ_j . If $q_1 = 0$, no lattice correction is needed: take $H = G$ and $h = v$ after Step 1, and proceed to Step 3.

(Step 1) *Normalize the constant term.* Using the centered decomposition from Section 3.3, write $g(0) = c\gamma_0$, where $\gamma_0 \in \Gamma$ and c belongs to the fixed fundamental box in G . Set

$$v(z) = c^{-1}g(z)\gamma_0^{-1}.$$

Then $v(0) = \text{id}_G$, $cv(z)\Gamma = g(z)\Gamma$, and $v(z)\Gamma$ is p -periodic. Conjugation by γ_0 preserves the filtration, so the Taylor form from Section 3.2 gives

$$v(z) = v_1^z v_2^{\binom{z}{2}}, \quad v_1 \in G, \quad v_2 \in G_2.$$

In particular, its real first-level coordinates are $\psi_1(v(z)) = z\alpha$, where $\alpha = \psi_1(v_1) \in \mathbb{R}^{d_1}$. Character constancy and periodicity give, respectively,

$$u = C\alpha \in \mathbb{Z}^{q_1}, \quad p\alpha \in \mathbb{Z}^{d_1}. \tag{52}$$

The first assertion follows from $\xi_j(g(1)) - \xi_j(g(0)) \in \mathbb{Z}$ for characters whose coefficient rows are from C . The second follows from $v(p)\Gamma = \Gamma$ which implies $v(p) \in \Gamma$ and $p\alpha = \psi_1(v(p)) \in \mathbb{Z}^{d_1}$.

(Step 2) *Replace integer equation $C\alpha = u$ by an exact vanishing equation.* We seek a lattice correction for $v(z)$ to make it lie in H . But membership in H requires $C\psi_1(v(z))$ to be zero. We therefore seek an integer vector $\zeta \in \mathbb{Z}^{d_1}$ satisfying $C\zeta = u$, so that

$$C\psi_1(v(z)) - C(z\zeta) = zC(\alpha - \zeta) = 0.$$

An invertible minor and Bézout's identity will produce such a vector.

Assume $q_1 > 0$. The matrix C has q_1 independent rows, so its rank is q_1 . Choose a nonsingular $q_1 \times q_1$ minor M of C and write $d = \det M$. The determinant expansion¹⁷ gives

$$0 < |d| \leq q_1! H_0^{q_1} \leq D! H_0^D < p.$$

¹⁷The determinant expansion is

$$d = \sum_{\sigma \in S_{q_1}} \text{sgn}(\sigma) \prod_{i=1}^{q_1} M_{i, \sigma(i)}$$

where S_{q_1} is the set of permutations of $\{1, \dots, q_1\}$. There are $q_1!$ summands and each product contains q_1 entries bounded by H_0 .

The adjugate identity $M \operatorname{adj}(M) = dI$ shows that $du = M(\operatorname{adj}(M)u) \in C\mathbb{Z}^{d_1}$,¹⁸ while Equation 52 gives $pu = C(p\alpha) \in C\mathbb{Z}^{d_1}$. Since p is prime and $0 < |d| < p$, Bézout's identity implies there exist integers a, b with $ad + bp = 1$ and

$$u = a(du) + b(pu) \in C\mathbb{Z}^{d_1}.$$

Choose $\zeta \in \mathbb{Z}^{d_1}$ with $C\zeta = u$, and choose the lattice correction $\gamma \in \Gamma$ with $\psi_1(\gamma) = \zeta$; the adapted lattice coordinates guarantee the latter choice. Put

$$h(z) = v(z)\gamma^{-z}.$$

Now, by additivity, $C\psi_1(h(z)) = zC(\alpha - \zeta) = 0$, so $h(z) \in H$ for every $z \in \mathbb{Z}$. With the convention $[u, v]vu = uv$, the commutator factor $[v_1, \gamma^{-1}]$ accounts for moving γ^{-1} factor past the v_1 factor. The degree-two product identity gives

$$h(z) = (v_1\gamma^{-1})^z(v_2[v_1, \gamma^{-1}])^{\binom{z}{2}}.$$

Its first coefficient $v_1\gamma^{-1} \in H$ and its second coefficient $v_2[v_1, \gamma^{-1}] \in G_2 \subseteq H$. Thus it is polynomial for the inherited filtration $H \supseteq G_2 \supseteq \{\operatorname{id}_G\}$. As $h(z)\Gamma = v(z)\Gamma$, it has the required ambient cosets. Moreover $h(z)^{-1}h(z+p) \in \Gamma \cap H$, which proves periodicity modulo Γ_H . No size bound on ζ or γ is required. The lattice correction therefore gives a normalized, p -periodic quadratic orbit on H/Γ_H representing the same cosets in G/Γ ; $h(z)$ is a valid replacement for $v(z)$.

(Step 3) *Take the abelian quotient and control its metric.* The Lie algebra of H consists of the vectors whose first-level coordinates lie in $\ker C$, with no restriction on their G_2 -coordinates. Cramer's rule supplies a rational basis of $\operatorname{Lie}(H)$ whose height is bounded by $D!H_0^D = H_0^{O_D(1)}$.

We will show that the subgroup $Z = [H, H]$ is also rational and central so we can apply Imported Lemma 7.2. Let $\mathfrak{h} = \operatorname{Lie}(H)$ be the Lie algebra of the Lie group H . Taking brackets of the rational basis vectors $U_1, \dots, U_\ell$ of $\mathfrak{h}$ gives a spanning family of

$$\mathfrak{c} = [\mathfrak{h}, \mathfrak{h}] = \operatorname{span}_{\mathbb{R}}\{[U_a, U_b] : a < b\}$$

whose coordinate height is at most $H_0^{c_g}$ for some constant c_g . Since the U_a have rational coordinates in the ambient Mal'cev basis, so does every bracket $[U_a, U_b]$. Selecting a linearly independent spanning subfamily of the brackets from existing vectors gives the required rational basis of $\mathfrak{c}$. Their height bound does not increase. In a simply connected two-step nilpotent group, we can identify this Lie algebra with subgroup

$$Z = [H, H] = \exp(\mathfrak{c}),$$

following from the fact that $[\exp X, \exp Y] = \exp([X, Y])$ and all such brackets are central. Thus, Z is connected as the continuous image of the vector space $\mathfrak{c}$, central because $\mathfrak{c} \subseteq \operatorname{Lie}(G_2)$, and rational because $\mathfrak{c}$ has rational basis constructed above.

Next, we will remove the commutator directions Z so that the projected orbits live on an additive torus, where we can later use Fourier approximation. Apply Imported Lemma 7.2 first to the subgroup H , using the quantitative rationality established above. With $\Gamma_H = H \cap \Gamma$, this gives an adapted Mal'cev basis for H/Γ_H and a controlled comparison with the ambient metric. Then, apply the central-quotient clause in Imported Lemma 7.2 to the connected, central, and quantitatively rational $Z = [H, H]$ and get a cocompact lattice $\Lambda := \pi_Z(\Gamma_H) \leq H/Z$ along with a controlled adapted basis for $(H/Z)/\Lambda$, where $\pi_Z : H \rightarrow H/Z$ is the quotient map. The quotient H/Z is abelian and connected, with a cocompact lattice Λ ; hence $H/(Z\Gamma_H)$ is a torus of dimension at most D ,

$$H/(Z\Gamma_H) \simeq (H/Z)/\Lambda \simeq \mathbb{T}^{d_H}, \quad 0 \leq d_H \leq D.$$

¹⁸ $\operatorname{adj}(M)$ is the transpose of the cofactor matrix; its entries are signed determinants of smaller integer submatrices, so $\operatorname{adj}(M)u \in \mathbb{Z}^{q_1}$. du is an integer combination of columns of M ; they are selected columns of C .

Since the readout $F: G/\Gamma \rightarrow [0, 1]$ is invariant under Z , forgetting the Z directions should not change its value. We want to express the original readout F exactly as a function on the torus. Its descended readout $A: H/(Z\Gamma_H) \rightarrow [0, 1]$ with $A(hZ\Gamma_H) := F(ch\Gamma)$ is well defined because $F(chw\Gamma) = F(wch\Gamma) = F(ch\Gamma)$ for $w \in Z$. Project the corrected orbit to the torus and get $q(z) := h(z)Z\Gamma_H$. By construction, $g(z)\Gamma = ch(z)\Gamma$, and hence

$$F(g(z)\Gamma) = F(ch(z)\Gamma) = A(q(z))$$

which is the desired representation of the readout F along the original orbit $g(z)$ as a torus function A evaluated along the projected orbit $q(z)$.

Imported Lemma 7.2 controls the subgroup and quotient metrics. Define $f_H(h\Gamma_H) = F(ch\Gamma)$. Since c lies in the fixed fundamental box, the bounded-translation and subgroup-metric clauses of **Imported Lemma 7.2** give

$$\|f_H\|_{\text{Lip}} \leq M_{\text{geom}}^{2C_D} \|F\|_{\text{Lip}}$$

where C_D depends only on the dimension bound D and $M_{\text{geom}} \geq 2$ is a common height bound used to describe the geometry. Since f_H is Z -invariant, the descent clause then gives

$$\|A\|_{\text{Lip}} \leq M_{\text{geom}}^{C_D} \|f_H\|_{\text{Lip}} \leq M_{\text{geom}}^{3C_D} \|F\|_{\text{Lip}}.$$

Combining the rationality bounds above with the metric conclusions gives

$$\|A\|_{\text{Lip}} \leq H_0^{C_G} \|F\|_{\text{Lip}},$$

for a sufficiently large constant C_G depending only on the fixed geometric data of $\mathcal{G}$.

Finally, we find the explicit form of the torus orbit $q(z)$. A quadratic polynomial in an additive torus has Taylor form $q(z) = \vartheta + z\lambda_1 + \binom{z}{2}\lambda_2$. Periodicity gives $p\lambda_2 = 0$ and $p\lambda_1 + \binom{p}{2}\lambda_2 = 0$ modulo the torus lattice. For odd p , $\binom{p}{2}$ is divisible by p , so $p\lambda_1 = 0$. Write $\lambda_j = \alpha_j/p$ for $\alpha_j \in \mathbb{F}_p^{d_H}$, $j = 1, 2$. Then the required coefficients are

$$\beta_2 = 2^{-1}\alpha_2, \quad \beta_1 = \alpha_1 - 2^{-1}\alpha_2 \quad \text{in } \mathbb{F}_p^{d_H}.$$

Here 2^{-1} is taken in $\mathbb{F}_p$. Therefore,

$$q(z) = \vartheta + \frac{\beta_1 z + \beta_2 z^2}{p}, \quad \beta_1, \beta_2 \in \mathbb{F}_p^{d_H}.$$

We have obtained the explicit quadratic torus model, with the exact identity

$$F(g(z)\Gamma) = A(q(z)).$$

Why connectedness matters in lattice correction. Character constancy modulo 1 only gives $C\psi_1(v(z)) \in \mathbb{Z}^{q_1}$ whereas membership in the connected subgroup H requires $C\psi_1(h(z)) = 0$ over $\mathbb{R}$. The subgroup defined by the first condition can be disconnected. The determinant-Bézout argument lets us correct the orbit representatives so that the second condition holds, while preserving their cosets. The explicit Taylor formula checks that it preserves degree two. This supplies the connected subgroup needed for **Imported Lemma 7.2** and the torus reduction. Only the defining character heights affect the quotient metric; the possibly large correction γ does not enter the readout.

Lemma 7.4 (Positive Approximation on a Torus). *Fix a dimension $0 \leq d_H \leq D$. For any $Q_0 \geq 1$, any Q_0 -Lipschitz probability function $A: \mathbb{T}^{d_H} \rightarrow [0, 1]$, and $0 < \varepsilon < 1$, positive product Fejér convolution gives a trigonometric polynomial $A_T: \mathbb{T}^{d_H} \rightarrow [0, 1]$ with*

$$\|A - A_T\|_\infty \leq \varepsilon, \quad \text{supp } \hat{A}_T \subseteq [-T, T]^{d_H} \cap \mathbb{Z}^{d_H}, \quad T \leq 1 + \left(\frac{2DQ_0}{\varepsilon} \right)^2.$$

It sends $1 - A$ to $1 - A_T$. Larger T retain the discrepancy, range, complement, and Fourier-support conclusions.

Proof of Positive Approximation Lemma 7.4. For an integer $T \geq 1$, the one-dimensional Fejér kernel is

$$K_T(t) = \frac{1}{T} \left| \sum_{j=0}^{T-1} e(jt) \right|^2 = \sum_{|k| < T} \left(1 - \frac{|k|}{T} \right) e(kt), \quad t \in \mathbb{T}.$$

It is nonnegative, has integral 1, and has Fourier support $|k| < T$. If $\|t\|_{\mathbb{T}}$ denotes the distance to the nearest integer, the geometric-series formula and $|\sin(\pi t)| \geq 2\|t\|_{\mathbb{T}}$ give

$$K_T(t) \leq \min \left\{ T, \frac{1}{4T\|t\|_{\mathbb{T}}^2} \right\}.$$

For $T \geq 4$, split the integral at $\|t\|_{\mathbb{T}} = T^{-1/2}$. The inner part contributes at most $T^{-1/2}$; bounding distance by $1/2$ on the outer part gives

$$\int_{\mathbb{T}} \|t\|_{\mathbb{T}} K_T(t) dt \leq T^{-1/2} + \frac{1}{4T \cdot T^{-1/2}} \leq 2T^{-1/2}.$$

The last bound also holds for $T < 4$. Convolve A with the product kernel $\prod_{j=1}^{d_H} K_T(t_j)$, and call the result A_T . Using the usual product torus metric, its discrepancy is at most

$$\|A - A_T\|_{\infty} \leq Q_0 \sum_{j=1}^{d_H} \int_{\mathbb{T}} \|t_j\|_{\mathbb{T}} K_T(t_j) dt_j \leq \frac{2DQ_0}{\sqrt{T}}.$$

Thus $T = \lceil (2DQ_0/\varepsilon)^2 \rceil$ suffices. Positivity and unit integral preserve $[0, 1]$, and linearity and $1 * K_T = 1$ preserve complements. The displayed Fourier expansion gives the required cutoff in every coordinate. If $d_H = 0$, the function is constant and the conclusions hold without approximation.

7.4. Conditioning and the Finite Fourier Calculation. After averaging over the local commutator tori, Lemma 7.3 gives probability maps on ordinary tori, evaluated along quadratic orbits. Before using Fourier analysis, we record how conditioning changes the code and how to combine torus frequencies that agree on its finite orbit. The latter step is needed for the coefficient bound that produces exponential decay.

Revealing a bounded set of shares. Let $J \subseteq [m]$ with $j = |J| < r$. The values $(P(a_i))_{i \in J}$ of a uniform polynomial of degree less than r are independent and uniform, by the Vandermonde property. Conditional on their values, the polynomial is uniform among

$$P(X) = P_0(X) + \prod_{j \in J} (X - a_j) Q(X), \quad \deg P_0 < |J|, \quad \deg Q < r' = r - |J|. \quad (53)$$

Here P_0 is the interpolating polynomial; when $J = \emptyset$, take $P_0 = 0$. Writing $I = J^c$, the input at each remaining share $i \in I$ is

$$P_0(a_i) + c_i \tilde{L}_i(x), \quad c_i = \prod_{j \in J} (a_i - a_j) \neq 0, \quad \tilde{L}_i(x) = \sum_{h=0}^{r'-1} a_i^h x_h.$$

In the independent-input experiment, condition on $U_j = u_j$ for $j \in J$, using the same revealed values as in the codeword experiment. The remaining inputs $(U_i)_{i \in I}$ are still independent and uniform in $\mathbb{F}_p$. The affine changes above are bijections, so they can be absorbed into the local maps.

Quadratic torus maps. After this absorption, write a local torus map as

$$\mathbb{F}_p \ni z \mapsto A_i \left(\vartheta_i + \frac{\beta_{i,1}z + \beta_{i,2}z^2}{p} \right) \in [0, 1], \quad (54)$$

$$A_i: \mathbb{T}^{d_i} \rightarrow [0, 1], \quad d_i \leq D, \quad \vartheta_i \in \mathbb{T}^{d_i}, \quad \beta_{i,1}, \beta_{i,2} \in \mathbb{F}_p^{d_i}.$$

More explicitly, precomposition by $z \mapsto P_0(a_i) + c_i z$ changes the constant position and sends the coefficient vectors to

$$\beta_{i,1} \mapsto c_i \beta_{i,1} + 2c_i P_0(a_i) \beta_{i,2}, \quad \beta_{i,2} \mapsto c_i^2 \beta_{i,2}. \quad (55)$$

The readout A_i and its Lipschitz bound are unchanged. The inputs to the modified maps are the forms $\tilde{L}_i(x)$ in the code experiment and independent uniform field elements in the comparison experiment.

Combine equal frequencies before using Parseval. Suppose A_i is a trigonometric polynomial supported in $[-T_{\text{tor}}, T_{\text{tor}}]^{d_i} \cap \mathbb{Z}^{d_i}$, as supplied by [Lemma 7.4](#). Define the finite-field space and its readout by

$$\mathbf{E}_i = \{s\beta_{i,1} + t\beta_{i,2} : s, t \in \mathbb{F}_p\} \subseteq \mathbb{F}_p^{d_i}, \quad A_i^*(v) = A_i(\vartheta_i + v/p) \quad (v \in \mathbf{E}_i).$$

The notation v/p means the natural embedding of $\mathbb{F}_p^{d_i}$ into the p -torsion of $\mathbb{T}^{d_i}$. The space $\mathbf{E}_i$ is spanned by the orbit coefficients; no assertion that the quadratic orbit is uniform on this space is needed. It is different from the real space E in [Equation 50](#).

Let $\mathbf{E}_i^* = \text{Hom}_{\mathbb{F}_p}(\mathbf{E}_i, \mathbb{F}_p)$ be the dual vector space, consisting of all $\mathbb{F}_p$ -linear maps from $\mathbf{E}_i$ to $\mathbb{F}_p$. The first equality below defines our normalized Fourier coefficients; the second is the Fourier inversion formula:

$$\hat{A}_i^*(\gamma) = \mathbb{E}_{v \in \mathbf{E}_i} A_i^*(v) e_p(-\gamma(v)), \quad A_i^*(v) = \sum_{\gamma \in \mathbf{E}_i^*} \hat{A}_i^*(\gamma) e_p(\gamma(v)), \quad (56)$$

where $e_p(t) = \exp(2\pi i t/p)$. An integer frequency $k \in \mathbb{Z}^{d_i}$ restricts to the character $\gamma_k(v) = (k \bmod p) \cdot v$. Thus

$$\hat{A}_i^*(\gamma) = \sum_{k: \gamma_k = \gamma} \hat{A}_i(k) e(k \cdot \vartheta_i). \quad (57)$$

The sum combines every integer frequency inducing the same finite character. Every nonzero finite Fourier coefficient therefore has a lift of maximum norm at most T_{tor} .

Set $\sigma = 2^{-1/2}$. For the two output bits, introduce the combined weights

$$\mu_i(0) = 1, \quad \mu_i(\gamma) = 2 \left| \hat{A}_i^*(\gamma) \right| \quad (\gamma \neq 0). \quad (58)$$

At each frequency, μ_i is exactly the sum of the absolute Fourier coefficients of A_i^* and $1 - A_i^*$. At zero, their means add to 1; elsewhere their coefficients are negatives of one another.

Lemma 7.5 (Fourier weights of a probability map). *The weights in [Equation 58](#) satisfy*

$$\sum_{\gamma} \mu_i(\gamma)^2 \leq 2, \quad \mu_i(\gamma) \leq \sigma \quad (\gamma \neq 0). \quad (59)$$

Their support has size at most $(2T_{\text{tor}} + 1)^D$.

Proof. Put $\bar{A}_i = \mathbb{E}_v A_i^*(v)$. Parseval and $0 \leq A_i^* \leq 1$ give

$$\sum_{\gamma \neq 0} |\hat{A}_i^*(\gamma)|^2 = \mathbb{E}_v (A_i^*(v))^2 - \bar{A}_i^2 \leq \bar{A}_i(1 - \bar{A}_i) \leq \frac{1}{4}.$$

This proves the first bound after multiplying by four and adding $\mu_i(0)^2 = 1$. Since A_i^* is real, the coefficients at γ and $-\gamma$ have equal magnitudes. For a nonzero γ these are distinct because p is odd. Consequently each has squared magnitude at most $1/8$, giving the second bound.

Equation (57) bounds the number of supported finite coefficients by the number of integer frequencies. If $A_i^* \neq 0$, positivity implies $\hat{A}_i^*(0) > 0$, so adding the weight at zero introduces no new frequency. If $A_i^* = 0$, the weight is supported only at zero. This proves the support bound in both cases. $\square$

For $\gamma \neq 0$, the sharp bound $\mu_i(\gamma) \leq [p \sin(\pi/(2p))]^{-1} = 2/\pi + \mathcal{O}(p^{-2})$ has interval indicators pulled back along γ as extremizers, echoing the extremality of intervals in Lev's majorization theorem [Lev01, Theorem 1, p. 244], and supplies the Fourier improvement used in Section 9 to obtain the rate $(2/\pi)^{(1-o(1))d}$.

Which Fourier tuples are relations? For $\gamma = (\gamma_i)_{i \in I}$, define the polynomial

$$Q_\gamma(x) = \sum_{i \in I} \left(\gamma_i(\beta_{i,1}) \tilde{L}_i(x) + \gamma_i(\beta_{i,2}) \tilde{L}_i(x)^2 \right).$$

Let $\mathcal{R}$ be the $\mathbb{F}_p$ -vector space of tuples for which Q_γ is identically zero as a polynomial. Only tuples supported by the Fourier expansions can contribute to the leakage probabilities. Accordingly, put

$$S_i = \text{supp } \mu_i = \{0\} \cup \text{supp } \hat{A}_i^*, \quad V_0 = \text{span}_{\mathbb{F}_p} \left(\mathcal{R} \cap \prod_{i \in I} S_i \right), \quad s_0 = \dim_{\mathbb{F}_p} V_0. \quad (60)$$

The full relation space $\mathcal{R}$ may have large dimension. The second certificate application will bound only s_0 , which is all the Fourier sum needs.

