Connecting Cryptanalytic Distinguishers through Divisibility

Hosein Hadipour[1]

Ruhr University Bochum, Bochum, Germany
hosein.hadipour@rub.de

Abstract. Integral distinguishers can reveal more than a zero-sum property: the number of ones in an output component may be divisible by a high power of 2, rather than merely being even. To exploit this divisibility, we relate linear, integral, differential-linear, and restricted differential sums through two-way conversions with explicit divisibility guarantees and data complexity, recovering established links as special cases. We derive exact random-permutation probabilities and rigorous bounds to quantify the distinguishing strength while accounting for dependencies between output conditions. To discover these properties, we develop automated MILP/SAT models that use complete-round Walsh coefficients or combine Walsh and ultrametric representations to obtain stronger bounds. These weighted models prove finite divisibility even when compatible trails exist, beyond what trail-existence tests alone can establish. For differential-linear sums, we show that combining differential and linear propagation with local zeros in a differential-linear connectivity table (DLCT) can give false zero-correlation claims. We address this limitation with exact mix-basis models that establish zero correlation and finite divisibility. To complement these symbolic methods, we provide a black-box algorithm using exact histograms and the fast Walsh–Hadamard transform to search all output masks for selected input conditions and identify zeros and finite divisibility properties. Our MILP/SAT models significantly improve the divisibility bounds reported by Beyne and Verbauwhede in 674 cases across the SIMON family, as well as for PRESENT, without trail enumeration and at unchanged data complexity. For example, we raise the bounds from 2^5 to 2^{11} on 9-round SIMON32 and from 2^6 to 2^{22} on 10-round SIMON48. To our knowledge, we are the first to prove higher divisibility for 8-round SPECK and for BEANIE, with new bounds for GIFT-64 and Ascon and exact counts for SKINNY-64. Our black-box searches reveal new properties on RC5, ChiLow-32, KeeLoq, and KATAN32, including output-bit combinations missed by single-bit searches, and lead us to introduce probabilistic divisibility: conditions that hold more often than for a random permutation.

Keywords: Integral cryptanalysis, Linear cryptanalysis, Differential-linear cryptanalysis, Divisibility, Automated cryptanalysis

1 Introduction

Classical integral distinguishers exploit predictable XOR-sums of selected output bits over structured sets of plaintexts [DKR97,KW02]. Todo introduced the division property, which derives integral distinguishers by tracking which products of state bits are guaranteed to sum to zero over the input structure [Tod15]. These are parity constraints over $\mathbb{F}_2$. For one output component, a zero-sum property says only that its number of ones, its one-count, is even. Divisibility by a higher power of 2 constrains this count further: 2^t fixes its lowest t binary digits to zero, not only its parity.

The same divisibility information appears elsewhere: in the integer sums underlying linear and differential-linear (DL) correlations, and in the counts underlying differential probabilities. Connections between these sums and counts offer a way to exploit this information: a bound established in a representation that is easier to analyze may yield a stronger property for another distinguisher. Such a transfer must account for the different input sets and normalizations.

Connections between Distinguishers. Bogdanov et al. linked zero-correlation linear (ZC) mask families to exactly balanced integrals [BLNW12, Proposition 1]. Sun et al. showed that a nontrivial ZC approximation of a block cipher implies an integral distinguisher [SLR$^+$15, Theorem 3]. They also connected impossible differentials (IDs) to ZC and integral properties through dual cipher structures [SLR$^+$15, Sections 3, 5]. Beyne described cryptanalytic properties through different representations of a linear operator [Bey21]. At ASIACRYPT 2024, Beyne and Verbauwhede studied finite integral divisibility and its relation to linear correlations [BV24]. At CRYPTO 2025, Hu et al. developed mix-basis descriptions and applied them to DL distinguishers [HZC$^+$25].

Questions We Address. These results motivate the following questions: how can divisibility of full-codebook linear or DL sums establish stronger integral properties with data complexity below the full codebook? To prove divisibility of the one-count on a chosen plaintext structure, which full-codebook sums must we bound? Conversely, what do sums over these plaintext structures tell us about the full-codebook sums? How do such conversions extend to differential counts when both the plaintexts and their output differences are restricted?

How unlikely are the resulting properties for a random permutation at the same data complexity? Answering this question requires accounting for dependencies between tested conditions and for candidate selection during a search.

Finally, these mathematical and statistical questions lead to a computational one: how can we exploit the connections to discover and prove stronger properties at unchanged data complexity? Different models of the same sum can yield different bounds because they account for different cancellations between trail contributions. We therefore seek methods that account for these cancellations, extract divisibility information even when nonzero trails exist, and search over linear combinations of output bits.

Our Approach and Contributions. We address these questions through explicit forward and inverse conversions between linear, integral, DL, and restricted differential sums. We use the conversions to guide automated and black-box

searches for stronger properties at the same data complexity, and derive random-permutation bounds to assess their distinguishing strength.

- **Conversion theorems.** We prove explicit forward and inverse conversions between full-codebook sums and sums over affine plaintext structures. For an n-bit permutation, a common linear-sum divisibility bound 2^t on a d-dimensional input-mask space gives signed integral divisibility by 2^{t-d} on every orthogonal coset, with data complexity 2^{n-d}, for $t \geq d$. For DL sums, the ordinary DL coefficient must also be bounded. Without this bound, the guarantee concerns differences between coset sums. Our two-sided conversions additionally handle simultaneous restrictions on plaintexts and output values or differences. The ZC links of Bogdanov et al. [BLNW12, Proposition 1] and Sun et al. [SLR+15, Corollary 4] follow as special cases (Subsection 3.4). We also extend the linear conversion to odd prime fields using additive characters.
- **Random-permutation bounds.** We derive exact probabilities and rigorous bounds for the distinguishing strength of the resulting divisibility properties, covering families of output masks, cosets, and differential and DL tests. These laws account for the dependencies between output masks and between cosets, so individual probabilities cannot simply be multiplied. We also bound the probability that any candidate in a large search passes by chance, which is higher than for a single test fixed in advance.
- **Automated search for divisibility distinguishers.** We develop weighted mixed-integer linear programming (MILP) models that prove finite divisibility even when compatible trails exist and trail-existence-based ZC searches cannot establish zero correlation. We strengthen these bounds through two constructions: a round-condensed (RC) linear model that evaluates a complete round before assigning a weight, and a Joint model that combines Walsh and ultrametric representations in one optimization.
- **Sound search for ZCDL and finite DL divisibility.** We give counterexamples showing that a local DL connectivity table (DLCT) entry alone cannot establish whether the complete DL sum is zero or nonzero when combined with differential and linear propagation. Our exact mix-basis DL models retain the discarded information through mask–difference pairs. Their Boolean satisfiability (SAT) models soundly identify zero-correlation differential-linear (ZCDL) properties, while their weighted MILP models prove finite DL divisibility (Subsection 5.3).
- **Black-box discovery.** We develop a multipurpose black-box algorithm using exact histograms and the fast Walsh–Hadamard transform (FWHT). For selected input conditions, it screens all linear output masks and identifies finite divisibility, ZC and ZCDL properties, IDs, and constant-parity integrals. Unlike parity-kernel searches [HKK+26], it also recovers exact counts. Our search for deterministic divisibility also leads us to introduce *probabilistic divisibility*: conditions that hold for only some keys or input offsets, but more often than for a random permutation. (Probabilistic zero-sum integrals are the special case of divisibility by 2.) Predicting these probabilities an-

alytically and developing tools to search for strong probabilistic divisibility distinguishers remain open problems (Subsection F.13).

– **Applications.** We apply these tools to block ciphers and permutations with substitution–permutation network (SPN), Feistel, addition–rotation–XOR (ARX), and nonlinear shift-register structures. Without trail enumeration, we significantly improve the one-count bounds reported by Beyne and Verbauwhede [BV24] in 674 cases across all five SIMON block sizes, at unchanged data complexity. These comprise 290 finite-bound improvements (Table 13) and 384 exact-balance results (Table 14). For example, we raise the proved divisibility from 2^5 to 2^{11} on 9-round SIMON32 and from 2^6 to 2^{22} on 10-round SIMON48 (Table 1). We also improve their proved bounds on PRESENT (Table 29).

  On the 8-round SPECK structures of [HKK+26,WHG26], the longest known for this family, we are the first to prove more than a zero sum: divisibility by 2^3 for SPECK32 (Table 21), 2^8 for SPECK48 (Table 24), and the exact one-count 2^{124} for SPECK128 at data complexity 2^{125} (Table 25).

  The finite-divisibility results also extend to S-box-based primitives. For 5-round BEANIE permutations, we prove divisibility by 2^{10} and 2^{12} for every key and tweak for the first time (Table 26). For the 5-round Ascon permutation, we prove divisibility by 2^{60} (Table 34), with further high finite bounds for PRESENT (Table 32) and GIFT-64 (Table 35). These are applications of our weighted models in cases where compatible trails remain. For 11-round SKINNY-64, support proofs and conversion instead strengthen a recent zero-sum property [ZYS+26] to the exact one-count 2^{59} at unchanged data complexity 2^{60} (Table 39). These stronger divisibility properties can give sharper wrong-key filters in key recovery [BV24, Section 7.4].

  Our black-box searches also yield properties not established by our symbolic models. To our knowledge they give the first reported integral distinguisher for 2.5-round RC5-16 and the first divisibility beyond evenness for 6-round BEANIE, by 4 on 1,000 tested configurations (Table 44 and Table 3). On ChiLow-32, KATAN32 and KeeLoq they find one-counts divisible by 2^3 to 2^{14} on all 1,000 tested keys, and RC5-16 differential counts divisible by 2^{10} and 2^{11} (Table 4 and Table 42). Many of these properties hold for combinations of output bits, which searches restricted to individual bits cannot find. Our black-box method also proves resistance to all-key constant-parity integral distinguishers using plaintext structures where one input bit is fixed and all others vary. For the 32-bit permutations in Table 51, complete sweeps provide counterexamples ruling out these properties, including zero sums, for every nonzero output mask. Section 6 gives the exact scope of each result and separates proofs from empirical observations.

Our code and Lean formalization are available at `github.com/hadipourh/divisible`. The formalization provides machine-checked proofs of the core conversion theorems, random-permutation laws, and general correctness theorems for our models.

Organization. Section 2 defines the quantities and distinguishing experiment. Section 3 gives the conversions, and Section 4 gives the random-permutation analysis. Section 5 develops the automated and black-box searches, and Section 6 presents applications.

2 Preliminaries

2.1 Notation and Definitions

In the binary setting, let $V = \mathbb{F}_2^n$, where n is the block size. Vectors, masks, and differences belong to V, and $\alpha \cdot x = \sum_{i=0}^{n-1} \alpha_i x_i$ denotes their dot product over $\mathbb{F}_2$. Addition in V is bitwise XOR. We write $+$ in vector-space formulas and $\oplus$ for explicit bit parities and cipher operations. The symbols $\lll$ and $\ggg$ denote cyclic word rotations to the left and right. We write $U \leq V$ when U is a subspace and reserve $\subseteq$ for ordinary set inclusion. The indicator $\mathbf{1}_S(x)$ equals 1 if $x \in S$ and 0 otherwise. Similarly, $\mathbf{1}[P]$ indicates whether a condition P holds. For $U \leq V$ of dimension d, let

$$A = U^{\perp} = \{x \in V : \alpha \cdot x = 0 \text{ for all } \alpha \in U\}.$$

Our plaintext structures are affine cosets $\mathcal{X} = a + A = \{a + z : z \in A\}$. Equivalently, they satisfy $\alpha_j \cdot x = \alpha_j \cdot a$ for a basis $\alpha_1, \ldots, \alpha_d$ of U. Each coset contains $M = 2^m$ plaintexts, where $m = \dim A = n - d$. Let $\mathcal{R}_A$ contain exactly one representative of each coset of A in V. When A is spanned by coordinates, we call $\mathcal{X}$ a coordinate cube and specify it by an active mask u through $A = u \wedge \mathbb{F}_2^n = \{u \wedge z : z \in \mathbb{F}_2^n\}$, where $\wedge$ is bitwise AND and $\mathrm{wt}(u)$ is the Hamming weight of u, so that $\dim A = \mathrm{wt}(u)$, rather than $\dim\langle u \rangle$. We index an n-bit state globally: x_0 is its least-significant bit, x_{n-1} its most-significant bit, and ciphertext bits are $c_0, \ldots, c_{n-1}$. The unit vector e_i selects bit i. The relation $v \preceq \beta$ means $v_i \leq \beta_i$ for every coordinate. We use $\mathcal{B} \subseteq V \setminus \{0\}$ for a family of output masks, which need not be a subspace.

A cipher is a family of permutations $\{E_K : K \in \mathcal{K}\}$. We first define each quantity for a fixed permutation E. Substituting E_K gives the fixed-key quantity. An *all-key bound* means that the same bound holds for every allowed key, not that the value of the sum is independent of the key. A bound proved for arbitrary round keys also holds for the tuples produced by a particular key schedule.

Definition 1 (Sums and output one-counts). *For masks $\alpha, \beta \in V$, a difference $\delta \in V$, and a structure $\mathcal{X} \subseteq V$, define*

$$L_{\alpha,\beta}(E) = \sum_{x \in V} (-1)^{\alpha \cdot x + \beta \cdot E(x)}, \qquad I_{\beta}(\mathcal{X}; E) = \sum_{x \in \mathcal{X}} (-1)^{\beta \cdot E(x)},$$
$$N_1(\mathcal{X}, \beta; E) = \#\{x \in \mathcal{X} : \beta \cdot E(x) = 1\}, \qquad D_\delta E(x) = E(x) + E(x + \delta).$$

The DL quantities are

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = L_{\alpha,\beta}(D_\delta E), \qquad I^{\mathrm{DL}}_{\delta,\beta}(\mathcal{X}; E) = I_{\beta}(\mathcal{X}; D_\delta E).$$

We similarly define $N_1^{\mathrm{DL}}(\mathcal{X}, \delta, \beta; E) = N_1(\mathcal{X}, \beta; D_\delta E)$*. This counts base inputs* $x \in \mathcal{X}$ *for which* $\beta \cdot E(x) \neq \beta \cdot E(x + \delta)$*, not unordered pairs.*

The normalized linear correlation is $\mathrm{Cr}_E(\alpha, \beta) = 2^{-n} L_{\alpha,\beta}(E)$. The ordinary normalized DL correlation is $2^{-n} L^{\mathrm{DL}}_{0,\delta,\beta}(E)$. Over $\mathbb{F}_2$, L and I denote integer signed sums, not normalized correlations. For $\beta \neq 0$, bijectivity gives $L_{0,\beta}(E) = 0$. There is no analogous automatic cancellation for $L^{\mathrm{DL}}_{0,\delta,\beta}(E)$.

For a nonzero integer z, its 2-adic valuation $v_2(z)$ is the largest integer t for which $2^t \mid z$. Set $v_2(0) = \infty$ and extend to dyadic rationals by $v_2(z/2^h) = v_2(z) - h$. We use

$$v_2(xy) = v_2(x) + v_2(y), \qquad v_2(x + y) \geq \min\{v_2(x), v_2(y)\}.$$

The second inequality is the ultrametric inequality. A common divisibility bound therefore survives addition even if the signs are unknown. The normalization changes valuations by the block size: $v_2(\mathrm{Cr}_E(\alpha, \beta)) = v_2(L_{\alpha,\beta}(E)) - n$. We distinguish thresholds by writing t_L for linear sums, t_I for signed integral sums, and t_N for one-counts.

2.2 Divisibility-Based Distinguishers

Our conversions concern signed sums, whereas we report integral properties through one-counts. For a plaintext set $\mathcal{X}$ of size M, these quantities satisfy

$$I_\beta(\mathcal{X}; E) = M - 2N_1(\mathcal{X}, \beta; E). \tag{1}$$

For $M = 2^m$ with $m \geq 1$ and $t_I \geq 1$, a bound $v_2(I_\beta(\mathcal{X}; E)) \geq t_I$ implies

$$v_2(N_1(\mathcal{X}, \beta; E)) \geq \min\{m - 1, t_I - 1\}. \tag{2}$$

For $1 \leq t_N \leq m - 1$, the condition $2^{t_N} \mid N_1$ is equivalent to $2^{t_N+1} \mid I$. This equivalence need not hold for $t_N \geq m$. In particular, $I = 0$ means $N_1 = M/2$: exact balance between zeros and ones. Conversely, divisibility of N_1 by $M/2$ alone allows $N_1 = 0, M/2, M$. It does not prove exact balance.

For $m \geq 2$, an ordinary zero-sum integral on one component means N_1 is even, equivalently $v_2(I) \geq 2$. Higher divisibility strengthens this parity constraint, but does not by itself establish exact balance. Section D illustrates these distinctions, including key-varying valuations in Subsection D.5.

To evaluate these properties as distinguishers, fix the plaintext structures, masks, and thresholds defining an event $\mathcal{D}$ before testing. The oracle is either E_K for a uniformly sampled secret key held fixed, or a uniformly random permutation Π on V. Write $p_1 = \Pr_K[\mathcal{D}(E_K)]$ and $p_0 = \Pr_\Pi[\mathcal{D}(\Pi)]$. An all-key proof gives $p_1 = 1$ and distinguishing advantage $1 - p_0$. Computing one signed sum on M plaintexts requires M oracle calls and a counter. For a DL statistic on $\mathcal{X} = a + A$, we query all plaintexts in $\mathcal{X} \cup (\mathcal{X} + \delta)$: M distinct plaintexts if $\delta \in A$, and $2M$ otherwise.

3 Conversion Theorems: From Spectra to Structured Counts

The common ingredient is the character expansion of a coset indicator.

Lemma 1 (Coset indicator). *Let $U \le V$ have dimension d and $A = U^\perp$. For all $a, x \in V$,*

$$\mathbf{1}_{a+A}(x) = 2^{-d} \sum_{\alpha \in U} (-1)^{\alpha \cdot (x+a)}. \tag{3}$$

Proof. The sum is 2^d when $x + a \in U^\perp$. Otherwise, choose $\alpha_0 \in U$ with $\alpha_0 \cdot (x + a) = 1$; the translation $\alpha \mapsto \alpha + \alpha_0$ pairs each summand with its negative.

The factor 2^{-d} is the source of the guaranteed divisibility loss. Subsection E.13 gives the coordinate-family screening argument. Subsection A.1 extends the conversion to higher-order derivatives.

3.1 Linear–Integral Conversion

Theorem 1 (Two-way mask–coset transform). *Let $U \le V$ have dimension d, with $A = U^\perp$, and fix $\beta \neq 0$. For every $a \in V$,*

$$I_\beta(a + A; E) = 2^{-d} \sum_{\alpha \in U} (-1)^{\alpha \cdot a} L_{\alpha,\beta}(E). \tag{4}$$

Conversely, for every $\alpha \in U$,

$$L_{\alpha,\beta}(E) = \sum_{a \in \mathcal{R}_A} (-1)^{\alpha \cdot a} I_\beta(a + A; E). \tag{5}$$

If $v_2(L_{\alpha,\beta}(E)) \ge t_L$ for all $\alpha \in U$, then $v_2(I_\beta(a+A; E)) \ge t_L - d$ on every coset. Conversely, if $v_2(I_\beta(a + A; E)) \ge t_I$ on every coset, then $v_2(L_{\alpha,\beta}(E)) \ge t_I$ for all $\alpha \in U$.

Proof. Multiply the identity in Lemma 1 by $(-1)^{\beta \cdot E(x)}$ and sum over x. Interchanging the two finite sums gives Equation 4. For the inverse, partition the full-codebook sum into cosets. On $a + A$, the character $(-1)^{\alpha \cdot x}$ equals $(-1)^{\alpha \cdot a}$ because $\alpha \in A^\perp$. The contribution of that coset is its signed sum times this constant sign. The valuation claims follow from the ultrametric inequality.

The theorem quantifies a loss in the guaranteed bound, not the exact change in valuation. With

$$\nu_L = \min_{\alpha \in U} v_2(L_{\alpha,\beta}(E)), \qquad \nu_I = \min_{a \in \mathcal{R}_A} v_2(I_\beta(a + A; E)),$$

we obtain $\nu_L - d \le \nu_I \le \nu_L$. Cancellation can make the forward bound strict. If $\nu_L = \infty$, both families vanish.

All Keys and Output-Mask Families. For each $\beta \in \mathcal{B}$, suppose $v_2(L_{\alpha,\beta}(E_K)) \geq t_L(\beta)$ for every key K and every nonzero $\alpha \in U$. Since $L_{0,\beta}(E_K) = 0$, Theorem 1 gives $v_2(I_\beta(a + A; E_K)) \geq t_L(\beta) - d$ for every key and every coset. We call the Cartesian product $U \times \mathcal{B}$ a mask rectangle. The divisibility bound $t_L(\beta)$ may differ between output masks.

For a fixed output mask, assume $d < n$ and $t_L \geq d + 1$. The induced one-count threshold is $t_N = \min\{n - d - 1, t_L - d - 1\}$, at data complexity 2^{n-d}. Enlarging U reduces the data complexity but requires a common bound over more input masks. Conversely, shrinking U can improve the divisibility bound at the cost of higher data complexity.

Corollary 1 (Lossless halfspace conversion). *For a permutation E, nonzero masks α, β, and the affine halfspace $\mathcal{X}_b = \{x \in V : \alpha \cdot x = b\}$ with $b \in \mathbb{F}_2$,*

$$I_\beta(\mathcal{X}_b; E) = \frac{(-1)^b}{2} L_{\alpha,\beta}(E), \qquad N_1(\mathcal{X}_b, \beta; E) = 2^{n-2} - \frac{(-1)^b}{4} L_{\alpha,\beta}(E). \tag{6}$$

In particular, $v_2(I_\beta(\mathcal{X}_b; E)) = v_2(L_{\alpha,\beta}(E)) - 1$.

Proof. Apply the forward transform with $U = \{0, \alpha\}$ and use $L_{0,\beta}(E) = 0$. Then apply Equation 1.

For an affine halfspace, one linear sum per output mask determines the signed integral sum exactly through $I = \pm L/2$. Any gap between the resulting valuation bound and the true valuation therefore comes from the bound on that linear sum. In higher codimension, taking the minimum valuation of several coefficients can also miss cancellations, although the transform remains exact and invertible. The example in Subsection D.2 illustrates a finite halfspace property through exact enumeration on 4,096 tested keys (Figure 2).

3.2 DL Conversion

The same transform applies to the derivative, but its zero-mask coefficient need not vanish.

Theorem 2 (DL cosets and their contrasts). *Fix $\delta \in V$ and $\beta \neq 0$. For $A = U^\perp$ with $\dim U = d$,*

$$I^{\mathrm{DL}}_{\delta,\beta}(a + A; E) = 2^{-d} \sum_{\alpha \in U} (-1)^{\alpha \cdot a} L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E), \tag{7}$$

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = \sum_{a \in \mathcal{R}_A} (-1)^{\alpha \cdot a} I^{\mathrm{DL}}_{\delta,\beta}(a + A; E) \quad (\alpha \in U). \tag{8}$$

If $v_2(L^{\mathrm{DL}}_{0,\delta,\beta}(E)) \geq t_0$ and $v_2(L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E)) \geq t_1$ for all $\alpha \in U \setminus \{0\}$, every coset sum has valuation at least $\min\{t_0, t_1\} - d$. Conversely, a common valuation bound s on all coset sums also holds for every coefficient indexed by U. Thus

$\nu_L - d \le \nu_I \le \nu_L$, where ν_L and ν_I are the minimum DL valuations over U and the cosets, respectively. The nonzero-mask bound alone gives, for any two cosets,

$$v_2\big(I^{\mathrm{DL}}_{\delta,\beta}(a+A;E) - I^{\mathrm{DL}}_{\delta,\beta}(b+A;E)\big) \ge t_1 - d + 1. \tag{9}$$

If $L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = 0$ for all $\alpha \in U \setminus \{0\}$, every coset sum equals $2^{-d}L^{\mathrm{DL}}_{0,\delta,\beta}(E)$.

Proof. Use Equation 3 with $(-1)^{\beta \cdot D_\delta E(x)}$. The inverse follows by partitioning the domain into cosets. The valuation bounds follow from the ultrametric inequality, including the $\alpha = 0$ term. Subtracting the forward identities for a and b cancels that term. Every remaining sign difference is 0 or ± 2, proving Equation 9. If the remaining coefficients vanish, only the zero-mask contribution survives.

For nonzero $\alpha \in U$, Equation 8 expresses the DL coefficient as a signed combination of coset sums, without requiring a bound on the zero-mask coefficient. If the hypotheses hold for every key, the guarantees hold for every key and output mask considered; thresholds may depend on the masks. Equal coset sums need not be zero or highly divisible. Subsection D.3 illustrates this distinction.

For $\delta \neq 0$, replacing x by $x + \delta$ gives

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = (-1)^{\alpha \cdot \delta} L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E).$$

Thus $\alpha \cdot \delta = 1$ forces zero even for every random permutation and cannot distinguish the cipher.

If $\delta \in A$, N^{DL}_1 counts each unordered pair twice and requires M distinct plaintexts. Otherwise, each pair joins two disjoint cosets, requiring $2M$ distinct plaintexts. Proposition 9 gives the corresponding random-permutation distributions. Proposition 7 extends the conversion to higher-order derivatives, retaining the zero-mask qualification and stating the same-coset condition.

3.3 Two-Sided Affine Differential Counts

Let $U, W \le V$ have dimensions d, e and let $A = U^\perp$, $B = W^\perp$. For a base-plaintext coset $a + A$ and output-difference coset $b + B$, define

$$N^{A,B}_{a,b}(\delta; E) = \#\{x \in a + A : D_\delta E(x) \in b + B\}. \tag{10}$$

This is an ordered count indexed by x, not the count of unordered pairs.

Theorem 3 (Two-sided count transform). *With this notation,*

$$N^{A,B}_{a,b}(\delta; E) = 2^{-d-e} \sum_{\alpha \in U} \sum_{\beta \in W} (-1)^{\alpha \cdot a + \beta \cdot b} L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E), \tag{11}$$

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = \sum_{a \in \mathcal{R}_A} \sum_{b \in \mathcal{R}_B} (-1)^{\alpha \cdot a + \beta \cdot b} N^{A,B}_{a,b}(\delta; E). \tag{12}$$

If $v_2(L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E)) \ge t$ for all $(\alpha, \beta) \in U \times W$, then $v_2(N^{A,B}_{a,b}(\delta; E)) \ge t - d - e$ for every pair of cosets. Conversely, if $v_2(N^{A,B}_{a,b}(\delta; E)) \ge s$ for every pair of cosets, then $v_2(L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E)) \ge s$ throughout $U \times W$.

Proof. Expand both indicators in $\sum_x \mathbf{1}_{a+A}(x)\mathbf{1}_{b+B}(D_\delta E(x))$. Their product supplies 2^{-d-e} and the two characters in Equation 11. For the inverse, partition inputs by $a + A$ and their derivative values by $b + B$. Both selected characters are constant on each part. Apply the ultrametric inequality.

All masks in the rectangle matter, including $\beta = 0$. Indeed $L^{\mathrm{DL}}_{0,\delta,0} = 2^n$, whereas $L^{\mathrm{DL}}_{\alpha,\delta,0} = 0$ for $\alpha \neq 0$. These terms encode the total population and cannot be discarded. If $\delta \in A \setminus \{0\}$, the involution $x \mapsto x + \delta$ makes Equation 10 even. Dividing by two gives the unordered count and decreases its finite valuation by one. This first factor of 2 is universal.