An *information set* for V_0 is a set of blocks $I_0 \subseteq I$ such that projection onto $\bigoplus_{i \in I_0} \mathbf{E}_i^*$ is injective on V_0 . The values on these blocks determine the entire relation. An information set with at most s_0 blocks always exists: keep adding a block that decreases the projection's kernel.

Lemma 7.6 (Summing the supported relations). *Suppose $2s_0 \leq r' + 1$. Then*

$$\sum_{\substack{\gamma \in \mathcal{R} \\ \gamma \neq 0}} \prod_{i \in I} \mu_i(\gamma_i) \leq 2^{s_0} \sigma^{r'+1-2s_0}. \quad (61)$$

Proof. First, every nonzero relation uses at least $r' + 1$ blocks. Otherwise, some nonzero relation would have support S of size at most r' , so the Vandermonde property would make the forms $(\tilde{L}_i)_{i \in S}$ linearly independent. Use these forms as part of a coordinate system, writing their values as t_i . The relation would become

$$\sum_{i \in S} (\gamma_i(\beta_{i,1}) t_i + \gamma_i(\beta_{i,2}) t_i^2) = 0.$$

Comparison of coefficients forces each γ_i to annihilate both generators of $\mathbf{E}_i$, so each $\gamma_i = 0$. This contradicts the choice of a nonzero relation.

Thus every nonzero relation has at least $r' + 1$ nonzero blocks. Choose an information set I_1 with $|I_1| \leq s_0$. Projection onto its complement is also injective: its kernel would otherwise contain a nonzero relation supported on at most $s_0 < r' + 1$ blocks. Inside that complement choose a second information set I_2 with $|I_2| \leq s_0$. Thus I_1, I_2 are disjoint. Every nonzero relation uses at least $r' + 1 - 2s_0$ blocks outside their union. By Lemma 7.5, the product of weights on these outside blocks is at most $\sigma^{r'+1-2s_0}$. Unsupported tuples have weight zero, and all supported relations belong to V_0 . Cauchy–Schwarz therefore gives

$$\sum_{\substack{\gamma \in \mathcal{R} \\ \gamma \neq 0}} \prod_i \mu_i(\gamma_i) \leq \sigma^{r'+1-2s_0} \prod_{j=1}^2 \left(\sum_{\gamma \in V_0} \prod_{i \in I_j} \mu_i(\gamma_i)^2 \right)^{1/2}$$

$$\leq \sigma^{r'+1-2s_0} \prod_{j=1}^2 \prod_{i \in I_j} \left(\sum_{\gamma_i \in \mathbf{E}_i^*} \mu_i(\gamma_i)^2 \right)^{1/2} \leq 2^{s_0} \sigma^{r'+1-2s_0}.$$

The second inequality uses the injectivity of the two projections. When $s_0 = 0$, there are no supported nonzero relations and the conclusion is immediate. $\square$

This is the two-information-set argument used for leakage Fourier sums in [BDIR19, Lemma 4.20 and its proof, pp. 31–32]; see also [BDIR21]. We included the proof to make clear which dimension and support bounds are required here.

8. PROOF OF THE STRUCTURED ESTIMATE

We now prove Lemma 7.1, using the preparations in Section 7. All notation from that section remains in force.

The probability event must work for all leakage maps, including maps chosen after a is published. We therefore choose the significance threshold and all height cutoffs first. The certificate is then required simultaneously for all bounded arrays and all bounded deletions. Only after fixing a tuple in this event do we choose the maps and their representations.

(Step 1) *Choose the cutoffs and a uniform good event.* All cutoffs below are fixed before the evaluation places and leakage maps are chosen.

(a) *Prepare the two certificate applications.* We will use Lemma 6.1 at rate $b/2$, so put

$$q_{\text{cert}} = \left\lceil \frac{8D}{b} \right\rceil. \quad (62)$$

The argument will remove fewer than q_{cert} shares. Since $r \geq bm$ and q_{cert} is fixed, take m sufficiently large that, for every $0 \leq j < q_{\text{cert}}$,

$$r - j \geq \frac{b}{2}(m - j), \quad r - j > 2q_{\text{cert}},$$

and the shortened length $m - j$ is large enough for the certificate lemma.

(b) *Choose the significance threshold and character cutoffs.* Choose a fixed constant $\kappa_{\text{sig}} \geq \log(2C_0) + 2$ and set

$$\delta = \exp(-\kappa_{\text{sig}} m).$$

This ensures $(2C_0)^m \delta \leq \exp(-2m)$, which will control the combined contribution of insignificant products.

Let $B \geq 2$ be the ceiling of the character-complexity bound in Equation 27, enlarged to 2 if necessary. Thus B bounds every character supplied by the product lemma at threshold δ . The certificate for arrays of height B gives the bound

$$H_B = (2B)^{C_{b/2,D}}.$$

Set

$$H_{\text{loc}} = \max\{B, DH_B\}.$$

The local cutoff therefore includes both the local rows of the product characters and the local characters obtained from their certificates: the latter have maximum norm at most H_B , hence ℓ^1 norm at most DH_B .

(c) *Choose the torus Fourier cutoff.* By Lemma 7.3 and the fixed C^R bound, choose a deterministic bound

$$1 \leq Q_{\text{tor}} \leq H_{\text{loc}}^{C_{\mathcal{G}}}$$

for the Lipschitz norm of every torus readout that can arise below. Here the exponent absorbs the fixed class constants.

Choose the approximation accuracy and frequency cutoff by

$$\varepsilon_{\text{tor}} = \frac{\delta}{2m}, \quad T_{\text{tor}} = \max \left\{ B, 2, \left\lceil \left(\frac{2DQ_{\text{tor}}}{\varepsilon_{\text{tor}}} \right)^2 \right\rceil \right\}.$$

By [Lemma 7.4](#), this cutoff gives probability-valued Fourier approximations within ε_{tor} . The total perturbation of the comparison will then be at most $2m\varepsilon_{\text{tor}} = \delta$. For the second certificate application, put

$$H_T = (2T_{\text{tor}})^{C_{b/2,D}}.$$

The cutoff T_{tor} concerns the torus readouts here; T_{der} concerns the derivative readouts inside the product lemma.

(d) *Check the field-size requirements.* The choices above give

$$\log B, \quad \log H_{\text{loc}}, \quad \log T_{\text{tor}} = \mathcal{O}_{b,\mathcal{G}}(m \log(2m)). \quad (63)$$

Consequently, choosing the constant C in [Lemma 7.1](#) sufficiently large ensures [Equation 26](#) and

$$p > \max \{ H_T, (mD)!T_{\text{tor}}^{mD}, m(m+1), D!H_{\text{loc}}^D \}. \quad (64)$$

Indeed, the logarithm of each quantity on the right is $\mathcal{O}_{b,\mathcal{G}}(m^2 \log(2m))$. These inequalities verify the certificate hypotheses at height B on the full tuple and at height T_{tor} on every required shortening. They also verify the prime-size condition in [Lemma 7.3](#).

(e) *Define the good event.* Let **Good** be the event on which the certificate conclusion holds

- at height B on the full tuple a , with code dimension r ;
- at height T_{tor} on every subtuple a_{J^c} with $|J| < q_{\text{cert}}$, with code dimension $r - |J|$.

At each height, the conclusion is required simultaneously for every eligible coefficient array and every solution of its moment equations. The two heights serve different purposes: the first gives the bound H_B used to choose the local cutoff, while the second will control the Fourier relations after revealing J .

Every fixed subtuple is uniform over ordered distinct nonzero places, and there are at most $q_{\text{cert}} m^{q_{\text{cert}}}$ deletion sets. Thus [Equation 38](#) and a union bound give

$$\Pr_a[\text{Good}^c] \leq p^{-1} \exp(\mathcal{O}_{b,D}(m \log(2T_{\text{tor}}) + \log(2m))). \quad (65)$$

Fix $a \in \text{Good}$. Only now choose the leakage maps f_i and their representations. The event already covers every coefficient array and exceptional set that these choices may produce.

(Step 2) *Bound the number of global constraints remaining after the local ones.* Form W, U_i, E, V as in [Equation 50](#). We claim that

$$s = \dim E - \dim V < q_{\text{cert}}. \quad (66)$$

Suppose otherwise. From the characters generating W , choose $\xi_1, \dots, \xi_{q_{\text{cert}}}$ whose differentials restrict independently to E . At each share i , collect their integer coefficient rows into a matrix $C_i \in \mathbb{Z}^{q_{\text{cert}} \times D}$, padding with zeros as needed. Each selected character has complexity at most B , so this array has height at most B . Since the characters are constant on the code orbit, [Equation 49](#) gives

$$\sum_i a_i^h C_i \beta_i = 0 \quad (0 \leq h < r),$$

where the slope vectors β_i are also padded with zeros.

Because $a \in \text{Good}$, the certificate supplies a nonzero integer vector $u_C \in \mathbb{Z}^{q_{\text{cert}}}$ such that

$$u_C^T C_i \beta_i = 0 \quad (i \in [m]).$$

Thus, at each share i , the coefficient row $u_C^\top C_i$ defines a character η_i that is constant on the local orbit. Moreover, Equation 37 gives

$$\|u_C^\top C_i\|_\infty \leq H_B, \quad \|u_C^\top C_i\|_1 \leq DH_B \leq H_{\text{loc}}.$$

Hence η_i is among the local characters used to define U_i , and its differential vanishes on $\text{Lie}(H_i)$.

Now take the combination of the selected global differentials with coefficients u_C . Its contribution at share i is exactly $d\eta_i$, so it vanishes on $E = \bigoplus_i \text{Lie}(H_i)$:

$$\left(\sum_{t=1}^{q_{\text{cert}}} (u_C)_t d\xi_t \right) \Big|_E = 0.$$

Since $u_C \neq 0$, this contradicts the independence of the selected differentials on E and proves Equation 66.

(Step 3) *Find one bounded set of exceptional shares.* We now show that the commutator contributions of all significant products are confined to one common set of at most s shares. To apply the common-support argument from Section 7.2, we first prove that

$$\omega_\nu|_{V \times V} = 0$$

for every significant product. Recall that a product is significant if its normalized mean has magnitude at least δ in either experiment, so there are two cases.

Significance in the code experiment. Apply Lemma 5.1 to the significant product. The characters it supplies have complexity at most B and are constant on the code orbit. Their differentials therefore belong to W . Since $V \subseteq W^\perp$, every direction in V preserves these character equations. The isotropy conclusion Equation 28 then gives $\omega_\nu|_{V \times V} = 0$.

Significance in the independent-input experiment. Apply the same lemma with input space $\mathbb{F}_p^m$ and coordinate forms $z \mapsto z_i$, as permitted by its arbitrary-linear-forms statement. Each supplied character is constant on the independent-input orbit. Holding all inputs except z_i fixed shows that its i th local character is constant on the corresponding local orbit. This local character has complexity at most $B \leq H_{\text{loc}}$, so its differential belongs to U_i and vanishes on $\text{Lie}(H_i)$. Consequently, the common kernel of the supplied character differentials contains E . The product lemma therefore gives $\omega_\nu|_{E \times E} = 0$, and hence $\omega_\nu|_{V \times V} = 0$.

Thus all significant forms vanish on the same $V \times V$. Let J be the union of their nonzero block supports. The common-support argument in Section 7.2 gives

$$J \subseteq [m], \quad |J| \leq s < q_{\text{cert}}. \quad (67)$$

For $i \notin J$, every significant vertical frequency ν_i is therefore trivial on the commutator torus Z_i . This is what allows the next step to average over these tori while preserving every significant product.

The set J may depend on the chosen representations, but it works simultaneously for all leakage profiles and all their significant products. This is why Step 1 required the good event to cover every deletion of fewer than q_{cert} shares.

(Step 4) *Average over the commutator tori outside J .* We now make each readout outside J invariant under its local commutator torus, while controlling the resulting perturbation of both leakage laws. For $i \notin J$, set

$$\tilde{F}_i(x\Gamma^{(i)}) = \int_{Z_i} F_i(tx\Gamma^{(i)}) dt,$$

where the integral uses the normalized Haar measure. For $i \in J$, keep $\tilde{F}_i = F_i$. Use $\tilde{F}_i$ for output one and $1 - \tilde{F}_i$ for output zero. By Section 4.4, averaging preserves the range $[0, 1]$, commutes with complements, and does not increase the C^R norm.

To understand its effect on the leakage laws, consider the vertical expansion of each readout. Averaging over Z_i preserves a component whose frequency ν_i is trivial on Z_i and deletes it otherwise. Thus a vertical product survives unchanged precisely when ν_i is trivial on Z_i at every averaged share. Step 3 guarantees this for every significant product. Consequently, every deleted product has a normalized mean of magnitude less than δ in both experiments.

Fix a leakage profile y . By Equation 46, the total coefficient weight of its vertical expansion is at most C_0^m . Summing the contributions of the deleted products therefore gives

$$\left| P_f(y) - \tilde{P}_f(y) \right| \leq C_0^m \delta, \quad \left| Q_f(y) - \tilde{Q}_f(y) \right| \leq C_0^m \delta,$$

where $\tilde{P}_f, \tilde{Q}_f$ are the laws defined by the averaged readouts. Absolute convergence justifies these term-wise comparisons.

The triangle inequality now yields

$$\text{SD}(P_f, Q_f) \leq \text{SD}(\tilde{P}_f, \tilde{Q}_f) + \frac{1}{2} \sum_{y \in \{0,1\}^m} \left(\left| P_f(y) - \tilde{P}_f(y) \right| + \left| Q_f(y) - \tilde{Q}_f(y) \right| \right).$$

There are 2^m profiles, so

$$\text{SD}(P_f, Q_f) \leq \text{SD}(\tilde{P}_f, \tilde{Q}_f) + (2C_0)^m \delta. \quad (68)$$

By our choice of δ , this perturbation is at most $\exp(-2m)$. Outside J , the averaged readouts are now invariant under $[H_i, H_i]$, as required for the local torus reduction in the next step.

(Step 5) *Reveal J and pass to quadratic torus maps.* Put $I = J^c$ and $r' = r - |J|$. We first condition on the field inputs at the exceptional shares, then express the remaining leakage probabilities using ordinary tori.

Reveal the underlying field inputs in J in both averaged experiments. Since $|J| < r$, the Vandermonde property makes the revealed code inputs uniform on $\mathbb{F}_p^J$, exactly as in the independent-input experiment. Conditional on any revealed vector u , the output bits in J have the same law in both experiments and are independent of the remaining output bits. They therefore contribute nothing to the conditional statistical distance. Writing $\tilde{P}_{I|u}$ and $\tilde{Q}_{I|u}$ for the conditional leakage laws on I , we obtain

$$\text{SD}(\tilde{P}_f, \tilde{Q}_f) \leq \mathbb{E}_{u \in \mathbb{F}_p^J} \text{SD}(\tilde{P}_{I|u}, \tilde{Q}_{I|u}).$$

Here the inequality comes from revealing additional information.

Fix the revealed vector u . By Equation 53, the remaining code inputs are an affine coordinate-wise rescaling of a code of dimension $r' = r - |J|$ on a_I . The independent comparison inputs remain uniform.

For each $i \in I$, the readout $\tilde{F}_i$ is invariant under $[H_i, H_i]$. The characters defining H_i have maximum norm at most H_{loc} , and Equation 64 verifies the prime-size hypothesis of Lemma 7.3. That lemma gives a probability map A_i on a torus of dimension $d_i \leq D$, with Lipschitz norm at most Q_{tor} .

Absorb the affine rescaling using Equation 55. The conditional comparison is then exactly Equation 54, with code dimension r' and places a_I . All bounds are uniform in u .

(Step 6) *Keep a finite Fourier expansion while preserving probabilities.* Fix the revealed vector from the preceding step. We now approximate each torus readout by a trigonometric polynomial whose frequencies have maximum norm at most T_{tor} . The approximation must remain in $[0, 1]$, since its values will still serve as output probabilities.

Apply [Lemma 7.4](#) to each A_i , using its Lipschitz bound Q_{tor} and the choices from [Step 1](#). This gives maps $\tilde{A}_i: \mathbb{T}^{d_i} \rightarrow [0, 1]$ satisfying

$$\|A_i - \tilde{A}_i\|_\infty \leq \varepsilon_{\text{tor}} = \frac{\delta}{2m}, \quad \text{supp } \widehat{\tilde{A}_i} \subseteq [-T_{\text{tor}}, T_{\text{tor}}]^{d_i} \cap \mathbb{Z}^{d_i}.$$

Use $\tilde{A}_i$ for output one and its complement $1 - \tilde{A}_i$ for output zero.

To bound the change in either leakage law, condition on all underlying field inputs. At each share, generate the original and approximating bits using the same uniform random number in $[0, 1]$, with these auxiliary random numbers independent across shares. The two bits disagree with probability equal to the absolute difference between their output-one probabilities, hence at most ε_{tor} . A union bound shows that the entire leakage profiles disagree with probability at most $|I|\varepsilon_{\text{tor}}$. Thus each leakage law changes in statistical distance by at most this amount, and the comparison changes by at most

$$2|I|\varepsilon_{\text{tor}} \leq 2m\varepsilon_{\text{tor}} = \delta.$$

This bound is uniform in the revealed vector.

Continue writing A_i for the approximating readouts. Use the finite-space readouts A_i^* and combined Fourier weights μ_i from [Section 7.4](#). The passage to [Equation 56](#) is exact: [Equation 57](#) combines equal finite characters. Every supported character has an integer representative of maximum norm at most T_{tor} , and [Lemma 7.5](#) applies.

(*Step 7*) *Bound the dimension of the supported Fourier relations.* Recall from [Equation 60](#) that V_0 is the $\mathbb{F}_p$ -span of the relations supported by the local Fourier weights, and $s_0 = \dim_{\mathbb{F}_p} V_0$. These are the relations that can contribute to the Fourier sum. We claim that

$$s_0 < q_{\text{cert}}. \tag{69}$$

The full relation space $\mathcal{R}$ may be larger; only the dimension of this span needs to be bounded.

Suppose otherwise. Choose q_{cert} linearly independent tuples

$$\gamma^{(1)}, \dots, \gamma^{(q_{\text{cert}})} \in \mathcal{R} \cap \prod_{i \in I} S_i.$$

Because these tuples are supported by the Fourier weights, [Step 6](#) supplies short integer representatives for their local characters. Specifically, choose $k_{t,i} \in \mathbb{Z}^{d_i}$ such that

$$\gamma_i^{(t)}(v) = (k_{t,i} \bmod p) \cdot v \quad (v \in \mathbb{E}_i), \quad \|k_{t,i}\|_\infty \leq T_{\text{tor}}.$$

Use $k_{t,i} = 0$ when $\gamma_i^{(t)} = 0$. At share i , place these representatives into the rows of $C_i \in \mathbb{Z}^{q_{\text{cert}} \times D}$, padding with zeros. The resulting array C has height at most T_{tor} . Pad the slope vectors to length D as well.

Each selected tuple is a relation, so $Q_{\gamma^{(t)}} = 0$ identically. Comparing the coefficients of x_h in these identities gives the moment equations for the linear orbit coefficients $\beta_{i,1}$. For the quadratic coefficients $\beta_{i,2}$, compare x_0^2 and, for $1 \leq h < r'$, $x_0 x_h$. Their coefficients in $\tilde{L}_i(x)^2$ are respectively 1 and $2a_i^h$. Since p is odd, we may divide the latter equations by 2. Thus we obtain

$$\sum_{i \in I} a_i^h C_i \beta_{i,j} = 0 \quad (0 \leq h < r', \quad j \in \{1, 2\}).$$

These are two solutions of the moment system for the *same coefficient array* C . The event [Good](#) includes the certificate conclusion at height T_{tor} on the shortened tuple a_I . Moreover, the assigned certificate u_C depends only on C and works for every solution of its moment equations. The same nonzero integer vector u_C therefore gives

$$u_C^\top C_i \beta_{i,1} = u_C^\top C_i \beta_{i,2} = 0 \quad (i \in I).$$

Since $\beta_{i,1}$ and $\beta_{i,2}$ span $\mathbf{E}_i$, the row $u_C^\top C_i$ represents the zero functional on $\mathbf{E}_i$. In terms of the selected local characters, this says

$$\sum_{t=1}^{q_{\text{cert}}} ((u_C)_t \bmod p) \gamma_i^{(t)} = 0 \quad \text{in } \mathbf{E}_i^* \quad (i \in I).$$

Hence the selected tuples satisfy a linear dependence over $\mathbb{F}_p$ with coefficient vector $u_C \bmod p$. This vector is nonzero: by Equation 37 and $p > H_T$,

$$0 < \|u_C\|_\infty \leq \frac{H_T}{q_{\text{cert}} T_{\text{tor}}} < p.$$

This contradicts their independence and proves Equation 69, the dimension bound needed for Lemma 7.6.

(Step 8) *Sum the relations and bound the remaining quadratic phases.* Fix the revealed inputs, and write $P^{\text{tor}}, Q^{\text{tor}}$ for the two conditional leakage laws after the approximation in Step 6. We compare these laws one Fourier tuple at a time.

For a tuple $\gamma = (\gamma_i)_{i \in I}$, let

$$M_\gamma = \mathbb{E}_{x \in \mathbb{F}_p^{r'}} e_p(Q_\gamma(x)),$$

$$N_\gamma = \mathbb{E}_{z \in \mathbb{F}_p^I} e_p \left(\sum_{i \in I} (\gamma_i(\beta_{i,1})z_i + \gamma_i(\beta_{i,2})z_i^2) \right).$$

These are its phase averages in the code and independent-input experiments, respectively. The local Fourier coefficients are the same in both experiments; only these averages differ.

For $\gamma = 0$, both averages equal 1, so this term cancels. For a nonzero relation $\gamma \in \mathcal{R}$, the code phase vanishes identically and $M_\gamma = 1$. In contrast, the independent-input phase vanishes identically only when $\gamma = 0$: its coefficients force each γ_i to annihilate both generators $\beta_{i,1}, \beta_{i,2}$ of $\mathbf{E}_i$.

We use the following elementary Gauss-sum bound to control every non-constant phase. For $\alpha \neq 0$,

$$\left| \mathbb{E}_{t \in \mathbb{F}_p} e_p(\alpha t^2 + \beta t) \right|^2 = \mathbb{E}_{h, t \in \mathbb{F}_p} e_p(2\alpha h t + \alpha h^2 + \beta h) = p^{-1},$$

because only $h = 0$ survives the average over t . Diagonalizing a quadratic form over the odd-characteristic field $\mathbb{F}_p$ gives the same upper bound $p^{-1/2}$ for the magnitude of any phase average with a nonzero quadratic part. A nonzero linear phase has mean zero.

Both phase polynomials defining M_γ, N_γ have zero constant term. Consequently,

$$|M_\gamma - N_\gamma| \leq \begin{cases} 0, & \gamma = 0, \\ 1 + p^{-1/2}, & \gamma \in \mathcal{R} \setminus \{0\}, \\ 2p^{-1/2}, & \gamma \notin \mathcal{R}. \end{cases}$$

We now sum over leakage profiles. At share i , the sum of the absolute Fourier coefficients for the two possible output bits is exactly $\mu_i(\gamma_i)$. Thus the triangle inequality gives

$$\text{SD}(P^{\text{tor}}, Q^{\text{tor}}) \leq \frac{1}{2} \sum_{\gamma} \left(\prod_{i \in I} \mu_i(\gamma_i) \right) |M_\gamma - N_\gamma|.$$

Substituting the phase bounds and separating the nonzero relations yields

$$\text{SD}(P^{\text{tor}}, Q^{\text{tor}}) \leq \frac{1}{2} \sum_{\substack{\gamma \in \mathcal{R} \\ \gamma \neq 0}} \prod_{i \in I} \mu_i(\gamma_i) + p^{-1/2} \prod_{i \in I} \sum_{\gamma_i} \mu_i(\gamma_i). \quad (70)$$

The first term is controlled by the dimension bound from Step 7. Indeed, $s_0 < q_{\text{cert}}$ and $r' > 2q_{\text{cert}}$ verify the hypotheses of Lemma 7.6, so

$$\sum_{\substack{\gamma \in \mathcal{R} \\ \gamma \neq 0}} \prod_{i \in I} \mu_i(\gamma_i) \leq 2^{s_0} \sigma^{r'+1-2s_0} \ll_{b,D} \sigma^{r'}.$$

The bounded dimension s_0 therefore incurs only a constant factor in the exponential bound.

For the second term, Cauchy–Schwarz and Lemma 7.5 give, at each share,

$$\sum_{\gamma_i} \mu_i(\gamma_i) \leq \left(|\text{supp } \mu_i| \sum_{\gamma_i} \mu_i(\gamma_i)^2 \right)^{1/2} \leq \sqrt{2}(2T_{\text{tor}} + 1)^{D/2}.$$

Multiplying over the remaining shares and using $|I| \leq m$, we conclude that every conditional distance is at most

$$\mathcal{O}_{b,D}(\sigma^{r'}) + p^{-1/2} \left[\sqrt{2}(2T_{\text{tor}} + 1)^{D/2} \right]^m. \quad (71)$$

All bounds are uniform in the revealed inputs.

(Step 9) *Restore the perturbations, take the supremum, and average.* Fix $a \in \text{Good}$ and arbitrary leakage maps f_i . We first transfer the conditional bound from the preceding step back to their original leakage laws.

Since $r' = r - |J|$ and $|J| < q_{\text{cert}}$,

$$\sigma^{r'} = \sigma^r \sigma^{-|J|} \leq \sigma^{-q_{\text{cert}}} \sigma^r.$$

Thus shortening the code changes the main bound by only a constant depending on b, D . The estimate Equation 71 is uniform in the revealed inputs, so averaging over those inputs as in Step 5 introduces no further factor.

We now restore the perturbations from Steps 4 and 6. Central averaging costs at most $(2C_0)^m \delta$, and the torus approximation costs at most δ . Consequently,

$$\begin{aligned} \text{SD}(P_f, Q_f) &\leq A_0 \sigma^r + (2C_0)^m \delta + \delta \\ &\quad + p^{-1/2} \left[\sqrt{2}(2T_{\text{tor}} + 1)^{D/2} \right]^m, \end{aligned}$$

where A_0 depends only on b and the fixed class data.

Crucially, the right-hand side is independent of the leakage maps and their representations. The event **Good** was chosen to cover every coefficient array and exceptional set that those choices could produce. Hence the same bound holds simultaneously for all leakage maps at this fixed a , and we may replace the left-hand side by their supremum $\Delta_{\mathcal{G}}(a; r)$.

We are left to bound the three additional terms. The choice $\delta = \exp(-\kappa_{\text{sig}} m)$ with $\kappa_{\text{sig}} \geq \log(2C_0) + 2$ gives

$$(2C_0)^m \delta \leq \exp(-2m), \quad \delta \leq \exp(-2m).$$

By Equation 63, the logarithm of the remaining Fourier term is at most

$$-\frac{1}{2} \log p + \mathcal{O}_{b,\mathcal{G}}(m^2 \log(2m)).$$

Similarly, Equation 65 bounds the probability of **Good**^c by

$$\exp(-\log p + \mathcal{O}_{b,\mathcal{G}}(m^2 \log(2m))).$$

Therefore, increasing the constant C in $\log p \geq C m^2 \log(2m)$ makes both quantities at most $\exp(-2m)$, while preserving all earlier field-size requirements. We have proved

$$\Delta_{\mathcal{G}}(a; r) \leq A_0 \sigma^r + 3 \exp(-2m) \quad \text{for every } a \in \text{Good},$$

and $\Pr_a[\text{Good}^c] \leq \exp(-2m)$.

Finally, use the trivial bound $\Delta_{\mathcal{G}}(a; r) \leq 1$ on the bad event. Splitting the expectation according to **Good** gives

$$\begin{aligned}\mathbb{E}_a \Delta_{\mathcal{G}}(a; r) &\leq A_0 \sigma^r + 3 \exp(-2m) + \Pr_a[\text{Good}^c] \\ &\leq A_0 \sigma^r + 4 \exp(-2m) \\ &\leq (A_0 + 4) \sigma^r.\end{aligned}$$

The last inequality uses $r < m$ and $\exp(-2m) \leq 2^{-r/2} = \sigma^r$. Taking $A = A_0 + 4$ proves [Lemma 7.1](#).

The reduction in [Section 2.1](#), applied with $b = c/2$, now gives [Theorem 3](#). Taking expectations in the pointwise inequality $\text{Insec}(a; d) \leq 2\Delta(a; d)$, established in [Equation 2](#) of [Section 1.1](#), gives [Theorem 4](#).

9. THE SHARP EXPONENTIAL RATE

The preceding proof gives a sharper bound after two changes: use the exact extremal Fourier coefficient in [Section 7](#), and make the truncation proportion in [Section 2](#) smaller. The resulting exponential base is $2/\pi$, matching the LSB lower bound in [Appendix A](#); see also [[MNP⁺22](#), Theorem 6, p. 16:7]. The product lemma and the moment certificate require no strengthening. Put

$$\tau = \frac{2}{\pi}, \quad M_p = \frac{1}{p \sin(\pi/(2p))}, \quad \lambda = \log_2(\pi/2) > 0.$$

Here M_p is the same finite-field constant as in [Proposition A.1](#).

9.1. The exact Fourier bound.

Lemma 9.1 (Sharp nonzero coefficient). *Let $\mathbf{E}$ be a finite-dimensional vector space over an odd prime field $\mathbb{F}_p$, and let $A: \mathbf{E} \rightarrow [0, 1]$. For every nonzero $\gamma \in \mathbf{E}^* = \text{Hom}_{\mathbb{F}_p}(\mathbf{E}, \mathbb{F}_p)$,*

$$2 \left| \mathbb{E}_{v \in \mathbf{E}} A(v) e_p(-\gamma(v)) \right| \leq M_p. \quad (72)$$

Equality is attained by $A(v) = \mathbb{1}_{\gamma(v) \in I}$ when I is a cyclic interval of $(p-1)/2$ or $(p+1)/2$ residues.

Proof. Since γ is nonzero, it is surjective, and its fibres have equal size. Set

$$a(t) = \mathbb{E}[A(v) \mid \gamma(v) = t] \in [0, 1].$$

The Fourier coefficient in [Equation 72](#) is $p^{-1} \sum_t a(t) e_p(-t)$. Rotate this sum so that it is real and nonnegative. For the resulting direction, its real part is maximized by taking $a(t) = 1$ where the projected root of unity is positive, and $a(t) = 0$ where it is negative. The positive roots occupy a cyclic interval of length $k = (p-1)/2$ or $(p+1)/2$. A root with zero projection may be assigned either value without changing the real part. The corresponding geometric sum has magnitude

$$\left| \sum_{j=0}^{k-1} e_p(j) \right| = \frac{\sin(k\pi/p)}{\sin(\pi/p)} = \frac{1}{2 \sin(\pi/(2p))}.$$

This proves the bound, and the interval indicators attain it. $\square$

For $h: \mathbb{F}_p \rightarrow [-1, 1]$, the signed form needed for the LSB lower bound follows by applying this lemma to $A = (1+h)/2$; see [Appendix A](#). Klein–Komargodski [[KK23](#), Claim 4.5 and Appendix A, equations (34)–(37)] showed that, for $h: \mathbb{F}_p \rightarrow [-1, 1]$ with mean μ , every nonzero Fourier coefficient satisfies

$$\left| \widehat{h}(\xi) \right| \leq \frac{2}{\pi} \cos\left(\frac{\pi\mu}{2}\right) + \mathcal{O}(p^{-2}) \quad (\xi \neq 0).$$

In particular, the bias-independent bound is $2/\pi + \mathcal{O}(p^{-2})$. The calculation above gives the exact finite-field constant M_p . Only this one-coefficient extremum is needed here, rather than the full majorization theorem of Lev [Lev01, Theorem 1, p. 244].

9.2. The sharper structured estimate.

Proposition 9.2 (Structured leakage with base $2/\pi$). *Under the hypotheses of Lemma 7.1, after changing its constants A, C , one has*

$$\mathbb{E}_a \Delta_{\mathcal{G}}(a; r) \leq A \left(\frac{2}{\pi} \right)^r. \quad (73)$$

The constants depend only on b and the fixed class data, independently of m, p, r .

Proof. Apply Lemma 9.1 to the finite-space readout $A_i^*: \mathbf{E}_i \rightarrow [0, 1]$ of Equation 56. This is done after combining the integer frequencies as in Equation 57. For the weights in Equation 58, the new estimates are

$$\sum_{\gamma} \mu_i(\gamma)^2 \leq 2, \quad \mu_i(\gamma) \leq M_p \quad (\gamma \neq 0).$$

The first inequality is unchanged. Replace σ by M_p in the proof of Lemma 7.6; it now gives

$$\sum_{\substack{\gamma \in \mathcal{R} \\ \gamma \neq 0}} \prod_{i \in I} \mu_i(\gamma_i) \leq 2^{s_0} M_p^{r'+1-2s_0}.$$

The certificate still gives $s_0 < q_{\text{cert}}$ and $|J| < q_{\text{cert}}$, with $r' = r - |J|$. Since

$$\tau \leq M_p \leq \frac{2}{3} < 1,$$

all the bounded exponent losses are absorbed into a constant independent of p . Thus the relation contribution is $O_{b,D}(M_p^r)$. The good event, torus approximation, and Gauss-sum estimate are unchanged. Their total contribution is at most $4 \exp(-2m)$ after increasing C , exactly as in Step 9. Since $r < m$, this is absorbed into AM_p^r .

Finally,

$$\log \frac{M_p}{\tau} = \log \frac{\pi/(2p)}{\sin(\pi/(2p))} = \mathcal{O}(p^{-2}).$$

The field condition can be enlarged to ensure $p > m(m+1)$, as already required in Equation 64. Hence $M_p^r = \tau^r \exp(\mathcal{O}(r/p^2)) \leq \mathcal{O}(1)\tau^r$, uniformly for $r < m$. Absorbing this factor proves Equation 73. $\square$

9.3. Make the outer reduction lose an arbitrarily small exponent.

Proof of Theorem 1. (Step 1) Choose the accuracy after fixing the desired exponent loss. Write $h_2(x) = -x \log_2 x - (1-x) \log_2(1-x)$, as in Section 2. Choose a fixed $\alpha > 0$ such that

$$\alpha < \min \left\{ \frac{1}{4}, \frac{c}{4}, \frac{3c-1}{4} \right\}, \quad h_2(\alpha) + \alpha \log_2 \pi \leq \frac{\varepsilon c \lambda}{4}. \quad (74)$$

This is possible because $h_2(\alpha) + \alpha \log_2 \pi \rightarrow 0$ as $\alpha \rightarrow 0$. Put

$$t = \lfloor \alpha m \rfloor, \quad \eta = 2^{-3/\alpha}, \quad b = c/2.$$

Repeat the construction in the proof of Proposition 4.2 with this fixed accuracy η in Equation 22. That construction works for every fixed positive accuracy: periodic regularity followed by positive smoothing gives an affine-closed class approximating every one-bit leakage function within η in U^3 . The resulting data $\mathcal{N}, D, Q$ now depend on (c, ε) , but remain independent of m, p and the leakage maps. Fix the constants A, C of Proposition 9.2 for this class.

(Step 2) Use the same simultaneous Cauchy–Schwarz reduction. For sufficiently large m , $1 \leq t < d$ and

$$3d - 2t - m \geq (3c - 2\alpha - 1)m - 3 > 0.$$

Thus Lemma 2.1 applies. For every $j < t$, we also have $d - j \geq (c/2)(m - j)$ for large m , because $\alpha < c/4$. The marginal tuple after deleting any fixed j places has the required uniform distribution, and the field bound at length m implies the corresponding bound at length $m - j$. Apply Equation 73 to these shortenings and average the outer reduction:

$$\begin{aligned} \mathbb{E}_a \Delta(a; d) &\leq A\tau^d \sum_{j < t} \binom{m}{j} (2/\tau)^j + 4^m \eta^t \\ &\leq A\tau^d 2^{m(h_2(\alpha) + \alpha \log_2 \pi)} + \eta^{-1} 2^{-m}. \end{aligned} \quad (75)$$

Here $\tau = 2/\pi$, the binomial entropy bound is the one used in Section 2, and $t \geq \alpha m - 1$ gives the last perturbation estimate.

(Step 3) Absorb the entropy cost and the fixed constants. For large m , $d \geq cm/2$, so Equation 74 gives

$$m(h_2(\alpha) + \alpha \log_2 \pi) \leq \frac{\varepsilon \lambda d}{2}.$$

Since $\tau = 2^{-\lambda}$, Equation 75 implies

$$\mathbb{E}_a \Delta(a; d) \leq A\tau^{(1-\varepsilon/2)d} + \eta^{-1} 2^{-m}.$$

Divide by $\tau^{(1-\varepsilon)d}$. The first term becomes $A\tau^{\varepsilon d/2} \rightarrow 0$. The second tends to zero as well, because $d \leq cm$ and $1 - (1 - \varepsilon)c\lambda > 0$. Increasing $m_{c,\varepsilon}$ therefore proves the first inequality in Equation 3. The second inequality follows by taking expectations in the pointwise comparison $\text{Insec}(a; d) \leq 2\Delta(a; d)$, from Equation 2 in Section 1.1.

(Step 4) Pass from every fixed loss to $o(1)$. Suppose $m \rightarrow \infty$ and $\log p/(m^2 \log(2m)) \rightarrow \infty$. For each fixed $\varepsilon > 0$, the field and length conditions just proved eventually hold. Consequently, the exponent loss can tend to zero; no uniform bound on the growth of $C_{c,\varepsilon}$ as $\varepsilon \rightarrow 0$ is needed for this conclusion. The factor 2 in the Shamir bound contributes $\mathcal{O}(1/d)$ to the exponent and is absorbed into $o(1)$.

In the stated n -bit regime, uniformly over the permitted primes,

$$\frac{\log p}{m^2 \log(2m)} \gg_u \frac{n^{1-2u}}{\log n} \rightarrow \infty,$$

because $u < 1/2$. This proves Equation 4. □

Scope of the asymptotic statement. For a prescribed fixed loss ε , the theorem has the same field-size form as before, with constant $C_{c,\varepsilon}$. The $o(1)$ conclusion uses the additional growth of $\log p/(m^2 \log(2m))$. A single fixed constant in that field-size ratio does not justify sending the approximation accuracy to zero by this argument. In particular, the proof gives no explicit rate for the $o(1)$ term without quantitative control of the approximation-class complexity.

Finally, Proposition A.1 gives

$$\mathbb{E}_a \Delta(a; d) \geq \frac{1}{2} \left(\frac{2}{\pi} \right)^d, \quad \mathbb{E}_a \text{Insec}(a; d) \geq \left(\frac{2}{\pi} \right)^d.$$

Together with Equation 4, these show that the exponential rate is sharp: in either expectation, the negative logarithm divided by d tends to $\log(\pi/2)$.

10. THE HIGHER-ORDER EXTENSION

This section proves [Theorem 2](#). For every fixed $\ell \geq 4$, we extend the relative-degree range from $c > 1/3$ to $c > 1/\ell$, while retaining the exponential base $2/\pi$ in the leakage bound. Higher-degree structure lets us reach smaller values of c , but requires a larger field. The condition $\log p \geq Cm^2 \log(2m)$ is replaced by $\log p \geq Cm^{\Theta_\ell} \log(2m)$, with a larger exponent Θ_ℓ . The threshold $c > 1/\ell$ enters through the outer Cauchy–Schwarz reduction; the structured estimate works at every fixed positive rate. For $c > 1/3$, [Theorem 1](#) retains the stronger field-size exponent 2.

Proof roadmap. We first record the higher-order outer reduction and smooth approximation. The main technical result is the structured estimate [Proposition 10.3](#). Its proof uses the moment certificate [Lemma 6.1](#) twice. First, the certificate bounds the number of global character equations that remain after imposing all short local equations. The common exception set [Lemma 10.4](#) then confines all significant top commutators to one bounded set of shares. Revealing this set of shares and taking exact local restrictions and central quotients lowers the nilpotency step. After at most $k - 1$ rounds, the remaining groups are tori. The second certificate application bounds the dimension of the supported Fourier relations, and the sharp coefficient bound supplies the base $2/\pi$. We finish by repeating the entropy calculation from [Section 9.3](#). Throughout, $k = \ell - 1$, and $\tau = 2/\pi$, $M_p = (p \sin(\pi/(2p)))^{-1}$, and $\lambda = \log_2(\pi/2)$ have the meanings fixed in [Section 9](#).

10.1. The outer reduction and the approximation class. For a function f on a finite abelian group A , set

$$\|f\|_{U^\ell(A)}^{2^\ell} = \mathbb{E}_{x, h_1, \dots, h_\ell \in A} \prod_{\omega \in \{0,1\}^\ell} \mathcal{C}^{|\omega|} f(x + \omega \cdot h).$$

The translation invariance and tensorization used in [Section 2](#) hold for every ℓ . Successive Cauchy–Schwarz in ℓ variables gives

$$\left| \mathbb{E}_{z_1, \dots, z_\ell \in A} H(z_1 + \dots + z_\ell) \prod_{j=1}^{\ell} B_j((z_i)_{i \neq j}) \right| \leq \|H\|_{U^\ell(A)}$$

whenever every B_j is one-bounded and independent of z_j . This is the higher-order form of the box inequality used in [Imported Lemma 2.2](#).

Lemma 10.1 (Higher-order simultaneous Cauchy–Schwarz). *Let $\mathcal{H}$ be an affine-closed class of probability maps such that every one-bit leakage function has a U^ℓ approximation in $\mathcal{H}$ within η . For every fixed tuple of distinct nonzero places, if*

$$1 \leq t < r, \quad m - t \leq \ell(r - t),$$

then

$$\Delta(a; r) \leq \sum_{|J| < t} 2^{|J|} \Delta_{\mathcal{H}}(a_{J^c}; r - |J|) + 4^m \eta^t. \quad (76)$$

Proof. Use the expansion in [Lemma 2.1](#); its treatment of $|J| < t$ is unchanged. For $|J| \geq t$, select $T \subseteq J$ of size t and partition $T^c = I_1 \sqcup \dots \sqcup I_\ell$ with $|I_j| \leq r - t$. Surjectivity of $L_{T \cup I_j}$ supplies maps W_j with $L_T W_j = \text{id}$ and $L_{I_j} W_j = 0$. Substitute

$$x = x^{(0)} + \sum_{j=1}^{\ell} W_j z_j, \quad z_j \in \mathbb{F}_p^T.$$

Average $x^{(0)}$ uniformly over $\mathbb{F}_p^r$, independently of all z_j , so that x is uniform on $\mathbb{F}_p^r$. For each fixed $x^{(0)}$, proceed as follows. The selected perturbations form $H(z_1 + \dots + z_\ell)$, where $H(w) = \prod_{i \in T} v_i(L_i(x^{(0)}) + w_i)$. The other factors indexed by I_j are independent of z_j . The box inequality bounds this average by $\|H\|_{U^\ell} = \prod_{i \in T} \|v_i\|_{U^\ell} \leq \eta^t$. The independent-input average has the same

bound. This bound is uniform in $x^{(0)}$, by translation invariance of the Gowers norm. Averaging over $x^{(0)}$ therefore gives the same bound for the codeword experiment. Summing over subsets and output profiles gives $4^m \eta^t$, as in the proof of [Lemma 2.1](#). $\square$

The fixed smooth class. Use the notation and smooth norms of [Section 4](#). Fix a finite collection $\mathcal{N}_\ell$ of filtered nilmanifolds of degree at most k , a dimension bound D , and a bound $Q \geq 2$. Put

$$R_0 = (k+1)(D+3).$$

The class $\mathcal{G}_{p,\ell}$ consists of the probability maps

$$\bar{z} \mapsto F(g(z)\Gamma), \quad g \in \text{poly}(\mathbb{Z}, G_\bullet), \quad g(z+p)\Gamma = g(z)\Gamma,$$

where $(G/\Gamma, G_\bullet, \mathcal{X}) \in \mathcal{N}_\ell$, $F: G/\Gamma \rightarrow [0, 1]$, and $\|F\|_{C^{R_0}(\mathcal{X})} \leq Q$. We require the orbit to return to the same point of the nilmanifold after p steps:

$$g(z+p)\Gamma = g(z)\Gamma \quad \text{for every } z \in \mathbb{Z}.$$

Thus $g(z)\Gamma$ depends only on $z \bmod p$, so every readout $F(g(z)\Gamma)$ defines a function on $\mathbb{F}_p$. Complements and bijective affine precompositions preserve this class, just as in [Proposition 4.2](#).