The same two-sided calculation also applies to ordinary input/output restrictions, without a derivative. Define the affine transition count

$$T^{A,B}_{a,b}(E) = \#\{x \in a + A : E(x) \in b + B\}.$$

Replacing the derivative by E in the proof of Theorem 3 gives

$$T^{A,B}_{a,b}(E) = 2^{-d-e} \sum_{\alpha \in U,\, \beta \in W} (-1)^{\alpha\cdot a+\beta\cdot b} L_{\alpha,\beta}(E), \tag{13}$$

$$L_{\alpha,\beta}(E) = \sum_{a \in \mathcal{R}_A,\, b \in \mathcal{R}_B} (-1)^{\alpha\cdot a+\beta\cdot b} T^{A,B}_{a,b}(E) \quad (\alpha \in U,\ \beta \in W).$$

The same $d + e$-unit forward loss and lossless inverse lower bound follow. Here bijectivity fixes both marginals: $\sum_b T^{A,B}_{a,b} = |A|$ and $\sum_a T^{A,B}_{a,b} = |B|$. The differential table has the first marginal but need not have the second, since a derivative need not be bijective. Thus both two-sided transforms arise from the same character calculation, while their structural constraints differ. The distinction is useful when deciding which terms vanish identically and which population terms must remain in a divisibility argument.

3.4 Special Cases and Comparison with Prior Work

We recover the classical ZC–integral links by setting the linear sums on U to zero in Theorem 1. For a fixed nonzero output mask, this is equivalent to exact balance on every coset of $U^\perp$, not merely one. Applying this equivalence to every nonzero mask in an output-mask space recovers Bogdanov et al.'s Proposition 1 [BLNW12]. The mask families form a Cartesian product (their *detached masks*), and their integral property requires uniformity of the selected output coordinates for all fixed input values, not merely a zero-sum. Keeping a single nonzero output mask instead recovers Sun et al.'s Corollary 4 [SLR+15] for arbitrary U. Taking $U = \langle\alpha\rangle$ for $\alpha \neq 0$ then gives their Theorem 3: one ZC approximation of a permutation suffices. The underlying correlation identity already holds beyond ZC in their Lemma 1 [SLR+15]. The ZC results are the vanishing-sum cases of our conversion; its finite-divisibility bounds additionally track the dimension-dependent loss.

We similarly recover the difference distribution table (DDT)–DLCT Walsh relation [BDKW19,HDE24] from Theorem 3. Taking $A = V$ and $B = \{0\}$ leaves

plaintexts unrestricted and fixes the output difference to γ. Using our unnormalized DL sums, we obtain

$$\mathrm{DDT}_E(\delta,\gamma) = 2^{-n}\sum_{\beta\in V}(-1)^{\beta\cdot\gamma}L^{\mathrm{DL}}_{0,\delta,\beta}(E). \tag{14}$$

An ID requires this entire sum to vanish, not just one DL coefficient. Without requiring vanishing, our theorem instead transfers divisibility bounds to differential counts. Table 42 gives finite RC5-16 examples.

The closest finite-divisibility comparison is Beyne and Verbauwhede's correlation/integral transition-matrix relation [BV24, Theorem 6.2]. It bounds entries after a change of basis using coordinatewise support conditions and Hamming-weight factors. Our theorems instead index additive mask spaces and affine cosets, with losses given by dimensions. Our theorems give exact inverse identities and divisibility bounds for the specified sums and counts. Hu et al. use mix-basis representations to search for multiple-of-2^t properties and evaluate exact DL correlations [HZC+25, Sections 5–7]. Our affine formulation makes the operational consequences explicit: one mask-space bound covers every translated structure, halfspace conversion has exact valuation scaling, and the two-sided identities handle simultaneous input and output restrictions.

3.5 Extension to Prime Fields

The linear–integral conversion extends to prime fields by replacing binary signs with additive characters. Let p be prime and $V_p = \mathbb{F}_p^\ell$, with dimension ℓ over $\mathbb{F}_p$. Let $\psi(z) = \zeta_p^z$ for a primitive pth root of unity ζ_p, and define

$$L^{(p)}_{\alpha,\beta}(E) = \sum_{x\in V_p}\psi(\alpha\cdot x+\beta\cdot E(x)), \qquad I^{(p)}_\beta(\mathcal{X};E) = \sum_{x\in\mathcal{X}}\psi(\beta\cdot E(x)).$$

Let ord_p denote the p-adic valuation on $\mathbb{Q}(\zeta_p)$, normalized by $\mathrm{ord}_p(p)=1$, with $\mathrm{ord}_p(0)=\infty$.

Proposition 1 (Prime-field mask–coset transform). *Let E be a permutation on V_p, $U \le V_p$ have dimension d, $A = U^\perp$, and $\beta \neq 0$. For $a \in V_p$ and $\alpha \in U$,*

$$I^{(p)}_\beta(a+A;E) = p^{-d}\sum_{\alpha\in U}\psi(-\alpha\cdot a)L^{(p)}_{\alpha,\beta}(E), \tag{15}$$

$$L^{(p)}_{\alpha,\beta}(E) = \sum_{a\in\mathcal{R}_A}\psi(\alpha\cdot a)I^{(p)}_\beta(a+A;E). \tag{16}$$

Let T_L and T_I be the minimum valuations of these linear and integral sums over U and $\mathcal{R}_A$, respectively. Then $T_L - d \le T_I \le T_L$.

Thus the conversion loses at most d valuation units at data complexity $p^{\ell-d}$. These valuations can be fractional and concern character sums, not integer

one-counts. This additive-character extension differs from the multiplicative-character emphasis of characteristic-p integral cryptanalysis [BV25]. Section B gives the proof and an illustration. Subsection B.2 and Subsection B.3 give the AES-prime application and its random-permutation law.

4 Random-Permutation Bounds

Given an all-key divisibility bound, the conversion yields a property that holds for every key. Its distinguishing strength depends on how often a uniformly random permutation satisfies the same property. We first derive this probability for fixed masks, plaintext structures, and thresholds, then account for selecting these parameters through a search.

4.1 Single-Mask Probabilities

Let $Q = 2^n$ and $M = |\mathcal{X}|$. For a uniformly random permutation Π, the image $\Pi(\mathcal{X})$ is a uniformly random M-subset of V. A nonzero mask β partitions V into two sets of size $Q/2$ according to $\beta \cdot y$. Consequently, the one-count is hypergeometric, not binomial.

Proposition 2 (Exact one-mask integral law). *For $\beta \neq 0$ and integers $0 \leq j \leq M$ and $t_I \geq 0$,*

$$\Pr[N_1(\mathcal{X}, \beta; \Pi) = j] = \frac{\binom{Q/2}{j}\binom{Q/2}{M-j}}{\binom{Q}{M}}, \tag{17}$$

$$\Pr[v_2(I_\beta(\mathcal{X}; \Pi)) \geq t_I] = \binom{Q}{M}^{-1} \sum_{\substack{0 \leq j \leq M \\ 2^{t_I} \mid M-2j}} \binom{Q/2}{j}\binom{Q/2}{M-j}. \tag{18}$$

We set $\binom{a}{b} = 0$ for $b \notin \{0, \ldots, a\}$. For an integer count threshold $t_N \geq 0$, replace the divisibility condition by $2^{t_N} \mid j$.

Proof. Choose j images with $\beta \cdot y = 1$ and $M - j$ with $\beta \cdot y = 0$. All $\binom{Q}{M}$ possible images of the structure are equally likely. Use Equation 1 and sum over the allowed counts.

The same hypergeometric distribution appears in wrong-key integral filtering [BLP25]. Here it evaluates divisibility on a fixed plaintext structure under a uniformly random permutation, accounting for sampling without replacement rather than assuming independent output values.

4.2 Certified Residue Bounds

Direct evaluation of Equation 18 can be expensive for large plaintext structures. The following bound approximates this probability by $2^{-(t_I-1)}$ with an explicit error.

Proposition 3 (Certified residue approximation). *Fix $\beta \neq 0$, suppose M is even, and let $t_I \geq 1$ be an integer. Set $q = 2^{t_I - 1}$ and*

$$p_{\mathrm{ctr}}(Q, M) = \frac{\binom{Q/2}{M/2}^2}{\binom{Q}{M}}.$$

Here $p_{\mathrm{ctr}} = \Pr[N_1 = M/2]$ is the probability of exact balance. Then

$$\left|\Pr[v_2(I_\beta(\mathcal{X}; \Pi)) \geq t_I] - q^{-1}\right| \leq p_{\mathrm{ctr}}(Q, M). \tag{19}$$

Proof. The event is $N_1 \equiv M/2 \pmod{q}$. The centered variable $X = N_1 - M/2$ has a unimodal distribution with mode zero and probability p_{ctr} at that mode. For such a law, the mass of the residue containing the mode differs from the mean residue mass $1/q$ by at most the modal probability. Subsection C.1 proves this last assertion by telescoping the two monotone tails.

Intersect the resulting interval with $[0, 1]$. Exact balance always satisfies the divisibility condition, so p_{ctr} also gives a lower bound on its probability. When $2^{t_I} > M$, exact balance is the entire event. The probability therefore need not keep halving as the threshold increases. The approximation is informative when its error is small relative to the probability being estimated. Figure 5 compares the bound with the exact law. Subsection C.4 gives rigorous numerical evaluations.

For fixed nonzero input and output masks, the random-permutation distribution follows from a halfspace integral. The identity $L_{\alpha,\beta}(\Pi) = 2I_\beta(\{x : \alpha \cdot x = 0\}; \Pi)$ converts a linear-sum threshold to a halfspace signed-sum threshold one unit lower. This gives the random-permutation baseline for a fixed linear coefficient without assuming independent signs over the full codebook.

4.3 Families of Output Masks

The masks in $\mathcal{B}$ are evaluated on the same ciphertext set, so their conditions must be assessed jointly. Let $b_1, \ldots, b_h$ be a basis of their span. The map $y \mapsto (b_1 \cdot y, \ldots, b_h \cdot y)$ partitions V into 2^h classes of size $Q/2^h$. For $z \in \mathbb{F}_2^h$, let N_z count ciphertexts in $\Pi(\mathcal{X})$ belonging to class z. These counts follow the multivariate hypergeometric distribution: for nonnegative integers n_z,

$$\Pr[(N_z)_z = (n_z)_z] = \binom{Q}{M}^{-1} \prod_{z \in \mathbb{F}_2^h} \binom{Q/2^h}{n_z}, \qquad \sum_z n_z = M. \tag{20}$$

For class counts $(n_z)_z$ and $\beta = \sum_j \lambda_j b_j$, the signed sum is $\sum_z (-1)^{\lambda \cdot z} n_z$. Summing Equation 20 over count vectors satisfying all prescribed divisibility conditions gives the exact joint probability.

The product counts output subsets realizing the specified class counts, not independent probabilities for individual masks. For linearly independent β_1, β_2, the masks $\beta_1, \beta_2, \beta_1 + \beta_2$ have rank 2 and share the same four class counts. The

ToySPN examples in Subsection D.4 show that joint evaluation can give a much smaller probability than the elementary upper bound

$$\Pr[\text{all conditions}] \le \min_{\beta \in \mathcal{B}} \Pr[\text{condition for } \beta].$$

The rank h, rather than the number of masks alone, controls the number of classes. Exact joint enumeration can still be combinatorially large even at modest rank. Probabilities for different cosets of one permutation cannot in general be multiplied either: their output sets are disjoint and therefore dependent (Remark 2). We use the joint law when feasible and the one-mask upper bounds otherwise.

4.4 Differential and DL Statistics

For $\delta \neq 0$, the images of the plaintext pairs $\{x, x+\delta\}$ under Π form a uniformly random pairing of all output values, a perfect matching. Each pair with difference $\gamma \neq 0$ contributes twice to $\mathrm{DDT}_\Pi(\delta,\gamma)$. Thus $\mathrm{DDT}_\Pi(\delta,\gamma)/2$ counts the pairs shared with the fixed matching $\{\{y, y+\gamma\} : y \in V\}$. Proposition 8 gives the exact distribution and proof. Divisibility tests include zero counts; requiring a nonzero count changes the probability.

For DL, fix $\mathcal{X} = a + A$, $M = |\mathcal{X}|$, and $\delta, \beta \neq 0$. Let $H_\mathcal{X}$ count the distinct unordered pairs used and κ the number of times each is counted:

$$(H_\mathcal{X}, \kappa) = \begin{cases} (M, 1), & \delta \notin A, \\ (M/2, 2), & \delta \in A. \end{cases}$$

Let J count the distinct pairs $\{x, x+\delta\}$ with $x \in \mathcal{X}$ for which $\beta \cdot \Pi(x) \neq \beta \cdot \Pi(x+\delta)$. Then $N_1^{\mathrm{DL}} = \kappa J$ and

$$I^{\mathrm{DL}}_{\delta,\beta}(\mathcal{X}; \Pi) = M - 2\kappa J. \tag{21}$$

Proposition 9 derives the distribution of J by assigning distinct outputs to the pairs. Together with κ, it accounts for sampling without replacement and for counting each pair twice when $\delta \in A$.

These results cover full-codebook DDT entries and one-mask DL sums on affine cosets. They do not give every joint distribution with restrictions on both inputs and output differences. Subsection C.5 evaluates our DL and DDT count tests.

5 Automated Search for Divisibility Distinguishers

To use the conversions, we must either prove divisibility bounds for the required mask families or compute the relevant sums exactly for selected cipher instances. We first develop linear models and their refinements, then combine linear and integral representations to strengthen the bounds. We also explain how to model DL sums correctly. Finally, we present a black-box algorithm that computes exact sums and counts for all linear output masks under a chosen input condition.

5.1 Linear and Round-Condensed Models

For a decomposition $E = F_r \circ \cdots \circ F_1$, matrix multiplication gives

$$L_{\alpha,\beta}(E) = 2^n \sum_{\alpha_1,\ldots,\alpha_{r-1}} \prod_{j=1}^{r} \mathsf{W}^{F_j}_{\alpha_j,\alpha_{j-1}}, \quad \mathsf{W}^F_{b,a} = 2^{-n} L_{a,b}(F), \tag{22}$$

Here $\alpha_0 = \alpha$, $\alpha_r = \beta$, and W^F is the normalized correlation matrix of F. Each product is a linear-trail contribution. Minimizing the sum of transition valuations over all supported trails, then adding n, gives a lower bound on $v_2(L_{\alpha,\beta}(E))$. A support-only SAT model checks whether any such trail exists: UNSAT proves ZC, whereas SAT does not prove nonzero correlation because trail contributions may cancel. Table 11 gives the local linear valuation table for the PRESENT S-box.

One Model for an Input Space. For a basis $u_1, \ldots, u_d$ of U, introduce binary variables λ_i and impose

$$\alpha_0 = \sum_{i=1}^{d} \lambda_i u_i \quad \text{over } \mathbb{F}_2, \qquad \sum_{i=1}^{d} \lambda_i \geq 1, \qquad \alpha_r = \beta.$$

The inequality is over the integers. The model minimizes over every nonzero input mask in U, giving a common divisibility bound in one optimization rather than a separate run for each mask. A certified objective lower bound therefore supplies the linear-sum bound required by Theorem 1. The zero input mask needs no constraint because $L_{0,\beta}(E) = 0$ for a permutation and $\beta \neq 0$. Subsection E.2 uses this model to search for input structures.

Why Contract a Round? A gate-level model can miss zeros or higher valuations caused by adding internal trail contributions. For each pair of input and output masks of a round, we instead sum these contributions before taking their valuation. We use the resulting exact round coefficient as one transition in the same linear MILP, giving the round-condensed (RC) model. Thus contributions within each round are combined exactly, but across rounds we still bound the complete sum using the minimum trail valuation.

Example 1 (A three-bit contraction). Let $G(x_0, x_1, x_2) = (x_0, x_1, x_2 + x_0x_1)$ and consider the two-gate block $F = G \circ G$. Exactly four intermediate masks give nonzero trails from e_2 to e_2 through F: $\gamma = e_2 + a_0e_0 + a_1e_1$, with $a_0, a_1 \in \mathbb{F}_2$. The corresponding coefficients of G have magnitude $1/2$. Since G is an involution, its normalized correlation matrix is symmetric, so each trail contributes $(\pm 1/2)^2 = 1/4$. Its valuation is $v_2(1/4) = -2$, giving only $v_2(L_{e_2,e_2}(F)) \geq 3 - 2 = 1$. Summing first gives $\mathsf{W}^F_{e_2,e_2} = 4 \cdot (1/4) = 1$, hence $L_{e_2,e_2}(F) = 8$ and $v_2(L_{e_2,e_2}(F)) = 3$. Indeed, F is the identity. Thus contraction raises the bound by 2 valuation units even though all contributions are positive: addition alone can increase the valuation. SIMON permits the same principle for an actual complete round without enumerating its gate-level trails.

Proposition 4 (Soundness of the linear model). *Suppose the model represents every supported trail for the covered input masks, with each transition cost at most its normalized coefficient's valuation. Let ℓ be a certified lower bound on its minimum total cost. Then $v_2(L_{\alpha,\beta}(E)) \geq n+\ell$ for every covered input mask. An infeasible model proves that every covered linear sum is zero. With exact transition valuations and fixed external masks, contracting consecutive transitions cannot decrease the minimum-trail valuation bound.*

Proof. Every supported trail has an assignment whose cost is at least ℓ and at most the valuation of its normalized contribution. Equation 22 and the ultrametric inequality give the bound; infeasibility excludes every supported trail. Contraction sums internal contributions before taking valuations, which cannot decrease their minimum; zero coefficients are removed. This comparison assumes otherwise identical constraints and concerns exact bounds, not time-limited solver outputs.

A feasible solution gives an upper bound on the minimum, not the lower bound ℓ required by Proposition 4. We use the certified objective lower bound, or the optimum after optimality is established. XOR round-key additions multiply Walsh coefficients by signs, preserving valuations and allowing bounds for arbitrary round keys.

SIMON's Complete-Round Relation. For SIMON, quadratic forms give complete-round support and valuations directly. Let $w = n/2$ and

$$F_k(x,y) = \big(y + f(x) + k, x\big), \qquad f(x) = ((x \lll 1) \wedge (x \lll 8)) + (x \lll 2).$$

Here x, y are the left and right words, with global bit indices $w+i, i$, respectively. Superscripts L, R denote their mask subvectors locally. For input mask $(a^{\mathrm{L}}, a^{\mathrm{R}})$ and output mask $(b^{\mathrm{L}}, b^{\mathrm{R}})$, summing over y forces $a^{\mathrm{R}} = b^{\mathrm{L}}$. Set $\gamma = a^{\mathrm{L}} + b^{\mathrm{R}} + (b^{\mathrm{L}} \ggg 2)$ and

$$\varphi_{b^{\mathrm{L}}}(x) = \sum_i (b^{\mathrm{L}})_i x_{i-1} x_{i-8},$$

with indices modulo w. Let $P_{b^{\mathrm{L}}}$ be the polar matrix of $\varphi_{b^{\mathrm{L}}}$, defined over $\mathbb{F}_2$ by

$$x^{\mathsf{T}} P_{b^{\mathrm{L}}} z = \varphi_{b^{\mathrm{L}}}(x+z) + \varphi_{b^{\mathrm{L}}}(x) + \varphi_{b^{\mathrm{L}}}(z).$$

For the SIMON word sizes $w \in \{16, 24, 32, 48, 64\}$, the support and nonzero-coefficient valuation are:

$$\mathsf{W}^{F_k}_{b,a} \neq 0 \iff a^{\mathrm{R}} = b^{\mathrm{L}},\ \gamma \in \operatorname{im} P_{b^{\mathrm{L}}}, \qquad v_2(\mathsf{W}^{F_k}_{b,a}) = -\tfrac{1}{2} \operatorname{rank} P_{b^{\mathrm{L}}}. \tag{23}$$

Subsection E.3 explains the support condition and the importance of the even-cycle structure. We use known quadratic Walsh theory; the contribution is its integration into the whole-subspace valuation model and conversion.

For each round j, we impose the support conditions in Equation 23 and assign cost $-\rho_j$, where ρ_j is half the polar rank. The objective is $n - \sum_j \rho_j$, as in

the linear model. The conversion then supplies the integral and count bounds. The MILP uses input and output mask bits for each round, an auxiliary vector z satisfying $P_{b^{\mathrm{L}}} z = \gamma$, binary products, and rank-counting variables. Products $h = uv$ use $h \le u$, $h \le v$, $h \ge u + v - 1$. A binary XOR $y = \sum_i x_i$ is encoded as $\sum_i x_i - y = 2c$ with a bounded integer c. The graph with edges $(i-1, i-8)$ for $(b^{\mathrm{L}})_i = 1$ consists of paths or an even cycle, allowing a finite-state recurrence to compute half the polar rank. The resulting formulation has $O(rw)$ variables and constraints. Thus, its size grows linearly with the number of rounds and the word size. For other ciphers, computing exact complete-round coefficients can be substantially harder.

An optional parity refinement considers all minimum-valuation trail contributions. After division by their common power of 2, these contributions are odd, so an even number of them proves one additional divisibility bit. We distinguish these refined bounds from those of the base RC model.

5.2 Joint Walsh–Ultrametric Models

RC strengthens valuation bounds within the Walsh representation. Joint combines Walsh signed sums with counts of states whose selected bits are all 1, as used in ultrametric integral cryptanalysis. It directly bounds output one-count divisibility for the chosen plaintext structure. Fixing the representations at component inputs and outputs gives one MILP with shared intermediate variables, rather than separately obtained bounds.

For a function $h : V \to \mathbb{Q}$, we define Walsh coordinates H_a and moment coordinates Z_u by

$$H_a = \sum_x h(x)(-1)^{a \cdot x}, \qquad Z_u = \sum_x h(x) x^u, \qquad x^u = \prod_{i : u_i = 1} x_i.$$

These sums and monomial products use ordinary arithmetic, while $a \cdot x$ is computed in $\mathbb{F}_2$. The moments Z_u form the ultrametric integral representation [BV24]. For the indicator $h = \mathbf{1}_{\mathcal{X}}$, the moments give the counts described above, and the Walsh input is

$$H_\eta = 2^m (-1)^{\eta \cdot a} \mathbf{1}_{A^\perp}(\eta), \qquad \mathcal{X} = a + A. \tag{24}$$

Thus the initial masks lie in $A^\perp$, and every nonzero input coefficient has valuation m.

Under a permutation F, the function becomes $h \circ F^{-1}$, with Walsh coordinates $H' = \mathsf{W}^F H$. To apply F and switch to moment coordinates in one step, we use $Z' = \mathsf{M}^F H$, where

$$\mathsf{M}^F_{v,a} = 2^{-n} \sum_{x \in V} (-1)^{a \cdot x} F(x)^v. \tag{25}$$

We then propagate moments by $Z' = \mathsf{A}^F Z$, with integer coefficients defined by

$$F(x)^v = \sum_u \mathsf{A}^F_{v,u} x^u.$$

This is the unique multilinear expansion over $\mathbb{Z}$ on Boolean inputs, not the algebraic normal form over $\mathbb{F}_2$. After the full cipher, $Z_{e_i} = N_1(\mathcal{X}, e_i; E)$.

We fix a round s at which to switch from Walsh to moment coordinates. Expanding the matrix product expresses the count as a sum of trail contributions. The MILP minimizes the valuation of a trail contribution over compatible masks a_j and exponents u_j, with objective

$$m + \sum_{j<s} v_2(\mathsf{W}^{F_j}_{a_j,a_{j-1}}) + v_2(\mathsf{M}^{F_s}_{u_s,a_{s-1}}) + \sum_{j>s} v_2(\mathsf{A}^{F_j}_{u_j,u_{j-1}}), \quad u_r = e_i. \tag{26}$$

The initial mask $a_0 \in A^\perp$ may be zero. Unlike RC, the resulting bound concerns the count directly, without a final linear-to-integral conversion.

Proposition 5 (Soundness of Joint). *For fixed intermediate representations, suppose the model includes every nonzero term of the mix-basis expansion and bounds each term's valuation from below. Any certified lower bound on the minimum objective is a lower bound on $v_2(N_1)$. Allowing additional transitions or lower transition costs preserves soundness.*

Proof. We insert Equation 24 into the product of the transition matrices and select the final moment coordinate. Each term's valuation is at least m plus its encoded transition costs. The bound follows by applying the ultrametric inequality to this exact expansion. Additional allowed transitions can lower the minimum, but cannot exclude an actual contribution.

Linear Output Combinations. A moment indexed by u selects a product of output bits, not their XOR. To handle output parities, we use the exact identity

$$(-1)^{\beta\cdot y} = \prod_{i:\beta_i=1} (1-2y_i) = \sum_{v\preceq\beta} (-2)^{\mathrm{wt}(v)} y^v. \tag{27}$$

For linear sums with $\alpha \neq 0$, we instead take $h(x) = (-1)^{\alpha\cdot x}$, giving $H_a = 2^n \mathbf{1}[a=\alpha]$. Since $\sum_x (-1)^{\alpha\cdot x} = 0$, the constant term in Equation 27 contributes zero. In the same mix-basis model, we fix the initial mask to α, allow a nonzero final exponent $v \preceq \beta$, replace the initial valuation m by n, and add $\mathrm{wt}(v)$ to the objective. This bounds $v_2(L_{\alpha,\beta}(E))$ using the same mix-basis transitions. We use this formulation to verify selected black-box discoveries in Subsection 6.1. An infeasible model proves zero. A finite bound alone does not establish a nonzero coefficient.

The Mix-Basis Transition for SIMON. We write the Walsh mask as (a, b) and the moment exponent as (p, q), in left–right word order. For a SIMON round, Equation 25 vanishes unless $b \preceq p$. Since SIMON copies the left input word x to the right output word, q fixes the selected coordinates of x to 1. Summing over the right word then leaves a quadratic sum on the free left-word coordinates. When the coefficient is nonzero, its valuation is

$$-\mathrm{wt}(p) - \mathrm{wt}(q) - \rho_s, \tag{28}$$

where ρ_s is half the rank of the polar matrix restricted to those free coordinates. Fixing coordinates can introduce linear terms, which we include in the support condition (Subsection E.3). RC's rank and binary-product constraints encode the support and valuation with $O(w)$ variables and constraints.

Core Construction. Our core construction uses Walsh transitions before the exact mix-basis round and ultrametric transitions afterwards, starting from Equation 24. Subsection 6.1 isolates the improvement from this construction in comparisons with the ultrametric model. We describe optional refinements separately in Subsection E.5.

Local MILP Encoding. For PRESENT, our Spatial Joint variant fixes a Walsh or moment representation for each intermediate bit before building the MILP. For each small component, we compute the exact transition matrix between these representations. Binary selector variables choose one nonzero coefficient per component and enforce its input and output indices. Components connected by a bit share its coordinate index, and the selected valuations add to the objective. Boolean minimization reduces the encoding size without changing the allowed transitions or costs. For bounded-size components and fixed representations, model size grows linearly with circuit size, although solving can remain difficult. XOR key addition changes only signs in Walsh coordinates. In moment coordinates, complementing a bit replaces x by $1 - x$, mixing constant and degree-one terms. For all-key bounds, we include every transition that is nonzero for some key, with a cost no larger than its valuation for every such key. Subsection E.5 gives the basis matrices, selector constraints, and key treatment. For SIMON, the quadratic formulas give exact round coefficients without constructing full-round transition tables.

5.3 Searching for ZCDL and Finite DL Divisibility

We ask whether techniques used to find IDs and impossible boomerangs can also prove ZCDL. Miss-in-the-middle searches use deterministic truncated differential or linear propagation to find incompatible intermediate conditions [HSE23]. For impossible boomerangs, probability-one truncated differential propagation proceeds forward from the input and backward from the output toward a middle S-box layer. If these propagations force local differences α, β at an S-box S whose boomerang connectivity table (BCT) satisfies $\mathrm{BCT}_S(\alpha, \beta) = 0$, no compatible quartet exists, giving an impossible-boomerang distinguisher [BCL$^+$24]. This suggests combining forward differential propagation, backward linear-mask propagation, and a local DLCT zero to prove ZCDL. However, the analogous inference is unsound in general: a DLCT zero expresses cancellation, not the absence of compatible values. Dependencies between intermediate state values and the surrounding rounds can prevent this local cancellation from holding in the complete cipher.

Our ToySPN counterexamples in Figure 6 have local DLCT zeros, yet their exact full-codebook DL sums are nonzero for all 1,000 sampled keys in the first

example and for 953 in the second. One nonzero instance refutes the zero claim. Subsection E.9 explains the propagation and local entries. Conversely, a nonzero local entry does not exclude cancellation in the complete sum. This limitation concerns the composition, not the DLCT of the complete permutation.

Exact Mix-Basis DL Models. A round's output difference $F(x) + F(y)$ is not generally determined by the input difference $x + y$ alone. We therefore use $F^{\times 2}(x, y) = (F(x), F(y))$ to apply the round to both values [BR22,HZC+25]. We use the quasidifferential basis [BR22, Definition 3.1]

$$q_{u,\delta}(x, y) = (-1)^{u \cdot x} \mathbf{1}[x + y = \delta].$$

Each mask–difference pair (u, δ) specifies a difference δ and a linear mask u on the first state value. Discarding u can merge contributions that acquire different signs in later rounds. At the output, we use linear characters on pairs of state values, $r_{u,v}(x, y) = (-1)^{u \cdot x + v \cdot (x+y)}$. These choices recover the DL sum exactly:

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = \sum_{x,y \in V} q_{\alpha,\delta}(x, y) r_{0,\beta}(E(x), E(y)).$$

For input and output bases $\{\phi_i\}$ and $\{\psi_j\}$, the transition coefficients are

$$\mathsf{T}^F_{j,i} = \sum_{x,y} \psi_j(F(x), F(y)) \phi^*_i(x, y),$$

where ϕ^*_i is dual to ϕ_i under the inner product $\langle f, g \rangle = \sum_{x,y} f(x, y) g(x, y)$. For our two orthogonal bases, $q^*_{u,\delta} = 2^{-n} q_{u,\delta}$ and $r^*_{u,v} = 2^{-2n} r_{u,v}$.

Proposition 6 (DL expansion, [BR22,HZC+25]). *For $E = F_R \circ \cdots \circ F_1$ with $F_j : V \to V$, choose a basis on pairs of state values at each component boundary, with q at the input and r at the output. Let $\mathsf{T}^{(j)}$ be the exact transition matrix of $F_j^{\times 2}$ between the adjacent bases. Then, for our unnormalized DL sum,*

$$L^{\mathrm{DL}}_{\alpha,\delta,\beta}(E) = 2^n \sum_{i_1,\ldots,i_{R-1}} \prod_{j=1}^{R} \mathsf{T}^{(j)}_{i_j, i_{j-1}}, \qquad i_0 = (\alpha, \delta), \quad i_R = (0, \beta). \tag{29}$$

The sum runs over all intermediate basis labels.

Proof. The only nonzero input coordinate is 2^n at (α, δ), and the final $(0, \beta)$ coordinate is the unnormalized DL sum. Multiplying the exact transition matrices gives the formula.

Each nonzero product in Equation 29 defines a compatible mix-basis trail. SAT and MILP use the same nonzero transitions and require adjacent components to agree on their shared basis indices. MILP also assigns each transition its valuation as a cost. For S-boxes, we enumerate the exact transitions and use Boolean minimization to obtain compact constraints. As in Proposition 4, a certified lower bound on the minimum of $n + \sum_{j=1}^{R} v_2(\mathsf{T}^{(j)}_{i_j, i_{j-1}})$ over these trails also

bounds the valuation of the DL sum, since cancellation can only increase it. If the support SAT model is unsatisfiable, then $L^{\mathrm{DL}}_{\alpha,\delta,\beta} = 0$ by Proposition 11. A satisfying assignment does not prove that the DL sum is nonzero, since trail contributions may cancel. For $\alpha = 0$, vanishing proves ordinary ZCDL. For nonzero α, the coefficient is a signed combination of coset sums given by Equation 8. We certify the $\alpha = 0$ coefficient and the nonzero-α family separately before applying Theorem 2. Subsection E.10 gives the local relations and soundness proofs. Subsection E.11 shows how our SboxAnalyzer extension generates their constraints.

5.4 Multipurpose Black-Box Discovery

The symbolic models exploit a cipher's internal structure. Our black-box approach instead evaluates all plaintexts in a chosen input set. For small block sizes, a weighted output histogram followed by an FWHT gives exact sums for all linear output masks simultaneously. We use this standard transform to search for finite divisibility and exact zeros, rather than only large correlations or parity constraints. This complements surrogate-differentiation searches for statistical properties [DDK+23] and empirical parity-kernel searches [HKK+26].

For each input specification j, we choose a finite domain S_j, a map $g_{K,j} : S_j \to V$ computed from the oracle, and weights $w_j : S_j \to \{\pm 1\}$. We form the weighted histogram and compute its unnormalized FWHT:

$$\begin{aligned} H_{K,j}(y) &= \sum_{\substack{x \in S_j \\ g_{K,j}(x) = y}} w_j(x), \\ \widehat{H}_{K,j}(\beta) &= \sum_{y \in V} H_{K,j}(y)(-1)^{\beta \cdot y} = \sum_{x \in S_j} w_j(x)(-1)^{\beta \cdot g_{K,j}(x)}. \end{aligned} \tag{30}$$

For $w_j = 1$, the output one-count for mask β is $(|S_j| - \widehat{H}_{K,j}(\beta))/2$.

Algorithm 1 in Subsection E.12 gives the generic pseudocode: construct the histograms, transform them, convert any requested mask families to coset sums, and retain properties that hold across the tested instances.

Three input specifications explain the scope of Algorithm 1. For $S_j = V$, $g_{K,j} = E_K$, and $w_j(x) = (-1)^{\alpha \cdot x}$, the coefficients are $L_{\alpha,\beta}(E_K)$. Replacing $g_{K,j}$ by $D_\delta E_K$ gives generalized DL coefficients. Taking $S_j = \mathcal{X}$, $g_{K,j} = E_K$, and $w_j = 1$ gives integral signed sums directly. For $S_j = V$, $g_{K,j} = D_\delta E_K$, and $w_j = 1$, the histogram is the DDT row $\mathrm{DDT}_{E_K}(\delta, \cdot)$. A zero histogram entry identifies an ID, whereas a zero transformed coefficient identifies ZCDL, for the tested instance. Neither event implies the other for an individual entry.

When the input specifications cover a complete d-dimensional mask space U, a length-2^d FWHT over its masks, divided by 2^d, gives every associated coset sum by Theorem 1 or Theorem 2. This includes the zero input mask, whose DL coefficient need not vanish. Then Equation 1 gives the exact one-count. Thus the same computation supports IDs, linear and DL zeros, finite signed-sum and count divisibility, integral parity, and DL coset contrasts. Different input specifications

require different histograms; one arbitrary histogram does not supply all these experiments simultaneously.

For properties found by the search, random-permutation bounds must account for candidate selection. Independent uniform-key tests separately give confidence bounds on the fraction of keys violating a fixed property, not an all-key proof (Subsection C.6).

Excluding All-Key Integral Properties. Black-box screening can also rule out specified families of all-key integral properties: one counterexample key suffices for each candidate. For a permutation and nonzero output mask, the two coordinate-halfspace offsets have equal signed-sum valuations and are balanced simultaneously (Equation 6). Thus, only n halfspace evaluations per key are needed to screen every nonzero output mask. A common divisibility bound or exact balance on every coset of a smaller coordinate subspace also holds on each enclosing coordinate halfspace. If, for every coordinate halfspace and nonzero output mask, the required condition fails for at least one tested key, no such all-key, all-offset property exists (Subsection E.13). This conclusion does not exclude probabilistic distinguishers, weak-key properties, isolated smaller cubes, nonlinear output predicates, or other affine orientations.

Cost. A length-2^n transform costs $O(n2^n)$ additions and $O(2^n)$ words, after constructing the histogram. One 64-bit array for $n = 32$ occupies 32 GiB; searching all output masks therefore uses suitable CPU/GPU hardware and batched input specifications. We use the GPU backend of LibFWHT [HMAV26] for our black-box FWHT computations. For a single fixed output mask, we can accumulate the sum without storing a histogram. We use exact arithmetic with sufficient integer precision to avoid overflow. For each sampled key, we evaluate every input in the chosen set and record the smallest valuation observed across the tested keys.

6 Applications and Comparative Results

We apply our methods to obtain stronger divisibility bounds and exact counts, analytically explain selected black-box discoveries, and identify properties missed by parity-only or single-output-bit searches. In the tables, t denotes the bound $2^t \mid N_1$. We label exact counts and other quantities explicitly. Analytical bounds cover all keys and the stated offsets, whereas empirical minima concern only the tested configurations. Bit indices follow Subsection 2.1. Each cipher subsection in Section F begins with its specification and round conventions.

6.1 Stronger Finite Bounds on SIMON

We use the standard 1-round input extension for SIMON [WLV+14], specified in Subsection F.1. Let R_0 denote a round without key addition and choose the plaintext set $\widetilde{\mathcal{X}} = R_0^{-1}(a + A)$. The first keyed round maps this set to a translate

of $a + A$, so a bound valid for every offset extends by 1 round without increasing the data complexity. The round counts in Table 1 already include this extension and match those of Beyne and Verbauwhede [BV24].

Table 1: SIMON one-count divisibility bounds for $\widetilde{\mathcal{X}} = R_0^{-1}(a + (u \wedge \mathbb{F}_2^n))$. RC and Joint use no trail enumeration. Each listed output bit defines a separate distinguisher.

Cipher	Rounds	Active mask u	Data	Output	t [BV24]	RC t	Joint t	Gain
SIMON32	9	`1fffffff`	2^{29}	c_5	5	10	11	+6
SIMON48	10	`0fffffffffff`	2^{44}	c_{20}, c_{21}	6	20	22	+16
SIMON32	13	`7fffffff`	2^{31}	$c_0, \ldots, c_{15}$	1	0	2	+1
SIMON48	15	`7fffffffffff`	2^{47}	$c_0, \ldots, c_{23}$	1	0	2	+1

The first two rows of Table 1 improve the divisibility exponents by 6 and 16 at unchanged data complexity. For the 8-round core of the 9-round SIMON32 result, RC proves $v_2(L_{\alpha,e_5}) \geq 14$ for every $\alpha \in A^\perp$. Since $\dim A^\perp = 3$, our conversion gives the one-count bound $t = 14 - 3 - 1 = 10$, whereas Joint proves $t = 11$ directly. For another 9-round SIMON32 structure and output parity, RC proves $t = 7$, whereas our tested Joint models prove $t = 6$ (Table 16).

We also evaluated both SIMON32 structures in Table 1 on 1,000 distinct random master keys, using the real key schedule and offset $a = 0$. Every tested key satisfied all listed bounds. At 9 rounds, the sampled minimum of $v_2(N_1)$ for c_5 was 13, and the signed integral sum was nonzero for 977 keys, confirming finite divisibility beyond ZC. At 13 rounds, the sampled minimum was 4 for $c_0, c_1, c_7, c_{13}, c_{14}$ and 3 for each remaining bit among $c_0, \ldots, c_{15}$. These values were observed on the tested keys, not proved for all keys. Table 1 reports the proved all-key bounds.

What Does Walsh Contribute? The published exponents for the first two rows of Table 1 are 5 and 6 [BV24, Tables 7–8]. Our implementation of the ultrametric model reproduces these bounds at c_5 and c_{20}, respectively. Our Joint model instead proves $t = 11$ and $t = 22$, respectively, by combining Walsh propagation in the initial rounds with ultrametric propagation in the remaining rounds. All four models were solved to optimality in under 3 seconds each using Gurobi 13.0.2 with one solver thread on an Apple M4 laptop. Neither trail enumeration nor special first-round cube propagation is needed. This comparison concerns the bounds, not the runtime of the authors' original implementation. Subsection E.6 gives the broader comparison and its limitations.

For the 9-round SIMON32 and 10-round SIMON48 bounds in Table 1, we also estimate how likely a random permutation is to satisfy the same divisibility condition. For the 9-round SIMON32 row, Proposition 3 bounds the random-permutation probability of divisibility by 2^5, as proved in [BV24], by $2^{-4.99}$. Our stronger guarantees of divisibility by 2^{10} and 2^{11} reduce these upper bounds to

$2^{-9.94}$ and $2^{-10.89}$, respectively. For the 10-round SIMON48 row, divisibility by 2^6 from the same work gives an upper bound of $2^{-5.99}$. Our guarantees of divisibility by 2^{20} and 2^{22} reduce it to $2^{-19.72}$ and $2^{-21.13}$, respectively. Thus our stronger divisibility guarantees give stronger distinguishers at unchanged data complexity. Each probability concerns one selected output bit, not a simultaneous test of several bits. Subsection C.4 gives the calculations. Table 8 and Table 9 summarize the finite-divisibility and exact-balance probabilities.

For 25 output bits of 10-round SIMON32 with active mask $u = \texttt{7fffffff}$ and data complexity 2^{31}, the published bounds give one-count divisibility by 2^3 or 2^4, depending on the bit [BV24]. Our linear MILP model proves that all linear sums required by the conversion are zero. The conversion therefore establishes $N_1 = 2^{30}$ for each of these bits: exactly half the outputs are ones (Table 14). At 15 rounds, Joint proves divisibility by 4, rather than 2, for 24 output bits of SIMON48 at unchanged data complexity (Table 1). Across the SIMON family, we improve the published bounds in 674 cases: 290 stronger finite-divisibility bounds and 384 exact-balance results (Subsection F.1).

At the longest round counts reported in [BV24], Joint also matches all positive one-count divisibility bounds, including those obtained with trail enumeration (Table 15). For these results, we combine the first two rounds after the input extension into one exact transition before assigning valuations. We additionally use trail enumeration for SIMON32.

6.2 SPECK: Black-Box Discovery and Analytical Proofs

SPECK (Subsection F.2) demonstrates the usefulness of our linear model for ARX ciphers. We prove finite divisibility and exact balance for multi-bit output parities, for arbitrary round keys and every affine offset.

Our black-box search over all output masks identifies the SPECK32 parity $c_2 \oplus c_9 \oplus c_{16} \oplus c_{18} \oplus c_{25}$. Table 2 gives analytical bounds for these discoveries and for SPECK48. Parenthesized rounds include the input extension $R_0^{-1}(\mathcal{X})$, preserving data complexity and the all-offset guarantee.

Table 2: SPECK finite one-count bounds from the linear model. Hexadecimal inputs denote the core active mask u. Fixed values are arbitrary. Parentheses in the bound column denote residue refinement by trail enumeration.

Cipher	Rounds	Core input	Data	Output mask β	t (refined)
SPECK32	6 (7)	$x_8 \oplus x_{10} = b$	2^{31}	02050204	10
SPECK32	7 (8)	ffffff9f	2^{30}	02050204	0 (3)
SPECK48	7 (8)	fffffffffff1f	2^{45}	000809000808	7 (8)

For the 6-round SPECK32 hyperplane, the linear model proves $t = 10$; a unique minimum-valuation trail establishes $v_2(N_1) = 10$ for every key (Table 21). This is genuinely finite divisibility, not ZC.

For the 7-round SPECK32 cube, we strengthen the bound by analyzing how the linear sums combine in our conversion. This proves $8 \mid N_1$ on all four cosets, whereas bounding the sums separately gives only $4 \mid N_1$ (Subsection E.4). All 1,000 validation keys satisfy the stronger bound. The same output parity appears in empirical kernel-based integral analysis [HKK+26] and in symbolic integral analysis using the split-and-cancel method [WHG26]. However, we establish a stronger count constraint.

For SPECK48, the linear model proves $t = 7$, and enumeration of only 2 leading trails raises it to 8 (Table 24). We also prove $N_1 = 2^{124}$ for 7 (8) rounds of SPECK128 with data complexity 2^{125} (Table 25). This exact-balance result uses an output mask of Hamming weight 5, beyond searches restricted to weight at most 3.

6.3 BEANIE: Black-Box Discovery and Analytical Proofs

BEANIE illustrates how black-box discoveries lead to all-key divisibility proofs. We study the 32-bit U-shaped permutation U_{r_1,r_2}: a forward encryption suffix under (K, T_1) followed by an inverse suffix under (K, T_2). Both are suffixes of the standard 5-round encryption and include its special final round. The key and tweaks are fixed while evaluating each structure. Subsection F.3 defines the forward and inverse suffixes.

Table 3: BEANIE halfspace counts on $\mathcal{X} = \{x : \alpha \cdot x = b\}$, for both $b \in \mathbb{F}_2$. Parentheses in the linear column denote exact U-turn contraction, also used by Joint. Neither model uses trail enumeration. Each empirical row uses 1,000 key/tweak configurations.

Rounds	Shape / input mask α	Data	Output mask β	Linear t	Joint t	Sample t
5	$U_{2,3}$ / `00004020`	2^{31}	`000440b0`	0 (8)	10	12
5	$U_{3,2}$ / `c0e00d00`	2^{31}	`00000001`	2 (10)	12	12
6	$U_{3,3}$ / `00000001`	2^{31}	`a0700601`	–	1	2

At the central U-turn, adjacent substitutions combine into $S^{-1}(S(x) + d)$ for a four-bit key-dependent shift d. Computing this complete local Walsh table before taking valuations accounts for cancellations lost by modelling the two substitutions separately. Taking the minimum valuation among nonzero coefficients over all 16 shifts gives a bound valid for every key and tweak. With this contraction, the linear model proves exponents 8 and 10 for $U_{2,3}$ and $U_{3,2}$, respectively (Subsection E.7). Joint uses the same contraction, uses Walsh coordinates throughout the first round, and switches to ultrametric coordinates at the next nonlinear layer. It raises these bounds to 10 and 12, respectively (Table 3), matching the sampled minimum for $U_{3,2}$. Both models cover every key/tweak configuration and both offsets, without trail enumeration.

For the 6-round case in Table 3, Joint proves evenness for every key/tweak configuration and both offsets. The stronger divisibility by 4 remains empirical; 468 of the 1,000 tested configurations attain exponent 2. Table 27 gives full-domain DL-count exponents 13 and 15 on the two 5-round shapes. Table 28 gives selected exact-balance observations. We evaluate these count constraints against a uniformly random 32-bit permutation, using the same random-permutation laws as for the other applications (Subsection C.4).

6.4 Further Integral and DL Results

PRESENT. For PRESENT (Subsection F.4), the weighted linear model proves finite divisibility even when its support-only counterpart admits trails. At 5 rounds, the hyperplane $x_{29} \oplus x_{30} = b$ gives $t = 33$ for $c_7 \oplus c_{39} \oplus c_{55}$ with data complexity 2^{63}; trail-enumeration refinement raises this to 34 (Table 32). For small coordinate cubes, we improve the bounds of [BV24]: Joint raises the 4-round exponent from 1 to 3 for three output bits. Exact 16-bit contraction raises the 5-round exponent from 1 to 2 for four output bits (Table 29). Across the 384 published coordinate-cube/output-bit pairs, our methods give 7 stronger bounds and 377 equal bounds (Table 30). Subsection F.4 gives the full comparison and further integral results. Table 33 gives our 5- and 6-round DL count bounds.

Ascon and GIFT-64. For 5 rounds of the 320-bit Ascon permutation (Subsection F.5), our linear MILP proves $2^{60} \mid N_1$ on the hyperplane $x_{256} = b$ at output c_{256} (Table 34). The data complexity is 2^{319}; we obtain the bound without enumerating the input structure. For 5-round GIFT-64 (Subsection F.6), the linear model with exact local input propagation proves $t = 53$ with data complexity 2^{60} for $u =$ `ffffffff1ffeffff` and output $c_{12} \oplus c_{29} \oplus c_{46} \oplus c_{55} \oplus c_{63}$ (Table 35). Our exact mix-basis DL model also proves $2^{26} \mid N_1^{\mathrm{DL}}$ for 5-round GIFT with data complexity 2^{63}, using active mask $u =$ `fffffbffffffffff` and input difference $\delta = e_0$ to count inputs for which $c_{59} \neq c'_{59}$ (Table 38). Subsection F.5 and Subsection F.6 give further results, including GIFT's longer-round integrals.

SKINNY-64. On 11 rounds of SKINNY-64-64 (Subsection F.7), active mask $u =$ `0fffffffffffffff` and output parity $c_{12} \oplus c_{28} \oplus c_{44}$ give $N_1 = 2^{59}$ at data complexity 2^{60}. One whole-subspace support check proves the required ZC family. The recent neural-inspired integral analysis [ZYS+26] gives a zero-sum property for this same structure and parity. We establish a much stronger statistical constraint at the same data complexity, not merely another proof of the zero-sum property. By Proposition 2, exact balance for this fixed test has random-permutation probability approximately $2^{-30.28}$, compared with approximately 1/2 for the zero-sum property. For the separate 8-round structure with data complexity 2^8, Joint with exact contraction of the final layers proves zero-sum, but exact balance does not hold (Table 40). Subsection F.7 gives both families.

6.5 Further Black-Box Results

Table 4 gives black-box observations for ciphers without our symbolic models. For the specifications of RC5-16, ChiLow-32, KATAN32, and KeeLoq, see Subsection F.8, Subsection F.9, Subsection F.10, and Subsection F.11, respectively. Subsection F.9 and Subsection F.11 give further ChiLow and KeeLoq results. The structures and masks are fixed within each experiment.

Table 4: Further black-box one-count observations. A condition describes $\mathcal{X}$. Unspecified coordinates vary. All rows cover both offsets $b \in \mathbb{F}_2$.

Cipher	Rounds	Input condition	Data	Output mask β	Sample t	Tests
RC5-16	2.5	$x_{23} = b$	2^{31}	00020000	1	1000
ChiLow-32^{-1}	4	$x_0 = b$	2^{31}	25c72ab9	3	1000
KATAN32	84	$x_0 = b$	2^{31}	00040000	3	1000
KeeLoq	64	$31200000 \cdot x = b$	2^{31}	00000003	14	1000

RC5's derivative histogram gives a separate differential-count application. For input difference 80008000, the DDT counts for 04100000 and 0c100000 have minimum valuations 11 and 10 over 1,000 sampled keys. These candidates were selected from the sampled counts, rather than evaluated on an independent holdout sample. Their minima exceed universal DDT evenness by 10 and 9 exponent bits, respectively. These counts may be zero for some keys, so the random-permutation probability includes zero counts (Equation 34). The same scans find 2.5-round RC5 ID candidates with output differences of Hamming weight 3 or 4, beyond the single-bit output differences reported by Hu et al. [HLJ+20]. Subsection F.8 gives the differential-count divisibility observations and the ID and integral candidates.

For the 84-round KATAN32 result in Table 4, the one-count is divisible by 8, but not 16, on every one of the 1,000 holdout keys. The signed sums are nonzero and vary across keys. For the same fixed structure and output mask, the probability that 1,000 independent uniformly random permutations all satisfy $8 \mid N_1$ is at most $2^{-2999.71}$ (Subsection E.14). This quantifies separation from the random-permutation model, not the probability of failure on an untested cipher key.

Subsection F.10 also reports constant-parity observations through 101 rounds of KATAN32, matching the round coverage of Hu and Wang [HW19]. The 101-round integral has constant parity 1.

Our black-box method also rules out all-key constant-parity integral properties on coordinate halfspaces. For 8-round SPECK32, 15-round SIMON32, 3-round RC5-16, and 6-round ChiLow-32 decryption, we evaluated all 32 coordinate halfspaces on 50 keys or key/tweak configurations. For each halfspace, the differences between its output-XOR vectors span $\mathbb{F}_2^{32}$. These evaluations give counterexamples to every constant-parity claim with a fixed nonzero output mask.

By Proposition 13, they also exclude all-key one-count divisibility by 2^t, $t \geq 1$, on every offset of a smaller coordinate cube. These exclusions concern coordinate cubes at the plaintext input, not the input-extended structures $R_0^{-1}(a + A)$ used in our positive SIMON and SPECK results. They do not exclude probabilistic distinguishers, special keys, isolated smaller cubes, nonlinear output functions, or other affine orientations. Subsection F.12 records the exact scope.

Probabilistic Observations. Our search for deterministic divisibility also reveals conditions that hold for only some keys or input offsets, but more often than for a random permutation. Subsection F.13 reports selected observations on PRESENT, GIFT, SPECK, Ascon, and SIMON, including their variation across keys and input offsets. Their analytical explanation is left for future work.

7 Conclusion and Future Works

We connected linear, integral, DL, and restricted differential sums through explicit conversions that enable stronger divisibility guarantees at unchanged data complexity. Our applications established finite divisibility beyond zero-sum properties and, in some cases, exact counts. We exploited these connections through automated models and black-box searches, and quantified the distinguishing strength of the resulting properties using random-permutation bounds.

Different representations of the same sum can yield different automated bounds. Accounting for more cancellations between trail contributions while keeping the search efficient remains an open problem. Our searches also led us to introduce probabilistic divisibility (Subsection F.13). This raises two open questions: how can we develop scalable methods to compute these probabilities over keys and input offsets analytically? How can we develop scalable tools to search for strong probabilistic divisibility distinguishers?

Acknowledgments. This work was carried out during the author's postdoctoral appointment at Ruhr University Bochum in 2025–2026. The author thanks Gregor Leander for his support throughout the project. The author also thanks Google for research credits used for GPU computations. Part of the calculations for this publication were also performed on the HPC cluster Elysium of Ruhr University Bochum, subsidised by the DFG (INST 213/1055-1). This work was funded by the European Union (ERC, SYMTRUST, 101097056). Views and opinions expressed are, however, those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Generative AI Disclosure. The research ideas, methodologies, results, and scientific contributions presented in this paper were developed by the author, reflecting his own scientific reasoning and judgment. To support this work, the author used OpenAI large language models to assist with writing parts of the

code, editorial proofreading of the manuscript, and checking the correctness of the proofs. The author substantially edited and revised the resulting code and text and takes full responsibility for the correctness, accuracy, and integrity of this work, including its mathematical proofs and experimental results.

References

BCL+24. Xavier Bonnetain, Margarita Cordero, Virginie Lallemand, Marine Minier, and María Naya-Plasencia. On impossible boomerang attacks application to simon and skinnyee. *IACR Trans. Symmetric Cryptol.*, 2024(2):222–253, 2024.

BDG+25. Yanis Belkheyar, Patrick Derbez, Shibam Ghosh, Gregor Leander, Silvia Mella, Léo Perrin, Shahram Rasoolzadeh, Lukas Stennes, Siwei Sun, Gilles Van Assche, and Damian Vizár. ChiLow and ChiChi: New constructions for code encryption. In *EUROCRYPT 2025, Part I*, volume 15601 of *LNCS*, pages 212–243. Springer, 2025.

BDKW19. Achiya Bar-On, Orr Dunkelman, Nathan Keller, and Ariel Weizman. DLCT: A new tool for differential-linear cryptanalysis. In *EUROCRYPT 2019*, volume 11476 of *LNCS*, pages 313–342. Springer, 2019.

Bey21. Tim Beyne. A geometric approach to linear cryptanalysis. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT 2021, Part I*, volume 13090 of *LNCS*, pages 36–66. Springer, 2021.

BJK+16. Christof Beierle, Jérémy Jean, Stefan Kölbl, Gregor Leander, Amir Moradi, Thomas Peyrin, Yu Sasaki, Pascal Sasdrich, and Siang Meng Sim. The SKINNY family of block ciphers and its low-latency variant MANTIS. In *CRYPTO 2016*, pages 123–153. Springer, 2016.

BKL+07. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: an ultra-lightweight block cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *CHES 2007*, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.

BLNW12. Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, and Meiqin Wang. Integral and multidimensional linear distinguishers with correlation zero. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, LNCS, pages 244–261. Springer, 2012.

BLP25. Christof Beierle, Gregor Leander, and Yevhen Perehuda. When the wrong key lives on: The key-recovery procedure in integral attacks. Cryptology ePrint Archive, Paper 2025/2011, 2025. Published in EUROCRYPT 2026.