For every fixed accuracy $\eta > 0$, the construction in [Section 4.2](#) gives data of this kind approximating every one-bit leakage function within η in U^ℓ . Use Candela–Sisask’s periodic regularity theorem at degree k [[CS14](#), Theorem 3.4, p. 10]. There is one change in the norm estimate: U^ℓ need not be bounded by L^2 when $\ell \geq 4$. The regularity theorem’s L^2 -small error term f_{smi} also satisfies $|f_{\text{smi}}(x)| \leq 1$ at every input. For any function h with $|h(x)| \leq 1$, take absolute values in the cube formula defining U^ℓ and bound all but one of its 2^ℓ factors by 1. This gives

$$\|h\|_{U^\ell}^{2^\ell} \leq \mathbb{E}_x |h(x)| \leq \|h\|_2.$$

Thus a sufficiently small L^2 error is also small in U^ℓ , although we lose the exponent $1/2^\ell$. Choose the regularity L^2 tolerance $(\eta/4)^{2^\ell}$ and the uniformity tolerance $\eta/4$. These parts together have U^ℓ norm at most $\eta/2$. Finally, replace the structured readout F_0 by a weighted local average F , using nonnegative smooth weights of total mass one. This preserves the range $[0, 1]$, gives a controlled C^{R_0} norm, and can be done with $|F - F_0|_\infty \leq \eta/2$. Consequently,

$$\|F(g(\cdot)\Gamma) - F_0(g(\cdot)\Gamma)\|_{U^\ell} \leq |F - F_0|_\infty \leq \eta/2.$$

Adding this to the preceding approximation error gives a smooth approximation within η in U^ℓ . All class data are fixed before m, p , and the leakage functions; for [Theorem 2](#) they depend only on c, ℓ, ε .

10.2. The structured statement and its quantitative input. We now turn to leakage from the fixed class $\mathcal{G}_{p,\ell}$. As in the quadratic argument, our goal is to remove the relevant non-commuting directions and reach a torus model where Fourier analysis applies. In higher degree, this may require several rounds of reduction. Each round lowers the nilpotency step of the underlying group, while the polynomial degree of the orbit may remain as large as $k = \ell - 1$. We therefore distinguish these two parameters: write $G_{(s)}$ for the s th term of the lower central series, and retain G_j for the degree filtration.

Each nilmanifold in $\mathcal{N}_\ell$ comes with a degree filtration: a descending chain of closed connected subgroups

$$G = G_0 = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_k \supseteq G_{k+1} = \{\text{id}_G\}, \quad [G_i, G_j] \subseteq G_{i+j},$$

where $G_j = \{\text{id}_G\}$ for $j > k$. This filtration controls the polynomial degree of the orbit.

The lower central series is determined by the group itself:

$$G_{(1)} = G, \quad G_{(s+1)} = [G, G_{(s)}] \quad (s \geq 1).$$

Here $[A, B]$ is the subgroup generated by the commutators $aba^{-1}b^{-1}$ with $a \in A$ and $b \in B$. The group has nilpotency step s when $G_{(s)} \neq \{\text{id}_G\}$ but $G_{(s+1)} = \{\text{id}_G\}$.

The filtration condition implies $G_{(s)} \subseteq G_s$, but these subgroups need not coincide. For example, an abelian group has $G_{(2)} = \{\text{id}_G\}$, while its degree filtration may have $G_2 \neq \{\text{id}_G\}$, allowing quadratic and higher-degree polynomial orbits.

Here a *horizontal character* is a continuous homomorphism $G \rightarrow \mathbb{R}$ that is integer-valued on Γ and annihilates $[G, G]$. It need not annihilate G_2 , so its composition with a degree- k orbit may have degree k . This is the change from the first-level characters used in [Section 7](#).

Imported Theorem 10.2 (Periodic reduction in nilpotency step [\[Len25\]](#)). *Fix k . There is a finite exponent $\rho_k \geq 1$ such that the following holds. Let G/Γ have degree at most k , nilpotency step $s \geq 2$, dimension h , and complexity at most M . Let $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ and let $g(z)\Gamma$ be p -periodic, with p prime. Suppose that F is a $G_{(s)}$ -vertical function of frequency ν , meaning that*

$$F(ux\Gamma) = e^{2\pi i\nu(u)}F(x\Gamma) \quad (u \in G_{(s)}, x \in G).$$

Assume its Lipschitz norm is at most M , its frequency ν has size at most M/δ , and

$$\left| \mathbb{E}_{z \in \mathbb{F}_p} F(g(z)\Gamma) \right| \geq \delta, \quad 0 < \delta < 1/10.$$

If $p > (M/\delta)^{C_k h^{\rho_k}}$, there are horizontal characters $\xi_1, \dots, \xi_t$ of height at most $(M/\delta)^{C_k h^{\rho_k}}$, constant modulo 1 on the orbit, such that

$$d\nu([\dots[v_1, v_2], v_3], \dots, v_s) = 0 \quad \text{for } v_1, \dots, v_s \in \bigcap_j \ker d\xi_j.$$

For $\nu = 0$ the conclusion holds without imposing any characters.

Source. Leng [\[Len25\]](#), Theorem 3.1.3, printed pp. 41–42; Section 3.5, pp. 62–63], with the Lipschitz hypothesis of its underlying Theorem 1.5.1, p. 13. Horizontal height is the integer coefficient norm of Definition 2.3.12, p. 34; vertical frequency size is the metric norm of Definition 2.3.14, p. 35. The source gives height $(M/\delta)^{O_k(h^{\rho_k(1)})}$; ρ_k denotes a permissible exponent. Its vanishing C^∞ norm implies constancy modulo 1. Differentiating its group commutator identity gives the displayed Lie algebra identity.

Fix an exponent $\rho_k \geq 1$ for which [Imported Theorem 10.2](#) holds, and set

$$\Theta_{k+1} = 2 + (k-1)(\rho_k + 2) \quad (k \geq 3), \quad \Theta_3 = 2. \quad (77)$$

The value $\Theta_3 = 2$ comes from the separate quadratic proof. For $k \geq 3$ this is a generous allowance for the successive restrictions and quotients below. The source does not specify a numerical ρ_3 , so this formula does not give a numerical value for Θ_4 .

Proposition 10.3 (Higher-degree structured leakage). *Fix $k \geq 3$, $\ell = k+1$ and the data $\mathcal{N}_\ell, D, Q$ above. Let $\mathcal{G}_{p,\ell}$ be the class define in [Section 10.1](#). Fix $0 < b < 1$. There are constants $A, C > 0$, depending only on these data and b , such that, for all sufficiently large m , every integer $bm \leq r < m$, and every odd prime satisfying*

$$\log p \geq Cm^{\Theta_{k+1}} \log(2m),$$

we have

$$\mathbb{E}_{a \in \mathcal{A}_{p,m}} \Delta_{\mathcal{G}_{p,\ell}}(a; r) \leq A \left(\frac{2}{\pi} \right)^r. \quad (78)$$

Here $\mathcal{A}_{p,m}$ is the set of ordered distinct nonzero evaluation tuples from [Section 6.1](#). The constants depend only on the fixed class, k , and b . The good event used in the proof is chosen before the leakage maps.

10.3. Multilinear localization and exact local restriction.

Lemma 10.4 (A common exceptional set for block-diagonal forms). *Fix integer $m \geq 1, s \geq 2$. Let $E = \bigoplus_{i=1}^m E_i$ be a finite-dimensional real vector space and let $V \leq E$ have codimension*

$$q = \dim E - \dim V$$

. For each $i \in [m]$, let $T_i: E_i^s \rightarrow \mathbb{R}$ be an s -linear form. Suppose an s -linear form

$$T(v_1, \dots, v_s) = \sum_{i=1}^m T_i(v_{i,1}, \dots, v_{i,s}), \quad v_j = (v_{1,j}, \dots, v_{m,j}) \in E$$

is block diagonal and vanishes on V^s . Then at most sq of the forms T_i are nonzero:

$$|\{i : T_i \neq 0\}| \leq sq.$$

Moreover, let $\mathcal{T}$ be a real linear space of block-diagonal s -linear forms on the same E , all vanishing on V^s . There exists one exceptional set $J \subseteq [m]$ with $|J| \leq sq$ containing the block support of every member of $\mathcal{T}$, equivalently,

$$T_i \equiv 0 \quad \text{for every } T \in \mathcal{T} \text{ and every } i \notin J.$$

Here, vanishing on V^s requires independently chosen arguments $v_1, \dots, v_s \in V$.

Each argument is restricted to the same subspace V while is defined by q independent linear constraints in E .

Proof. Let N be the number of nonzero blocks T_i . Fix one s -linear form T and relabel its N nonzero blocks by $1, \dots, N$. We will replace each nonzero local form T_i by a scalar expression whose nonzero coefficient records that the original block was nonzero. For each such block, choose nonzero vectors $v_{i,1}, \dots, v_{i,s} \in E_i$ such that

$$c_i := T_i(v_{i,1}, \dots, v_{i,s}) \neq 0.$$

(Step 1) *Scalarize the nonzero blocks.* For each argument position $j \in \{1, \dots, s\}$, define the linear map

$$\iota_j: \mathbb{R}^N \rightarrow E, \quad \iota_j(x_1, \dots, x_N) = (x_1 v_{1,j}, \dots, x_N v_{N,j}, 0, \dots, 0).$$

The map is injective; if the output equals zero, then $x_i v_{i,j} = 0$ and each $x_i = 0$. Define

$$W_j := \{x \in \mathbb{R}^N : \iota_j(x) \in V\}.$$

Let $\pi: E \rightarrow E/V$ be the quotient projection such that $\pi(u) = u + V$. Then, $W_j = \ker(\pi \circ \iota_j)$ and rank-nullity gives

$$\text{codim}_{\mathbb{R}^N} W_j = \text{rank}(\pi \circ \iota_j) \leq \dim E/V = q.$$

For $x^{(j)} = (x_i^{(j)})_{i=1}^N \in \mathbb{R}^N$, define the scalar expression $\tilde{T}$,

$$\tilde{T}(x^{(1)}, \dots, x^{(s)}) := T(\iota_1(x^{(1)}), \dots, \iota_s(x^{(s)})) = \sum_{i=1}^N c_i x_i^{(1)} \dots x_i^{(s)}.$$

If $x^{(j)} \in W_j$ for every j , all arguments of T belong to V and by our assumption that T vanishes on V^s , we have

$$\tilde{T} = 0 \quad \text{on } W_1 \times \dots \times W_s, \quad \text{codim } W_j \leq q.$$

(Step 2) *Contract to a bilinear form.* When $s = 2$, $B = \tilde{T}$. For $s \geq 3$, we will fix the first $s - 2$ arguments to obtain a bilinear form, while keeping enough nonzero coefficients to count the original blocks. For every $j \leq s - 2$, let F_j records the coordinates that stays zero throughout W_j and $\mathcal{C}_j$ is a coordinate subspace,

$$F_j = \{i \in \{1, \dots, N\} : x_i = 0, \forall x \in W_j\}, \quad \mathcal{C}_j := \{x \in \mathbb{R}^N : x_i = 0, \forall i \in F_j\}.$$

Since $W_j \subseteq \mathcal{C}_j$ and $\dim \mathcal{C}_j = N - |F_j|$, we have

$$|F_j| \leq N - \dim W_j = \text{codim } W_j \leq q.$$

Thus, at most q coordinates vanish identically on W_j . We next choose a vector $w^{(j)} \in W_j$ with no additional zero coordinates using a successive-perturbation construction analogous to the construction from [Section 7.2](#). Take $w = 0 \in W_j$ and address the coordinates outside F_j one at a time. Whenever a coordinate that can be nonzero is still zero, add a small multiple of a vector in W_j that is nonzero. For the next coordinate i , choose $u \in W_j$ with $u_i \neq 0$, which exists by the definition of F_j and replace $w \mapsto w + tu$. The new coordinate and each previously made nonzero coordinate exclude at most one real value of t . Repeating this activates every coordinate that can be nonzero, leaving only the coordinates that are forced to be zero. The resulting vector belongs to W_j and has exactly the forced zero coordinates, hence at most q zeroes.

Fix $w^{(1)}, \dots, w^{(s-2)}$ as above, substitute vectors and count the surviving coefficients. Define a bilinear form $B : \mathbb{R}^N \times \mathbb{R}^N \rightarrow \mathbb{R}$,

$$B(x, y) := \tilde{T}(w^{(1)}, \dots, w^{(s-2)}, x, y) = \sum_{i=1}^N \lambda_i x_i y_i, \quad \lambda_i = c_i \prod_{j=1}^{s-2} w_i^{(j)}.$$

Given $c_i \neq 0$ and our construction above, a coefficient λ_i vanishes precisely when

$$i \in F := \bigcup_{j=1}^{s-2} F_j.$$

Consequently,

$$|F| \leq (s-2)q, \quad \text{rank } B = N - |F| \geq N - (s-2)q$$

where rank is the rank of the diagonal matrix $\text{diag}(\lambda_1, \dots, \lambda_N)$. Moreover, the fixed vector belong to their respective W_j , so the original vanishing assumption gives

$$B(x, y) = 0, \quad x \in W_{s-1}, \quad y \in W_s.$$

Set $a := N - \dim W_{s-1} \leq q$, and $b := N - \dim W_s \leq q$ so $a, b \leq q$. Extend the bases of W_{s-1} and W_s to bases $e_1, \dots, e_N$ and $f_1, \dots, f_N$ of $\mathbb{R}^N$ respectively. We will use the basis e_i as the first argument for the bilinear form B and f_j as the second argument. Define the matrix M as $M_{ij} = B(e_i, f_j)$. Whenever $i \leq N - a$ and $j \leq N - b$, we have $M_{ij} = 0$. Then,

$$M = \begin{pmatrix} 0 & A \\ C & D \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & A \\ 0 & D \end{pmatrix}}_{\text{supported on last } b \text{ columns}} + \underbrace{\begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}}_{\text{supported on last } a \text{ rows}}.$$

where the bottom blocks have a rows and the right blocks have b columns. The block decomposition gives

$$\text{rank } B = \text{rank } M \leq b + a \leq 2q.$$

A bilinear form vanishing on subspaces of codimensions at most q in its two slots has rank at most $2q$. Combining the two rank condition, we get

$$N - (s-2)q \leq \text{rank } B \leq 2q \implies N \leq sq.$$

Fixing the first arguments can eliminate at most $(s-2)q$ blocks and at most $2q$ blocks can survive, so there were at most sq blocks originally. Thus $N \leq sq$.

(Step 3) *Obtain one exceptional set.* Finally, we define the exceptional set for the s -form T . For every $T \in \mathcal{T}$, put

$$J_T = \{i : T_i \neq 0\}.$$

The single-form case gives $|J_T| \leq sq$. If $\mathcal{T} = \{0\}$, take $J = \emptyset$. Otherwise, choose a basis $T^{(1)}, \dots, T^{(d)}$ with $d \geq 1$ for the given space of forms $\mathcal{T}$. Set

$$J = \bigcup_{j=1}^d J_{T^{(j)}}.$$

Every form in $\mathcal{T}$ is a linear combination of the basis form and every block outside J vanishes in every form in $\mathcal{T}$.

Recall our reasoning from [Section 7.2](#) that individual size bounds alone do not bound their union. We will construct one s -linear form whose block support is exactly J following the same argument.

Start with $\Theta_1 = T^{(1)}$. For $j = 2, \dots, d$, choose a real number t and set $\Theta_j = \Theta_{j-1} + tT^{(j)}$. For each share i in the union of the supports of Θ_{j-1} and $T^{(j)}$, the new block is $A_i + tB_i$, where $A_i = (\Theta_{j-1})_i$ and $B_i = T_i^{(j)}$. These two blocks are not both zero.

- If $B_i = 0$, then $A_i \neq 0$ and $A_i + tB_i = A_i \neq 0$ for every real number t .
- If $B_i \neq 0$, at most one value of t can make the block zero; otherwise, if $A_i + tB_i = 0$ and $A_i + t'B_i = 0$, subtraction gives $(t - t')B_i = 0$ and therefore $t = t'$.

There are finitely many blocks and each excludes at most one $t \in \mathbb{R}$. Choose t avoiding all these values. Every block occurring in either form Θ_{j-1} or $T^{(j)}$ remains nonzero while blocks absent from both remain zero. Then

$$J_{\Theta_j} = J_{\Theta_{j-1}} \cup J_{T^{(j)}}.$$

Repeating the steps gives a linear combination Θ_* supported on J satisfies

$$J_{\Theta_*} = J, \quad |J| = |J_{\Theta_*}| \leq sq$$

because Θ_* still vanishes on V^s , previous argument applies. This is the desired common exceptional set. □

The following lemma impose the local character equations exactly, then remove the last commutator subgroup while preserving the relevant readout along the orbit. It extends the local reduction to quadratic torus argument of [Lemma 7.3](#).

Lemma 10.5 (Exact local restriction and central quotient). *Fix integers $k \geq 3$ and $D \geq 1$. There exists a constant $C_{k,D} \geq 1$, depending only on k, D , with the following property.*

Let $Q, H \geq 2$ and let $(G/\Gamma, G_\bullet, \mathcal{X})$ be a filtered nilmanifold of degree at most k , dimension at most D , and complexity at most Q . Let p be an odd prime satisfying

$$p > (QH)^{C_{k,D}}.$$

Suppose that $g \in \text{poly}(\mathbb{Z}, G_\bullet)$ such that $g(z)\Gamma$ is p -periodic and that a finite family of horizontal characters $\xi_1, \dots, \xi_t$ of height at most H are constant modulo 1 on this orbit. Put

$$H_* = \exp\left(\bigcap_j \ker d\xi_j\right), \quad \Gamma_* = H_* \cap \Gamma, \quad (H_*)_j = H_* \cap G_j.$$

For an empty character family, take the intersection to be $\text{Lie}(G)$. The subgroup H_ is connected and rational, and Γ_* is a cocompact lattice in H_* . There exists c in a fixed Mal'cev fundamental box and $h \in \text{poly}(\mathbb{Z}, (H_*)_\bullet)$ such that for every $z \in \mathbb{Z}$,*

$$g(z)\Gamma = ch(z)\Gamma, \quad h(z+p)\Gamma_* = h(z)\Gamma_*.$$

The restricted filtered nilmanifold $(H_*/\Gamma_*, (H_*)_\bullet)$ admits an adapted Mal'cev basis giving complexity at most $(QH)^{C_{k,D}}$. Suppose additionally that $2 \leq s \leq k$ is an integer and the group G has nilpotency step at most s . Set

$$Z = (H_*)_{(s)}.$$

Then Z is connected, rational, and central in G . The quotient nilmanifold $H_*/(Z\Gamma_*)$ with acting group H_*/Z and the filtration induced from $(H_*)_\bullet$ has nilpotency step at most $s-1$, while its filtration degree remains at most k and dimension at most D . It admits an adapted Mal'cev basis $\tilde{\mathcal{X}}$ giving complexity at most $(QH)^{C_{k,D}}$. The orbit $z \mapsto h(z)(Z\Gamma_*)$ is polynomial for the induced filtration and is p -periodic.

For every integer $R \geq 0$, there exists a constant $C_{k,D,R} \geq 1$, depending only on k, D, R such that the following holds. If $F \in C^R(G/\Gamma; \mathbb{C})$ is Z -invariant, i.e. $F(w\Gamma) = F(\Gamma)$ where $w \in Z$ and $x \in G$, then

$$\tilde{F}(u(Z\Gamma_*)) := F(cu\Gamma), \quad u \in H_*,$$

define a C^R function on $H_*/(Z\Gamma_*)$ satisfying for every $z \in \mathbb{Z}$,

$$F(g(z)\Gamma) = \tilde{F}(h(z)(Z\Gamma_*)), \quad \left\| \tilde{F} \right\|_{C^R(\tilde{\mathcal{X}})} \leq (QH)^{C_{k,D,R}} \|F\|_{C^R(\mathcal{X})}$$

where $\tilde{\mathcal{X}}$ is a controlled adapted basis supplied by the quotient conclusion. If F takes value in $[0, 1]$, then so does $\tilde{F}$.

Note that $(H_*)_j = H_* \cap G_j$ denotes the inherited degree filtration, whereas $(H_*)_{(s)}$ denotes the lower central series.

Proof. We follow the rational correction strategy of Leng [Len25, Lemmas 3.7.1 and 3.7.3, printed pp. 94–96]. After normalizing the sequence, we multiply it by the inverse of a rational polynomial sequence so that all the character equations vanish exactly. We then clear the denominators and use p -periodicity to recover the original orbit modulo Γ . The construction below verifies that these operations preserve the degree filtration.