BPP+17. Subhadeep Banik, Sumit Kumar Pandey, Thomas Peyrin, Yu Sasaki, Siang Meng Sim, and Yosuke Todo. GIFT: A small present - towards reaching the limit of lightweight encryption. In Wieland Fischer and Naofumi Homma, editors, *CHES 2017*, volume 10529 of *LNCS*, pages 321–345. Springer, 2017.

BR22. Tim Beyne and Vincent Rijmen. Differential cryptanalysis in the fixed-key model. In Yevgeniy Dodis and Thomas Shrimpton, editors, *CRYPTO 2022, Part III*, volume 13509 of *LNCS*, pages 687–716. Springer, 2022.

BSS+15. Ray Beaulieu, Douglas Shors, Jason Smith, Stefan Treatman-Clark, Bryan Weeks, and Louis Wingers. The SIMON and SPECK lightweight block ciphers. In *DAC 2015*, pages 175:1–175:6. ACM, 2015.

BV24. Tim Beyne and Michiel Verbauwhede. Ultrametric integral cryptanalysis. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024*, LNCS, pages 392–423. Springer, 2024.

BV25. Tim Beyne and Michiel Verbauwhede. Integral cryptanalysis in characteristic p. In Goichiro Hanaoka and Bo-Yin Yang, editors, *ASIACRYPT 2025*, LNCS, pages 66–96. Springer, 2025.

CHNE24. Debasmita Chakraborty, Hosein Hadipour, Phuong Hoa Nguyen, and Maria Eichlseder. Finding complete impossible differential attacks on AndRX ciphers and efficient distinguishers for ARX designs. *IACR Trans. Symmetric Cryptol.*, 2024(3):84–176, 2024.

DDK09. Christophe De Cannière, Orr Dunkelman, and Miroslav Knežević. KATAN and KTANTAN – a family of small and efficient hardware-oriented block ciphers. In *CHES 2009*, volume 5747 of *LNCS*, pages 272–288. Springer, 2009.

DDK+23. Itai Dinur, Orr Dunkelman, Nathan Keller, Eyal Ronen, and Adi Shamir. Efficient detection of high probability statistical properties of cryptosystems via surrogate differentiation. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023*, LNCS, pages 98–127. Springer, 2023.

DEMS21. Christoph Dobraunig, Maria Eichlseder, Florian Mendel, and Martin Schläffer. Ascon v1.2 (Submission to NIST). Finalist submission to the NIST lightweight cryptography standardization process, 2021.

DF20. Patrick Derbez and Pierre-Alain Fouque. Increasing precision of division property. *IACR Trans. Symmetric Cryptol.*, 2020(4):173–194, 2020.

DKR97. Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The block cipher Square. In *FSE 1997*, volume 1267 of *LNCS*, pages 149–165. Springer, 1997.

FWG+16. Kai Fu, Meiqin Wang, Yinghua Guo, Siwei Sun, and Lei Hu. MILP-based automatic search algorithms for differential and linear trails for SPECK. In *FSE 2016*, volume 9783 of *LNCS*, pages 268–288. Springer, 2016.

GHM+25. Simon Gerhalter, Samir Hodzic, Marcel Medwed, Marcel Nageler, Artur Folwarczny, Ventzislav Nikov, Jan Hoogerbrugge, Tobias Schneider, Gary McConville, and Maria Eichlseder. BEANIE – a 32-bit cipher for cryptographic mitigations against software attacks. *IACR Trans. Symmetric Cryptol.*, 2025(4):31–69, 2025.

HDE24. Hosein Hadipour, Patrick Derbez, and Maria Eichlseder. Revisiting differential-linear attacks via a boomerang perspective with application to aes, ascon, clefia, skinny, present, knot, twine, warp, lblock, simeck, and SERPENT. In Leonid Reyzin and Douglas Stebila, editors, *CRYPTO 2024*, LNCS, pages 38–72. Springer, 2024.

HGSE24. Hosein Hadipour, Simon Gerhalter, Sadegh Sadeghi, and Maria Eichlseder. Improved search for integral, impossible differential and zero-correlation attacks application to ascon, forkskinny, skinny, mantis, PRESENT and qarmav2. *IACR Trans. Symmetric Cryptol.*, 2024(1):234–325, 2024.

HKK+26. Yunjae Hwang, Insung Kim, Sunyeop Kim, Myung-Kyu Lee, Hanbeom Shin, Deukjo Hong, Seokhie Hong, Dongjae Lee, Jaechul Sung, and Byoungjin Seok. Improving neural-inspired integral distinguishers via a linear-algebraic approach. *IACR Cryptol. ePrint Arch.*, 2026:340, 2026.

HLJ+20. Xichao Hu, Yongqiang Li, Lin Jiao, Shizhu Tian, and Mingsheng Wang. Mind the propagation of states: New automatic search tool for impossible differentials and impossible polytopic transitions. In *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 415–445. Springer, 2020.

HMAV26. Hosein Hadipour, Saleh Khalaj Monfared, Jens Alich, and Jan Vorloeper. LibFWHT: From exact walsh spectra to key dependence in differential-linear correlations. Cryptology ePrint Archive, Paper 2026/1997, 2026.

HNE22. Hosein Hadipour, Marcel Nageler, and Maria Eichlseder. Throwing boomerangs into feistel structures application to CLEFIA, WARP, LBlock, LBlock-s and TWINE. *IACR Trans. Symmetric Cryptol.*, 2022(3):271–302, 2022.

HSE23. Hosein Hadipour, Sadegh Sadeghi, and Maria Eichlseder. Finding the impossible: Automated search for full impossible-differential, zero-correlation, and integral attacks. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023*, volume 14007 of *LNCS*, pages 128–157. Springer, 2023.

HW19. Kai Hu and Meiqin Wang. Automatic search for a variant of division property using three subsets. In *CT-RSA 2019*, volume 11405 of *LNCS*, pages 412–432. Springer, 2019.

HZC+25. Kai Hu, Chi Zhang, Chengcheng Chang, Jiashu Zhang, Meiqin Wang, and Thomas Peyrin. Unlocking mix-basis potential: Geometric approach for combined attacks. In Yael Tauman Kalai and Seny F. Kamara, editors, *CRYPTO 2025*, LNCS, pages 293–334. Springer, 2025.

IKD+08. Sebastiaan Indesteege, Nathan Keller, Orr Dunkelman, Eli Biham, and Bart Preneel. A practical attack on KeeLoq. In *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 1–18. Springer, 2008.

KLT15. Stefan Kölbl, Gregor Leander, and Tyge Tiessen. Observations on the SIMON block cipher family. In *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 161–185. Springer, 2015.

KW02. Lars R. Knudsen and David A. Wagner. Integral cryptanalysis. In *FSE 2002*, volume 2365 of *LNCS*, pages 112–127. Springer, 2002.

LGCX19. Mingming Li, Jiansheng Guo, Jingyi Cui, and Linhong Xu. Impossible differential cryptanalysis of SPECK. In *CTCIS 2018*, volume 960 of *Communications in Computer and Information Science*, pages 16–31. Springer, 2019.

MMMS23. Loïc Masure, Pierrick Méaux, Thorben Moos, and François-Xavier Standaert. Effective and efficient masking with low noise using small-mersenne-prime ciphers. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023*, LNCS, pages 596–627. Springer, 2023.

PKA+26. Shuo Peng, Akram Khalesi, Zahra Ahmadian, Hosein Hadipour, Jiahui He, Kai Hu, Zhongfeng Niu, Shahram Rasoolzadeh, and Meiqin Wang. Cube and integral attacks on ChiLow-32. *IACR Trans. Symmetric Cryptol.*, 2026(1):253–292, 2026.

Riv95. Ronald L. Rivest. The RC5 encryption algorithm. In *FSE 1994*, volume 1008 of *LNCS*, pages 86–96. Springer, 1995.

SLR+15. Bing Sun, Zhiqiang Liu, Vincent Rijmen, Ruilin Li, Lei Cheng, Qingju Wang, Hoda Alkhzaimi, and Chao Li. Links among impossible differential, integral and zero correlation linear cryptanalysis. In Rosario Gennaro and Matthew Robshaw, editors, *CRYPTO 2015*, pages 95–115. Springer, 2015.

SWLW16. Ling Sun, Wei Wang, Ru Liu, and Meiqin Wang. Milp-aided bit-based division property for arx-based block cipher. *IACR Cryptol. ePrint Arch.*, 2016:1101, 2016.

Tod15. Yosuke Todo. Structural evaluation by generalized integral property. In *EUROCRYPT 2015*, volume 9056 of *LNCS*, pages 287–314. Springer, 2015.

WHG+20. Senpeng Wang, Bin Hu, Jie Guan, Kai Zhang, and Tairong Shi. Exploring secret keys in searching integral distinguishers based on division property. *IACR Trans. Symmetric Cryptol.*, 2020(3):288–304, 2020.

WHG26. Dachao Wang, Hosein Hadipour, and Simon Gerhalter. On extending integral distinguishers. Cryptology ePrint Archive, Paper 2026/1402, 2026.

WLV+14. Qingju Wang, Zhiqiang Liu, Kerem Varici, Yu Sasaki, Vincent Rijmen, and Yosuke Todo. Cryptanalysis of reduced-round SIMON32 and SIMON48. In *INDOCRYPT 2014*, pages 143–160. Springer, 2014.

ZDY+19. Wentao Zhang, Tianyou Ding, Bohan Yang, Zhenzhen Bao, Zejun Xiang, Fulei Ji, and Xuefeng Zhao. KNOT: Algorithm specifications and supporting document. *Submission to NIST lightweight cryptography project*, 2019.

ZYS+26. Liu Zhang, Yiran Yao, Danping Shi, Dongchen Chai, Jian Guo, and Zilong Wang. Neural-inspired advances in integral cryptanalysis. In Joan Daemen and Emmanuel Thomé, editors, *EUROCRYPT 2026*, LNCS, pages 451–481. Springer, 2026.

A Supplementary Conversion Details

A.1 Higher-Order Derivatives

We extend Theorem 2 by replacing the first derivative with a higher-order derivative. Fix linearly independent directions $\delta_1, \ldots, \delta_s \in V$, with $1 \le s \le n$, and let $H = \langle \delta_1, \ldots, \delta_s \rangle$. Define

$$D_H E(x) = \sum_{h \in H} E(x+h) = \sum_{\epsilon \in \mathbb{F}_2^s} E\left(x + \sum_i \epsilon_i \delta_i \right),$$

where the additions are over $\mathbb{F}_2$. For a nonzero output mask β, define the full-codebook and affine-coset signed sums by

$$L^{\mathrm{HDL}}_{\alpha,H,\beta}(E) = \sum_{x \in V} (-1)^{\alpha \cdot x + \beta \cdot D_H E(x)},$$
$$I^{\mathrm{HDL}}_{H,\beta}(a+A;E) = \sum_{x \in a+A} (-1)^{\beta \cdot D_H E(x)}.$$

Proposition 7 (Higher-order DL affine-coset conversion). *Let $A = U^\perp$ with $\dim U = d$, and let $\mathcal{R}_A$ contain one representative of each coset of A. Then*

$$I^{\mathrm{HDL}}_{H,\beta}(a+A;E) = 2^{-d} \sum_{\alpha \in U} (-1)^{\alpha \cdot a} L^{\mathrm{HDL}}_{\alpha,H,\beta}(E), \tag{31}$$

$$L^{\mathrm{HDL}}_{\alpha,H,\beta}(E) = \sum_{a \in \mathcal{R}_A} (-1)^{\alpha \cdot a} I^{\mathrm{HDL}}_{H,\beta}(a+A;E) \quad (\alpha \in U). \tag{32}$$

If $v_2(L^{\mathrm{HDL}}_{\alpha,H,\beta}(E)) \ge T$ for all $\alpha \in U$, including $\alpha = 0$, then $v_2(I^{\mathrm{HDL}}_{H,\beta}(a+A;E)) \ge T - d$ on every coset. If $H \subseteq A$, every evaluation $E(x+h)$ needed for $x \in a+A$ remains within the same input set $a+A$.

Proof. Multiply the coset-indicator identity in Equation 3 by $(-1)^{\beta \cdot D_H E(x)}$ and sum over x to obtain Equation 31. Partitioning V into cosets gives Equation 32, since $\alpha \cdot x = \alpha \cdot a$ on $a + A$ for $\alpha \in U$. The ultrametric inequality and the factor 2^{-d} give the valuation bound. Finally, $x + h \in a + A$ whenever $x \in a + A$ and $h \in H \subseteq A$.

The independence assumption makes the iterated derivative equal to the sum over distinct elements of H. For dependent directions, the iterated sum over $\mathbb{F}_2^s$ repeats each element of H an even number of times and is identically zero. That cancellation is universal, not a cipher distinguisher. The conversion is algebraic; an efficient higher-order search model and the corresponding random-permutation laws remain separate questions.

B Prime-Field Extension and Illustration

Proof (of Proposition 1). For $U \leq \mathbb{F}_p^\ell$, orthogonality gives

$$\mathbf{1}_{a+U^\perp}(x) = p^{-d} \sum_{\alpha \in U} \psi(\alpha \cdot (x-a)).$$

Multiplying by $\psi(\beta \cdot E(x))$ proves Equation 15. Partitioning the domain into cosets proves Equation 16. The minus sign in the forward phase matters when $p > 2$.

In $\mathbb{Z}[\zeta_p]$, the unique prime ideal above p is generated by $1 - \zeta_p$. Normalize its valuation by $\mathrm{ord}_p(p) = 1$, so $\mathrm{ord}_p(1-\zeta_p) = 1/(p-1)$. The valuation group is contained in $(p-1)^{-1}\mathbb{Z}$ and every root of unity is a unit. The ultrametric inequality applied to the forward transform gives $T_I \geq T_L - d$; applied to the inverse, it gives $T_L \geq T_I$.

The two transforms also give the exact-zero equivalence

$$L^{(p)}_{\alpha,\beta}(E) = 0 \quad \forall \alpha \in U \quad \Longleftrightarrow \quad I^{(p)}_{\beta}(a + A; E) = 0 \quad \forall a \in V_p.$$

Applying the forward bound to each output mask and key extends the result to mask families: a common bound $t(\beta)$ for all $\alpha \in U$ and all keys gives $t(\beta) - d$ on every coset for every key. For a permutation and $\beta \neq 0$, the coefficient with input mask zero vanishes.

Remark 1 (Prime-field derivative sums). The coset expansion applies to any function $g : V_p \to \mathbb{Q}(\zeta_p)$. Define $L_\alpha(g) = \sum_{x \in V_p} \psi(\alpha \cdot x) g(x)$ and $I_g(a + A) = \sum_{x \in a+A} g(x)$. Then

$$I_g(a + A) = p^{-d} \sum_{\alpha \in U} \psi(-\alpha \cdot a) L_\alpha(g),$$
$$L_\alpha(g) = \sum_{a \in \mathcal{R}_A} \psi(\alpha \cdot a) I_g(a + A) \qquad (\alpha \in U).$$

In particular, take $g(x) = \psi(\beta \cdot (E(x+\delta) - E(x)))$. A common valuation bound t on all $L_\alpha(g)$, including $\alpha = 0$, gives $t - d$ on every coset sum. This identity extends the conversion to prime-field derivatives; search models and random-permutation laws for these sums require separate analysis.

Prime-Field Linear Model. We use the same minimum-trail-valuation principle as in Subsection 5.1, with exact local additive-character valuation tables. For an S-box $S : \mathbb{F}_p \to \mathbb{F}_p$, define $W_S(a, b) = \sum_x \psi(ax + bS(x))$. We compute local valuations exactly from integer coefficients modulo the cyclotomic polynomial Φ_p, not from approximate complex absolute values. To compose rounds, the normalized character matrix uses $p^{-1} W_S(-a, b)$; the minus sign accounts for character orthogonality in odd characteristic. Its input and output indices for $L^{(p)}_{\alpha,\beta}$ are $-\alpha$ and β. A local model retains exactly the pairs for which $W_S(-a, b) \neq$

0, with cost $\operatorname{ord}_p W_S(-a, b)$. General finite tables can be encoded by selecting one supported pair and its cost, clearing a common denominator to obtain integer weights. Each S-box layer on ℓ cells contributes a normalization factor $p^{-\ell}$; recovering the unnormalized sum contributes p^{ℓ}. For r layers, the full-codebook bound is therefore the minimum total local cost minus $(r-1)\ell$. Round-key translations contribute only roots of unity, so this bound holds for arbitrary round keys.

The local additive-character valuation table in Table 5 illustrates why general local models must distinguish more than S-box activity. For the permutation $S : \mathbb{F}_{13} \to \mathbb{F}_{13}$ given by

$$(S(0), \ldots, S(12)) = (5, 9, 4, 6, 8, 3, 7, 2, 1, 0, 12, 11, 10),$$

nonzero input and output masks give either a zero coefficient or one of two finite valuations. This is a local modelling example, distinct from the power S-box used in the ToySPN below. The table classifies all 169 pairs for $W_S(a, b)$. The model uses the same entries with input mask $-a$.

Table 5: Local additive-character valuation table for $W_S(a, b)$ over $\mathbb{F}_{13}$, with $\operatorname{ord}_{13}(13) = 1$. An entry ∞ means that the coefficient is zero and the transition is excluded.

Mask condition	Number of pairs	$\operatorname{ord}_{13} W_S(a, b)$
$a = b = 0$	1	1
Exactly one of a, b is zero	24	∞
$a, b \neq 0$, $ab^{-1} = 1$	12	1/2
$a, b \neq 0$, $ab^{-1} = 10$	12	∞
$a, b \neq 0$, $ab^{-1} \notin \{1, 10\}$	120	1/3

For AES-prime's S-box $S(x) = x^5 + 2$ over $\mathbb{F}_{127}$, the local valuation is 1 at $(0, 0)$ and 13/63 whenever both masks are nonzero; pairs with exactly one zero mask are unsupported. We encode this local support and valuation with one activity bit η per S-box:

$$\eta \leq a, b \leq 126\eta, \qquad a, b \in \{0, \ldots, 126\}, \quad \eta \in \{0, 1\}, \qquad \omega = 63 - 50\eta.$$

Here ω is the local valuation multiplied by 63; changing a to $-a$ does not change this support or cost. For a linear layer with matrix B, vector input and output masks a, b satisfy $a = \mathsf{B}^{\mathsf{T}} b$ over $\mathbb{F}_{127}$. The MILP minimizes $\sum_{j,i} \omega_{j,i} - 63(r-1)\ell$; dividing a certified objective lower bound by 63 gives the full-codebook valuation bound. For $U = \langle u_1, \ldots, u_d \rangle$, coefficients $\lambda_i \in \mathbb{F}_{127}$ impose $\alpha = \sum_i \lambda_i u_i \neq 0$, so one model covers the entire nonzero input-mask family. If the support model is infeasible, every covered linear sum vanishes. Otherwise, the conversion gives the certified full-codebook bound minus d on every coset.

B.1 Prime-Field ToySPN

The prime-field examples use a one-column AES-prime analogue over $\mathbb{F}_{13}^4$. Its state is (x_0, x_1, x_2, x_3), and its permutation S-box is $S(x) = x^5 + 2$ over $\mathbb{F}_{13}$. Since $\gcd(5, 12) = 1$, this substitution is bijective. Reducing the AES-prime column matrix modulo 13 gives

$$M_{13} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 \\ 1 & 4 & 3 & 2 \\ 1 & 3 & 2 & 4 \end{pmatrix}.$$

Let T_k denote coordinatewise translation by $k \in \mathbb{F}_{13}^4$. The 2-round permutation used in the examples is

$$E_{k_0,k_1} = S^{\times 4} \circ T_{k_1} \circ M_{13} \circ S^{\times 4} \circ T_{k_0}.$$

Outputs are taken immediately after the second S-box layer. The two round keys are independent uniform elements of $\mathbb{F}_{13}^4$ in the sampled experiments.

Example 2 (Prime-field integral sums). We apply the conversion to this prime-field ToySPN. For $U = \langle e_0 \rangle$ and $\beta = (1, 9, 0, 2)$, the support model proves vanishing on the full mask family and hence on all thirteen cosets. For the control $U = \langle e_0 + e_3 \rangle$, $\beta = (0, 1, 8, 6)$, exact checks instead exhibit nonzero coset sums. The linear model gives a valuation bound of $2/3$, which converts to $-1/3$ and is weaker than $\operatorname{ord}_{13} I \geq 0$, since I is an algebraic integer. Figure 1 compares the exact coset-sum valuations for sampled keys with those for sampled random permutations.

The AES-prime application and its comparison with characteristic-p integral cryptanalysis [BV25] follow in Subsection B.2.

B.2 AES-prime: Additive-Character Bounds and Exact Uniformity

Specification. AES-prime [MMMS23] acts on $\mathbb{F}_{127}^{16}$. We pack the coordinates $x_0, \ldots, x_{15}$ row-wise into a 4×4 array; they are field elements, not bits. The S-box is $S(x) = x^5 + 2$ over $\mathbb{F}_{127}$. It is a permutation because $\gcd(5, 126) = 1$. The round architecture follows AES: coordinatewise key addition, parallel substitution, row shifts, and column mixing. Row j is rotated left by j cells. The column matrix is

$$M_{127} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 16 \\ 1 & 4 & 16 & 2 \\ 1 & 16 & 2 & 4 \end{pmatrix} \quad \text{over } \mathbb{F}_{127}.$$

The final round omits column mixing. Our reported output masks are at the last S-box output, before the final row shift and key addition. The additive-character masks are vectors over $\mathbb{F}_{127}$. The analytical bounds allow arbitrary round-key translations and therefore cover the specified AES-prime key schedule.

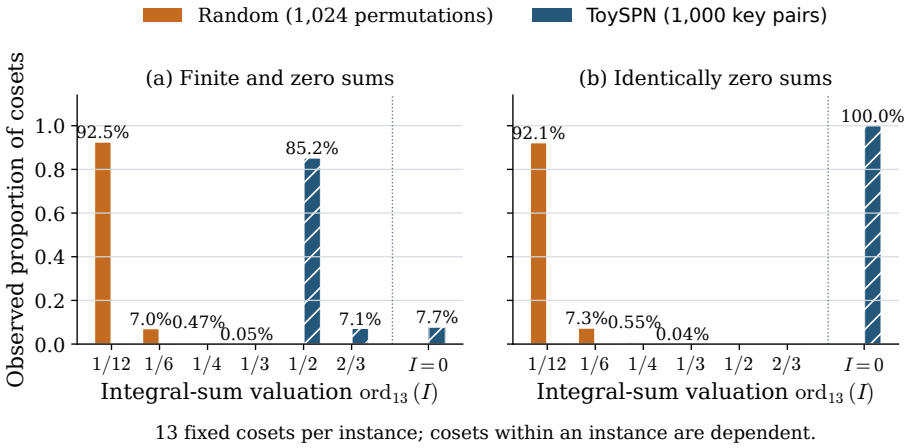


Fig. 1: Prime-field ToySPN versus random permutations. Both panels use the same independently sampled round-key pairs and random permutations. The separation in (a) is empirical. The vanishing in (b) also has a symbolic proof.

Results. AES-prime illustrates how our conversion applies to additive-character sums beyond binary fields. Let $F_{K,r}$ denote its r-round map for round-key tuple K, ending at the last S-box layer. All results hold for arbitrary round keys and every fixed value $\lambda \in \mathbb{F}_{127}$, with data complexity 127^{15}. For output mask β, define

$$N_z = \#\{x \in \mathcal{X} : \beta \cdot F_{K,r}(x) = z\}, \qquad I_\beta^{(127)}(\mathcal{X}; F_{K,r}) = \sum_{z=0}^{126} N_z \zeta_{127}^z.$$

We report $\mathrm{ord}_{127}\, I_\beta^{(127)}$, normalized by $\mathrm{ord}_{127}(127) = 1$. These fractional valuations are not binary one-count exponents. Let $u = e_0 + e_5 + e_{10} + e_{15}$ and $b = e_3 + e_7 + e_{11} + e_{15}$. Table 6 compares converted linear bounds with structural bounds from repeated contributions to the character sum.

Table 6: AES-prime character-sum valuation bounds from linear conversion and input multiplicities.

Rounds	Affine input structure	Data	β	Converted linear	Structural
1	$x_0 = \lambda$	127^{15}	e_0	895/63	15 (exact)
1	$x_0 + x_5 + x_{10} + x_{15} = \lambda$	127^{15}	u	745/63	12
2	$x_3 = \lambda$	127^{15}	b	695/63	12

For the rows of Table 6, the full-codebook bounds are respectively 958/63, 808/63, and 758/63; conversion subtracts 1 since $\dim U = 1$. We obtain the first two by exact local character arithmetic and the third by the prime-field linear valuation MILP. For the structural bounds, let D contain the input coordinates on which the selected output depends. If the projection of A onto D has rank q, each assignment in the projected structure has 127^{15-q} preimages in $\mathcal{X}$. Factoring out this multiplicity leaves an algebraic integer, giving $\operatorname{ord}_{127} I_\beta^{(127)} \geq 15 - q$. The three rows have $q = 0, 3, 3$. For the first row of Table 6, the selected output is constant, giving valuation exactly 15. For the 2- and 3-round cases in Table 7, whole-subspace support proofs establish vanishing character sums, which imply uniform output counts.

Table 7: AES-prime exact output uniformity from whole-subspace support proofs and conversion. The count holds for each $z \in \mathbb{F}_{127}$, every round-key tuple, and every scalar λ.

Rounds	Affine input structure	Data	β	Exact N_z
2	$x_0 = \lambda$	127^{15}	e_0	127^{14}
3	$x_0 = \lambda$	127^{15}	e_0	127^{14}

For both rows of Table 7, our support model proves that the linear character sums vanish for $U = \langle e_0 \rangle$ and output mask e_0. By Proposition 1, the integral character sum is therefore zero on every coset. The minimal polynomial $1 + T + \cdots + T^{126}$ then implies equality of all 127 output counts. This is exact uniformity, not merely an output sum equal to zero in $\mathbb{F}_{127}$. The 2-round vanishing result also holds for every input-mask space $U = \langle e_i \rangle$ and output mask $\beta = e_j$.

Beyne and Verbauwhede study AES-prime using characteristic-p integral cryptanalysis [BV25]. Their integer-count divisibility results concern a different quantity from our additive-character sums. Our examples illustrate the prime-field conversion rather than improve AES-prime round coverage.

B.3 Prime-Field Random-Permutation Law

We now evaluate how often a uniformly random permutation satisfies the prime-field conditions established above. Let $Q = p^\ell$, $M = p^{\ell-d}$, and Π be a uniform random permutation on $\mathbb{F}_p^\ell$. For fixed $\mathcal{X}$ of size M and $\beta \neq 0$, let $I = I_\beta^{(p)}(\mathcal{X}; \Pi)$. Let N_j count images in output class j, defined by $\beta \cdot y = j$; each class has size Q/p. Summing over nonnegative integer counts gives

$$\Pr[\operatorname{ord}_p I \geq s] = \binom{Q}{M}^{-1} \sum_{\substack{n_0 + \cdots + n_{p-1} = M \\ \operatorname{ord}_p(\sum_j n_j \zeta_p^j) \geq s}} \prod_{j=0}^{p-1} \binom{Q/p}{n_j}. \tag{33}$$

The proof is the random-subset argument of Proposition 2, now with p output classes. Exact uniformity is equivalent to $I = 0$; the same counting argument gives

$$\Pr[N_j = M/p \text{ for all } j] = \frac{\binom{Q/p}{M/p}^p}{\binom{Q}{M}},$$

when $p \mid M$. For the AES-prime rows in Table 7, this is $\binom{127^{15}}{127^{14}}^{127} / \binom{127^{16}}{127^{15}}$.

To avoid summing over all count vectors, we use a necessary congruence on a single output count. Assume $\ell \geq 2$, $d = 1$, $s \geq 1$, and let $k = \lceil s \rceil$. Write $\pi = \zeta_p - 1$ and

$$I = \sum_{j=0}^{p-2} (N_j - N_{p-1}) \zeta_p^j = \sum_{t=0}^{p-2} a_t \pi^t, \qquad a_0 = M - pN_{p-1}.$$

The valuations $v_p(a_t) + t/(p-1)$ of the nonzero summands have distinct fractional parts. When $I \neq 0$, their minimum is attained uniquely and equals $\operatorname{ord}_p I$; when $I = 0$, all a_t vanish. Thus $\operatorname{ord}_p I \geq s$ requires $v_p(a_0) \geq k$, hence $N_{p-1} \equiv M/p \pmod{p^{k-1}}$. The count N_{p-1} is hypergeometric with a mode at M/p. Lemma 2 gives

$$\Pr[\operatorname{ord}_p I \geq s] \leq p^{-(k-1)} + \frac{\binom{Q/p}{M/p}\binom{Q-Q/p}{M-M/p}}{\binom{Q}{M}}.$$

This bounds the valuation event, not exact uniformity. For output masks of rank h, the exact joint law uses p^h output classes of size Q/p^h and imposes the conditions on their character sums using the same count vector. This retains dependencies between masks rather than assuming independence.

C Probability Proofs and Pair Laws

C.1 The Residue Deviation Bound

Lemma 2 (Residue probability for a unimodal distribution). *Let X be integer-valued, with probabilities $p_j = \Pr[X = j]$ nondecreasing for $j \leq 0$ and nonincreasing for $j \geq 0$. For every integer $q \geq 1$,*

$$\left| \Pr[X \equiv 0 \pmod{q}] - \frac{1}{q} \right| \leq p_0.$$

Proof. The case $q = 1$ is immediate. Assume $q > 1$ and let $\pi_r = \sum_{k \in \mathbb{Z}} p_{r+kq}$. For $0 < r < q$,

$$\pi_0 - \pi_r = \sum_{k \geq 0} (p_{kq} - p_{kq+r}) - \sum_{k \geq 1} (p_{-kq+r} - p_{-kq}).$$

Both sums are nonnegative and at most p_0: for the first, sum

$$0 \leq p_{kq} - p_{kq+r} \leq p_{kq} - p_{(k+1)q}$$

over $k \geq 0$ and telescope; the negative tail follows similarly. Their difference therefore has absolute value at most p_0. Since q^{-1} is the average of the residue masses,

$$\left|\pi_0 - \frac{1}{q}\right| \leq \frac{1}{q}\sum_{r=1}^{q-1} |\pi_0 - \pi_r| \leq p_0.$$

The centered balanced hypergeometric law has mode zero and modal mass p_{ctr}. Taking $X = N_1 - M/2$ and $q = 2^{t_I - 1}$ proves Proposition 3. The lemma does not require symmetry and also gives the prime-field estimate in Subsection B.3.

C.2 Distribution of a DDT Entry

Proposition 8 (Exact one-entry DDT law). *Let Π be a uniform random permutation on V, with $Q = |V| = 2^n$ and $H = Q/2$, and fix $\delta, \gamma \neq 0$. Let a_j count perfect matchings (partitions into disjoint unordered pairs) on $2j$ vertices that avoid every edge of a fixed perfect matching:*

$$a_j = \sum_{i=0}^{j} (-1)^i \binom{j}{i} (2(j-i)-1)!!, \qquad (-1)!! = 1.$$

For each integer $0 \leq j \leq H$,

$$\Pr[\mathrm{DDT}_\Pi(\delta, \gamma) = 2j] = \frac{\binom{H}{j} a_{H-j}}{(2H-1)!!}. \tag{34}$$

Odd counts have probability zero.

Proof. The images of the input pairs $\{x, x+\delta\}$ under Π form a uniform random perfect matching. The pairs $\{y, y+\gamma\}$ form a fixed matching with H edges, and $\mathrm{DDT}_\Pi(\delta, \gamma)/2$ counts the common edges. There are $(2H-1)!!$ perfect matchings in total. To intersect the fixed matching in exactly j edges, first choose these edges in $\binom{H}{j}$ ways. On the remaining $2(H-j)$ vertices, forbid the remaining $H-j$ fixed edges. Inclusion–exclusion over these forbidden edges gives a_{H-j} allowed matchings. Thus $\binom{H}{j} a_{H-j}$ matchings have exactly j common edges; dividing by $(2H-1)!!$ gives the probability. Each common edge contributes the two inputs x and $x+\delta$, so every DDT entry is even.

C.3 Exact One-Mask DL Law

Let $(a)_b = a(a-1)\cdots(a-b+1)$, with $(a)_0 = 1$ and $(a)_b = 0$ for $b > a \geq 0$.

Proposition 9 (Exact one-mask DL law). *Let Π be a uniform random permutation on V, $\mathcal{X} = a + A$, and $\delta, \beta \neq 0$. Let $Q = |V|$ and $M = |\mathcal{X}|$. As in Subsection 4.4, let $H_{\mathcal{X}}$ count distinct plaintext pairs and κ their multiplicity.*

The number J of distinct unordered pairs with different output parities has, for integer $0 \le j \le H_{\mathcal{X}}$, probability

$$\Pr[J = j] = \frac{\binom{H_{\mathcal{X}}}{j} 2^j}{(Q)_{2H_{\mathcal{X}}}} \sum_{k=0}^{H_{\mathcal{X}}-j} \binom{H_{\mathcal{X}} - j}{k} (Q/2)_{j+2k} (Q/2)_{2H_{\mathcal{X}}-j-2k}. \tag{35}$$

Consequently $N_1^{\mathrm{DL}} = \kappa J$ and $I^{\mathrm{DL}} = M - 2\kappa J$.

Proof. The $H_{\mathcal{X}}$ pairs use $2H_{\mathcal{X}}$ distinct plaintexts; the output classes $\beta \cdot y = 0$ and 1 each have size $Q/2$. Choose the j mixed pairs (those with different output parities) and orient each in two ways. Among the remaining pairs, whose outputs have equal parity, choose k for class 0. The pattern uses $j + 2k$ outputs from class 0 and $2H_{\mathcal{X}} - j - 2k$ from class 1. The falling factorials count injective assignments of actual outputs. The restriction of Π to the involved inputs is uniform among $(Q)_{2H_{\mathcal{X}}}$ injections.

We evaluate the DL signed-sum event by summing Equation 35 over j satisfying $2^{t_I} \mid M - 2\kappa j$. The one-count event uses $2^{t_N} \mid \kappa j$. When $\delta \in A$, each unordered pair is counted twice, explaining $\kappa = 2$. When $\mathcal{X}$ is a halfspace and $\delta \notin A$, the pairs cover the full input space, although each pair is counted only once.

Proposition 10 (Cross-coset half-codebook DL bound). *Let $Q = 2^n$ with $n \ge 3$, $M = Q/2$, and $\delta \notin A$. For a nonzero output mask and $2 \le t_N \le n - 2$, define*

$$h_{\mathrm{DL}}(Q) = \frac{\binom{Q/2}{Q/4}\binom{Q/4}{Q/8} 2^{Q/4}}{\binom{Q}{Q/2}}.$$

Then

$$\left| \Pr[2^{t_N} \mid N_1^{\mathrm{DL}}] - 2^{-(t_N - 1)} \right| \le h_{\mathrm{DL}}(Q), \qquad \Pr[2^{t_N} \mid N_1^{\mathrm{DL}}] \ge h_{\mathrm{DL}}(Q). \tag{36}$$

For every integer $t_I \ge 3$, the same statements hold for the event $v_2(I^{\mathrm{DL}}) \ge t_I$, with reference probability $2^{-(t_I - 2)}$. Within the count-threshold range above, $t_I = t_N + 1$ gives the equivalent event.

Proof. Let $Z = Q/2 - J$ count pairs whose outputs have equal parity. Each output class contains exactly $Q/2$ elements, so both Z and J are even, with

$$\Pr[Z = z] = \frac{\binom{Q/2}{z}\binom{z}{z/2} 2^{Q/2 - z}}{\binom{Q}{Q/2}} \quad (z \text{ even}).$$

Indeed, choose the z pairs with equal output parities, assign half to each class, and orient the mixed pairs. The ratio of successive nonzero masses is $(Q/2 - z)(Q/2 - z - 1)/(z + 2)^2$, giving a mode at $z = Q/4$. For $X = (Z - Q/4)/2$, we have $N_1^{\mathrm{DL}} = J = Q/4 - 2X$ and $I^{\mathrm{DL}} = 4X$. Since $2^{t_N} \mid Q/4$ in the stated range, the count event is $X \equiv 0 \pmod{2^{t_N - 1}}$; the reference probability thus accounts for the evenness of J. The signed-sum event uses modulus $2^{t_I - 2}$ for any $t_I \ge 3$. Apply Lemma 2 to obtain the deviation bounds. The mode $X = 0$ satisfies both divisibility conditions and has probability $h_{\mathrm{DL}}(Q)$, giving the lower bounds.

C.4 Concrete Random-Permutation Bounds

We apply the random-permutation laws to the count conditions reported in our application tables. Each probability concerns one fixed input structure and one fixed output mask under a uniform random permutation, with the structure and mask chosen independently of that permutation. For a proved all-key property, the cipher passes with probability one; the corresponding distinguishing advantage of this test is $1 - p_0$. For an empirical property, p_0 is the random-permutation probability, not a guarantee over untested cipher keys. The formulas apply equally to a fixed non-coordinate affine structure or a fixed inverse-round extension of the same size.

Certified Numerical Evaluation. For even $0 < M < Q$, the central hypergeometric mass satisfies

$$\sqrt{\frac{2Q}{\pi(M+1)(Q-M+1)}} < p_{\mathrm{ctr}}(Q,M) < \sqrt{\frac{2(Q+1)}{\pi M(Q-M)}} =: h(Q,M). \quad (37)$$

To verify this without evaluating enormous factorials, set $b_k = 4^{-k}\binom{2k}{k}$. Then $p_{\mathrm{ctr}} = b_{M/2}b_{(Q-M)/2}/b_{Q/2}$. The integral recurrence for $J_r = \int_0^{\pi/2} \sin^r x\,dx$, together with $J_{2k+1} < J_{2k} < J_{2k-1}$, gives $1/\sqrt{\pi(k+1/2)} < b_k < 1/\sqrt{\pi k}$. Substitution proves Equation 37. Consequently, for $M = 2^m$ and $0 \le t \le m-1$,

$$p_0 = \Pr[2^t \mid N_1] \le \min\{1, 2^{-t} + h(Q,M)\}. \quad (38)$$

For exact balance, use $p_0 = p_{\mathrm{ctr}}(Q,M) < h(Q,M)$ instead. These are rigorous inequalities, not the approximation $p_0 \approx 2^{-t}$. The numerical exponents below are rounded down: an entry a certifies $p_0 \le 2^{-a}$. We use rational bounds and directed rounding to certify the displayed exponents. For $M = 16$, we sum the exact hypergeometric probabilities instead.

Table 8 covers selected finite-count applications. The referenced result tables specify the masks and distinguish proof from observation. Table 9 treats exact balance separately. Each bound applies to one output mask. It is not a joint bound for several masks.

Table 8: Certified one-permutation bounds $p_0 \le 2^{-a}$ for $2^t \mid N_1$. Here n is the block size and the data complexity is 2^m. The result tables specify the refinements and empirical evidence.

Cipher	Rounds	n	m	t	a	Method / evidence	Results
SIMON32	9	32	29	5	4.99	Ultrametric [BV24]	Table 1
SIMON32	9	32	29	10	9.94	RC	Table 1
SIMON32	9	32	29	11	10.89	Joint	Table 1
SIMON48	10	48	44	6	5.99	Ultrametric [BV24]	Table 1
SIMON48	10	48	44	20	19.72	RC	Table 1
SIMON48	10	48	44	22	21.13	Joint	Table 1

Table 8 (continued)

Cipher	Rounds	n	m	t	a	Method / evidence	Results
SIMON32	13	32	31	2	1.99	Joint	Table 1
SPECK32	7 (8)	32	30	3	2.99	RC + residues	Table 21
SPECK48	7 (8)	48	45	8	7.99	RC + residues	Table 24
BEANIE $U_{2,3}$	5	32	31	10	9.96	Joint	Table 3
BEANIE $U_{3,2}$	5	32	31	12	11.86	Joint	Table 3
BEANIE $U_{3,3}$	6	32	31	1	0.99	Joint	Table 3
BEANIE $U_{3,3}$	6	32	31	2	1.99	Empirical	Table 3
PRESENT	4	64	4	1	0.99	Ultrametric [BV24]	Table 29
PRESENT	4	64	4	3	2.34	Joint	Table 29
PRESENT	5	64	12	3	2.86	Exact contraction	Table 29
PRESENT	5	64	62	26	25.95	Linear	Table 32
PRESENT	5	64	63	34	31.11	Linear + parity	Table 32
PRESENT	6	64	62	7	6.99	Linear	Table 32
PRESENT	6	64	63	18	17.99	Linear + parity	Table 32
Ascon	5	320	319	60	59.99	Linear	Table 34
Ascon	5	320	318	53	52.99	Linear	Table 34
Ascon	5	320	317	50	49.99	Linear	Table 34
Ascon	5	320	316	47	46.99	Linear	Table 34
GIFT-64	5	64	60	53	30.27	Linear / Joint	Table 35
GIFT-64	6	64	63	7	6.99	Linear	Table 35
GIFT-64	7	64	60	4	3.99	Joint	Table 36
GIFT-64	8	64	60	2	1.99	Spatial Joint	Table 36
GIFT-64	9	64	61	1	0.99	Spatial Joint	Table 36
RC5-16	2.5	32	31	1	0.99	Empirical	Table 44
ChiLow-32^{-1}	4	32	31	3	2.99	Empirical	Table 45
KATAN32	84	32	31	3	2.99	Empirical	Table 47
KeeLoq	64	32	31	14	13.51	Empirical	Table 49
KeeLoq	64	32	31	10	9.96	Empirical	Table 49

Table 9: Certified one-permutation bounds $p_0 \le 2^{-a}$ for exact $N_1 = M/2$. The exact probability is $p_{\mathrm{ctr}}(2^n, 2^m)$, not $2^{-(m-1)}$.

Cipher	Rounds	n	m	a	Results
SIMON32	10	32	31	15.32	Table 14
SPECK64	7 (8)	64	61	30.72	Table 25
SPECK96	7 (8)	96	93	46.72	Table 25
SPECK128	7 (8)	128	125	62.72	Table 25
SKINNY-64	8	64	12	6.32	Table 39
SKINNY-64	10	64	48	24.32	Table 39
SKINNY-64	11	64	60	30.27	Table 39
KeeLoq	64	32	31	15.32	Table 50

The GIFT-64 exponent 53 illustrates why the central mass matters: exact balance already belongs to this divisibility event and has probability approximately $2^{-30.28}$. Thus 2^{-53} would be an invalid upper bound. Conversely, Ascon's central mass is far smaller than 2^{-60}, making the residue estimate informative at the claimed exponent. For KeeLoq, the two-point count sets in Subsection F.11 permit sharper tests than divisibility alone: for $d \in \{2^{14}, 2^{10}\}$, their exact null probability is

$$\frac{2\binom{2^{31}}{2^{30}-d}\binom{2^{31}}{2^{30}+d}}{\binom{2^{32}}{2^{31}}}.$$

The prime-field exact-uniformity probability for AES-prime is given separately in Subsection B.3. For BEANIE, we compare each analyzed U-shaped permutation with a uniformly random 32-bit permutation. In particular, for distinct tweaks, replacing the two branches by independent uniform permutations Π_1, Π_2 makes $\Pi_2^{-1} \circ \Pi_1$ uniform as well: conditional on Π_2, composition is a bijection on the permutation space. For the halfspace count events in Table 26 and Table 3, Proposition 3 therefore gives $p_0 \le 2^{-t} + p_{\mathrm{ctr}}(2^{32}, 2^{31})$. This is a one-permutation probability; repeated-instance bounds additionally require independent ideal instances.

C.5 Numerical DL and DDT Bounds

For cross-coset half-codebook DL, the central-binomial inequalities give

$$h_{\mathrm{DL}}(Q) < \sqrt{\frac{16(Q+1)}{\pi Q^2}}.$$

We substitute this bound into Proposition 10. When the pairs use fewer than Q plaintexts, we bound the DL pass probability without evaluating the full distribution. Write $H = H_{\mathcal{X}}$, with $0 < 2H < Q$ and $2^{t_N} \ge \kappa$. As an auxiliary construction, assign independent fair bits to the Q plaintexts. Conditioning on exactly $Q/2$ ones gives the distribution of the output parities $(\beta \cdot \Pi(x))_{x \in V}$. Before conditioning, $J \sim \mathrm{Bin}(H, 1/2)$ and $2^{t_N} \mid \kappa J$ means $J \equiv 0 \pmod{2^{t_N}/\kappa}$. For even H with $2^{t_N} \mid \kappa H/2$, this residue contains the mode; Lemma 2 bounds its probability by $\kappa/2^{t_N} + \sqrt{2/(\pi H)}$. For any assignment on the paired plaintexts, the remaining $Q - 2H$ bits complete the balance with probability at most their central binomial mass. Dividing by the central binomial mass on all Q bits gives

$$\Pr[2^{t_N} \mid N_1^{\mathrm{DL}}] \le \min\left\{1, \sqrt{\frac{Q+1}{Q-2H}}\left(\frac{\kappa}{2^{t_N}} + \sqrt{\frac{2}{\pi H}}\right)\right\}, \quad 2^{t_N} \ge \kappa. \tag{39}$$

The conditioning accounts for sampling without replacement, without assuming independent permutation outputs. Table 10 applies Equation 36 and Equation 39 to the PRESENT and GIFT-64 results in Table 33 and Table 38.

Table 10: Certified random-permutation probabilities for DL count divisibility on 64-bit permutations. The exponent a certifies $\Pr[2^t \mid N_1^{\mathrm{DL}}] \le 2^{-a}$, not equality. The base structure has size 2^m. Queries count distinct plaintexts.

Cipher	Rounds	Pair setting	m	t	Queries	a
PRESENT	5	Cross-coset	60	41	2^{61}	30.22
PRESENT	5	Cross-coset	62	43	2^{63}	30.82
PRESENT	5	Cross-coset	63	44	2^{64}	30.82
PRESENT	6	Cross-coset	63	9	2^{64}	7.99
GIFT-64	5	Same-coset	60	15	2^{60}	13.95
GIFT-64	5	Same-coset	62	21	2^{62}	19.79
GIFT-64	5	Same-coset	63	26	2^{63}	24.48

DDT Counts and Zero Entries. The DDT law has substantial mass at zero, so count divisibility must not be assigned probability 2^{-t}. We evaluate Proposition 8 using binomial moments with a certified truncation error. Let $J = \mathrm{DDT}_{\Pi}(\delta,\gamma)/2$ and $H = Q/2$. Its binomial moments are

$$\mu_k = \mathbb{E}\binom{J}{k} = \frac{\binom{H}{k}}{(Q-1)(Q-3)\cdots(Q-2k+1)}, \qquad \mu_0 = 1,$$

because fixing k edges leaves a uniform matching on the remaining vertices. Binomial inversion gives $\Pr[J=j] = \sum_{k=j}^{H}(-1)^{k-j}\binom{k}{j}\mu_k$. For a truncation at $L < H$, the error for $j \le L$ is at most $\binom{L+1}{j}\mu_{L+1}$, and $\Pr[J > L] \le \mu_{L+1}$. Indeed, for a fixed value $J = z > L$, the truncated alternating sum has magnitude $\binom{z}{j}\binom{z-j-1}{L-j} \le \binom{L+1}{j}\binom{z}{L+1}$; averaging proves the first bound. Thus any event on J can be evaluated with total error at most $2^{L+1}\mu_{L+1}$, by summing its truncated masses with $j \le L$. For $J = 0$, an even truncation is already an upper bound by inclusion–exclusion.

With $Q = 2^{32}$ and $L = 32$, rational evaluation gives

$$\begin{aligned} &\Pr[\mathrm{DDT} = 0] < 0.606531, \\ &\Pr[4 \mid \mathrm{DDT}] \ < 0.683940, \\ &\Pr[2^t \mid \mathrm{DDT}] < 0.606531, \quad t \in \{10, 11\}. \end{aligned}$$

Zero counts dominate the high-divisibility events in the last line, which evaluates the RC5 predicates in Table 42 without requiring positive counts. These are fixed-cell probabilities, not search-adjusted bounds for the selected candidates. As an illustrative scan bound, the probability that at least one of the $32(2^{32}-1)$ cells with Hamming-weight-1 input differences is zero for all 100 independent ideal permutations is at most $2^{-35.13}$, by Equation 40. This illustrative sample size differs from the 1,000-key RC5 experiment. The union bound requires no independence between cells and gives no all-key cipher guarantee.

Cost of Exact Evaluation. For $M = 2^m \le Q/2$ and $0 \le t \le m-1$, the one-count event has $M/2^t + 1$ admissible hypergeometric terms; exact balance has one. For

an output family of rank h, the joint model has at most $\binom{M+2^h-1}{2^h-1}$ nonnegative class-count vectors before population and event constraints. These are summand counts, not unit-cost arithmetic estimates: the binomial integers can themselves be very large. This explains the use of certified bounds for large instances and exact joint evaluation for the toy examples.

C.6 Repeated Tests and Dependent Structures

For the black-box search in Subsection 5.4 (Algorithm 1), we assess how often random permutations could satisfy the discovered properties and what successful tests imply about untested cipher keys.

Accounting for the Search. Testing more candidates can increase the chance of finding one that passes. For a fixed searched family $\mathcal{H}$, let q_h be candidate h's pass probability for one uniform random permutation. For r independent such permutations, the union bound gives

$$\Pr[\exists h \in \mathcal{H} \text{ passing all } r] \le \min\left\{1, \sum_{h\in\mathcal{H}} q_h^r\right\}. \tag{40}$$

The candidate events may be dependent, but the r permutations must be independent.

Here $\mathcal{H}$ includes every searched candidate, not only those reported. If candidates are selected using separate data, we condition on that selection: $\mathcal{H}$ may contain only those selected, and r counts only fresh independent tests under the random-permutation model.

Evidence across Cipher Keys. A small random-permutation probability does not prove validity for every cipher key. Fix a candidate before testing r independent uniform cipher keys. If it fails on a fraction θ of keys, all tests pass with probability $(1-\theta)^r$. With no observed failures, the one-sided confidence bound at level $1-\eta$, for $0 < \eta < 1$, is

$$\theta \le 1 - \eta^{1/r} \tag{41}$$

This confidence level applies to one candidate, not simultaneously to all reported candidates. It concerns the failure fraction, not the probability of an all-key property. Correlated keys and constrained key/tweak samples require separate models.

Remark 2 (Dependent cosets). Different cosets of one permutation are not independent trials: their images are disjoint subsets of one output population. For complementary halfspaces and $\beta \neq 0$, the two one-counts sum to $Q/2$. Whenever $2^t \mid Q/2$, divisibility by 2^t on one halfspace is equivalent to divisibility on the other; exact balance is likewise equivalent. Their joint pass probability is the one-halfspace probability, not its square. For larger coset families, use an established joint law or the minimum of the marginal upper bounds, never an unsupported product. The same caution applies to different output masks, equivalent linear/integral tests, and shift-related cipher configurations.

Validation Budget. For a fixed nonempty family $\mathcal{H}$, suppose each candidate has pass probability at most $\bar{q}$, where $0 < \bar{q} < 1$. Then Equation 40 gives the upper bound $|\mathcal{H}|\bar{q}^r$ after r independent ideal instances. For $0 < \varepsilon < 1$,

$$r \geq \left\lceil \frac{\log(|\mathcal{H}|/\varepsilon)}{-\log \bar{q}} \right\rceil \tag{42}$$

is sufficient for a family-wise probability at most ε. For one preselected candidate, set $|\mathcal{H}| = 1$; after discovery, apply the fresh-sample convention above. If a candidate must pass all r_{d} discovery and r_{v} validation instances, we may instead use $|\mathcal{H}|\bar{q}^{r_{\mathrm{d}}+r_{\mathrm{v}}}$, provided $\mathcal{H}$ is the entire predetermined search family and all instances are independent. Repeated evaluations on the same keys and tests on dependent offsets do not increase the number of independent instances.

For the 84-round KATAN32 event in Table 47, Equation 38 gives a 1,000-instance bound below $2^{-2999.71}$ for one fixed candidate. A hypothetical predetermined sweep over all 32 coordinate halfspaces and all $2^{32} - 1$ nonzero output masks at that round and threshold would instead have bound

$$32(2^{32} - 1)\left(2^{-3} + h(2^{32}, 2^{31})\right)^{1000} < 2^{-2962.71}.$$

This example states the search family explicitly; it does not assume the reported rows comprise the entire discovery search or that the two offsets are independent. The corresponding real-key zero-failure statement remains Equation 41: at 95% confidence, 1,000 independent uniform-key tests of a fixed candidate bound its failure fraction by about 0.299%, not by $2^{-2999.71}$.

D Illustrative Examples and Validation

The binary examples use the ToySPN specified in Subsection D.1. They illustrate the conversions of Section 3 and the probability analysis of Section 4. Each 4,096-key sample used for the binary distribution figures contains the fixed keys 0 and 1 and 4,094 pseudorandom distinct 80-bit keys.

D.1 ToySPN Specification

ToySPN is the adjustable small-state SPN used for exact validation. An instance with m parallel four-bit S-boxes and r rounds has a $4m$-bit state. Each round applies a round-key XOR, the parallel S-box layer, and

$$P_m(i) = mi \bmod (4m - 1) \quad (0 \leq i < 4m - 1), \qquad P_m(4m - 1) = 4m - 1.$$

A final whitening-key XOR follows the last permutation. All masks are full-state hexadecimal words with global least-significant-bit indexing. Unless stated otherwise, outputs are taken after the last permutation and whitening. The PRESENT-S-box instantiation reduces to the usual PRESENT round structure when $m = 16$ [BKL+07]. The DL counterexamples also use the KNOT S-box [ZDY+19]. Both substitutions appear in Table 12. Any independently sampled round keys in an exact-validation example are part of its stated toy model, not a claim about a real cipher key schedule.

D.2 A Finite Single-Mask Property

We consider the 3-round, 16-bit ToySPN with four parallel PRESENT S-boxes. For input mask $\alpha = \texttt{7000}$ and ciphertext output mask $\beta = \texttt{000e}$, we use the halfspaces $\mathcal{X}_b = \{x : \alpha \cdot x = b\}$, $b \in \mathbb{F}_2$, each with data complexity 2^{15}. The fixed-mask linear MILP proves $v_2(L_{\alpha,\beta}) \geq 10$ for every key. By Equation 6, both signed integral sums have valuation at least 9, equivalently $2^8 \mid N_1$ by Equation 1. Exact input enumeration on each of 4,096 tested keys gave linear sums $\pm 2^{10}$ and coset sums $\pm 2^9$, so none of these sampled sums was zero. These exact values are observed. The model certifies the all-key lower bound. For a random permutation, the probability of $2^9 \mid I$ on either fixed halfspace is approximately $2^{-7.32}$ by Equation 18. Figure 2 compares the sampled signed-sum valuations with this exact law. Each key contributes one valuation because its two halfspace sums are opposite.