Write $g(0) = c\gamma_0$, with c in the fundamental box and $\gamma_0 \in \Gamma$, and put

$$v(z) = c^{-1}g(z)\gamma_0^{-1}.$$

Then $v(0) = \text{id}_G$, $v(z)\Gamma$ is p -periodic, and $\xi_i(v(z)) \in \mathbb{Z}$ for every i, z . Conjugation preserves each G_j , so v is polynomial for the original filtration. Write its filtered Taylor expansion as

$$v(z) = \prod_{j=1}^k v_j^{(z)}, \quad v_j \in G_j.$$

Applying a character turns this product into a sum, and finite differences show that $\xi_i(v_j) \in \mathbb{Z}$. First choose at most D characters from the defining family whose differentials span the same real row space. It suffices to solve their equations: every remaining equation is a real linear combination, and the same combination holds for its prescribed right-hand side. Thus the rationality bounds are independent of the number of defining characters. For each j , use rational linear algebra to choose $b_j \in G_j$ with bounded denominator and

$$\xi_i(b_j) = \xi_i(v_j) \quad \text{for every } i, \quad b_j \in G_j.$$

In an adapted rational basis of $\text{Lie}(G_j)$ these are consistent linear equations with integer right-hand sides and bounded rational coefficients. Cramer's rule gives a Lie algebra solution B_j whose coordinate denominators are bounded by $(QH)^{C_{k,D}}$; its numerators need not be bounded. Take $b_j = \exp(B_j)$. Set

$$\gamma(z) = \prod_{j=1}^k b_j^{(z)}, \quad h_1(z) = v(z)\gamma(z)^{-1}.$$

Both sequences are polynomial for the original filtration, and $\xi_i(h_1(z)) = 0$ for every i . Thus h_1 is polynomial for the inherited filtration on H_* : its multiplicative derivatives belong both to H_* and to the corresponding G_j .

The rational correction γ makes $h_1 = v\gamma^{-1}$ satisfy the character equation exactly, but it need not preserve the orbit modulo Γ , since $\gamma(z)$ need not lie in Γ .

Its bounded coordinate denominators allow us to apply Leng [Len25, Lemmas 7.2.16–7.2.17, printed pp. 317–318]. There is a positive integer $L \leq (QH)^{C_{k,D}}$ with $\gamma(Lz) \in \Gamma$ for every z , regardless of the sizes of the numerators. Consequently, the correction disappears modulo the lattice on multiples of L ;

$$h_1(Lz)\Gamma = v(Lz)\gamma(Lz)^{-1}\Gamma = v(Lz)\Gamma.$$

Since $p > L$, Chinese remainder theorem can choose one dilation a_0 that both removes the rational correction and preserves the original orbit. a_0 is divisible by L and congruent to 1 modulo p . Then $h(z) = h_1(a_0z)$ remains polynomial for the inherited filtration, and

$$h(z)\Gamma = v(a_0z)\gamma(a_0z)^{-1}\Gamma = v(a_0z)\Gamma = v(z)\Gamma$$

where divisibility by L removes the rational correction and $a_0 \equiv 1 \pmod{p}$ preserves the original orbit by p -periodicity. Moreover, $h(z) \in H_*$ and $v(z)\Gamma$ is p -periodic, so

$$h(z)^{-1}h(z+p) \in H_* \cap \Gamma = \Gamma_*.$$

Thus, $h(z+p)\Gamma_* = h(z)\Gamma_*$. Neither the sizes of the Taylor coefficients nor the size of the dilation a_0 enters the complexity estimate. The geometry is determined by the bounded character equations, while the construction preserves the orbit exactly modulo the lattice.

Suppose G has nilpotency step at most $s \geq 2$. Let $Z = (H_*)_{(s)}$ be the subgroup generated by s -fold commutators in H_* . It is rational and central in G , and H_*/Z has step at most $s - 1$. For the rational subgroup basis and metric comparison, use Leng [Len25, Lemma 7.2.15, p. 317; Lemma 7.2.11, pp. 312–313]. For a Z -invariant readout F , the function $u \mapsto F(cu\Gamma)$ descends to the quotient; changing u by an element of Z does not change its value. The Lie algebra of Z is spanned by brackets of length s in a bounded rational basis of $\text{Lie}(H_*)$. Its rationality is therefore bounded by

$$(QH)^{C_{k,D}}.$$

It is central in G because G has step at most s . The quotient with its induced filtration is covered by [Len25, Lemma 2.3.11, pp. 33–34].

Finally, the chain rule in bounded-degree Mal'cev coordinates gives the stated finite smoothness bounds; bounded translation is also covered by [Len25, Lemma 7.2.6, pp. 308–309]. These are local estimates in dimension at most D . □

10.4. Successive central reductions.

Proof roadmap. The proof successively lowers the nilpotency step of the group, estimates the resulting torus experiment, and restores the accumulated errors. Repeat until the remaining groups are tori after at most $k - 1$ rounds. On these tori, positive Fourier approximation through Lemma 7.4 reduces the comparison to a finite sum, where the certificate and sharp local Fourier bounds control the exact relations. The nonconstant polynomial phases are upper bounded by a Weyl-type bound Equation 100.

The induction invariant and the one-round claim. Put $q_{\text{cert}} = \lceil 8D/b \rceil$ and fix $\kappa > 2$. After t rounds, let S_t be the revealed coordinates and put

$$m_t = m - |S_t|, \quad r_t = r - |S_t|, \quad R_t = R_0 - t(D + 2). \quad (79)$$

The remaining local groups have nilpotency step at most $k - t$ and dimension at most D ; their orbits are p -periodic and polynomial of filtration degree at most k . The readouts are probability

valued, and the geometry and C^{R_t} norms are bounded by a deterministic Q_t . Also, $|S_t| < k^2 q_{\text{cert}}$ and $r_t/m_t \geq b/2$. All estimates hold on one event chosen before the leakage maps. Write P_t, Q_t for the current output laws. One round, with current step $s \geq 2$, reveals fewer than $s q_{\text{cert}}$ inputs and produces models of step at most $s - 1$ satisfying

$$\text{SD}(P_t, Q_t) \leq e^{-\kappa m} + \mathbb{E}_u \text{SD}(P_{t+1,u}^{\text{new}}, Q_{t+1,u}^{\text{new}}). \quad (80)$$

Only central averaging contributes to this error; restriction and quotient preserve the conditional laws exactly. Steps 2–5 verify this claim. The later steps estimate the terminal tori and close the quantitative bounds.

Proof of Proposition 10.3. (Step 1) Prepare one uniform good event and quantitative allowances that supports every round. Put $q_{\text{cert}} = \lceil 8D/b \rceil$.

(a) *There will be at most $k - 1$ rounds of reduction.* The degree k -filtration ensures that every group has nilpotency step at most k ,

$$G_{(k+1)} \subseteq G_{k+1} = \{\text{id}_G\}.$$

Suppose the current maximum nilpotency step $s \geq 2$. At every remaining share, the round restricts to a subgroup H_i and quotients by the local central series $Z_i = (H_i)_{(s)}$. The resulting group has step at most $s - 1$ because $(H_i/Z_i)_{(s)} = \{\text{id}_G\}$. Each round decreases the maximum step by at least one. Starting with nilpotency at most k , we reach step one nilpotency, the abelian case, after at most $k - 1$ rounds. Each round reveals fewer than $k q_{\text{cert}}$ share values, hence fewer than $k^2 q_{\text{cert}}$ in total. The polynomial orbit's degree may remain k throughout the reduction. The nilpotency step s tells us what to remove and when to step while the filtration degree k controls the polynomial structure that survive each reduction.

(b) *In each round, the remaining share vectors still have sufficient rate to use the certificate lemma.* Each local readout is evaluated at a field input $z_i \in \mathbb{F}_p$, where

$$F_i(g^{(i)}(z_i)\Gamma^{(i)}).$$

Consider the two experiments from Section 7.1. In the code experiment, the inputs are the shares $z_i = L_i(x)$ and in the comparison experiment, they are independent uniform field elements. Initially, dimension of the space of share vectors is $r \geq bm$ and the number of shares is m . After revealing d inputs, the remaining number of shares is $m - d$ and dimension of the space of share vectors is $r - d$. We reveal at most a fixed number of $k^2 q_{\text{cert}}$ shares, independent of m . When $m \geq 2k^2 q_{\text{cert}}/b$,

$$r - d \geq bm - k^2 q_{\text{cert}} \geq \frac{b}{2}m \geq \frac{b}{2}(m - d).$$

For large m , every remaining share vectors has relative dimension/rate at least $b/2$. Therefore, we can apply Lemma 6.1 at rate $b/2$ for every deletion of fewer than $k^2 q_{\text{cert}}$ shares throughout the reduction. The required number of rows $q = q_{\text{cert}}$.

(c) *Choose deterministic complexity budget for every round.* We will specify in advance which coefficient heights the good event must cover. Step 2 – 4 later verify that the actual quantities fit these budgets. At the beginning of round t , let $Q_t \geq 2$ bound the local geometry and the required smoothness norms. Starting from fixed Q_0 , use

$$\log Q_t \leq C_t m^{v_t} \log(2m), \quad v_0 = 0, \quad v_{t+1} = v_t + \rho_k + 2 \quad (81)$$

where ρ_k follows from Imported Theorem 10.2. All constants depend only on the fixed data. The bounds determine the character cutoffs and the final Fourier cutoff.

(d) *Define the certificate event for one fixed set of revealed shares and fixed coefficient cutoff.* Fix a set S with $|S| < k^2 q_{\text{cert}}$, and a required height cutoff $B \geq 2$. Apply Lemma 6.1 to the

remaining evaluation tuple a_{S^c} with rate parameter $b/2$. For every integer array

$$C = (C_i)_{i \in S^c}, \quad C_i \in \mathbb{Z}^{q_{\text{cert}} \times D}, \quad \|C\|_\infty \leq B,$$

[Lemma 6.1](#) provides a certificate vector $u_C \in \mathbb{Z}^{q_{\text{cert}}}$, depending on C alone, with the height bound stated in the lemma. The certificate assignments are fixed before sampling the evaluation places. We say that the evaluation tuple $a = (a_1, \dots, a_m)$ belongs to $\text{Good}_{S,B}$ if for every bounded array C above and every collection of solution vectors $\beta_i \in \mathbb{F}_p^D$,

$$\left[\sum_{i \in S^c} a_i^h C_i \beta_i = 0 \quad (0 \leq h < r - |S|) \right] \implies u_C^\top C_i \beta_i = 0 \quad (i \in S^c). \quad (82)$$

The β_i are auxiliary coefficient vectors; later they will be the horizontal polynomial coefficients. On the $\text{Good}_{S,B}$ event, every permitted system of global moment equations has its assigned local certificate, simultaneously for all its solutions. For uniformly random distinct nonzero evaluation places, [Lemma 6.1](#) gives

$$\Pr[a \notin \text{Good}_{S,B}] \leq \frac{2Dm'(r' - 1)(2B + 1)^{q_{\text{cert}}m'D}}{p} \quad (83)$$

where $m' = m - |S|$ and $r' = r - |S|$. This applies because the retained tuple a_{S^c} is uniformly distributed among distinct nonzero evaluation tuples.

(e) *Define the good event.* Make the certificate assignments consistently across cutoffs. For each array C , choose its certificate using the intrinsic cutoff $B(C) = \max\{2, \|C\|_\infty\}$, so that

$$q_{\text{cert}} B(C) \|u_C\|_\infty \leq (2B(C))^{C_{b/2,D}}.$$

The maximal-minor event in the proof of [Lemma 6.1](#) enforces these same assigned certificates for every solution; its stated union bound is unchanged. This choice retains the smaller height bound whenever the actual array has height less than the master cutoff.

Choose a deterministic master cutoff $T_* \geq 2$ covering every character and Fourier cutoff used below. The recurrence [Equation 81](#) and the later estimates allow us to fix it with

$$\log T_* \leq Cm^{v_{k-1}} \log(2m).$$

Define

$$\text{Good} = \bigcap_{\substack{S \subseteq [m] \\ |S| < k^2 q_{\text{cert}}}} \text{Good}_{S,T_*}.$$

This event depends only on the evaluation places and the predetermined parameters. It covers every later bounded array and every solution, including choices made after earlier inputs have been revealed. There are at most $k^2 q_{\text{cert}} m^{k^2 q_{\text{cert}} - 1}$ deletion sets. Thus [Equation 83](#) gives

$$\Pr[\text{Good}^c] \leq p^{-1} \exp\{Cm \log(2T_*) + C \log(2m)\}. \quad (84)$$

The field-size assumption

$$\log p \geq Cm^{v_{k-1}+2} \log(2m) = Cm^{\Theta_{k+1}} \log(2m)$$

covers all certificate, restriction, and line-selection conditions checked below. Increasing C makes the failure probability and the second term of [Equation 103](#) at most e^{-2m} .

There is no additional union bound over revealed values; the universal event and the deterministic bounds apply in every conditional branch. The significance thresholds and the positive approximations can also be chosen so that their total error over all rounds is $O_k(e^{-2m})$.

(f) *Distinguishing advantage.* Since the distinguishing advantage is at most 1, we have

$$\mathbb{E}_a \Delta_{\mathcal{G}_{p,\ell}}(a; r) = \mathbb{E}_a [\Delta_{\mathcal{G}_{p,\ell}}(a; r) \mathbb{1}_{\text{Good}}(a)] + \mathbb{E}_a [\Delta_{\mathcal{G}_{p,\ell}}(a; r) \mathbb{1}_{\text{Good}^c}(a)]$$

$$\begin{aligned} &\leq \mathbb{E}_a[\Delta_{\mathcal{G}_{p,\ell}}(a; r) \mathbb{1}_{\text{Good}}(a)] + \Pr[\text{Good}^c] \\ &\leq \mathbb{E}_a[\Delta_{\mathcal{G}_{p,\ell}}(a; r) \mathbb{1}_{\text{Good}}(a)] + e^{-2m}. \end{aligned}$$

It remains to bound the average for good evaluation tuples, uniformly over all admissible leakage maps in $\mathcal{G}_{p,\ell}$.

(Step 2) *Decompose the readouts and isolate significant products.* Fix an evaluation tuple $a \in \text{Good}$ and the original leakage maps. At the beginning of round t , fix a revealed history and put

$$I = [m] \setminus S_t, \quad m' = |I|, \quad r' = r - |S_t|.$$

Let $s \geq 2$ be the current maximum nilpotency step. Absorb the affine changes from shortening into the local orbits, so that the current share inputs are

$$L_i(x) = \sum_{h=0}^{r'-1} a_i^h x_h, \quad x \in \mathbb{F}_p^{r'}, \quad i \in I.$$

For $i \in I$, write $G^{(i)}/\Gamma^{(i)}$, $g^{(i)}$, and F_i for the current local nilmanifold, polynomial orbit, and probability readout. The local dimensions are at most D , the polynomial degrees are at most k , and each orbit is p -periodic modulo its lattice. Set the derivative budget as Equation 79. Let $Q_t \geq 2$ bound the current local geometry and the C^{R_t} norms. All smooth norms below use the current Mal'cev bases and include the supremum norm. Put

$$F_{i,1} = F_i, \quad F_{i,0} = 1 - F_i, \quad \text{with} \quad \|F_{i,0}\|_{C^{R_t}}, \|F_{i,1}\|_{C^{R_t}} \leq Q_t.$$

We will expand each local readout into Fourier components along its last central layer $G_{(s)}$ and normalize these components.

For each remaining index $i \in I$, set

$$Z_i = G_{(s)}^{(i)}, \quad T_i = Z_i / (Z_i \cap \Gamma^{(i)}), \quad (85)$$

where $G_{(1)} = G$ and $G_{(j+1)} = [G, G_{(j)}]$. The subgroup Z_i is central and rational, so T_i is a torus. If the local nilpotency step is smaller than s , this torus is trivial. Apply the vertical Fourier decomposition of Section 4.4 along the torus T_i . Explicitly, for a continuous vertical character $\nu: T_i \rightarrow \mathbb{R}/\mathbb{Z}$, define

$$F_{i,\varepsilon,\nu}(z) = \int_{T_i} F_{i,\varepsilon}(u \cdot z) e^{-2\pi i \nu(u)} du, \quad \varepsilon \in \{0, 1\}$$

using Haar probability measure. The local rational-coordinate bounds, as used in Lemma 10.5, cost at most a fixed power of Q_t . Consequently, the integration by parts argument proving Equation 24, spending $D + 2$ derivatives, gives

$$F_{i,\varepsilon} = \sum_{\nu} F_{i,\varepsilon,\nu}, \quad \sum_{\nu} \|F_{i,\varepsilon,\nu}\|_{C^{K_t}} \leq Q_t^C, \quad K_t \geq 1$$

for a fixed constant $C = C_{k,D}$. This exponent includes the coordinate changes and the fixed summation constants; dependence on the derivative orders is absorbed because they are bounded by R_0 .

We normalize exactly as in Section 7.1. Omit the zero components and put

$$c_{i,\varepsilon,\nu} = \|F_{i,\varepsilon,\nu}\|_{C^{K_t}} > 0, \quad F_{i,\varepsilon,\nu} = c_{i,\varepsilon,\nu} \Phi_{i,\varepsilon,\nu}, \quad w_{y,\nu} = \prod_{i \in I} c_{i,y_i,\nu_i}, \quad \sum_{\nu} w_{y,\nu} \leq Q_t^{Cm'}$$

where each normalized readout $\Phi_{i,\varepsilon,\nu}$ has vertical frequency ν , $\|\Phi_i\|_{C^{K_t}} = 1$, and is one-bounded.

With the fixed error exponent $\kappa > 2$ chosen above, set

$$\delta_t = \exp\{-C_* m(1 + \log Q_t)\}, \quad C_* \geq C + \kappa + 2.$$

Fix an output profile $y \in \{0, 1\}^I$ and a frequency tuple $\nu = (\nu_i)_{i \in I}$ such that every component F_{i, y_i, ν_i} is nonzero. Write

$$\Phi_i = \Phi_{i, y_i, \nu_i} \quad (i \in I).$$

Define

$$A_{y, \nu} := \mathbb{E}_{x \in \mathbb{F}_p^{r'}} \prod_{i \in I} \Phi_i(g^{(i)}(L_i(x))\Gamma^{(i)}), \quad B_{y, \nu} := \mathbb{E}_{z \in \mathbb{F}_p^I} \prod_{i \in I} \Phi_i(g^{(i)}(z_i)\Gamma^{(i)}).$$

We call the choice of normalized components *significant* if

$$\max\{|A_{y, \nu}|, |B_{y, \nu}|\} \geq \delta_t. \quad (86)$$

Define the significant index set

$$S_y = \{\nu: \max\{|A_{y, \nu}|, |B_{y, \nu}|\} \geq \delta_t\}.$$

Consider the two experiments from [Section 7.1](#). For the share-input experiment, we have

$$\begin{aligned} P_f(y) &= \mathbb{E}_x \prod_{i \in I} F_{i, y_i} (g^{(i)}(L_i(x))\Gamma^{(i)}) \\ &= \mathbb{E}_x \prod_{i \in I} \left(\sum_{\nu_i} c_{i, y_i, \nu_i} \Phi_{i, y_i, \nu_i} (g^{(i)}(L_i(x))\Gamma^{(i)}) \right) \\ &= \sum_{\nu} \underbrace{\prod_{i \in I} c_{i, y_i, \nu_i}}_{w_{y, \nu}} \underbrace{\mathbb{E}_x \prod_{i \in I} \Phi_{i, y_i, \nu_i} (g^{(i)}(L_i(x))\Gamma^{(i)})}_{A_{y, \nu}} \\ &= \underbrace{\sum_{\nu \in S_y} w_{y, \nu} A_{y, \nu}}_{\text{significant product}} + \underbrace{\sum_{\nu \notin S_y} w_{y, \nu} A_{y, \nu}}_{R_P(y)} \end{aligned}$$

where $\nu = (\nu_i)_{i \in I}$ ranges over the frequency tuples and $R_P(y)$ is the insignificant remainder in the share-input experiment. Similarly, one can also get

$$Q_f(y) = \underbrace{\sum_{\nu \in S_y} w_{y, \nu} B_{y, \nu}}_{\text{significant product}} + \underbrace{\sum_{\nu \notin S_y} w_{y, \nu} B_{y, \nu}}_{R_Q(y)},$$

where $R_Q(y)$ is the insignificant remainder in the random-input experiment. Every term outside S_y has both normalized averages smaller than δ_t . Therefore,

$$|R_P(y)|, |R_Q(y)| \leq \delta_t \sum_{\nu} w_{y, \nu} \leq \delta_t Q_t^{Cm'},$$

where $m' = |I|$ and the final inequality follows from our bound on the total component weight. There are $2^{m'}$ output profiles, so

$$\sum_y (|R_P(y)| + |R_Q(y)|) \leq 2^{m'+1} \delta_t Q_t^{Cm'} \leq e^{-\kappa m}.$$

We can directly control contribution of the remainders, leaving the significant products for the structural argument.

(Step 3) *Convert significant products into short global equations.* After the vertical decomposition and normalization, fix a significant product $\Psi: \mathbb{F}_p^d \rightarrow \mathbb{C}$, so

$$\left| \mathbb{E}_x \Psi(x) \right| \geq \delta_t$$

where d is the number of input parameters in the experiment under consideration. We will convert the significant product averages into short character equations and commutator constraints that the later certificate and central reduction arguments can use.

(a) *Parameters preparation.* Set

$$G/\Gamma = \prod_{i \in I} G^{(i)}/\Gamma^{(i)}, \quad \mathcal{F}((v_i)_{i \in I}) = \prod_{i \in I} \Phi_i(v_i), \quad \nu((u_i)_{i \in I}) = \sum_{i \in I} \nu_i(u_i) \pmod{1}.$$

The product has filtration degree at most k and dimension $h = \sum_{i \in I} \dim G^{(i)} \leq Dm$. Use integral coordinates on the horizontal tori, whose dual lattices are the horizontal characters. The changes from the current Mal'cev coordinates cost $Q_t^{O_{k,D}(1)}$ and are performed locally, before taking products. For a sufficiently large constant $C = C(k, D)$, set

$$M_t = (2m)^C Q_t^C.$$

The product dimension is at most Dm , and its geometry, Lipschitz norm, and frequency bound

$$\text{complexity}(G/\Gamma) \leq M_t, \quad \|\mathcal{F}\|_{\text{Lip}} \leq M_t, \quad \|\nu\|_{\text{vert}} \leq \frac{M_t}{\delta_t}.$$

Here is the frequency bound used above. Significance and one-boundedness imply

$$\delta_t \leq \prod_j \|\Phi_j\|_{\infty} \leq \|\Phi_i\|_{\infty} \quad (i \in I).$$

For a unit vertical direction X , differentiating the vertical transformation law gives $X\Phi_i = 2\pi i \nu_i(X)\Phi_i$. The normalized first derivative bounds, with the local metric comparison, therefore give $\|\nu_i\|_{\text{vert}} \leq Q_t^{O_{k,D}(1)}/\delta_t$. Passing to the product costs only a polynomial factor in m ; the chosen M_t absorbs it. For the direct-product basis, see [CS14, Lemma A.2, p. 27].