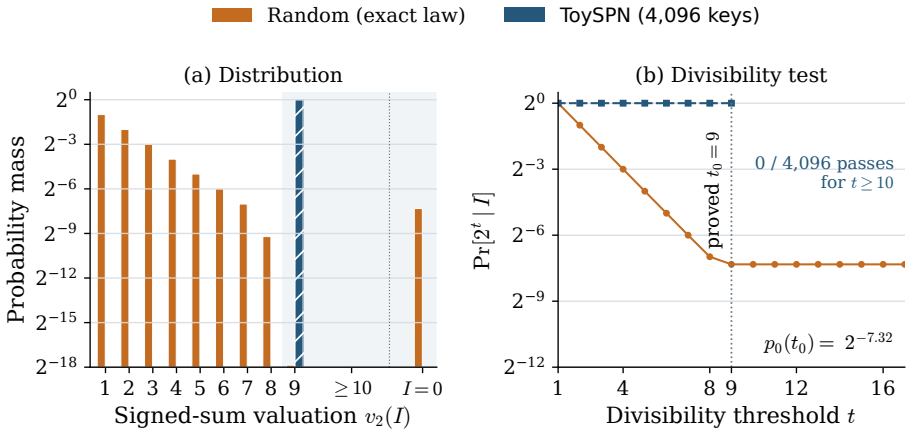


Fig. 2: Signed integral sums for the 3-round ToySPN. The bound holds for every key. The cipher distribution is empirical. Zero sums satisfy every divisibility threshold and are excluded from the pooled finite bin.

D.3 DL Equalization and Contrasts

For a 16-bit, 3-round ToySPN, set $\delta = \texttt{0001}$, $\beta = \texttt{0002}$ at the last S-box output, and $U = \langle \texttt{0002} \rangle$. The support model proves $L^{\mathrm{DL}}_{\texttt{0002},\delta,\beta} = 0$. The two coset sums are therefore equal to half the zero-mask DL sum by Theorem 2. All 1,000 exact sampled-key evaluations give full sum -2^{14} and coset sums -2^{13}. The equality is proved; these particular values are observed.

For $\delta = \texttt{0001}$, $\beta = \texttt{0020}$, and $U = \langle \texttt{0020} \rangle$, the contrast identity instead links a generally nonzero coefficient to the difference of the two coset sums. At the

zero key the exact coset sums are 4704 and 1440. Their difference 3264 equals the generalized DL sum. This example makes the distinction between a coset value and a relation among cosets explicit.

D.4 Joint Probabilities for Multiple Output Masks

For the same 3-round ToySPN, take

$$U = \langle \texttt{0060} \rangle, \qquad \mathcal{B} = \{\texttt{000b}, \texttt{0005}, \texttt{000e}\}.$$

The three output masks have rank two. The respective linear bounds $(10, 9, 8)$ give signed-coset bounds $(9, 8, 7)$. The exact joint probability from Equation 20 is approximately $2^{-19.30}$, whereas the minimum of the one-mask probabilities gives only the upper bound approximately $2^{-7.32}$. The gain comes from evaluating the divisibility conditions jointly, not multiplying individual probabilities.

For a five-S-box, 4-round instance, we use

$$U = \langle \texttt{00b00} \rangle, \qquad \mathcal{B} = \{\texttt{10000}, \texttt{00800}, \texttt{10800}\}$$

and prove signed-coset thresholds $(12, 12, 12)$. Its exact joint probability is approximately $2^{-26.98}$, compared with the one-mask value approximately $2^{-9.33}$. These rounded exponents describe exact-law evaluations, not asserted upper bounds after rounding. The event is too rare to validate its probability directly with the stated sample sizes. For the representative mask $\beta = \texttt{10000}$, exact enumeration on 4096 tested keys and both cosets gives signed-sum valuation 12 throughout. Figure 3 compares these observations with the exact one-mask random-permutation law. It does not plot the joint three-mask event.

D.5 Variable Valuations and the Baseline Approximation

A four-S-box, 4-round example with $U = \langle \texttt{000f} \rangle$ and last-S-box mask `2000` has a refined linear bound 3 and hence signed integral bound 2. The base linear bound is 2; complete trail-parity enumeration gives the extra factor of 2 (Subsection E.4). On 4096 tested keys, the coset valuations vary and their minimum is 3. The model bound is therefore not the exact empirical minimum. Figure 4 contrasts this distribution with the single-point pattern in Figure 2.

Both examples share the baseline in Figure 5: the approximation of Proposition 3 is accurate at $t = 2$, whereas the exact-balance probability dominates at $t = 9$. Agreement with Monte Carlo does not establish a rare tail below the resolution of the sample size; the formula and its proof do.

E Search Models and Supporting Proofs

We give the local encodings and refinements for the linear and Joint models, followed by the DL counterexamples and exact mix-basis models. We then present the black-box algorithm and the scope of its complete coordinate-cube searches.

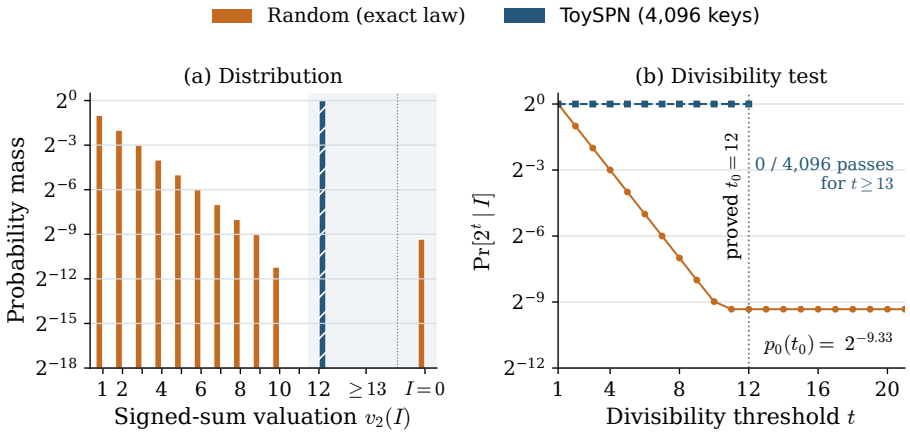


Fig. 3: 4-round, five-S-box ToySPN comparison for $U = \langle 00b00 \rangle$ and $\beta = 10000$. The data complexity is 2^{19} per coset. Each key contributes one observation because the two cosets have equal signed-sum valuations.

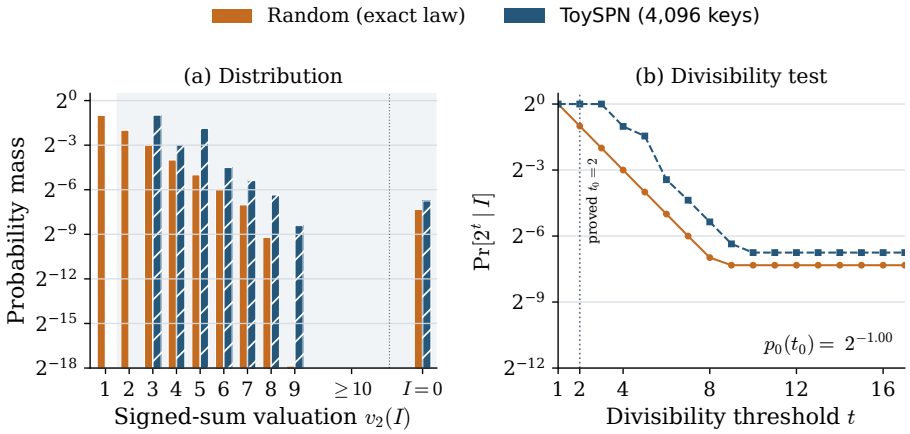


Fig. 4: Signed integral sums for the 4-round, four-S-box ToySPN. The cipher plateau includes the 38 sampled zero sums, which satisfy every divisibility threshold.

E.1 Local Linear Valuation Tables

A local linear valuation table records $v_2(L_{\alpha,\beta}(S))$ for each mask pair, whereas its support table records only whether the sum is nonzero. Table 11 gives the valuation table for the PRESENT S-box (Subsection F.4), used by the linear models of Subsection 5.1. Its entries are valuations of unnormalized sums. Since the S-box has 4 input bits, subtracting 4 gives the normalized transition cost

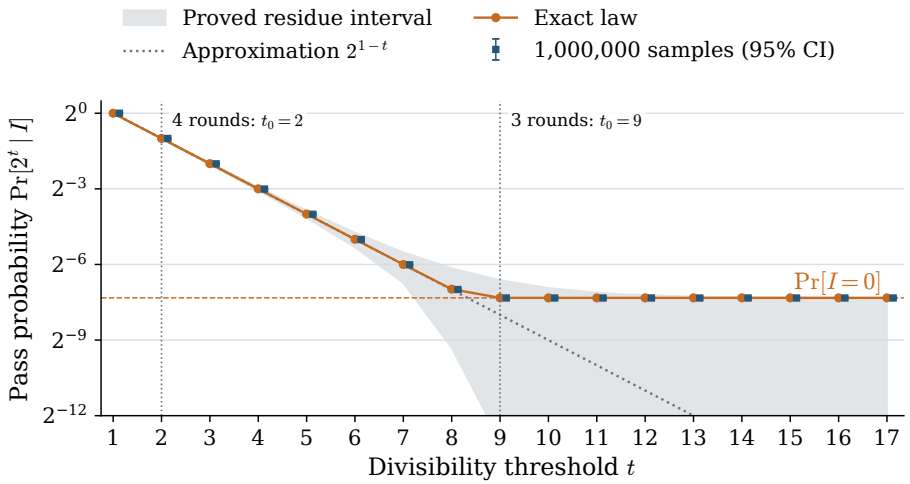


Fig. 5: Shared random-permutation baseline for the 3- and 4-round four-S-box examples: $n = 16$ and data complexity 2^{15}. Samples are independent draws from the one-coset hypergeometric law. Error bars are pointwise 95% Wilson intervals, not simultaneous confidence bands.

in Equation 22: finite entries $2, 3, 4$ give costs $-2, -1, 0$. Entries ∞ denote zero coefficients; both the support SAT model and the weighted MILP exclude these transitions. The SAT model uses only the remaining mask pairs, whereas the MILP also uses their costs. Listing 1 shows how to generate constraints encoding the allowed transitions and their costs.

Table 11: Local linear valuation table $v_2(L_{\alpha,\beta}(S))$ for the PRESENT S-box. Row and column labels are complete 4-bit masks in hexadecimal. ∞ denotes a zero coefficient.

$\alpha\backslash\beta$	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	4	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞
1	∞	∞	∞	∞	∞	3	∞	3	∞	∞	∞	∞	∞	3	∞	3
2	∞	∞	2	2	2	2	∞	∞	2	2	∞	3	∞	3	2	2
3	∞	∞	2	2	2	2	3	∞	2	2	3	∞	∞	∞	2	2
4	∞	∞	2	2	2	2	∞	3	2	2	∞	3	∞	∞	2	2
5	∞	∞	2	2	2	2	∞	∞	2	2	3	∞	3	∞	2	2
6	∞	∞	∞	3	∞	∞	3	∞	∞	3	∞	∞	3	∞	∞	∞
7	∞	∞	∞	3	3	∞	∞	∞	∞	3	∞	∞	∞	∞	3	∞
8	∞	∞	2	2	∞	∞	2	2	2	2	∞	∞	2	2	3	3
9	∞	3	2	2	∞	∞	2	2	2	2	3	∞	2	2	∞	∞
a	∞	∞	3	∞	2	2	2	2	∞	∞	∞	3	2	2	2	2
b	∞	3	∞	∞	2	2	2	2	3	∞	∞	∞	2	2	2	2
c	∞	∞	∞	∞	2	2	2	2	3	∞	∞	3	2	2	2	2
d	∞	3	3	∞	2	2	2	2	∞	∞	∞	∞	2	2	2	2
e	∞	∞	2	2	3	3	2	2	2	2	∞	∞	2	2	∞	∞
f	∞	3	2	2	∞	∞	2	2	2	2	3	∞	2	2	∞	∞

E.2 Searching for Input Structures

We search for affine plaintext structures whose one-count divisibility reaches a prescribed bound within a given data budget. We fix the cipher family $\{E_K\}$, consisting of permutations, its round count, an output mask $\beta \neq 0$, and an initial mask space U_0. Sparse masks and coordinate spaces provide possible starting points. We certify every retained space with the whole-subspace model of Subsection 5.1, not by checking only its basis vectors.

For $U \leq U_0$ of dimension $1 \leq d < n$, let $T(U)$ be a proved common lower bound on the unnormalized linear-sum valuation for all nonzero masks in U and all allowed keys. The structures $\mathcal{X} = a + U^{\perp}$ give data complexity 2^{n-d}, and Theorem 1 gives the one-count bound

$$t(U) = \max\{0, \min\{n - d - 1, T(U) - d - 1\}\}.$$

Here $t(U) = 0$ means that no nontrivial divisibility is certified. The zero input mask requires no check because E_K is a permutation and $\beta \neq 0$. An infeasible support model instead proves exact balance on every coset. We use a minimization solver's objective lower bound, with the normalization in Proposition 4, to obtain $T(U)$. A feasible solution gives an upper bound on the model optimum and cannot replace that lower bound.

Excluding a Limiting Mask. Suppose solving the model to optimality returns a minimizing input mask, $\alpha^* = \sum_{i=1}^{d} \lambda_i^* u_i \neq 0$ in a basis $u_1, \ldots, u_d$ of U.

This mask limits the model bound. It need not attain the true valuation after cancellations. For each $c \in \mathbb{F}_2^d$ satisfying $c \cdot \lambda^* = 1$, form

$$U_c = \left\{ \sum_{i=1}^{d} \lambda_i u_i : c \cdot \lambda = 0 \right\}. \tag{43}$$

There are exactly 2^{d-1} such vectors c, each defining a distinct codimension-1 subspace. Each subspace in Equation 43 excludes α^* while preserving linearity, which deleting one mask would not do. We solve the whole-subspace model for each candidate and select the one with the largest proved $t(U_c)$, breaking ties by $T(U_c)$. The bound for U remains valid on each U_c and is retained if the new optimization gives a weaker lower bound.

We repeat this step until the requested count bound is reached or the prescribed data budget would be exceeded. We also stop if optimality is not established or $d = 1$; a zero-dimensional mask space would give the full codebook. Shrinking U by one dimension enlarges $U^\perp$ by one dimension and doubles the data, while decreasing the conversion loss by one. Thus the search uses higher data complexity to seek stronger divisibility bounds. This greedy search need not find the best subspace; its 2^{d-1} optimizations per step make it suitable for modest mask-space dimensions. It changes the input structure, unlike the residue refinement for fixed masks in Subsection E.4.

E.3 Quadratic Sums and SIMON Rank Encoding

We use the quadratic-sum argument underlying exact SIMON linear analysis [KLT15, Section 4]. Let $\varphi : \mathbb{F}_2^w \to \mathbb{F}_2$ be quadratic with $\varphi(0) = 0$ and polar matrix P, defined by $\varphi(x + z) + \varphi(x) + \varphi(z) = x^\mathsf{T} P z$. All ranks are over $\mathbb{F}_2$; the radical is $\ker P$. For $G(\gamma) = \sum_x (-1)^{\varphi(x)+\gamma\cdot x}$, squaring and substituting the difference variable gives

$$|G(\gamma)|^2 = 2^w \sum_{z \in \ker P} (-1)^{\varphi(z)+\gamma\cdot z}.$$

Since φ is linear on $\ker P$, the sum is nonzero exactly when $\gamma \cdot z = \varphi(z)$ for every $z \in \ker P$, and then $|G(\gamma)| = 2^{w-\operatorname{rank}(P)/2}$. Its support is an affine translate of $\operatorname{im} P$, equal to $\operatorname{im} P$ precisely when φ vanishes on $\ker P$.

For SIMON, the rotation offsets differ by 7, which is odd and coprime to every word size $w \in \{16, 24, 32, 48, 64\}$, all of which are even. Thus the edges $(i-1, i-8)$ form one even cycle, and since every edge joins indices of opposite parity, every selected-edge graph is bipartite. Partitioning coordinates by index parity writes its quadratic form as $x_{\text{even}}^\mathsf{T} B x_{\text{odd}}$. On $\ker P$, we have $B x_{\text{odd}} = 0$, so the form vanishes, giving the image condition in Equation 23.

Write $\rho = \operatorname{rank}(P)/2$ for the half-rank. A path of j selected edges contributes $\lceil j/2 \rceil$ to ρ; the complete even cycle gives $\rho = w/2 - 1$. After reordering edge bits b_i along the cycle, binary markers satisfying $z_i = b_i(1 - z_{i-1})$, with cyclic indices, count alternating edges in each run. If all $b_i = 1$, both alternating marker

assignments give $\sum_i z_i = w/2$. Thus $\rho = \sum_i z_i - h$, where h is a binary variable indicating whether every edge is selected. Binary-product constraints and the constraints defining h use $O(w)$ variables and constraints.

For the mix-basis round, we use Walsh mask (a, b) and moment exponent (p, q) in left–right word order. We substitute $z = y + f(x) + k$ in Equation 25 and sum over z. The coefficient vanishes unless $b \preceq p$; otherwise it equals

$$2^{-w-\mathrm{wt}(p)}(-1)^{\mathrm{wt}(b)+b\cdot k} \sum_{\substack{x\in\mathbb{F}_2^w \\ x_i=1 \text{ whenever } q_i=1}} (-1)^{a\cdot x+b\cdot f(x)}.$$

Fixing these coordinates to 1 leaves $w - \mathrm{wt}(q)$ free variables and a quadratic phase plus an affine term. The induced quadratic graph remains bipartite; the affine term's linear part shifts the image condition, while its constant part changes only the sign. With ρ_s half the restricted polar rank, every nonzero residual sum has magnitude $2^{w-\mathrm{wt}(q)-\rho_s}$. The coefficient therefore has magnitude $2^{-\mathrm{wt}(p)-\mathrm{wt}(q)-\rho_s}$, proving Equation 28. The factor $(-1)^{b\cdot k}$ shows that the key changes only the sign.

E.4 Residue Refinement by Trail Enumeration

We refine the valuation bound by summing the lowest-valuation trail contributions exactly. For XOR round-key additions, we write an unnormalized linear sum as

$$L(K) = 2^T \sum_\tau 2^{d_\tau}\sigma_\tau(-1)^{\kappa_\tau\cdot K}, \qquad d_\tau \geq 0, \quad \sigma_\tau \in 2\mathbb{Z}+1.$$

Here K concatenates the round-key bits, T is the base lower bound on trail valuations, and d_τ is a nonnegative integer valuation gap. The odd signed factor is σ_τ, and κ_τ selects the key bits determining the trail's sign. For $h \geq 1$, we compute $2^{-T}L(K)$ modulo 2^h using only trails with $d_\tau < h$. We enumerate every distinct sequence of round-boundary masks in these layers exactly once. For $h = 1$, an even number of odd terms proves one additional divisibility bit, independently of signs. For SPECK, the odd factors are ± 1; general local tables may have larger odd factors, whose residues must be retained. For larger h, we expand each character as

$$(-1)^{\kappa\cdot K} = \prod_{j:\kappa_j=1}(1-2K_j)$$

and sum the resulting multilinear polynomials modulo 2^h. A degree-j term from trail τ contains the factor $2^{d_\tau+j}$, so terms with $j \geq h - d_\tau$ vanish. Setting all key bits to zero gives the constant coefficient. We recover the remaining coefficients by setting one, two, and successively more key bits to one, leaving the others zero, and subtracting the terms already determined. Consequently the normalized residue vanishes for all round-key tuples if and only if every

multilinear coefficient is zero modulo 2^h, proving $v_2(L(K)) \geq T + h$. Since the round-key bits are unrestricted, this guarantee holds for every key schedule.

For 7-round SPECK32, let $\beta_* = \texttt{02050204}$. Complete enumeration of the three lowest valuation layers for input masks 00000040 and 00000060 gives 63 and 51 trails, respectively. Writing $k_{r,i}$ for bit i of round key r, with both indices starting at 0 and bits numbered from the LSB, we obtain

$$L_{\texttt{00000040},\beta_*} \equiv L_{\texttt{00000060},\beta_*} \equiv 32(k_{2,8} \oplus k_{3,3} \oplus k_{3,10}) \pmod{64}. \tag{44}$$

Together with the support proof $L_{\texttt{00000020},\beta_*} = 0$, the exact conversion for the cube with $x_5 = b_5, x_6 = b_6$ and data complexity 2^{30} yields

$$N_1 = 2^{29} - \frac{(-1)^{b_6} L_{\texttt{00000040},\beta_*} + (-1)^{b_5+b_6} L_{\texttt{00000060},\beta_*}}{8}.$$

Both the sum and difference of the two coefficients in Equation 44 are multiples of 64. Thus $8 \mid N_1$ for every key and all four offsets, whereas separate coefficient bounds give only divisibility by 4. For the $x_6 = b$ halfspaces in Table 21, two valuation layers containing 34 trails raise the one-count exponent from 1 to 3 at 7 rounds; four layers containing 24 trails raise it from 6 to 10 at 6 rounds. These refinements are distinct from the base linear-model bounds.

E.5 Local Encodings for the Joint Model

The per-bit choice of representation, called Spatial Joint in the result tables, uses the same local construction across ciphers. We choose Walsh or moment coordinates (see Subsection 5.2) for each bit in advance and solve a separate MILP for each choice. On one bit, the two coordinate systems are related by

$$\begin{pmatrix} H_0 \\ H_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} Z_0 \\ Z_1 \end{pmatrix}, \qquad \begin{pmatrix} Z_0 \\ Z_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} H_0 \\ H_1 \end{pmatrix}. \tag{45}$$

For a local s-bit component, we assign either coordinate system to each input and output bit. The evaluation matrices have basis functions $(1, (-1)^x)$ or $(1, x)$ as rows and bit values $0, 1$ as columns. Equation 45 gives the change between these bases. We form P_{in} and P_{out} as tensor products of the selected evaluation matrices. With state-transition matrix $(Q_F)_{y,x} = \mathbf{1}[y = F(x)]$, the exact local coefficient matrix is

$$\mathsf{T}^F = P_{\text{out}} Q_F P_{\text{in}}^{-1}. \tag{46}$$

An S-box has a fixed small s, so this matrix can be computed exactly even when the complete cipher state is large. For PRESENT, $s = 4$; no table on the full 64-bit state is required. A bit permutation only changes bit positions. If we also change between Walsh and moment coordinates, we use Equation 45.

For each input–output index pair τ with $\mathsf{T}^F_\tau \neq 0$, we introduce a binary selector z_τ. We impose $\sum_\tau z_\tau = 1$, set the component's input and output indices to the selected pair, and assign cost $\sum_\tau v_2(\mathsf{T}^F_\tau) z_\tau$. Components joined by a bit share

its coordinate index, and their costs add to the objective. Truth-table minimization can compact this encoding while preserving the same allowed transitions and costs. For bounded-size components and fixed choices of representation, the number of local relations grows linearly with the circuit size.

We represent XOR key additions explicitly because their action depends on the chosen coordinates. In Walsh coordinates an XOR key bit only changes a character's sign. In moment coordinates, complementing a bit uses $1 - x$ and therefore mixes its constant and degree-one coordinates. For arbitrary-round-key bounds, we include every transition that is nonzero for some key, assigning a cost no larger than its valuation for every such key. The model still includes every nonzero trail contribution. Ignoring cancellations or relations between key bits may weaken the bound, but it remains valid for every key. For SIMON, the exact quadratic formulas avoid constructing a full-word version of Equation 46. The global product argument remains the one in Proposition 5.

Additional Local Bounds. For fixed input and output indices, two valid valuation bounds g and h on the same suffix coefficient give the stronger bound $\max\{g, h\}$. We encode this maximum by imposing $\lambda \geq g$ and $\lambda \geq h$ and minimizing an objective containing λ with positive weight; the coefficient's cost is counted only once. For coordinate cubes, each active input bit that does not affect the selected output contributes a factor of 2 to the one-count. We factor out these powers of 2 before building the model. We distinguish these enhancements from the core results. Changing the representation alone does not guarantee a stronger bound.

E.6 Contribution of the Walsh Representation

We examine whether the SIMON gains come from the core Joint construction or require additional refinements. For the 9-round SIMON32 output c_5 and 10-round SIMON48 output c_{20} in Table 1, our implementation reproduces the published one-count exponents 5 and 6 [BV24, Tables 7–8]. The core Joint construction of Subsection 5.2 proves exponents 11 and 22, respectively, without additional refinements or trail enumeration.

Preserving cancellations in the input representation is essential to these bounds. If we convert the initial moment contributions to Walsh coordinates separately and take valuations before summing, we lose the cancellations that eliminate coefficients outside $A^{\perp}$ in Equation 24. Keeping the rest of the model unchanged then gives only the trivial one-count bound $t = 0$ in both cases.

Special first-round cube propagation is unnecessary for all 674 SIMON improvements in Subsection F.1. For 481 targets, including all 290 finite-bound improvements, this follows from the fully active right input word. The first Walsh round then has zero quadratic rank on the allowed input masks and gives the same relation as exact first-round cube propagation. For the remaining 193 targets, optimizations with and without special first-round propagation give the same optimum, with all other representation choices and local bounds fixed. This broader comparison retains the additional ultrametric bounds of Subsection E.5 and tests only the need for special first-round propagation in SIMON; it does

not establish that other refinements are unnecessary or address the PRESENT applications.

E.7 Exact Local Contraction beyond SIMON

Exact contraction also applies to small keyed components. At BEANIE's forward/inverse junction, the bit permutations cancel after applying the inverse permutation to the intervening key difference. Each S-box pair then becomes

$$F_d(x) = S^{-1}(S(x) + d), \qquad d \in \mathbb{F}_2^4,$$

where d is the resulting 4-bit shift. For each local mask pair, we retain the minimum valuation among its nonzero coefficients over all 16 shifts, excluding pairs unsupported for every shift. The eight contracted components leave the remaining linear MILP and its endpoint conversion unchanged. For $U_{2,3}$ and $U_{3,2}$, this gives one-count exponents 8 and 10, respectively (Table 26). For Joint, we compute the coefficients in the chosen bases before taking this minimum, obtaining one-count exponents 10 and 12, respectively. Neither construction enumerates complete cipher trails.

For SKINNY's 8-round, two-active-cell example, the selected three-bit parity reduces to a terminal function involving only two S-boxes. We contract this function, including all relevant local key classes, before assigning its mix-basis coefficient. A prefix with 5 rounds in ultrametric coordinates followed by 1 in Walsh coordinates then gives one-count exponent 1 in Table 40. An all-ultrametric prefix gives the same bound, so the improvement comes from the exact terminal contraction.

For SPECK's w-bit addition, we use an exact two-state relation for input masks a, b and output mask g, with carries summed out before valuation. Its auxiliary state starts at $s_w = 0$, and we read (a_i, b_i, g_i) from most to least significant bit. When $s_{i+1} = 0$, support requires $a_i = b_i = g_i$; otherwise all triples are allowed. The next state is $s_i = s_{i+1} \oplus a_i \oplus b_i \oplus g_i$, and the nonzero normalized coefficient has valuation $-\sum_i s_{i+1}$. This is the established addition-correlation relation [FWG+16], used here with a valuation objective rather than a maximum-correlation objective. The residue-refined gains in our SPECK comparisons come from the calculation in Subsection E.4, not from a new addition automaton.

E.8 The 5-Round PRESENT Exact Contraction

We obtain the refined 5-round bounds in Table 29 by reducing the counts to an exact 16-bit convolution. For the least-significant unit vector e_0, the PRESENT S-box satisfies

$$S(x)_0 + S(x + e_0)_0 = 1 \quad \text{over } \mathbb{F}_2.$$

Equivalently, its least-significant output bit is x_0 plus a function of the other three input bits. This identity is checked on the sixteen S-box inputs.

The active nibbles of the cube `000000000000fff0` are 1, 2, and 3. After their first S-boxes, each ranges over all sixteen values exactly once. The bit permutation distributes the twelve independent bits into positions 1, 2, and 3 of second-round S-boxes 0, 4, 8, and 12. Their remaining input bits are constants $b_0, \ldots, b_3$ after key addition. Thus the four varying second-round S-box outputs y_k range independently, for each fixed key, over $Y_{b_k} = \{S(x) : x_0 = b_k\}$. There are $8^4 = 4096$ tuples, each appearing once.

We consider ciphertext bits $i = 16q + 4v$, where $q, v \in \{0, 1, 2, 3\}$. Only bit zero of each third-round S-box output contributes to these ciphertext bits, and only input bit zero varies. The S-box identity carries that varying bit through, up to a complement. We absorb these complements and the next key addition into a translation $T \in \mathbb{F}_2^{16}$. We write $b = (b_0, \ldots, b_3) \in \mathbb{F}_2^4$ and arrange $y_0, \ldots, y_3$ as the rows of a four-by-four bit matrix. The indicator H_b selects their allowed values; z_j and $F_{v,q,h}$ describe the relevant fourth- and fifth-round computations:

$$H_b(y) = \prod_{k=0}^{3} \mathbf{1}[(S^{-1}(y_k))_0 = b_k],$$

$$z_j(y) = S\left(\sum_{k=0}^{3} (y_k)_j \, 2^k\right)_v,$$

$$F_{v,q,h}(y) = S\left(\left(\sum_{j=0}^{3} z_j(y) \, 2^j\right) \oplus h\right)_q.$$

Here $h \in \mathbb{F}_2^4$ is the remaining key translation; final whitening can complement the selected output. Before that possible complement, the count is exactly

$$N_{b,T,h}^{v,q} = \sum_{y \in \mathbb{F}_2^{16}} H_b(y) F_{v,q,h}(y + T). \tag{47}$$

Every key and inactive-input assignment induces some (b, T, h). Proving a bound for all parameter tuples is sound even when the real key schedule makes them dependent.

The XOR-convolution theorem computes the counts in Equation 47 for all T simultaneously, with $\odot$ denoting entrywise multiplication:

$$N_{b,\cdot,h}^{v,q} = 2^{-16} \, \mathrm{FWHT}\big(\mathrm{FWHT}(H_b) \odot \mathrm{FWHT}(F_{v,q,h})\big).$$

We evaluate this identity in exact integer arithmetic; the normalization is exact and every count lies between 0 and 4096. For $i = 0, 4, \ldots, 60$, the minimum one-count valuations over all (b, T, h) are

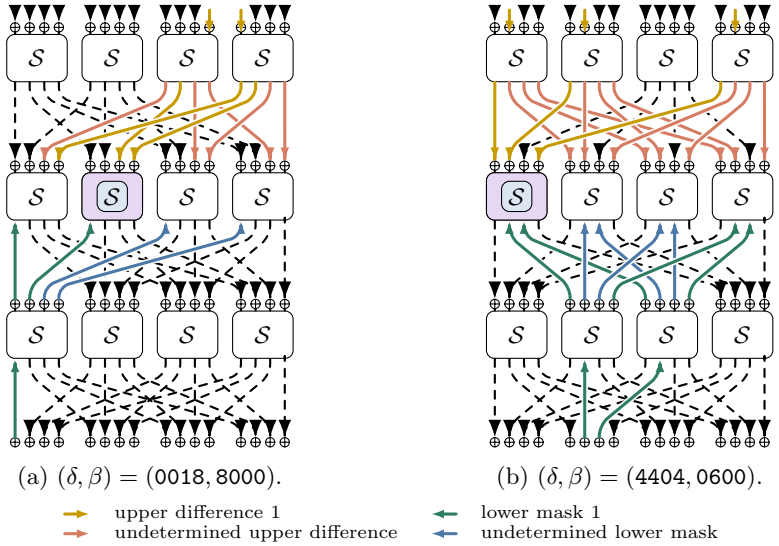
$$(5, 3, 3, 3,\ 4, 2, 2, 2,\ 4, 2, 2, 2,\ 4, 2, 2, 2).$$

A final complement changes N to $4096 - N$ and preserves these divisibility bounds. The bounds need not be tight for the actual key schedule, which may not realize every parameter tuple. This exhaustive evaluation proves the stated bounds for all keys and inactive-input assignments at these 16 output positions, without enumerating cipher trails.

E.9 Why a Local DLCT Zero Does Not Imply ZCDL

We use a 16-bit, 3-round ToySPN (Subsection D.1) with the KNOT S-box S from Table 12. In Figure 6, we propagate truncated differences forward and linear masks backward. At each step, we fix only those bits whose values agree across all nonzero DDT or LAT transitions compatible with the current conditions. Both examples force a zero DLCT entry at the highlighted S-box of the second round, yet the full-codebook DL sum $L^{\mathrm{DL}}_{0,\delta,\beta}(E_K)$ need not vanish.

For $(\delta, \beta) = (\mathtt{0018}, \mathtt{8000})$, the local difference and mask are $\mathtt{3}$ and $\mathtt{8}$, with $\mathrm{DLCT}_S(\mathtt{3}, \mathtt{8}) = 0$. Exact full-codebook evaluation on 1,000 sampled keys gives only nonzero DL sums in $\{\pm 1024, \pm 3072\}$. For $(\delta, \beta) = (\mathtt{4404}, \mathtt{0600})$, the local difference and mask are $\mathtt{d}$ and $\mathtt{6}$, with $\mathrm{DLCT}_S(\mathtt{d}, \mathtt{6}) = 0$. On the same keys, 953 DL sums are nonzero and 47 are zero. The samples use distinct 80-bit keys and the low $4m$ bits of each PRESENT round key. The all-zero master key gives sums 1024 and -512, respectively, so each example already refutes the universal zero claim. A local DLCT zero expresses cancellation when summing over all S-box inputs. In the complete cipher, dependencies involving the surrounding rounds can change the weights of these contributions, so the local cancellation need not survive.



(a) $(\delta, \beta) = (\mathtt{0018}, \mathtt{8000})$. (b) $(\delta, \beta) = (\mathtt{4404}, \mathtt{0600})$.

Fig. 6: Local DLCT zeros do not imply fixed-key ZCDL in these 3-round KNOT-S-box ToySPN examples. Wires run from the most significant bit on the left to the least significant on the right, consistent with the hexadecimal values in Subsection D.1.

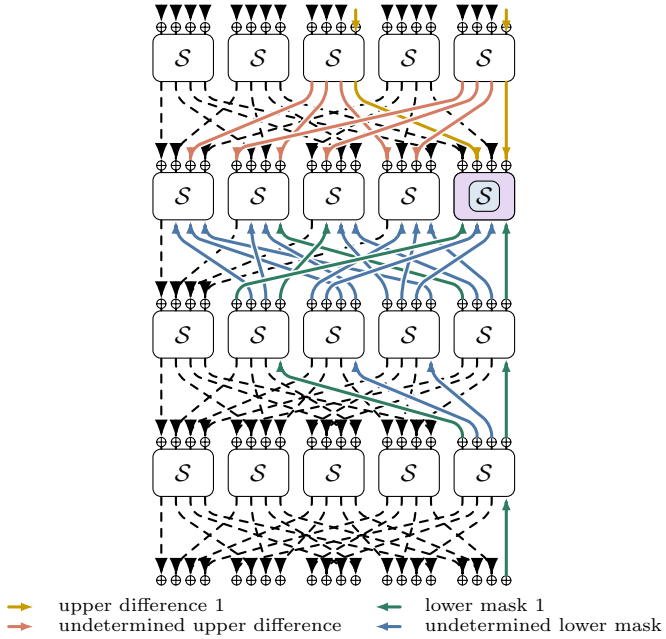


Fig. 7: A 4-round, 20-bit PRESENT-S-box ToySPN counterexample with $(\delta, \beta) = (\texttt{00101}, \texttt{00001})$. Every local DLCT entry allowed at the highlighted S-box is zero, yet the full-codebook DL sum need not vanish. Wire conventions follow Figure 6.

The five-S-box example in Figure 7 illustrates the same issue on a 20-bit state. The indicated S-box has local input difference pattern `0101` and truncated output-mask pattern `1**1`, where each `*` denotes an undetermined bit. Thus the local difference is `5` and the compatible masks are $\texttt{9}, \texttt{b}, \texttt{d}, \texttt{f}$, all with zero DLCT entries. Complete-domain checks on the same 1,000 keys give 990 nonzero sums, with observed sums between -3968 and 4320. The all-zero master key gives the nonzero sum -544. These examples invalidate the inference from a local DLCT zero to ZCDL, not the correctly computed DLCT of the whole permutation. Subsection E.10 gives the exact mix-basis DL model, and Subsection E.11 shows its SboxAnalyzer constraint generation.

E.10 Exact Mix-Basis DL Relations and Constraint Generation

For an s-bit S-box, we enumerate all 2^{2s} input pairs and compute the exact transitions between the bases defined in Subsection 5.3. For a nonzero coefficient R/D with integer numerator R, the cost is $v_2(R) - v_2(D)$. The input-basis normalization is $D = 2^s$ for quasidifferential coordinates and $D = 2^{2s}$ for linear coordinates on pairs of state values. The support relation contains precisely the transition labels with $R \neq 0$. Among these transitions, a valuation subtable selects $v_2(R) = t$, while a threshold subtable selects $v_2(R) \geq t$. A signed-numerator subtable fixes

R itself and may omit other entries of the same valuation. Such restricted subtables cannot replace the full support relation in a zero proof unless every nonzero contribution is still represented. Boolean minimization gives smaller constraints for the same finite relation. We check these constraints against every entry of the exact local table.

Proposition 11 (Mix-basis DL support-SAT soundness). *If the support SAT model includes every nonzero term of Proposition 6, UNSAT for its endpoints proves that the corresponding DL sum is zero.*

Proof. Every nonzero term would induce a satisfying assignment of local support constraints and shared intermediate indices. UNSAT excludes all such terms, so the exact expansion vanishes. The implication holds for any fixed choice of intermediate bases that includes every nonzero contribution.

Difference-only models also give sound UNSAT conclusions, but SAT may leave incompatible intermediate values unresolved.

Proposition 12 (Difference-only trail models). *Fix a permutation and its component decomposition. A difference-only model retaining all nonzero local DDT transitions and exact difference-wiring constraints gives necessary conditions for a right pair: a plaintext pair realizing the prescribed endpoint differences. UNSAT for fixed endpoint differences proves an ID; SAT alone does not prove the existence of a right pair for that permutation.*

Proof. Every right pair induces a supported transition at each component and consistent differences on every connecting wire. It therefore gives a satisfying assignment. Conversely, the local transitions may require incompatible intermediate values, which a difference-only model does not retain. Thus satisfiability of these necessary conditions alone does not establish a right pair.

An exact SAT model of both encryptions retains their intermediate values and avoids this loss of information. With independently chosen full-state XOR keys between rounds, supported round transitions can be aligned by choosing those keys. This establishes existence for some round-key tuple, not for a fixed key or a prescribed key schedule.

E.11 S-box Constraint Generation with S-box Analyzer

The global MILP and SAT models use only the local relations of the kind defined in Section 5. We extended S-box Analyzer [HNE22] to generate these relations automatically from the S-box lookup table. It rewrites each target local relation as a Boolean function, minimizes it, and exports equivalent SAT formulas in conjunctive normal form (CNF), MILP constraints, and constraint programming (CP) constraints. The original tool already covered ordinary binary character-sum tables. Our extension adds exact transition tables for mix-basis DL models, including quasidifferential-to-quasidifferential, quasidifferential-to-linear, and

linear-to-linear transitions on pairs of state values. The interface names these bases `quasidifferential` and `linear`, and provides signed-numerator and valuation-weighted relations. Listing 1 illustrates the existing linear interface. Listing 2 generates the nonzero mix-basis transitions, while Listing 3 selects a signed-numerator subtable. Each call returns three equivalent encodings of its local relation.

```
from sage.all import QQ
from sboxanalyzer import SboxAnalyzer
from sage.crypto.sboxes import PRESENT

sa = SboxAnalyzer(PRESENT)
cnf_support, milp_support, cp_support = sa.minimized_linear_constraints(
    subtable="star", mode=6
)
cnf_sqcorr, milp_sqcorr, cp_sqcorr = sa.minimized_linear_constraints(
    subtable=QQ("1/16"), mode=6
)
```

Listing 1: Generating the CNF/SAT, MILP, and CP encodings of the character-sum support relation and one squared-correlation subtable for the PRESENT S-box.

```
from sboxanalyzer import SboxAnalyzer
from sage.crypto.sboxes import PRESENT

sa = SboxAnalyzer(PRESENT)
cnf_pair_support, milp_pair_support, cp_pair_support = sa.minimized_exact_pair_constraints(
    input_basis="quasidifferential", output_basis="linear",
    subtable="star", mode=6
)
```

Listing 2: Generating CNF/SAT, MILP, and CP constraints for nonzero quasidifferential-to-linear transitions on pairs of state values.

```
from sboxanalyzer import SboxAnalyzer
from sage.crypto.sboxes import PRESENT

sa = SboxAnalyzer(PRESENT)
cnf_pair_weight, milp_pair_weight, cp_pair_weight = sa.minimized_exact_pair_constraints(
    input_basis="quasidifferential", output_basis="linear",
    subtable=8, mode=6
)
```

Listing 3: Generating the quasidifferential-to-linear subtable with signed raw numerator $+8$.

In Listing 3, `subtable=8` selects numerator $+8$, not valuation 8 and not entries with numerator -8. For the quasidifferential input basis of a four-bit S-box, normalization divides by 16, giving transition cost $3 - 4 = -1$. To generate the full valuation-weighted relation, use `subtable=None` with `weight_model="v2"`. Changing the basis arguments gives the analogous transitions within either basis. In the DL model, we use quasidifferential coordinates before the mix-basis round and linear coordinates on pairs of state values afterwards.

E.12 Generic Black-Box Search

Algorithm 1 gives the common procedure underlying Subsection 5.4, independently of the cipher and implementation. An input specification $j \in \mathcal{J}$ consists of the domain S_j, the oracle-derived map $g_{K,j}$, and weights w_j, as in Equation 30. Each index $\lambda \in \Lambda$ specifies an input condition, an output mask or difference, and a statistic z_λ: a histogram entry, signed sum, associated one-count, or coset contrast. The set $\mathcal{P}$ lists requested properties (λ, P), such as $z_\lambda = 0$, $2^t \mid z_\lambda$, or $z_\lambda \bmod 2 = b$ for a fixed b. The input specifications must supply every requested statistic, including a complete input-mask family whenever mask-to-coset conversion is requested. Both FWHTs are unnormalized, and all arithmetic is exact.

For a complete d-dimensional input-mask family U, let $j(\alpha)$ denote the specification with domain V, weight $(-1)^{\alpha \cdot x}$, and a common oracle map. Within each key iteration, this map is $g = E_{K_\ell}$ or $g = D_\delta E_{K_\ell}$ for a fixed δ. Let $A = U^\perp$ and choose a basis $u_1, \ldots, u_d$ of U. Index masks $\alpha = \sum_i \xi_i u_i$ by $\xi \in \mathbb{F}_2^d$ and cosets $a + A$ by $(u_1 \cdot a, \ldots, u_d \cdot a)$, using the same binary coordinate order. Applying $2^{-d}\,\mathrm{FWHT}_\alpha$ to the coefficients gives $I_\beta(a + A; g)$ by Theorem 1 or Theorem 2. The family includes $\alpha = 0$, whose DL coefficient need not vanish. For $\mathcal{X} = a + A$ of size M, the one-count is $N_1(\mathcal{X}, \beta; g) = (M - I_\beta(\mathcal{X}; g))/2$, also written $N_1^{\mathrm{DL}}(\mathcal{X}, \delta, \beta; E_{K_\ell})$ in the DL case. Direct unit-weight structure sums use the same count conversion without the mask-to-coset step.

For linear sums, choose $S_j = V$, $g_{K,j} = E_K$, and $w_j(x) = (-1)^{\alpha \cdot x}$. For DL sums, replace E_K by $D_\delta E_K$; for direct integral sums, choose $S_j = \mathcal{X}$, $g_{K,j} = E_K$, and $w_j = 1$. For $g_{K,j} = D_\delta E_K$ and $w_j = 1$, $H_{K,j}(\gamma)$ counts output differences: it equals $\mathrm{DDT}_{E_K}(\delta, \gamma)$ when $S_j = V$ and a restricted count otherwise. We retain these entries separately from the transformed coefficients. Thus a zero histogram entry and a zero transformed entry are tested as different properties.

The returned properties and minimum valuations concern only the s tested instances; they are not all-key guarantees. The exact values are available during each property test and may also be recorded. Independent validation and selection-aware interpretation follow Subsection C.6. No independence between output masks or cosets is assumed by the algorithm. Specifications can be processed in batches subject to the complete-family requirement; this changes storage, not the mathematical computation.

E.13 Complete Coordinate-Family Screening

Here we consider properties required to hold on every coset of a coordinate subspace, as in the complete screening of Subsection 5.4.

Proposition 13 (Reduction to coordinate-halfspace screening). *Fix a permutation E, an output mask $\beta \neq 0$, and an integer signed-sum threshold $t \geq 0$. Let $A_J = \langle e_j : j \in J \rangle$ for a proper subset J of the coordinates. If $v_2(I_\beta(a + A_J; E)) \geq t$ for every coset, then the same bound holds on both coordinate halfspaces $x_i = b$ for every $i \notin J$. The assertion also holds for exact balance and for a common one-count divisibility bound. For fixed β and threshold, such*

Algorithm 1 Black-box discovery of sum and count properties

Input: Oracle instances $E_{K_1}, \ldots, E_{K_s}$; input specifications $\mathcal{J}$; statistics $(z_\lambda)_{\lambda \in \Lambda}$ and requested properties $\mathcal{P}$

Output: Properties $\mathcal{R}$ holding on every tested instance; minimum observed valuations $(t_\lambda)_{\lambda \in \Lambda}$; sample size s

$\mathcal{R} \leftarrow \mathcal{P}$; $t_\lambda \leftarrow \infty$ for all $\lambda \in \Lambda$
for $\ell = 1, \ldots, s$ **do**
  ▷ *Histograms and output-mask sums*
  for each $j \in \mathcal{J}$ **do**
    $H_{K_\ell,j}(y) \leftarrow 0$ for all $y \in V$
    for each $x \in S_j$ **do**
      $y \leftarrow g_{K_\ell,j}(x)$
      $H_{K_\ell,j}(y) \leftarrow H_{K_\ell,j}(y) + w_j(x)$
    end for
    $\widehat{H}_{K_\ell,j} \leftarrow \mathrm{FWHT}(H_{K_\ell,j})$
  end for
  ▷ *Optional mask-to-coset conversion*
  for each complete family U with common map g requiring coset sums **do**
    $d \leftarrow \dim U$; $A \leftarrow U^\perp$
    for each requested nonzero output mask β **do**
      $(I_\beta(a + A; g))_{a \in \mathcal{R}_A} \leftarrow 2^{-d}\,\mathrm{FWHT}_\alpha((\widehat{H}_{K_\ell,j(\alpha)}(\beta))_{\alpha \in U})$
    end for
  end for
  ▷ *Property tests across instances*
  for each $\lambda \in \Lambda$ **do**
    Obtain $z_\lambda(E_{K_\ell})$ from the histograms or signed sums, forming a one-count or coset contrast if requested
    $t_\lambda \leftarrow \min\{t_\lambda, v_2(z_\lambda(E_{K_\ell}))\}$
    $\mathcal{R} \leftarrow \mathcal{R} \setminus \{(\lambda, P) : P(z_\lambda(E_{K_\ell})) = \text{false}\}$
  end for
end for
return $\mathcal{R}$, $(t_\lambda)_{\lambda \in \Lambda}$, s

a property therefore holds on every coset of some proper coordinate subspace if and only if it holds on both halves of some coordinate halfspace partition.

Proof. Each halfspace $x_i = b$ is a disjoint union of cosets of A_J. Add their signed sums and apply the ultrametric inequality. One-counts also add over these disjoint cosets, preserving any common divisibility bound. Conversely, a coordinate-halfspace family is itself a complete proper coordinate family. For exact balance, all summands are zero.

The threshold is absolute: it is not a threshold normalized by the size of the structure. For a permutation, the two complementary halfspace sums add to zero, so their signed valuations coincide. Their one-counts add to 2^{n-1}, so divisibility by 2^{t_N} is equivalent on the two halves for $0 \le t_N \le n-1$. Testing coordinate halfspaces does not cover arbitrary affine orientations, isolated smaller cubes, or

nonlinear output functions. For a keyed family, a common all-key claim keeps the selected family and output mask fixed across keys.

E.14 Interpreting the Key Samples

We apply the random-permutation bounds to the specific events tested in our applications. For the 84-round KATAN32 row with $x_0 = b$ in Table 47, let p_0 be the probability that $8 \mid N_1$ for its fixed structure and output mask under a uniform random permutation. Then Proposition 3 gives

$$p_0 \leq 2^{-3} + p_{\mathrm{ctr}}(2^{32}, 2^{31}).$$

For 1,000 independent ideal permutations, the probability that this condition holds for all of them is at most the 1,000th power of that expression, hence at most $2^{-2999.71}$ by Equation 37. For several candidates, Equation 40 sums their individual bounds; multiplication by $|\mathcal{H}|$ requires a common upper bound. Subsection C.6 specifies the candidate family to use after search or independent validation and gives the required sample size. Dependent tests on the same instance do not provide additional independent permutations (Remark 2).

The real-key statement is different. For a candidate fixed before validation, zero failures in s independent uniformly sampled real keys give the confidence bound in Equation 41 on its failure fraction, not a probability that it holds for every key. For BEANIE, the real-sample bound concerns the failure fraction under the sampled key/tweak distribution; the uniform-permutation baseline is separate (Subsection C.4). Zero DDT entries satisfy every divisibility threshold, so their probability must be included when evaluating a DDT-count test (Equation 34).

F Detailed Cipher Results

Each cipher subsection begins with its specification, followed by its results. For an n-bit state, $x_i^{(r)}$ is bit i at the input of round r, with bit 0 least significant and bit $n-1$ most significant. The output bits c_i use the same global indexing. Word and cell names in the round descriptions abbreviate slices of this state; they do not restart the bit numbering in our results. Unless stated otherwise, a reported round includes its complete linear layer and affine additions. The four-bit substitutions used below are collected in Table 12.

MILP bounds use Gurobi, SAT checks use CaDiCaL or Kissat, and black-box sums use exact FWHT arithmetic with LibFWHT [HMAV26], including its GPU backend. All binary integral tables use the one-count N_1 and the conventions of Section 6. A numerical lower bound t means $2^t \mid N_1$. A separate exact-count table is used when balance is proved. In a cell $t\,(t')$, the parenthesized number is obtained by the refinement named in that table, not by the base model. A dash denotes no reported result, not a zero count. Hexadecimal masks and differences retain all leading zeros to show the full block width. Long masks may wrap

Table 12: Four-bit substitutions, with inputs and outputs in hexadecimal.

Input	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
PRESENT	c	5	6	b	9	0	a	d	3	e	f	8	4	7	1	2
GIFT	1	a	4	c	6	f	3	9	2	d	b	7	5	0	8	e
SKINNY-64	c	6	9	0	1	a	2	b	3	8	5	d	4	e	7	f
BEANIE	0	4	2	b	a	c	9	8	5	f	d	3	7	1	6	e
KNOT	4	0	a	7	b	e	1	d	9	f	6	8	5	2	c	3

between digit groups without omitting any digits. Comma-separated output bits are separate tests, whereas $\oplus$ denotes one parity. Numerical observations over sampled keys are explicitly separated from all-key bounds.

F.1 SIMON

Specification. SIMON is a family of Feistel block ciphers [BSS+15] with block size $n = 2w$. We consider $w \in \{16, 24, 32, 48, 64\}$. Write the left word as $s = \sum_{i=0}^{w-1} x_{w+i}2^i$ and the right word as $t = \sum_{i=0}^{w-1} x_i 2^i$, so the packed state is $2^w s + t$. For a w-bit round key k_r, a round is

$$(s, t) \longmapsto \big(t \oplus ((s \lll 1) \wedge (s \lll 8)) \oplus (s \lll 2) \oplus k_r,\ s\big).$$

All rotations are within a word. Figure 8 shows this round structure for $w = 16$. Our analytical bounds allow arbitrary round keys and therefore cover the standard key schedules of the corresponding block-size variants. The SIMON32 black-box evaluations use the standard SIMON32/64 key schedule, with a fixed 64-bit master key throughout each evaluation. Their reduced-round windows start at round 0. The comparison input structures involving the inverse of a keyless first round are specified with their result tables; they are not ordinary plaintext coordinate cubes.

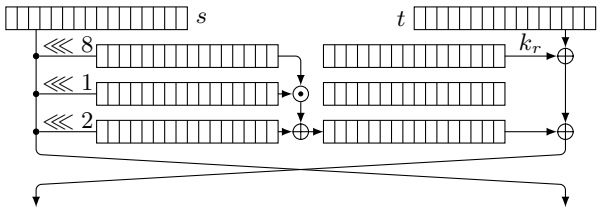


Fig. 8: The SIMON round function, drawn for $w = 16$. s is the left word and t the right word.

Results. The published-structure comparisons in Table 13 and Table 14 use $R_0^{-1}(a+A)$ and already include the input round. They cover the 290 finite-bound targets and 384 exact-balance targets for which our methods improve the published one-count bound. We compare against the bounds reported in [BV24] for the same input structure and output bit. Parenthesized RC bounds use complete enumeration of the minimum-valuation trails; the base RC and Joint bounds use no trail enumeration. A finite lower bound need not be the exact valuation and does not by itself establish a nonzero correlation.