(b) *Assuming a suitable line exists, apply [Imported Theorem 10.2](#).* Fix the deterministic character cutoff

$$B_t = \left\lceil Q_t^C \left(\frac{M_t}{\delta_t} \right)^{C_k(Dm)^{\rho_k}} \right\rceil, \quad (87)$$

where the additional power of Q_t covers coordinate conversion. Assume

$$\log p \geq Cm^{v_t + \rho_k + 1} \log(2m) \geq Cm^{\rho_k} (\log(2m) + \log Q_t + \log(1/\delta_t)) \geq \log B_t, \quad (88)$$

which ensures the theorem's field-size condition. Suppose an affine line $x = x_0 + nv$ satisfies the following two properties:

- The restricted product average along the line still has magnitude at least δ_t .
- Every horizontal character within cutoff B_t that is constant on the line is also constant on the whole orbit.

Write $\tilde{g}(n)$ for the product orbit restricted to this line. Affine restriction preserves polynomiality and p -periodicity, and

$$\left| \mathbb{E}_{z \in \mathbb{F}_p} \mathcal{F}(\tilde{g}(z)\Gamma) \right| \geq \delta_t.$$

Apply [Imported Theorem 10.2](#) with parameters k, s, h, M_t, δ_t gives horizontal characters $\xi_1, \dots, \xi_{q_{\text{hor}}}$ of height within the cutoff B_t , constant modulo 1 on the line's orbit, such that

$$d\nu([\cdots [[v_1, v_2], v_3], \dots, v_s]) = 0 \quad \text{for } v_1, \dots, v_s \in \bigcap_{j=1}^{q_{\text{hor}}} \ker d\xi_j. \quad (89)$$

(c) *Line selection.* For every horizontal character ξ_i of height at most B_t , its composition with the product orbit has the form

$$c_{\xi_i} + \frac{P_{\xi_i}(x)}{p} \pmod{1}, \quad \deg P_{\xi_i} \leq k$$

where c_{ξ_i} is a constant. There are at most $(2B_t + 1)^{Dm}$ candidate characters, so we assume the stronger field size condition

$$p > k(2B_t + 1)^{Dm}.$$

One may follow the same line-selection argument from Step 1 in the proof of [Lemma 5.1](#) from [Section 5](#). Replace the previous hyperplane-avoidance argument with polynomial zero sets avoidance. If the polynomial P_{ξ_i} is non-constant of degree j , let $P_{\xi_i,j}$ be its highest homogeneous part. For every starting point x_0 ,

$$P_{\xi_i}(x_0 + ze) = P_{\xi_i,j}(e)z^j + \text{lower-degree terms in } z.$$

The restriction remains non-constant whenever $P_{\xi_i,j}(e) \neq 0$. There are at most $(2B_t + 1)^{Dm}$ candidate characters, where D bounds each local dimension and m is the initial number of shares. Each polynomial excludes at most kp^{d-1} directions. Therefore, there exists a direction e with $P_{\xi_i,j}(e) \neq 0$ for all these polynomials when $p > k(2B_t + 1)^{Dm}$. For the chosen direction,

$$\delta_t \leq \left| \mathbb{E}_x \Psi(x) \right| \leq \mathbb{E}_{x_0} \left| \mathbb{E}_{z \in \mathbb{F}_p} \Psi(x_0 + ze) \right|.$$

Some affine line has bias at least the full average. Any character supplied by [Imported Theorem 10.2](#) that is constant on the line must be constant on the whole orbit, by the choice of e . The same argument applies to the independent-input orbit.

(Step 4) *Apply one certificate to all polynomial coefficients.* [Equation 89](#) from previous step supplies character equations for each significant product. Now, we want to show that, after imposing the local equations, fewer than q_{cert} independent global constraints remain. The prepares the bounded-support argument of [Lemma 10.4](#).

Consider the Good event defined as in [1](#). Let I index the remaining shares, and suppose the remaining number r' of polynomial coefficients satisfies $r'/|I| > b/2$. Set

$$H_t = \max\{B_t, D(2B_t)^{C_{b/2,D}}\},$$

where $C_{b/2,D}$ is the certificate exponent from [Lemma 6.1](#). In horizontal coordinates, the i -th local orbit has the form

$$\vartheta_i + \frac{1}{p} \sum_{j=1}^k \beta_{i,j} z^j \pmod{1}, \quad \beta_{i,j} \in \mathbb{F}_p^{h_i}, \quad h_i \leq D.$$

Prime periodicity and $p > k$ makes the relevant factors invertible and imply this form, by successive finite differences of the torus polynomial. For a global (product) horizontal character with integer local coefficient rows c_i , constancy on the share orbit gives

$$\sum_{i \in I} a_i^h c_i \beta_{i,j} = 0 \quad (1 \leq j \leq k, 0 \leq h < r'). \quad (90)$$

For a fixed degree j , compare the coefficient of x_0^j in $\sum_{i \in I} (c_i \beta_{i,j}) L_i(x)^j$, and then those of $x_0^{j-1} x_h$ for $1 \leq h < r'$. The later coefficients are j times the required moments. Since $p > k$, j is invertible. Homogeneous degrees separate, and degree less than p makes the polynomial-function identity a formal polynomial identity.

Define two families of equations

- Let U_i be the real span of differentials of local horizontal characters of height at most H_t that are constant modulo 1 on $g^{(i)}$.

- Let W be the real span of differentials of global (product) horizontal characters of height at most B_t (as in Equation 87) that are constant modulo 1 on the full share orbit.

Put

$$E_i = \{v_i \in \text{Lie}(G^{(i)}) : \lambda(v_i) = 0, \forall \lambda \in U_i\}, \quad E = \bigoplus_i E_i,$$

and

$$V = \{v \in E : \lambda(v) = 0, \forall \lambda \in W\} = E \cap W^\perp.$$

Thus, E consists of directions satisfying all the chosen local equations, and $V \subseteq E$ additionally satisfies $d\xi_i(v) = 0$ for every chosen global character ξ_i . Some global equations already follow from the local ones, so they vanish automatically on E . Others may become dependent after restriction to E . On the good event, we want

$$\text{codim}_E V < q_{\text{cert}}, \quad \text{codim}_E V = \dim\{\lambda|_E : \lambda \in W\},$$

which measure the additional global constraints inside E .

Next, we want to construct one bounded integer array C to which the certificate applies. Assume the contrary that $\text{codim}_E V \geq q_{\text{cert}}$, choose q_{cert} global characters $\xi_1, \dots, \xi_{q_{\text{cert}}}$ whose differentials restricted to E , $d\xi_1|_E, \dots, d\xi_{q_{\text{cert}}}|_E$ are linearly independent over $\mathbb{R}$. At each share, collect the local rows of ξ_i into

$$C_i \in \mathbb{Z}^{q_{\text{cert}} \times D},$$

padding with zeroes and each entry has height at most B_t . Apply Lemma 6.1 to the r' evaluation tuple moments in Equation 90 and get one nonzero certificate $u_C \in \mathbb{Z}^{q_{\text{cert}}}$ such that

$$u_C^T C_i \beta_{i,j} = 0 \quad (i \in I, 1 \leq j \leq k). \quad (91)$$

The certificate is uniform over all solutions of the same integer array C . Write $\xi_{\ell,i}$ for the restriction of the global character ξ_ℓ to the i -th group. For each i , the row $u_C^T C_i$ defines a locally constant character η_i whose height is at most H_t and is constant modulo 1 on the local orbit,

$$\eta_i = \sum_{\ell=1}^{q_{\text{cert}}} (u_C)_\ell \xi_{\ell,i}, \quad d\eta_i = \sum_{\ell=1}^{q_{\text{cert}}} (u_C)_\ell d\xi_{\ell,i}.$$

For any $v = (v_i)_{i \in I} \in E = \bigoplus_{i \in I} E_i$, we therefore have

$$\left(\sum_{\ell=1}^{q_{\text{cert}}} (u_C)_\ell d\xi_\ell \right) (v) = \sum_{\ell=1}^{q_{\text{cert}}} (u_C)_\ell \sum_{i \in I} d\xi_{\ell,i}(v_i) = \sum_{i \in I} d\eta_i(v_i) = 0,$$

since $d\eta_i \in U_i$, while E_i is defined to annihilate every functional in U_i . Consequently,

$$\sum_{\ell=1}^{q_{\text{cert}}} (u_C)_\ell d\xi_\ell \Big|_E = 0.$$

The linear independence between characters $d\xi_1|_E, \dots, d\xi_{q_{\text{cert}}}|_E$ forces every $(u_C)_\ell = 0$. But the certificate guarantees that $u_C \neq 0$, giving the contradiction. Therefore,

$$\text{codim}_E V < q_{\text{cert}}. \quad (92)$$

This is Section 8 Step 2 with k solutions for the same array.

(Step 5) *Lower the nilpotency step outside an exceptional set of coordinates.* We want to turn the dimension constraints from Equation 92 into simpler local nilsequence representations, so that we can repeat the argument until the remaining groups are abelian. Suppose the current maximum nilpotency step $s \geq 2$, and I indexes the remaining shares.

(a) *Find one bounded set of exceptional shares.* Analogous to Section 7.2 and Section 8 Step 3, we want to find an exceptional set

$$J \subseteq I, \quad |J| \leq s \cdot q_{\text{cert}} \leq k \cdot q_{\text{cert}},$$

such that, after a small modification of the readouts and revealing the field inputs in J , every remaining local group has nilpotency step at most $s - 1$. The filtration degree remains at most k , and the local dimensions remain at most D .

We begin by considering significant vertical product of the normalized component Φ_i and its vertical frequencies ν_i , for $i \in I$, to define the block-diagonal s -linear form

$$T_\nu(v_1, \dots, v_s) = \sum_i d\nu_i([\dots[[v_{i,1}, v_{i,2}], v_{i,3}], \dots, v_{i,s}]) \quad \text{on } E$$

where $v_j = (v_{i,j})_{i \in I}$ is a vector in the product Lie algebra. Here, $d\nu_i$ is the linear functional associated with the vertical frequency on the last lower-central Lie algebra. The commutator $T_\nu(v_1, \dots, v_s)$ measures the combined vertical phase detected by the frequencies when we take s -fold commutators. It is block diagonal because each summand involves directions from just one share. For $s = 2$, it is the familiar form from [Lemma 5.1](#),

$$T_\nu(v, w) = \sum_{i \in I} d\nu_i([v_i, w_i]).$$

Continuing from the dimension conclusion of the previous step, we now show that

$$T_\nu(v_1, \dots, v_s) = 0 \quad \text{whenever } v_1, \dots, v_s \in V,$$

for every significant product, using the same V . The preceding application of [Imported Theorem 10.2](#) supplies horizontal characters $\xi_1, \dots, \xi_{q_{\text{hor}}}$ of height at most B_t , such that $T_\nu(v_1, \dots, v_s)$ vanishes on

$$K^s, \quad K = \bigcap_{\ell=1}^{q_{\text{hor}}} \ker d\xi_\ell.$$

We will show that $V \subseteq K$.

When the significant product is in the share experiment. Each ξ_ℓ is constant modulo 1 on the full share orbit and has height at most B_t and $d\xi_\ell \in W$. Since V annihilates every functional in W , then

$$v \in V \implies d\xi_\ell(v) = 0 \quad \text{for every } \ell.$$

Therefore, [Imported Theorem 10.2](#) makes the s -linear form vanish on V^s ,

$$V \subseteq K, \quad T_\nu|_{V^s} = 0.$$

The same holds when the significant product is in the independent-input experiment. Decompose each global (product) character into its local restrictions, for $(u_i)_{i \in I} \in \prod_{i \in I} G^{(i)}$ where $u_i \in G^{(i)}$,

$$\xi_\ell((u_i)_{i \in I}) = \sum_{i \in I} \xi_{\ell,i}(u_i),$$

where the local restriction $\xi_{\ell,i}(u_i)$ is applying ξ_ℓ on a vector with u_i at the i -th position and 1 everywhere else. Constancy on the independent-input orbit means

$$\sum_{i \in I} \xi_{\ell,i}(g^{(i)}(z_i)) \equiv \text{constant} \pmod{1} \quad \text{for every } z \in \mathbb{F}_p^I.$$

Fix all inputs except z_i for some $i \in I$. Comparing z_i with 0 gives

$$\xi_{\ell,i}(g^{(i)}(z_i)) - \xi_{\ell,i}(g^{(i)}(0)) \equiv 0 \pmod{1}.$$

Each local character $\xi_{\ell,i}$ is constant on its local orbit. Since the height of $\xi_{\ell,i}$ is at most $B_t \leq H_t$, within the cutoff defining U_i , then $d\xi_{\ell,i} \in U_i$. For every $v = (v_i)_{i \in I} \in E$, $d\xi_\ell(v) = \sum_{i \in I} d\xi_{\ell,i}(v_i) = 0$ and hence,

$$V \subseteq E \subseteq K.$$

The span of all commutator forms, $T_\nu(v_1, \dots, v_s)$, also vanishes on V^s . Apply [Lemma 10.4](#) to the linear span, because the commutator forms are block diagonal and $\text{codim}_E V < q_{\text{cert}}$, there exists one exceptional set J with

$$|J| \leq s \text{codim}_E V < sq_{\text{cert}} \leq kq_{\text{cert}}$$

coordinates containing every nonzero local block, simultaneously for all output profiles and significant frequencies. We have selected an exceptional set J such that outside J , each significant frequency individually detects *no* s -fold commutators within E_i . The same J works for all output profiles and significant frequency tuples.

(b) *Average over the commutator tori outside J .* Analogous to [Section 4.4](#), [Section 7.2](#) and [Section 8 Step 4](#), we begin by identifying the central subgroups that are invisible to all significant frequency tuples. For $i \notin J$, the previous step guarantees

$$d\nu_i([\dots[[v_1, v_2], v_3], \dots, v_s]) = 0 \quad \text{for all } v_1, \dots, v_s \in E_i. \quad (93)$$

Set $H_i = \exp(E_i)$, $Z_i = (H_i)_{(s)}$, where Z_i is the s -th lower central subgroup of H_i . The group H_i is rational and contains $[G^{(i)}, G^{(i)}]$. The subgroup Z_i is central in the ambient group, and its Lie algebra is spanned by the s -fold brackets from $\text{Lie}(H_i)$,

$$\text{Lie}(Z_i) = \text{span}_{\mathbb{R}}\{[\dots[[v_1, v_2], v_3], \dots, v_s] : v_1, \dots, v_s \in E_i\}.$$

Since $H_i \subseteq G^{(i)}$, $Z_i \subseteq G_{(s)}^{(i)}$, the vertical character ν_i can be restricted to Z_i . Here, we regard ν_i as an $\mathbb{R}/\mathbb{Z}$ -valued character of the ambient last central subgroup. By linearity, [Equation 93](#) implies that

$$d\nu_i|_{\text{Lie}(Z_i)} = 0.$$

The group Z_i is connected, and every $z \in Z_i$ has the form $z = \exp(w)$ with $w \in \text{Lie}(Z_i)$. Then, every significant frequency annihilates Z_i at these coordinates, i.e.

$$\nu_i(z) = d\nu_i(w) \pmod{1} = 0.$$

We have identified the central motions we can average away while preserving every significant product Φ_i at the current round t ,

$$\mathcal{D}_t = \{\nu = (\nu_i)_{i \in I} : \nu_i|_{Z_i} = 0, \forall i \in I \setminus J\}.$$

We will then average the readout over $Z_i/(Z_i \cap \Gamma^{(i)})$ outside J to make them invariant under the subgroups we will quotient out, while changing the two output laws by at most $e^{-\kappa m}$. For $i \notin J$, define

$$\tilde{F}_i(x) = \int_{Z_i/(Z_i \cap \Gamma^{(i)})} F_i(zx) dz,$$

where dz is normalized Haar measure. For $i \in J$, keep $\tilde{F}_i = F_i$.

By definition of the significant frequency of Φ_i , for $z \in Z_i$ and $x \in G^{(i)}/\Gamma^{(i)}$,

$$\Phi_i(zx) = e^{2\pi w_i(z)} \Phi_i(x) = \Phi_i(x)$$

holds for every significant product Φ_i and every $i \notin J$. Translation invariance of Haar measure makes $\tilde{F}_i$ invariant under Z_i . Moreover,

$$0 \leq \tilde{F}_i \leq 1, \quad \int (1 - F_i(zx)) dz = 1 - \tilde{F}_i(x).$$

Define two output-bit distribution obtained after replacing each readout F_i by the averaged $\tilde{F}_i$. For $(u_i)_{i \in I} \in \mathbb{F}_p^I$, generate the bits independently with

$$\Pr[Y_i = 1 | u_i] = \tilde{F}_i(g^{(i)}(u_i)\Gamma^{(i)}).$$

Write $\tilde{F}_{i,1} = \tilde{F}_i$ and $\tilde{F}_{i,0} = 1 - \tilde{F}_i$. For an output profile $y \in \{0,1\}^I$, define the modified experiments after averaging the readouts as follows.

- Share inputs: Take $u_i = L_i(x)$, where x is the uniform coefficient vector in the current share experiment. Thus,

$$\tilde{P}_f(y) = \mathbb{E}_{x \in \mathbb{F}_p^{r'}} \prod_{i \in I} \tilde{F}_{i,y_i}(g^{(i)}(L_i(x))\Gamma^{(i)}).$$

- Independent inputs: Take $u_i = z_i$ independently and uniformly from $\mathbb{F}_p$. Thus,

$$\tilde{Q}_f(y) = \mathbb{E}_{z \in \mathbb{F}_p^I} \prod_{i \in I} \tilde{F}_{i,y_i}(g^{(i)}(z_i)\Gamma^{(i)}).$$

Recall the original experiments from Step 2 where $w_{y,\nu} \geq 0$ are the normalization weights, A, B are the corresponding normalized product averages in the two experiment,

$$\begin{aligned} P_f(y) &= \sum_{\nu} w_{y,\nu} A_{y,\nu}, & Q_f(y) &= \sum_{\nu} w_{y,\nu} B_{y,\nu} & (\text{Step 2}) \\ \implies P_f(y) - \tilde{P}_f(y) &= \sum_{\nu \notin \mathcal{D}_t} w_{y,\nu} A_{y,\nu}, & Q_f(y) - \tilde{Q}_f(y) &= \sum_{\nu \notin \mathcal{D}_t} w_{y,\nu} B_{y,\nu} \\ & & & (\text{Averaging restricts the sums to } \mathcal{D}_t) \\ \implies |P_f(y) - \tilde{P}_f(y)| + |Q_f(y) - \tilde{Q}_f(y)| &\leq 2\delta_t Q_t^{Cm'} & (|A|, |B| < \delta_t \text{ for } \nu \notin \mathcal{D}_t, \sum_{\nu} w_{y,\nu} \leq Q_t^{Cm'}) \end{aligned}$$

The statistical error is upper bounded by

$$\begin{aligned} \text{SD}(P_f, Q_f) &\leq \text{SD}(\tilde{P}_f, \tilde{Q}_f) + \frac{1}{2} \sum_{y \in \{0,1\}^I} \left(|P_f(y) - \tilde{P}_f(y)| + |Q_f(y) - \tilde{Q}_f(y)| \right) \\ & & & (\text{Triangle inequality}) \\ &\leq \text{SD}(\tilde{P}_f, \tilde{Q}_f) + 2^{m'} \delta_t Q_t^{Cm'} & (\text{Summing over } 2^{m'} \text{ output profiles}) \\ &\leq \text{SD}(\tilde{P}_f, \tilde{Q}_f) + e^{-\kappa m} & (m' \leq m \text{ and } \delta_t = \exp(-C_* m(1 + \log(Q_t)))) \end{aligned}$$

The total change in the two laws is at most $e^{-\kappa m}$ and averaging preserves the range $[0, 1]$ and complements.

(c) *Reveal J , restrict the remaining orbits, and take central quotients.* Following exactly the same conditioning argument from Step 5 in the proof of Lemma 7.1 based on Equation 53. The revealed field inputs are uniform in both experiments. Conditional on those inputs, the bits in J form a common independent factor and can be removed. For the current round, start with current coordinate set I and number of free coefficients r' . Let

$$I' = I \setminus J, \quad m'' = m' - |J|, \quad r'' = r' - |J|$$

and earlier bounds ensures $|J| < r'$. Interpolation expresses the remaining shares as affine transformations of evaluations of a smaller uniform polynomial. The same affine transformations preserve independent uniform inputs. Combining with part (b) of the current step gives

$$\text{SD}(P_f, Q_f) \leq e^{-\kappa m} + \mathbb{E}_{u \in \mathbb{F}_p^J} \text{SD}(\tilde{P}_{I'|u}, \tilde{Q}_{I'|u}).$$

Finally, we want to represent each remaining local map on a group of nilpotency step at most $s - 1$, without changing either of the conditional output law $\tilde{P}_{I'|u}$ or $\tilde{Q}_{I'|u}$. Conditioning replaces the input at each remaining coordinate by an affine bijection of $\mathbb{F}_p$. Hence the set of its local orbit points is unchanged. Every character defining H_i remains constant on the reparametrized orbit, so the same subgroups H_i and Z_i apply in every conditional branch. Apply Lemma 10.5

with geometry bound Q_t and character height bound H_t to restrict each remaining orbit to a translate of H_i . Under the prepared condition $p > (Q_t H_t)^{C_{k,D}}$, [Lemma 10.5](#) gives a bounded representative $c_{i,u}$ in a fixed fundamental box and a polynomial orbit $h_{i,u}$ in H_i such that

$$g_u^{(i)}(z)\Gamma^{(i)} = c_{i,u}h_{i,u}(z)\Gamma^{(i)}, \quad h_{i,u}(z+p)\Gamma_{H_i} = h_{i,u}(z)\Gamma_{H_i}, \quad \Gamma_{H_i} = H_i \cap \Gamma^{(i)}.$$

The local character equations defining H_i now hold exactly along the restricted orbit. Because the averaged readout is Z_i -invariant, it descends to the induced nilmanifold with group

$$Z_i = (H_i)_{(s)}, \quad \tilde{H}_i = H_i/Z_i, \quad (\tilde{H}_i)_{(s)} = \{1\}.$$

Let $\pi_i : H_i \rightarrow \tilde{H}_i$ be the quotient map. The quotient lattice is $\tilde{\Gamma}_i = \pi_i(\Gamma_{H_i})$. Define the new readout and projected orbit as

$$F_{i,u}^{\text{new}}(\pi_i(h)\tilde{\Gamma}_i) = \tilde{F}_i(c_{i,u}h\Gamma^{(i)}), \quad \tilde{h}_{i,u}(z) = \pi_i(h_{i,u}(z)).$$

where $\tilde{F}_i$ is the previous averaged readout. Then,

$$F_{i,u}^{\text{new}}(\tilde{h}_{i,u}(z)\tilde{\Gamma}_i) = \tilde{F}_i(g_u^{(i)}(z)\Gamma^{(i)}). \quad (94)$$

The new readout still takes values in $[0, 1]$. The maximum nilpotency step decreases by at least one so $\tilde{H}_i$ has nilpotency step at most $s - 1$. The projected orbit remains p -periodic, polynomial of filtration degree at most k , and the local dimension at most D . [Lemma 10.5](#) controls the geometry and retained smoothness norms uniformly in u ,

$$Q_{t+1} \leq (Q_t H_t)^{C_{k,D}, R_0}$$

where R_0 is the initial smoothness order. Together with [Equation 88](#), this is covered by [Equation 81](#). The derivative budget satisfies

$$R_{k-1} = (k+1)(D+3) - (k-1)(D+2) = 2D + k + 5 > D + 2,$$

so all expansions and the terminal torus approximation are justified.

Let P_u^{new} and Q_u^{new} be the output laws using these readouts and projected orbits in the two remaining experiments. [Equation 94](#) gives

$$P_u^{\text{new}} = \tilde{P}_{I'|u}, \quad Q_u^{\text{new}} = \tilde{Q}_{I'|u}$$

and

$$\text{SD}(P_f, Q_f) \leq e^{-\kappa m} + \mathbb{E}_{u \in \mathbb{F}_p^J} \text{SD}(P_u^{\text{new}}, Q_u^{\text{new}})$$

The error arise from previous part (b); restriction and quotient step introduce no additional error.

In the proof of [Lemma 7.1](#), [Lemma 7.3](#) performs the analogous restriction and quotient by $Z_i = [H_i, H_i] = (H_i)_{(2)}$ which immediately produces an abelian group. However, here, quotienting by $(H_i)_{(s)}$ produces a group of nilpotency step at most $s - 1$, which may require further rounds of reduction.

(Step 6) *Use the sharp coefficient bound on the terminal tori.* After repeating the central reduction at most $k - 1$ rounds, the remaining local nilmanifolds are tori of dimension at most D . Let I be its remaining coordinates and r' be the number of free polynomial coefficients. The local probability maps now have the form

$$z \mapsto A_i \left(\vartheta_i + \frac{1}{p} \sum_{j=1}^k \beta_{i,j} z^j \right), \quad A_i : \mathbb{T}^{d_i} \rightarrow [0, 1], \quad 0 \leq d_i \leq D. \quad (95)$$

We will largely reuse the same Fourier argument as Step 6 in the proof of [Lemma 7.1](#) and [Proposition 9.2](#). The main generalization is that the relations now involve k polynomial coefficients instead of two. We will show that on the common Good event, after a finite Fourier

approximation costing at most $e^{-\kappa m}$ error, the total weight of all nonzero exact relations is at most

$$C_{b,D} M_p^{r'}, \quad M_p = \frac{1}{p \sin(\pi/(2p))} = \frac{2}{\pi} + O(p^{-2}).$$

Apply [Lemma 7.4](#) with uniform discrepancy $e^{-\kappa m}/(2m)$ to replace each A_i by a trigonometric polynomial $A_{i,T}$ with

$$0 \leq A_{i,T} \leq 1, \quad \|A_i - A_{i,T}\|_\infty \leq \varepsilon_{\text{tor}} := \frac{e^{-\kappa m}}{2m} \quad (96)$$

and its deterministic frequency cutoff T_{tor} satisfies

$$\log T_{\text{tor}} \leq C m^{v_{k-1}} \log(2m).$$

For each i , define

$$\mathbf{E}_i^{\text{fin}} = \text{span}_{\mathbb{F}_p} \{\beta_{i,1}, \dots, \beta_{i,k}\}, \quad A_i^*(v) = A_{i,T} \left(\vartheta_i + \frac{v}{p} \right). \quad (97)$$

We expand $A_i^*(v)$ and combine all integer frequencies $q \in \mathbb{Z}^{d_i}$ that induces the same linear functional

$$\gamma_q(v) = (q \bmod p) \cdot v \quad (v \in \mathbf{E}_i^{\text{fin}}).$$

As in [Equation 56](#)–[Equation 57](#), summing the corresponding coefficients gives the normalized finite Fourier coefficients of A_* ,

$$\mu_i(0) = 1, \quad \mu_i(\gamma_i) = 2 \left| \widehat{A}_i^*(\gamma_i) \right| \quad (\gamma_i \neq 0), \quad \widehat{A}_i^*(\gamma) = \sum_{\substack{q \in \mathbb{Z}^{d_i} \\ \|q\|_\infty \leq T_{\text{tor}} \\ \gamma_q = \gamma}} \widehat{A}_{i,T}(q) e(q \cdot \vartheta_i)$$

where $e(t) = e^{2\pi i t}$. The resulting readout takes values in $[0, 1]$. For the weights of [Equation 58](#), [Lemma 7.5](#) and [Lemma 9.1](#) give

$$\mu_i(0) = 1, \quad \sum_{\gamma_i} \mu_i(\gamma_i)^2 \leq 2, \quad \mu_i(\gamma_i) \leq M_p \quad (\gamma_i \neq 0), \quad |\text{supp } \mu_i| \leq (2T_{\text{tor}} + 1)^D. \quad (98)$$

A frequency tuple $\gamma \in \bigoplus_{i \in I} (\mathbf{E}_i^{\text{fin}})^*$ is an exact relation when

$$P_\gamma(x) := \sum_{i \in I} \sum_{j=1}^k \gamma_i(\beta_{i,j}) L_i(x)^j \equiv 0.$$

Let R be the space of these relations, and let V_0 be the $\mathbb{F}_p$ -span of those relations that actually occur in the Fourier support

$$V_0 = \text{span}_{\mathbb{F}_p} \left(R \cap \prod_{i \in I} \text{supp } \mu_i \right).$$

Lift any q_{cert} independent supported relations to integer frequencies of height at most T_{tor} and apply the same assigned certificate to their k coefficient vectors. The certificate gives

$$s_0 = \dim V_0 < q_{\text{cert}}.$$

It gives a nontrivial dependence of the relation tuples because the $\beta_{i,j}$ span $\mathbf{E}_i^{\text{fin}}$; the field condition ensures that $u_C \neq 0$ modulo p . This repeats Step 7 from the proof of [Lemma 7.1](#).

Choose a j for which some $\gamma_i(\beta_{i,j}) \neq 0$ and use Vandermonde independence in [Equation 90](#), we can show that every nonzero relation has at least $r' + 1$ nonzero blocks;

$$|\{i: \gamma_i \neq 0\}| \geq r' + 1 \quad (\gamma \in R \setminus \{0\}).$$

The two-information-set argument of [Lemma 7.6](#), with $\sigma = M_p$, therefore gives

$$\sum_{\gamma \in R \setminus \{0\}} \prod_{i \in I} \mu_i(\gamma_i) \leq 2^{s_0} M_p^{r'+1-2s_0} \leq C_{b,D} M_p^{r'}. \quad (99)$$

Here $r' > 2q_{\text{cert}}$ for large m and $\tau \leq M_p \leq 2/3$, so the bounded exponent losses are uniform in p .

(Step 7) *Bound the nonconstant polynomial phases.* We will check cancellation for the nonconstant Fourier phases and completes the comparison of the two conditional torus output laws by generalizing the argument from Step 8 in the proof of [Lemma 7.1](#). The quadratic Gauss estimate is replaced by the elementary bound

$$\left| \mathbb{E}_x e_p(P(x)) \right| \leq \rho_{k,p} := \left(\frac{k-1}{p} \right)^{1/2^{k-1}}, \quad e_p(t) = e^{2\pi i t/p}. \quad (100)$$

for every nonconstant polynomial P of degree at most $k < p$. We will replace diagonalization of the quadratic forms from Step 8 with a direction-selection argument that resembles the line-selection argument in Step 3. Let P have degree $d \leq k$, with highest homogeneous part P_d . Choose a direction e satisfying $P_d(e) \neq 0$. On every parallel line with direction e , $P(x_0 + te)$ has degree d in terms on t and the same nonzero leading coefficient $P_d(e)$. We have reduced [Equation 100](#) to a uniform one-variable estimate. For a one-variable polynomial with degree $d \geq 2$ and leading coefficient a_d , take $d-1$ additive differences. The resulting polynomial is linear with coefficient

$$d! a_d h_1 \cdots h_{d-1}$$

where $h_1, \dots, h_{d-1} \in \mathbb{F}_p$ are the auxiliary shifts of the one-variable input t . For $d=1$, character orthogonality gives mean zero. $d-1$ steps of repeated Cauchy-Schwarz and linear-character orthogonality leaves only shifts where the linear coefficients vanishes. That coefficient is a nonzero constant times the product of the $d-1$ shifts, which vanishes with probability at most $(d-1)/p$. Then,

$$\left| \mathbb{E}_t e_p(P(t)) \right|^{2^{d-1}} \leq \Pr[h_1 \cdots h_{d-1} = 0] \leq \frac{d-1}{p}$$

implying [Equation 100](#).

Fix a revealed history u . Let $P_u^{\text{new}}, Q_u^{\text{new}}$ be the terminal laws before the Fourier approximation of Step 6, and let $P_u^{\text{tor}}, Q_u^{\text{tor}}$ be the laws after it. Within each experiment, couple the original and approximating readouts using the same field inputs and the same independent auxiliary uniform variables. The coupling inequality and [Equation 96](#) give

$$\text{SD}(P_u^{\text{new}}, P_u^{\text{tor}}), \quad \text{SD}(Q_u^{\text{new}}, Q_u^{\text{tor}}) \leq |I| \varepsilon_{\text{tor}} \leq \frac{1}{2} e^{-\kappa m}. \quad (101)$$

Consequently,

$$\text{SD}(P_u^{\text{new}}, Q_u^{\text{new}}) \leq \text{SD}(P_u^{\text{tor}}, Q_u^{\text{tor}}) + e^{-\kappa m}.$$

Recall the finite-space readout $A_i^* : \mathbb{E}_i^{\text{fin}} \rightarrow [0, 1]$ from [Equation 97](#). Set $A_{i,1}^* = A_i^*$, $A_{i,0}^* = 1 - A_i^*$, $v_i(t) = \sum_{j=1}^k \beta_{i,j} t^j$. For the fixed field-input vector $w = (w_i)_{i \in I}$, conditional independence and Fourier inversion implies

$$A_{i,y_i}^*(v_i(w_i)) = \sum_{\gamma_i \in (\mathbb{E}_i^{\text{fin}})^*} \widehat{A_{i,y_i}^*}(\gamma_i) e_p \left(\sum_{j=1}^k \gamma_i(\beta_{i,j}) w_i^j \right) \quad (\text{Fourier inversion})$$

$$\Pr[Y^{\text{tor}} = y | w] = \prod_{i \in I} A_{i,y_i}^*(v_i(w_i)) = \sum_{\gamma} \left(\prod_{i \in I} \widehat{A_{i,y_i}^*}(\gamma_i) \right) \cdot e_p \left(\sum_{i \in I} \sum_{j=1}^k \gamma_i(\beta_{i,j}) w_i^j \right)$$

For frequency tuple $\gamma = (\gamma_i)_{i \in I}$, define

$$M_\gamma = \mathbb{E}_{x \in \mathbb{F}_p^{r'}} e_p \left(\sum_{i \in I} \sum_{j=1}^k \gamma_i(\beta_{i,j}) L_i(x)^j \right), \quad N_\gamma = \mathbb{E}_{z \in \mathbb{F}_p^I} e_p \left(\sum_{i \in I} \sum_{j=1}^k \gamma_i(\beta_{i,j}) z_i^j \right),$$

where $L_i(x) = \sum_{h=0}^{r'-1} a_i^h x_h$. Then,

$$P^{\text{tor}}(y) = \sum_{\gamma} \left(\prod_{i \in I} \widehat{A_{i,y_i}^*}(\gamma_i) \right) M_\gamma, \quad Q^{\text{tor}}(y) = \sum_{\gamma} \left(\prod_{i \in I} \widehat{A_{i,y_i}^*}(\gamma_i) \right) N_\gamma.$$

Let R be the exact relation from the previous step. Then, similar as the three case estimate from Step 8 in Section 8, replacing $p^{-1/2}$ with $\rho_{k,p}$, then

$$|M_\gamma - N_\gamma| \leq \begin{cases} 0 & \gamma = 0, \\ 1 + \rho_{k,p} & \gamma \in R \setminus \{0\}, \\ 2\rho_{k,p} & \gamma \notin R. \end{cases} \quad (102)$$

For every nonzero tuple, the independent-input phase is nonconstant. Otherwise, each frequency γ_i would annihilate all generators $\beta_{i,j}$, forcing $\gamma_i = 0$. For every local frequency γ ,

$$\sum_{\varepsilon \in \{0,1\}} \left| \widehat{A_{i,\varepsilon}^*}(\gamma_i) \right| = \mu_i(\gamma_i).$$

At frequency zero, it follows because the two means are nonnegative and sum to one. At nonzero frequencies, the two Fourier coefficients are negative of each other. Following the same derivation as Equation 70, with weight μ_i , triangle inequality over all output profiles gives

$$\begin{aligned} \text{SD}(P^{\text{tor}}, Q^{\text{tor}}) &\leq \frac{1}{2} \sum_y \sum_{\gamma} \prod_{i \in I} \left| \widehat{A_{i,y_i}^*}(\gamma_i) \right| |M_\gamma - N_\gamma| \\ &= \frac{1}{2} \sum_{\gamma} \left(\prod_{i \in I} \mu_i(\gamma_i) \right) |M_\gamma - N_\gamma| \\ &\leq \frac{1}{2} \sum_{R \setminus \{0\}} \prod_{i \in I} \mu_i(\gamma_i) + \rho_{k,p} \prod_{i \in I} \sum_{\gamma_i} \mu_i(\gamma_i) && \text{(Equation 102)} \\ &\leq C_{b,D} M_p^{r'} + \rho_{k,p} \prod_{i \in I} \sum_{\gamma_i} \mu_i(\gamma_i) && \text{(Equation 99)} \\ &\leq C_{b,D} M_p^{r'} + \rho_{k,p} \prod_{i \in I} \left(\sqrt{|\text{supp } \mu_i|} \left(\sum_{\gamma_i} \mu_i(\gamma_i)^2 \right)^{1/2} \right) && \text{(Cauchy-Schwarz)} \\ &\leq C_{b,D} M_p^{r'} + \rho_{k,p} \left[\sqrt{2}(2T_{\text{tor}} + 1)^{D/2} \right]^m && \text{(Equation 98)} \end{aligned}$$

This accounts for all Fourier coefficients and all output profiles uniformly over the revealed values. The conditional statistical distance is at most

$$C_{b,D} M_p^{r'} + \rho_{k,p} \left[\sqrt{2}(2T_{\text{tor}} + 1)^{D/2} \right]^m. \quad (103)$$

(Step 8) Close the quantitative estimates and return to the original structured laws. Finally, we combine bounds from previous steps for each conditional torus model into a bound for the original structured experiments

$$\mathbb{E}_{a \in \mathcal{A}_{p,m}} \Delta_{\mathcal{G}_{p,\ell}}(a; r)$$

under the field assumption

$$\log p \geq Cm^{\Theta_{k+1}} \log(2m).$$

Recall from step 1 that T_* bound all character and Fourier cutoffs used throughout the proof. The parameter recurrences gives

$$\log T_* \leq Cm^{v_{k-1}} \log(2m), \quad v_{k-1} = (k-1)(\rho_k + 2),$$

and the proposition uses $\Theta_{k+1} = v_{k-1} + 2$ where ρ_k is the exponent from [Imported Theorem 10.2](#). We spell out why the ledger covers a complete round. The definitions give

$$\begin{aligned} \log(1/\delta_t) &\leq Cm(1 + \log Q_t), \\ \log M_t &\leq C(\log(2m) + \log Q_t), \\ \log B_t, \log H_t &\leq Cm^{v_t + \rho_k + 1} \log(2m), \\ \log Q_{t+1} &\leq C(\log Q_t + \log H_t). \end{aligned}$$

The logarithmic cost of line selection and of the certificate field-size conditions is at most

$$C(m \log B_t + m \log(2m)) \leq Cm^{v_t + \rho_k + 2} \log(2m).$$

Thus $v_{t+1} = v_t + \rho_k + 2$ covers one full round. At the terminal stage, all remaining error bounds cost at most $Cm \log(2T_*) + C \log(2m)$ in the exponent. The choice $\Theta_{k+1} = v_{k-1} + 2$ leaves ample room for these costs and for making the errors exponentially small in m .

The error from the previous step satisfies

$$\log\left(\rho_{k,p} \left[\sqrt{2}(2T_{\text{tor}} + 1)^{D/2}\right]^m\right) \leq -\frac{\log p}{2^{k-1}} + Cm \log(2T_*). \quad (104)$$

Summing the certificate failure bound over the permitted deletion sets gives

$$\Pr[\text{Good}^c] \leq p^{-1} \exp\{Cm \log(2T_*) + C \log(2m)\} \leq 2e^{-2m}.$$

Iterate the one round inequality through at most $k-1$ reductions, then restore the terminal torus approximation. The chosen thresholds make the accumulated error $O_k(e^{-2m})$. Writing u for the complete revealed history and r_u for the remaining number of coefficients in that branch, we obtain

$$\text{SD}(P_f, Q_f) \leq C_0 \mathbb{E}_u M_p^{r_u} + O_k(e^{-2m}).$$

The total number of revealed inputs is bounded by $k^2 q_{\text{cert}}$. Then, $0 \leq r - r_u \leq k^2 q_{\text{cert}}$ and $M_p \geq 2/\pi$ implies that

$$M_p^{r_u} = M_p^r M_p^{-(r-r_u)} \leq \left(\frac{2}{\pi}\right)^{-k^2 q_{\text{cert}}} M_p^r$$

where $(2/\pi)^{-k^2 q_{\text{cert}}}$ is a fixed constant, uniform over all rounds. Average over their values, restore the perturbations, and take the supremum over the original maps. All estimates hold uniformly on the same good event. Following from Step 1, we have

$$\begin{aligned} \mathbb{E}_a \Delta_{\mathcal{G}_{p,\ell}}(a; r) &\leq \mathbb{E}_a [\Delta_{\mathcal{G}_{p,\ell}}(a; r) \mathbb{1}_{\text{Good}}(a)] + \Pr[\text{Good}^c] \\ &\leq A_0 M_p^r + O_k(e^{-2m}) \end{aligned}$$

Since $M_p = \frac{2}{\pi} + O(p^{-2}) = \left(\frac{2}{\pi}\right) e^{O(p^{-2})}$, then

$$M_p^r = \left(\frac{2}{\pi}\right)^r e^{O(rp^{-2})} \leq C_0 \left(\frac{2}{\pi}\right)^r$$

uniformly for $r < m$ and $p > m(m+1)$; absorb C_0 into A . With $e^{-2m} \leq (\frac{2}{\pi})^r$, enlarging the constant gives

$$\mathbb{E}_a \Delta_{\mathcal{G}_{p,\ell}}(a; r) \leq A \left(\frac{2}{\pi} \right)^r. \quad (105)$$

□

10.5. Finish the outer reduction.

Proof roadmap. [Proposition 10.3](#) gives the same structured bound as in the quadratic case, with a larger field-size exponent. We now pass from structured maps to arbitrary one-bit leakage. The cutoff t controls the number of terms with few approximation errors; the accuracy η suppresses the remaining terms.

Proof of Theorem 2. Fix ℓ, c, ε as in the theorem, and put $\tau = 2/\pi$ and $\lambda = \log_2(\pi/2)$. Write $h_2(x) = -x \log_2 x - (1-x) \log_2(1-x)$. Choose a fixed $\alpha > 0$ such that

$$\alpha < \min \left\{ \frac{1}{4}, \frac{c}{4}, \frac{\ell c - 1}{2(\ell - 1)} \right\}, \quad h_2(\alpha) + \alpha \log_2 \pi \leq \frac{\varepsilon c \lambda}{4}. \quad (106)$$

This is possible because $\ell c > 1$ and the left-hand side of the second inequality tends to zero with α . Set

$$t = \lfloor \alpha m \rfloor, \quad \eta = 2^{-3/\alpha}, \quad b = c/2.$$

Fix the affine-closed class $\mathcal{G}_{p,\ell}$ from [Section 10.1](#) at accuracy η , and let A, C be the constants in [Proposition 10.3](#) for this class and b . The class data and these constants depend only on c, ℓ, ε ; the exponent Θ_ℓ depends only on ℓ .

Apply the reduction to every shortening. For sufficiently large m , we have $1 \leq t < d$ and

$$\ell d - (\ell - 1)t - m \geq (\ell c - (\ell - 1)\alpha - 1)m - \ell > 0.$$

Thus [Lemma 10.1](#) applies. Moreover, for every $j < t$,

$$d - j \geq \frac{c}{2}(m - j)$$

once m is sufficiently large. Every fixed deletion of j coordinates leaves a uniformly distributed evaluation tuple of length $m - j$. The field-size hypothesis at length m implies the corresponding hypothesis at length $m - j$, so [Proposition 10.3](#) gives

$$\mathbb{E}_a \Delta_{\mathcal{G}_{p,\ell}}(a_{J^c}; d - |J|) \leq A \tau^{d - |J|} \quad (|J| < t).$$

Averaging [Lemma 10.1](#) and grouping the deletion sets by size, we obtain

$$\begin{aligned} \mathbb{E}_a \Delta(a; d) &\leq A \tau^d \sum_{j < t} \binom{m}{j} (2/\tau)^j + 4^m \eta^t \\ &\leq A \tau^d 2^{m(h_2(\alpha) + \alpha \log_2 \pi)} + \eta^{-1} 2^{-m}. \end{aligned}$$

Here $2/\tau = \pi$, the binomial sum is bounded by the usual entropy estimate, and $t \geq \alpha m - 1$ gives $4^m \eta^t \leq \eta^{-1} 2^{-m}$.

Absorb the entropy cost. For large m , we have $d \geq cm/2$, so [Equation 106](#) implies

$$m(h_2(\alpha) + \alpha \log_2 \pi) \leq \varepsilon \lambda d / 2.$$

Since $\tau = 2^{-\lambda}$, it follows that

$$\mathbb{E}_a \Delta(a; d) \leq A \tau^{(1-\varepsilon/2)d} + \eta^{-1} 2^{-m}.$$

Divide by $\tau^{(1-\varepsilon)d}$. The two terms on the right become

$$A \tau^{\varepsilon d / 2} \quad \text{and} \quad \eta^{-1} 2^{-m + (1-\varepsilon)\lambda d},$$

and both tend to zero: for the second, use $d \leq cm$ and $(1 - \varepsilon)c\lambda < 1$. Increasing the lower threshold for m therefore gives

$$\mathbb{E}_a \Delta(a; d) \leq \tau^{(1-\varepsilon)d}.$$

The pointwise comparison $\text{Insec}(a; d) \leq 2\Delta(a; d)$ from [Equation 2](#) in [Section 1.1](#) gives the other inequality in [Equation 5](#).

Let the exponent loss tend to zero. Suppose $m \rightarrow \infty$ and

$$\frac{\log p}{m^{\Theta_\ell} \log(2m)} \rightarrow \infty.$$

For every fixed $\varepsilon > 0$, the field-size and length conditions just proved eventually hold. A diagonal choice of $\varepsilon \rightarrow 0$ therefore gives the claimed $o(1)$ loss; the factor 2 in the insecurity bound costs only $\mathcal{O}(1/d)$ in the exponent. This uses no bound on the growth of the class data as $\varepsilon \rightarrow 0$. For n -bit primes and $m = \lfloor (n)^u \rfloor$,

$$\frac{\log p}{m^{\Theta_\ell} \log(2m)} \gg_u \frac{(n)^{1-u\Theta_\ell}}{\log n} \rightarrow \infty \quad \text{if } 0 < u < 1/\Theta_\ell.$$

The estimate is uniform over these primes. □

The resulting tradeoff. Taking $\varepsilon = 1/2$ recovers $\mathbb{E}_a \Delta(a; d) \leq 2^{-d/4}$ and $\mathbb{E}_a \text{Insec}(a; d) \leq 2 \cdot 2^{-d/4}$, since $\frac{1}{2} \log_2(\pi/2) > 1/4$. For example, $\ell = 4$ covers the new range $1/4 < c \leq 1/3$. The admissible party exponent is $u_0(\ell) = 1/\Theta_\ell$. The dependence on Leng's exponent ρ_k is explicit in [Equation 77](#); no numerical value of $u_0(4)$ is asserted. When $c > 1/3$, the separate quadratic argument retains the stronger range $0 < u < 1/2$.

APPENDIX A. THE LSB LOWER BOUND AND ITS INTERVAL FORM

Exponential decay in d is the right scale for insecurity. For additive and Shamir sharing, the base $2/\pi$ already appears in the lower bounds proved in [MNP⁺22, Theorems 1 and 6, pp. 16:5 and 16:7]. The same reconstruction-multiplier LSB leakage gives the following formulation with our degree convention. Here LSB means the least significant bit of the canonical binary representative in $\{0, \dots, p-1\}$. We include the Fourier calculation to make the connection with structured leakage explicit.

Proposition A.1 (The LSB lower bound). *For every odd prime p , every tuple a of distinct nonzero evaluation places, and every $1 \leq d < m$, put $M_p = 1/(p \sin(\pi/(2p)))$. Then*

$$\text{Insec}(a; d) \geq M_p^d \geq \left(\frac{2}{\pi}\right)^d, \quad \Delta(a; d) \geq \frac{1}{2} M_p^d \geq \frac{1}{2} \left(\frac{2}{\pi}\right)^d.$$

Proof. Choose $k = d + 1$ shares. Their nonzero reconstruction coefficients $\lambda_1, \dots, \lambda_k$ satisfy $s = \sum_{i=1}^k \lambda_i P_s(a_i)$. Thus $Z_i = \lambda_i P_s(a_i)$ are uniform subject to $\sum_i Z_i = s$. Party $i \leq k$ leaks the LSB of Z_i ; the remaining parties leak constants. This is the parity-of-parity construction after multiplication by the reconstruction coefficients [MNP⁺22, Theorem 6 and its proof, p. 16:7].

Write $[z]_p \in \{0, \dots, p-1\}$ for the canonical representative of $z \in \mathbb{F}_p$, and set

$$\ell(z) = (-1)^{[z]_p}.$$

For a real function h on $\mathbb{F}_p$, use the normalized Fourier coefficient $\widehat{h}(\xi) = \mathbb{E}_x h(x) \exp(-2\pi i \xi x/p)$. A geometric sum gives

$$\widehat{\ell}(\xi) = \frac{2}{p(1 + \exp(-2\pi i \xi/p))}, \quad \left| \widehat{\ell}(\xi_*) \right| = M_p, \quad \xi_* = (p-1)/2.$$

For every nonzero ξ , Lemma 9.1 gives

$$\sup_{h: \mathbb{F}_p \rightarrow [-1,1]} \left| \widehat{h}(\xi) \right| = M_p.$$

Indeed, apply the lemma to $A = (1 + h)/2$, using $\widehat{A}(\xi) = \widehat{h}(\xi)/2$. Equality is attained by $h(x) = \ell((\xi/\xi_*)x)$, by the preceding calculation.

The parity of the leaked bits has signed expectation

$$H(s) = \mathbb{E}_{\sum_i Z_i = s} \prod_{i=1}^k \ell(Z_i).$$

With our normalized Fourier convention,

$$H(s) = \frac{1}{p^{k-1}} \sum_{\substack{z_1, \dots, z_k \in \mathbb{F}_p \\ z_1 + \dots + z_k = s}} \prod_{i=1}^k \ell(z_i), \quad \widehat{H}(\xi) = \widehat{\ell}(\xi)^k.$$

Let

$$R = \max_s H(s) - \min_s H(s), \quad c_0 = \frac{\max_s H(s) + \min_s H(s)}{2}.$$

In particular, $|\widehat{H}(\xi_*)| = M_p^k$. Since $\widehat{H}(\xi_*) \neq 0$, H is non-constant, and we have $R > 0$. The normalized function

$$h(s) = \frac{H(s) - c_0}{R/2}$$

takes values in $[-1, 1]$. Subtracting c_0 does not change the Fourier coefficient at the nonzero frequency ξ_* , so the extremal estimate gives

$$M_p^k = \left| \widehat{H}(\xi_*) \right| = \frac{R}{2} \left| \widehat{h}(\xi_*) \right| \leq \frac{R}{2} M_p.$$

Thus $R/2 \geq M_p^{k-1} = M_p^d$.

Choose secrets s_+ and s_- at which H attains its maximum and minimum. The signed parity of the leaked bits has expectations $H(s_+)$ and $H(s_-)$ under their respective leakage laws. These expectations differ by R , so the two laws have statistical distance at least $R/2$. Consequently,

$$\text{Insec}(a; d) \geq \frac{R}{2} \geq M_p^{k-1} = M_p^d \geq (2/\pi)^d,$$

where the last inequality follows from $\sin t \leq t$. Finally, the pointwise comparison $\text{Insec}(a; d) \leq 2\Delta(a; d)$ from Equation 2 of Section 1.1 gives

$$\Delta(a; d) \geq \frac{1}{2}M_p^d \geq \frac{1}{2}(2/\pi)^d.$$

□

Lev's intervals and the LSB leakage. The LSB partition becomes an interval partition after an invertible rescaling:

$$\ell(2y) = 2\mathbf{1}_I(y) - 1, \quad I = \{0, 1, \dots, (p-1)/2\}.$$

Indeed, $[2y]_p = 2[y]_p$ on I , whereas $[2y]_p = 2[y]_p - p$ on its complement; these representatives are even and odd, respectively. Thus leaking the LSB of $\lambda_i P_s(a_i)$ is exactly the interval test $\lambda_i P_s(a_i)/2 \in I$. For arbitrary evaluation places, this is one-bit leakage after a public multiplication; it need not be a physical bit of the original share [MNP⁺22, Theorem 6 and the following discussion, p. 16:7].

Lev's majorization theorem supplies the extremal additive-combinatorial perspective: replacing sets in $\mathbb{F}_p$ by intervals of the same sizes can only increase the sums of the largest convolution representation counts [Lev01, Theorem 1, p. 244]. At the half-field densities above, these extremal intervals are precisely the LSB fibres after dilation. Lev's theorem concerns the concentration of a single additive relation; the LSB construction turns this structure into a leakage distinguisher. It does not follow that these maps maximize the full Reed–Solomon leakage distance.

REFERENCES

- [ADN⁺19] Divesh Aggarwal, Ivan Damgård, Jesper Buus Nielsen, Maciej Obremski, Erick Purwanto, João Ribeiro, and Mark Simkin. Stronger leakage-resilient and non-malleable secret sharing schemes for general access structures. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019, Part II*, volume 11693 of *Lecture Notes in Computer Science*, pages 510–539, Santa Barbara, CA, USA, August 18–22, 2019. Springer, Cham, Switzerland. doi:10.1007/978-3-030-26951-7_18. ↑5
- [BDIR18] Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018, Part I*, volume 10991 of *Lecture Notes in Computer Science*, pages 531–561, Santa Barbara, CA, USA, August 19–23, 2018. Springer, Cham, Switzerland. doi:10.1007/978-3-319-96884-1_18. ↑1, 2
- [BDIR19] Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. Cryptology ePrint Archive, Report 2019/653, 2019. URL: <https://eprint.iacr.org/2019/653>. ↑9, 57
- [BDIR21] Fabrice Benhamouda, Akshay Degwekar, Yuval Ishai, and Tal Rabin. On the local leakage resilience of linear secret sharing schemes. *Journal of Cryptology*, 34(2):10, April 2021. doi:10.1007/s00145-021-09375-2. ↑1, 2, 5, 57
- [BGW88] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 1–10, Chicago, IL, USA, May 2–4, 1988. ACM Press. doi:10.1145/62212.62213. ↑2
- [BHM⁺26] Aniruddha Biswas, Jihun Hwang, Hemanta K. Maji, Ilya D. Shkredov, and Xiuyu Ye. Beyond all-or-nothing corruption: Additive secret sharing under hamming-weight leakage. Manuscript available at <https://www.cs.purdue.edu/homes/hmaji/papers/BHMSY26b.pdf>, 2026. Preprint. ↑5
- [CKOS21] Nishanth Chandran, Bhavana Kanukurthi, Sai Lakshmi Bhavana Obbattu, and Sruthi Sekar. Adaptive extractors and their application to leakage resilient secret sharing. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021, Part III*, volume 12827 of *Lecture Notes in Computer Science*, pages 595–624, Virtual Event, August 16–20, 2021. Springer, Cham, Switzerland. doi:10.1007/978-3-030-84252-9_20. ↑5
- [CKOS22] Nishanth Chandran, Bhavana Kanukurthi, Sai Lakshmi Bhavana Obbattu, and Sruthi Sekar. Short leakage resilient and non-malleable secret sharing schemes. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022, Part I*, volume 13507 of *Lecture Notes in Computer Science*, pages 178–207, Santa Barbara, CA, USA, August 15–18, 2022. Springer, Cham, Switzerland. doi:10.1007/978-3-031-15802-5_7. ↑5
- [CS14] Pablo Candela and Olof Sisask. Convergence results for systems of linear forms on cyclic groups and periodic nilsequences. *SIAM Journal on Discrete Mathematics*, 28(2):786–810, 2014. Definition and page references use arXiv:1212.3681v2. URL: <https://epubs.siam.org/doi/abs/10.1137/130935677>, doi:10.1137/130935677. ↑7, 18, 19, 20, 21, 23, 25, 27, 28, 51, 68, 79
- [DLM⁺26] Dean Doron, Tal Leonov, Jonathan Mosheiff, Henrique Navas, Nicolas Resch, and João Ribeiro. Discrepancy for random linear codes. *arXiv e-prints*, page arXiv:2606.24471, 2026. arXiv:2606.24471. ↑3, 5
- [ET12] Tanja Eisner and Terence Tao. Large values of the Gowers–Host–Kra seminorms. *Journal d’Analyse Mathématique*, 117(1):133–186, 2012. Equation and page references use arXiv:1012.3509v2. URL: <https://arxiv.org/abs/1012.3509>, doi:10.1007/s11854-012-0018-2. ↑27
- [FMM⁺24] Sebastian Faust, Loïc Masure, Elena Micheli, Maximilian Ortl, and François-Xavier Standaert. Connecting leakage-resilient secret sharing to practice: Scaling trends and physical dependencies of prime field masking. In Marc Joye and Gregor Leander, editors, *Advances in Cryptology – EUROCRYPT 2024, Part IV*, volume 14654 of *Lecture Notes in Computer Science*, pages 316–344, Zurich, Switzerland, May 26–30, 2024. Springer, Cham, Switzerland. doi:10.1007/978-3-031-58737-5_12. ↑5
- [GK18] Vipul Goyal and Ashutosh Kumar. Non-malleable secret sharing. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *50th Annual ACM Symposium on Theory of Computing*, pages 685–698, Los Angeles, CA, USA, June 25–29, 2018. ACM Press. doi:10.1145/3188745.3188872. ↑1
- [GT10] Ben Green and Terence Tao. Linear equations in primes. *Annals of Mathematics. Second Series*, 171(3):1753–1850, 2010. doi:10.4007/annals.2010.171.1753. ↑14, 16
- [GT12] Ben Green and Terence Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Annals of Mathematics*, 175(2):465–540, 2012. URL: <https://annals.math.princeton.edu/2012/175-2/p02>, doi:10.4007/annals.2012.175.2.2. ↑6, 18, 23, 25, 26

- [KA25] Hiroki Koga and Hiroto Abe. New tight bounds on the local leakage resilience of the additive (n, n) -threshold scheme determined by the eigenvalues of circulant matrices. In Goichiro Hanaoka and Bo-Yin Yang, editors, *Advances in Cryptology – ASIACRYPT 2025, Part VIII*, volume 16252 of *Lecture Notes in Computer Science*, pages 34–66, Melbourne, VIC, Australia, December 8–12, 2025. Springer, Singapore, Singapore. doi:10.1007/978-981-95-5125-5_2. ↑5
- [Kas24] Dustin Kasser. An improvement upon the bounds for the local leakage resilience of shamir’s secret sharing scheme. In Elette Boyle and Mohammad Mahmoody, editors, *TCC 2024: 22nd Theory of Cryptography Conference, Part IV*, volume 15367 of *Lecture Notes in Computer Science*, pages 395–422, Milan, Italy, December 2–6, 2024. Springer, Cham, Switzerland. doi:10.1007/978-3-031-78023-3_13. ↑2
- [KK23] Ohad Klein and Ilan Komargodski. New bounds on the local leakage resilience of Shamir’s secret sharing scheme. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023, Part I*, volume 14081 of *Lecture Notes in Computer Science*, pages 139–170, Santa Barbara, CA, USA, August 20–24, 2023. Springer, Cham, Switzerland. doi:10.1007/978-3-031-38557-5_5. ↑2, 64
- [KMS19] Ashutosh Kumar, Raghu Meka, and Amit Sahai. Leakage-resilient secret sharing against colluding parties. In David Zuckerman, editor, *60th Annual Symposium on Foundations of Computer Science*, pages 636–660, Baltimore, MD, USA, November 9–12, 2019. IEEE Computer Society Press. doi:10.1109/FOCS.2019.00045. ↑5
- [KPS01] Teresa Krick, Luis Miguel Pardo, and Martín Sombra. Sharp estimates for the arithmetic Nullstellensatz. *Duke Mathematical Journal*, 109(3):521–598, 2001. Page references use the linked author version: Corollary 2.11 and its affine-space specialization, p. 25. URL: <https://www.maia.ub.es/~sombra/publications/arith-nss/maino.pdf>, doi:10.1215/S0012-7094-01-10934-4. ↑3, 10, 11, 41, 42
- [Len25] James Leng. *Szemerédi’s Theorem, Primes, and Nilsequences*. PhD thesis, University of California, Los Angeles, 2025. Page references use the printed thesis pagination. PDF: <https://escholarship.org/content/qt7z39f2qk/qt7z39f2qk.pdf>. URL: <https://escholarship.org/uc/item/7z39f2qk>. ↑3, 6, 10, 11, 12, 13, 21, 25, 32, 33, 36, 38, 50, 51, 69, 73, 74
- [Lev01] Vsevolod F. Lev. Linear equations over $\mathbb{F}_p$ and moments of exponential sums. *Duke Mathematical Journal*, 107(2):239–263, 2001. doi:10.1215/S0012-7094-01-10722-9. ↑10, 56, 65, 93
- [LMS25a] Matan Levi, Jonathan Mosheiff, and Nikhil Shagrirhaya. Random Reed–Solomon codes and random linear codes are locally equivalent. arXiv:2406.02238v7, 2025. Revision of 14 November 2025. Proposition 6.1, equations (23)–(24), p. 28; quotient construction, pp. 28–29; Lemma 7.7, p. 35. The conference version appears in FOCS 2025, pp. 2097–2131. URL: <https://arxiv.org/abs/2406.02238v7>. ↑3, 9, 10, 11, 39, 41, 42
- [LMS25b] Matan Levi, Jonathan Mosheiff, and Nikhil Shagrirhaya. Random reed-solomon codes and random linear codes are locally equivalent. In *66th Annual Symposium on Foundations of Computer Science*, pages 2097–2131, Sydney, Australia, December 14–17, 2025. IEEE Computer Society Press. doi:10.1109/FOCS63196.2025.00112. ↑10, 41, 42
- [MNP⁺21] Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, Tom Suad, and Mingyuan Wang. Leakage-resilience of the Shamir secret-sharing scheme against physical-bit leakages. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology – EUROCRYPT 2021, Part II*, volume 12697 of *Lecture Notes in Computer Science*, pages 344–374, Zagreb, Croatia, October 17–21, 2021. Springer, Cham, Switzerland. doi:10.1007/978-3-030-77886-6_12. ↑2, 5
- [MNP⁺22] Hemanta K. Maji, Hai H. Nguyen, Anat Paskin-Cherniavsky, Tom Suad, Mingyuan Wang, Xiuyu Ye, and Albert Yu. Tight estimate of the local leakage resilience of the additive secret-sharing scheme & its consequences. In Dana Dachman-Soled, editor, *ITC 2022: 3rd Conference on Information-Theoretic Cryptography*, volume 230 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 16:1–16:19, Cambridge, MA, USA, July 5–7, 2022. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik. doi:10.4230/LIPIcs.ITC.2022.16. ↑2, 5, 64, 92, 93
- [MPSW21] Hemanta K. Maji, Anat Paskin-Cherniavsky, Tom Suad, and Mingyuan Wang. Constructing locally leakage-resilient linear secret-sharing schemes. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021, Part III*, volume 12827 of *Lecture Notes in Computer Science*, pages 779–808, Virtual Event, August 16–20, 2021. Springer, Cham, Switzerland. doi:10.1007/978-3-030-84252-9_26. ↑5
- [Ngu24] Hai H. Nguyen. Towards breaking the half-barrier of local leakage-resilient shamir’s secret sharing. In Leonid Reyzin and Douglas Stebila, editors, *Advances in Cryptology – CRYPTO 2024, Part V*, volume 14924 of *Lecture Notes in Computer Science*, pages 257–285, Santa Barbara, CA, USA, August 18–22, 2024. Springer, Cham, Switzerland. doi:10.1007/978-3-031-68388-6_10. ↑2
- [Ngu26] Hai H. Nguyen. Local leakage resilience of Shamir secret sharing: Below half threshold rate. arXiv:2610.17711v1, 2026. URL: <https://arxiv.org/abs/2610.17711v1>, doi:10.48550/arXiv.2610.17711. ↑2, 4
- [Rud62] Walter Rudin. *Fourier Analysis on Groups*. Interscience Publishers, New York, 1962. ↑18

- [Sha79] Adi Shamir. How to share a secret. *Communications of the Association for Computing Machinery*, 22(11):612–613, November 1979. doi:[10.1145/359168.359176](https://doi.org/10.1145/359168.359176). ↑1
- [SV19] Akshayaram Srinivasan and Prashant Nalini Vasudevan. Leakage resilient secret sharing and applications. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019, Part II*, volume 11693 of *Lecture Notes in Computer Science*, pages 480–509, Santa Barbara, CA, USA, August 18–22, 2019. Springer, Cham, Switzerland. doi:[10.1007/978-3-030-26951-7_17](https://doi.org/10.1007/978-3-030-26951-7_17). ↑5
- [SW26] Yihang Sun and Mary Wootters. Improved local leakage resilience of shamir secret sharing and worst-case optimal polynomial intersection, 2026. URL: <https://arxiv.org/abs/2610.12357>, arXiv:2610.12357. ↑2, 4
- [TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. doi:[10.1017/CB09780511755149](https://doi.org/10.1017/CB09780511755149). ↑9