Table 13: All improved finite SIMON targets, grouped only when the compared bounds coincide. Input $R_0^{-1}(a + (u \wedge \mathbb{F}_2^n))$; RC parentheses denote parity refinement. Gain is the larger of the base RC and Joint bounds minus the published bound. Parenthesized refinements are not counted.

n	r	u	Data	Output	t [BV24]	RC	Joint	Gain
32	8	`01ffffff`	2^{25}	c_1, c_2	6	5 (6)	7	+1
32	9	`1fffffff`	2^{29}	c_5	5	10 (11)	11	+6
32	9	`1fffffff`	2^{29}	c_{17}, c_{18}, c_{30}	3	2	5	+2
32	9	`1fffffff`	2^{29}	c_{19}, c_{20}	3	4 (5)	5	+2
32	9	`1fffffff`	2^{29}	c_{27}, c_{28}	3	2 (3)	5	+2
32	9	`1fffffff`	2^{29}	c_{29}	3	3 (4)	6	+3
32	10	`7fffffff`	2^{31}	c_{17}	3	4 (5)	8	+5
32	10	`7fffffff`	2^{31}	c_{18}	3	6 (7)	7	+4
32	10	`7fffffff`	2^{31}	$c_{19}, c_{27}, c_{29}, c_{31}$	3	4 (5)	7	+4
32	10	`7fffffff`	2^{31}	c_{28}	3	6 (7)	8	+5
32	11	`7fffffff`	2^{31}	c_1	3	4 (5)	8	+5
32	11	`7fffffff`	2^{31}	c_2	3	6 (7)	7	+4
32	11	`7fffffff`	2^{31}	$c_3, c_{11}, c_{13}, c_{15}$	3	4 (5)	7	+4
32	11	`7fffffff`	2^{31}	c_{12}	3	6 (7)	8	+5
32	11	`7fffffff`	2^{31}	$c_{16}, \ldots, c_{21}, c_{23}, c_{25}, \ldots, c_{31}$	2	0	4	+2
32	12	`7fffffff`	2^{31}	$c_0, \ldots, c_5, c_7, c_9, \ldots, c_{15}$	2	0	4	+2
32	12	`7fffffff`	2^{31}	$c_{16}, \ldots, c_{31}$	1	0	2	+1
32	13	`7fffffff`	2^{31}	$c_0, \ldots, c_{15}$	1	0	2	+1
48	8	`00001fffffff`	2^{29}	c_9, c_{10}	8	7 (8)	9	+1
48	9	`007fffffffff`	2^{39}	c_{11}	7	9 (10)	9	+2
48	9	`007fffffffff`	2^{39}	c_{13}	7	11 (12)	11	+4
48	9	`007fffffffff`	2^{39}	c_{15}	6	8 (9)	9	+3
48	9	`007fffffffff`	2^{39}	c_{16}	7	9 (10)	10	+3
48	9	`007fffffffff`	2^{39}	c_{17}	7	8 (9)	10	+3
48	9	`007fffffffff`	2^{39}	c_{21}	7	8 (9)	11	+4
48	9	`007fffffffff`	2^{39}	c_{44}	4	0	5	+1
48	10	`0fffffffffff`	2^{44}	c_{20}, c_{21}	6	20 (21)	22	+16
48	10	`0fffffffffff`	2^{44}	c_{22}	6	13 (14)	15	+9
48	10	`0fffffffffff`	2^{44}	c_{23}	6	11	14	+8
48	10	`0fffffffffff`	2^{44}	$c_{26}, \ldots, c_{28}, c_{40}, c_{47}$	4	0	6	+2
48	10	`0fffffffffff`	2^{44}	c_{29}	4	1	6	+2
48	10	`0fffffffffff`	2^{44}	c_{32}	4	3	6	+2

Table 13 (continued)

n	r	u	Data	Output	t [BV24]	RC	Joint	Gain
48	10	`0fffffffffff`	2^{44}	c_{33}	4	0	5	+1
48	10	`0fffffffffff`	2^{44}	c_{34}, c_{35}	4	1	5	+1
48	10	`0fffffffffff`	2^{44}	c_{41}	3	0	6	+3
48	10	`0fffffffffff`	2^{44}	c_{44}	4	13 (14)	14	+10
48	10	`0fffffffffff`	2^{44}	c_{45}	4	7	9	+5
48	10	`0fffffffffff`	2^{44}	c_{46}	4	2	8	+4
48	11	`3fffffffffff`	2^{46}	c_{10}	5	9	11	+6
48	11	`3fffffffffff`	2^{46}	c_{11}	5	10 (11)	11	+6
48	11	`3fffffffffff`	2^{46}	c_{22}	5	15 (16)	16	+11
48	11	`3fffffffffff`	2^{46}	c_{23}	5	9	13	+8
48	11	`3fffffffffff`	2^{46}	$c_{26}, \ldots, c_{33}, c_{37}, c_{41}, \ldots, c_{43}$	3	0	6	+3
48	11	`3fffffffffff`	2^{46}	$c_{34}, \ldots, c_{36}$	3	0	5	+2
48	11	`3fffffffffff`	2^{46}	c_{38}, c_{39}	3	5	8	+5
48	11	`3fffffffffff`	2^{46}	c_{40}	3	5	9	+6
48	11	`3fffffffffff`	2^{46}	c_{46}	3	4	10	+7
48	11	`3fffffffffff`	2^{46}	c_{47}	3	0	7	+4
48	12	`7fffffffffff`	2^{47}	c_3, c_{15}	4	6	9	+5
48	12	`7fffffffffff`	2^{47}	c_4, c_{18}	4	9	10	+6
48	12	`7fffffffffff`	2^{47}	$c_5, c_7, c_9, c_{13}, c_{17}, c_{19}$	4	6	10	+6
48	12	`7fffffffffff`	2^{47}	c_{10}, c_{12}	4	7	9	+5
48	12	`7fffffffffff`	2^{47}	c_{11}	4	8	9	+5
48	12	`7fffffffffff`	2^{47}	c_{23}	4	5	11	+7
48	12	`7fffffffffff`	2^{47}	$c_{25}, c_{42}, c_{45}, c_{47}$	2	0	6	+4
48	12	`7fffffffffff`	2^{47}	$c_{26}, \ldots, c_{30}, c_{33}, \ldots, c_{35}, c_{40}$	3	0	5	+2
48	12	`7fffffffffff`	2^{47}	c_{31}	3	0	6	+3
48	12	`7fffffffffff`	2^{47}	$c_{32}, c_{36}, \ldots, c_{39}, c_{41}, c_{43}, c_{44}$	2	0	5	+3
48	13	`7fffffffffff`	2^{47}	$c_1, c_{18}, c_{21}, c_{23}$	2	0	6	+4
48	13	`7fffffffffff`	2^{47}	$c_2, \ldots, c_6, c_9, \ldots, c_{11}, c_{16}$	3	0	5	+2
48	13	`7fffffffffff`	2^{47}	c_7	3	0	6	+3
48	13	`7fffffffffff`	2^{47}	$c_8, c_{12}, \ldots, c_{15}, c_{17}, c_{19}, c_{20}$	2	0	5	+3
48	13	`7fffffffffff`	2^{47}	$c_{24}, \ldots, c_{47}$	2	0	3	+1
48	14	`7fffffffffff`	2^{47}	$c_0, \ldots, c_{23}$	2	0	3	+1
48	14	`7fffffffffff`	2^{47}	$c_{24}, \ldots, c_{47}$	1	0	2	+1
48	15	`7fffffffffff`	2^{47}	$c_0, \ldots, c_{23}$	1	0	2	+1

Table 14: SIMON exact-balance improvements from support proofs and conversion. The input convention is the same as in Table 13. Published exponent ranges describe the individual listed bits, not a joint event.

n	r	u	Data	Output	t [BV24]	Exact N_1
32	7	`0001ffff`	2^{17}	$c_0, \ldots, c_{15}, c_{24}$	3–7	2^{16}
32	8	`01ffffff`	2^{25}	$c_0, c_4, \ldots, c_{15}, c_{23}, c_{24}$	3–7	2^{24}
32	9	`1fffffff`	2^{29}	$c_0, \ldots, c_4, c_6, \ldots, c_{16}, c_{21}, \ldots, c_{26}, c_{31}$	3–5	2^{28}
32	10	`7fffffff`	2^{31}	$c_0, \ldots, c_{16}, c_{20}, \ldots, c_{26}, c_{30}$	3–4	2^{30}
32	11	`7fffffff`	2^{31}	$c_0, c_4, \ldots, c_{10}, c_{14}, c_{22}, c_{24}$	2–3	2^{30}
32	12	`7fffffff`	2^{31}	c_6, c_8	2	2^{30}
48	7	`00000001ffff`	2^{17}	$c_0, \ldots, c_{16}$	7–10	2^{16}
48	8	`00001fffffff`	2^{29}	$c_0, \ldots, c_8, c_{16}, c_{18}$	8–10	2^{28}
48	9	`007fffffffff`	2^{39}	$c_0, \ldots, c_{10}, c_{12}, c_{14}, c_{18}, \ldots, c_{20}, c_{22}, c_{23}$	7–8	2^{38}
48	10	`0fffffffffff`	2^{44}	$c_0, \ldots, c_{19}, c_{24}, c_{25}, c_{30}, c_{31}, c_{36}, \ldots, c_{39}, c_{42}, c_{43}$	4–6	2^{43}
48	11	`3fffffffffff`	2^{46}	$c_0, \ldots, c_9, c_{12}, \ldots, c_{21}, c_{24}, c_{25}, c_{44}, c_{45}$	3–5	2^{45}
48	12	`7fffffffffff`	2^{47}	$c_0, \ldots, c_2, c_6, c_8, c_{14}, c_{16}, c_{20}, \ldots, c_{22}, c_{24}, c_{46}$	3–4	2^{46}
48	13	`7fffffffffff`	2^{47}	c_0, c_{22}	3	2^{46}
64	7	`000000000001ffff`	2^{17}	$c_8, \ldots, c_{24}$	6–13	2^{16}
64	8	`00000001ffffffff`	2^{33}	c_8	13	2^{32}
96	7	`000000000000000000001ffff`	2^{17}	$c_9, \ldots, c_{24}, c_{50}, \ldots, c_{66}$	6–15	2^{16}
96	8	`000000000000000100ffffffff`	2^{33}	$c_2, \ldots, c_{34}$	13–25	2^{32}
128	7	`00000000000000000000000001ffff`	2^{17}	$c_9, \ldots, c_{24}, c_{72}, \ldots, c_{82}$	6–15	2^{16}
128	8	`00000000000000000000000001ffffffff`	2^{33}	$c_2, \ldots, c_{34}, c_{64}, \ldots, c_{96}$	9–31	2^{32}

Longest-Round Coverage. Table 15 compares all nontrivial output bounds at the longest round counts reported in [BV24, Tables 7–11]. We use Joint with exact contraction of the first 2 core rounds before assigning valuations. The resulting bounds cover every key and offset; the final Feistel copy transfers a left-word output bound to the corresponding right-word bit one round later. All round counts already include the input extension.

For the SIMON32 row, we refine the count expansion modulo 2, following the complete-layer principle of Subsection E.4. After exact prefix contraction, all transition coefficients are integers, so only odd trail contributions can affect the parity of N_1. For each selected output bit, complete enumeration gives 16 such trails, each contributing 1 modulo 2 for every key and offset. Their sum is even, proving $2 \mid N_1$. The same calculation proves $t = 1$ for 14-round SIMON32

Table 15: SIMON bounds at the longest published round counts, for $\widetilde{\mathcal{X}} = R_0^{-1}(a + (u \wedge \mathbb{F}_2^n))$. Joint includes exact 2-round prefix contraction. Parentheses give the additional complete parity refinement.

n	Rounds	Active mask u	Data	Output	t [BV24]	Joint t
32	15	`7fffffff`	2^{31}	c_0, c_7, c_{14}	1	0 (1)
48	16	`7fffffffffff`	2^{47}	$c_0, \ldots, c_{23}$	1	1
64	18	`7fffffffffffffff`	2^{63}	$c_0, \ldots, c_9, c_{15}, c_{21}, \ldots, c_{31}$	1	1
96	22	`7fffffffffffffff ffffffff`	2^{95}	$c_1, c_6, c_{40}, c_{45}, c_{47}$	1	1
128	25	`7fffffffffffffff ffffffffffffffff`	2^{127}	c_1, c_{61}, c_{63}	1	1

at c_{16}, c_{23}, c_{30}, before the final Feistel copy. These are matches to the published enumeration-enhanced bounds, separate from the strict improvements above.

We also evaluated the 15-round SIMON32 counts on 1,000 distinct random master keys with the real key schedule and offset $a = 0$. Every count was even, and each component attained $v_2(N_1) = 1$; divisibility by 4 failed for 510, 492, and 520 keys at c_0, c_7, c_{14}, respectively. Thus the experiments confirm the proved bound on the tested keys and rule out an all-key improvement to $t = 2$ for these structures and components.

The black-box discoveries in Table 16 use ordinary coordinate cubes and 1,000 distinct validation master keys. Both complementary offsets are covered. Parenthesized round counts include the input extension. The analytical bounds cover arbitrary round keys and therefore extend; the empirical minima concern only the directly tested rounds.

Table 16: SIMON32 discovery-to-proof comparison, with data complexity 2^{31} per distinguisher. The sample column gives the minimum observed $v_2(N_1)$ over 1,000 keys. RC and Joint are all-key lower bounds.

Rounds	Active mask u	Output parity	Sample	RC	Joint
9 (10)	`fffdffff`	$c_1 \oplus c_{15} \oplus c_{17}$	12	**7**	6
10 (11)	`fffeffff`	c_{14}	10	4	**7**
11 (12)	`ffefffff`	c_4	8	0	**4**
12 (13)	`fffeffff`	c_8	4	0	**2**
14 (15)	`fffeffff`	$c_1 \oplus c_{15}$	1	0	0

For the separate 11-round linear approximation with masks $(\alpha, \beta) = (\texttt{00010000}, \texttt{00000080})$, the support model proves ZC, consistent with all 1,000 tested keys. For either halfspace offset its count is therefore exactly 2^{30}. The candidates in Table 17, Table 18, and Table 19 illustrate multi-bit output differences, multi-bit linear masks, and nontrivial DL-count divisibility, respec-

tively. Their domains are evaluated exhaustively, but the positive claims remain empirical over keys. The 11-round ID and ZC rows match the round coverage of [CHNE24, Section 5.1, Tables 8–9]. An integral input extension does not automatically produce a fixed-mask ZC or fixed-difference ID extension.

Table 17: Selected 11-round SIMON32 ID candidates. Each ordered DDT entry is zero on 1,000 tested keys. All 2^{32} base inputs are evaluated per key.

Input difference δ	Output differences γ		
00000001	00810000	02050000	40800000
00000002	01020000	040a0000	81000000
00000080	02810000	40800000	40a00000
00008000	01400000	81020000	a0400000

Table 18: Additional SIMON32 ZC candidates, with full-codebook sums zero on 1,000 tested keys.

Rounds	α	β
10	00810000	0281c1e2
10	00810000	4081f160
11	00050000	00000205
11	000a0000	0000040a

Table 19: 10-round SIMON32 DL one-count divisibility over V. Each row uses 1,000 keys and 2^{32} base inputs per key.

δ	β	Sample minimum $v_2(N_1^{\mathrm{DL}})$
00000001	00000400	5
00000040	00000001	5
00002000	00000080	5

F.2 SPECK

Specification. SPECK [BSS+15] is a family of ARX block ciphers with the same block sizes $n = 2w$. We use the packing $2^w s + t$, where s contains global bits $w, \ldots, 2w-1$ and t contains bits $0, \ldots, w-1$. Its round function is

$$s' = \big((s \ggg \sigma) + t \bmod 2^w\big) \oplus k_r, \qquad t' = (t \lll \tau) \oplus s'.$$

The rotation amounts are $(\sigma, \tau) = (7, 2)$ for $w = 16$ and $(8, 3)$ for the larger word sizes. Figure 9 shows the data path. Our black-box SPECK32/64 experiments use the first encryption rounds, with round keys derived from a 64-bit key using the standard key schedule. The analytical results for larger variants allow arbitrary round keys. The parenthesized input-side extensions in the result tables change the queried input set as explained there; the masks in those tables describe the directly analyzed core.

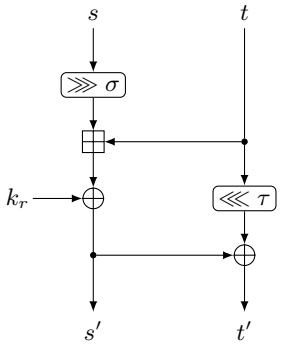


Fig. 9: The SPECK round function.

Results. Table 20 gives our SPECK32 black-box discoveries, and Table 21 gives the corresponding analytical bounds. Let $\beta_* =$ 02050204, selecting $c_2 \oplus c_9 \oplus c_{16} \oplus c_{18} \oplus c_{25}$. All affine offsets are covered in these finite-count rows. For the dimension-30 structure, the 1,000-key experiment checks all four cosets. Those 4,000 counts are not 4,000 independent instances. Earlier 6-round SPECK32 integrals have data complexity 2^{30} [WHG+20]. The established 7-round parity in [HKK+26] is strengthened to divisibility by 8 through the residue relation in Equation 44.

Table 20: SPECK32 finite-count observations. Parenthesized rounds denote candidate integral extensions. Samples concern the unparenthesized rounds. A hexadecimal input denotes the active mask u.

Rounds	Input structure	Data	β	Sample t	Keys
7 (8)	`ffffff9f`	2^{30}	β_*	**3**	1000
7 (8)	`ffffffbf`	2^{31}	β_*	**3**	1000
7 (8)	`ffffff7f`	2^{31}	β_*	3	1000
7 (8)	`fffffeff`	2^{31}	β_*	1	1000
6 (7)	`ffffffbf`	2^{31}	00300030	10	1000
6 (7)	$x_8 \oplus x_{10} = b$, $b \in \mathbb{F}_2$	2^{31}	β_*	10	1000

Table 21: All-round-key linear-model bounds for the SPECK32 observations. Parentheses in the bound column denote complete residue refinement of RC. The local linear and RC models give the same base values.

Rounds	Input structure	Data	β	Base t (refined)	Sample t
6 (7)	$x_8 \oplus x_{10} = b$	2^{31}	β_*	**10**	10
6 (7)	`ffffffbf`	2^{31}	00300030	6 (**10**)	10
7 (8)	`ffffffbf`	2^{31}	β_*	1 (**3**)	3
7 (8)	`ffffff7f`	2^{31}	β_*	0 (**1**)	3
7 (8)	`fffffeff`	2^{31}	β_*	0	1
7 (8)	`ffffff9f`	2^{30}	β_*	0 (**3**)	3

For the 6-round hyperplane $x_8 \oplus x_{10} = b$ in Table 21, a unique minimum-valuation trail proves $v_2(N_1) = 10$ for every key, and hence a nonzero linear sum. By contrast, Table 22 records exact balance. The 7-round row has an all-round-key support proof; the remaining displayed rows are sampled observations. These are the same events as their halfspace linear zeros under Equation 6, not two independent discoveries.

Table 22: SPECK32 halfspace counts for $\mathcal{X} = \{x : \alpha \cdot x = b\}$, either b, with data complexity 2^{31}. The 7-round row is proved for all round keys. The remaining rows are empirical.

Rounds	α	β	Exact N_1	Keys
7 (8)	00000020	β_*	2^{30}	1000
6 (7)	00000001	β_*	2^{30}	1000
6 (7)	00000021	β_*	2^{30}	1000
6 (7)	00000041	β_*	2^{30}	1000
6 (7)	00000081	β_*	2^{30}	1000

Table 23: 7-round SPECK32 ID candidates. All entries are zero on 1,000 keys after evaluating the complete 2^{32}-input domain.

δ	γ
00200000	80008002
00400000	80008002

Both rows of Table 23 use an output difference of Hamming weight 3, which a search restricted to single-bit output differences would miss. IDs covering 7

rounds were reported in [LGCX19]; our black-box search finds the displayed multi-bit output difference at the same round count.

For the integral configurations of [WHG26], Table 24 and Table 25 give automated proofs covering arbitrary round keys and all affine offsets. Define

$$\beta_n \cdot c = c_3 \oplus c_{11} \oplus c_{n/2} \oplus c_{n/2+3} \oplus c_{n/2+11}.$$

Parenthesized round counts include the input extension without changing the data complexity. Exponent 1 reproduces the earlier zero-sum property; higher exponents and exact counts strengthen it.

Table 24: Proved SPECK count-divisibility bounds. n is the block size. Parentheses in the bound column denote residue refinement by trail enumeration.

n	Rounds	Active mask u	Data	Output	t (refined)
48	6 (7)	`fffffffffff1f`	2^{45}	$c_3 \oplus c_{27}$	**6 (7)**
48	7 (8)	`fffffffffff1f`	2^{45}	$\beta_{48} \cdot c$	**7 (8)**
64	6 (7)	`ffffffffffffff1f`	2^{61}	$c_3 \oplus c_{35}$	1
128	6 (7)	`ffffffffffffffff`	2^{125}	$c_3 \oplus c_{67}$	1
		`ffffffffffffff1f`			

Table 25: Exact SPECK counts from linear support proofs and conversion. These rows require no trail enumeration.

n	Rounds	Active mask u	Data	Output	Exact N_1
32	6 (7)	`ffffff1f`	2^{29}	$c_2 \oplus c_{18}$	2^{28}
32	6 (7)	`ffffff9f`	2^{30}	$c_2 \oplus c_{18}$	2^{29}
64	7 (8)	`ffffffffffffff1f`	2^{61}	$\beta_{64} \cdot c$	2^{60}
96	6 (7)	`ffffffffffffffffffffe9f`	2^{93}	$c_6 \oplus c_{54}$	2^{92}
96	7 (8)	`ffffffffffffffffffffe9f`	2^{93}	$\beta_{96} \cdot c$	2^{92}
128	**7 (8)**	`ffffffffffffffff`	2^{125}	$\beta_{128} \cdot c$	2^{124}
		`ffffffffffffff1f`			

For SPECK48 in Table 24, the refined exponents 7 and 8 require enumeration of only 13 and 2 trails, respectively. The exact counts in Table 25 follow from ordinary linear support proofs and conversion, without residue refinement.

F.3 BEANIE

Specification. BEANIE is a 32-bit tweakable block cipher with a 128-bit key and a 128-bit tweak [GHM+25]. The state consists of eight nibbles. A regular round

applies a round-key XOR, eight parallel BEANIE S-boxes, a nibble permutation, and a column-wise linear map. If nibble j contains bits $4j, \ldots, 4j+3$, the permutation exchanges nibbles 0 and 4, and nibbles 2 and 6. The linear map acts separately on each 16-bit column, whose nibbles are ordered from most to least significant, by

$$M_{\mathrm{B}} = \begin{pmatrix} 2 & 1 & 1 & 9 \\ 1 & 4 & \mathtt{f} & 1 \\ \mathtt{d} & 9 & 4 & 1 \\ 1 & \mathtt{d} & 1 & 2 \end{pmatrix} \quad \text{over } \mathbb{F}_{16} = \mathbb{F}_2[z]/(z^4 + z + 1).$$

The entries are hexadecimal polynomial representations. The final round omits this column map and ends with whitening.

Our applications use suffixes of the standard 5-round cipher, retaining its tweakey schedule. For fixed K, T_1, T_2, let $E_{K,T}[i:5]$ denote the encryption suffix beginning at round i and including the special final round and whitening. The analyzed 32-bit permutation is

$$U_{r_1,r_2} = \big(E_{K,T_2}[5-r_2:5]\big)^{-1} \circ E_{K,T_1}[5-r_1:5].$$

Both suffixes end at the ciphertext of the 5-round cipher. The master key is shared, while the tweaks may differ; all three remain fixed within each integral, linear, or DL evaluation. The total number of traversed rounds is $r_1 + r_2$.

Results. For each mask α, we use $\mathcal{X} = \{x : \alpha \cdot x = b\}$, with arbitrary b and data complexity 2^{31}. Each sampled configuration consists of one master key and two distinct tweaks, fixed throughout the evaluation. Table 26 compares the linear model before and after exact contraction at the forward/inverse junction with Joint using the same contraction. Joint uses a whole-round Walsh prefix, then ultrametric propagation. All analytical bounds cover every key, tweak, and offset, without trail enumeration. Joint matches the sampled minimum 12 for $U_{3,2}$; for $U_{2,3}$, the sampled minimum remains 2 exponent bits higher. Table 27 gives the separate full-domain DL-count minima, and Table 28 records halfspace exact-balance observations.

Table 26: BEANIE 5-round count bounds and empirical minima. Linear parentheses and all Joint bounds use exact contraction at the forward/inverse junction. Each sample uses 1,000 key/tweak configurations.

U-shape	α	Data	β	Linear t (refined)	Joint t	Sample t
$U_{2,3}$	`00004020`	2^{31}	`000440b0`	0 (8)	**10**	12
$U_{3,2}$	`c0e00d00`	2^{31}	`00000001`	2 (10)	**12**	12

For 6-round $U_{3,3}$, take $x_0 = b$ and $\beta =$ `a0700601`. Joint proves $2 \mid N_1$ for every key/tweak configuration and both offsets, using Walsh propagation through the

Table 27: BEANIE full-domain DL-count observations. The domain and distinct-query count are both 2^{32}. Each row uses 1,000 configurations.

U-shape	δ	β	Sample minimum $v_2(N_1^{\mathrm{DL}})$
$U_{2,3}$	`00800400`	`e070000c`	**13**
$U_{3,2}$	`00000080`	`20000700`	**15**

Table 28: 5-round BEANIE $U_{4,1}$ exact-balance observations on 1,000 configurations. Each mask defines either halfspace, with data complexity 2^{31}.

α	β	Exact N_1
e_i, $0 \le i \le 5$	`00000001`	2^{30}
`00000001`	`0000cafe`	2^{30}

first S-box layer, ultrametric propagation thereafter, and exact contraction at the forward/inverse junction, without trail enumeration. The stronger condition $4 \mid N_1$ holds on all 1,000 tested configurations but remains empirical. Exactly 468 attain valuation 2. These configurations use distinct tweaks $T_1 \neq T_2$. The design [GHM+25] includes longer or lower-data ID/ZC and integral constructions; our results add count-divisibility information for the stated U-shaped permutations.

F.4 PRESENT

Specification. PRESENT [BKL+07] is a 64-bit SPN with 80- or 128-bit master keys. A round applies a round-key XOR, sixteen parallel PRESENT S-boxes, and the bit permutation

$$P(i) = 16i \bmod 63 \quad (0 \le i < 63), \qquad P(63) = 63.$$

A final whitening-key XOR follows the last round. Figure 10 shows the bit permutation. Our result tables use global output bits after the final permutation. Where a local model ends at the last S-box, we transport its output mask through this permutation before reporting it. The analytical bounds allow every round-key value and hence apply to both standard key schedules.

Results. The coordinate-cube comparisons match the rounds, input sets, and ciphertext outputs of [BV24]. Their bit index i corresponds to our global index $i - 1$. Table 29 gives the selected small-data comparisons.

The 4-round improvements in Table 29 use Walsh propagation through the initial rounds. At 5 rounds, exact contraction to a 16-bit convolution supplies an extra factor of 2 beyond the Joint bounds (Subsection E.8). This proves the stronger bounds for $c_{20}, c_{28}, c_{52}, c_{60}$ previously observed experimentally in [BV24, Table 5]; the bounds for c_4, c_{12} match that work's enumeration-enhanced values.

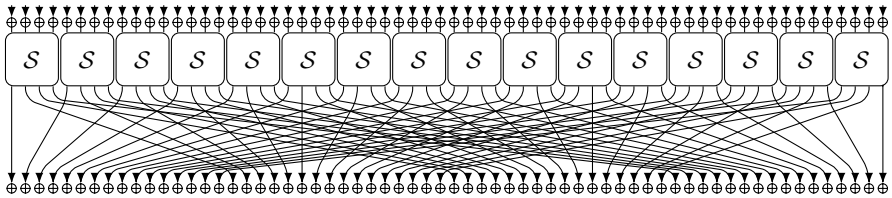


Fig. 10: A round of PRESENT.

Table 29: PRESENT count bounds on $\mathcal{X} = a + (u \wedge \mathbb{F}_2^{64})$. Parentheses denote trail enumeration in [BV24] and exact 16-bit contraction for our bounds. Each listed bit defines a separate output component.

Rounds	Active mask u	Data	Output bits	t [BV24]	Joint t
4	`000000000000000f`	2^4	c_0	2 (3)	3
4	`000000000000000f`	2^4	c_{16}, c_{32}, c_{48}	1	3
5	`0000000000000fff0`	2^{12}	c_4, c_{12}	2 (3)	2 (3)
5	`0000000000000fff0`	2^{12}	$c_{20}, c_{28}, c_{52}, c_{60}$	1	1 (2)

For c_4, a real PRESENT-80 key gives $N_1 = 1944$, attaining $t = 3$ with a nonzero signed sum. The tested Joint models, covering 30 round-wide basis choices and 15 choices near the output for each listed bit, are optimal at their unrefined bounds. Spatial Joint also reproduces 55 published bounds at 5 and 6 rounds without enumeration.

Table 30 summarizes the complete coordinate-cube comparison, and Table 31 gives all six additional low-data targets.

Table 30: Coverage of all 384 published PRESENT coordinate-cube/output targets. Equal bounds include exponent zero. The choice of model varies by output.

Rounds	Active mask u	Data	Stronger / equal	Methods
4	`000000000000000f`	2^4	3 / 61	Round-wide Joint
5	`0000000000000fff0`	2^{12}	4 / 60	Round-wide/Spatial Joint, exact contraction
6	`00000000ffffffff`	2^{32}	0 / 64	Round-wide and Spatial Joint
7	`fffffffffffff000`	2^{52}	0 / 64	Round-wide Joint
8	`fffffffffffffffe`	2^{63}	0 / 64	Round-wide Joint
9	`fffffffffffffffe`	2^{63}	0 / 64	Round-wide Joint

The longest reported coordinate-cube example has 9 rounds and data complexity 2^{63}: $t = 2$ for c_0, c_4, c_8, c_{12}, and $t = 1$ for 24 further individual bits. These

Table 31: All six low-data PRESENT targets from Section 7.4 of [BV24]. Each listed bit is a separate test. The two-bit row covers two targets.

Rounds	Active mask u	Data	Output	t [BV24]	Joint t
6	`0000000000000ff0`	2^8	c_0	1	1
6	`000000000000eff0`	2^{11}	c_0	2	2
6	`000000000000fff0`	2^{12}	c_0	4	4
6	`000000000000fff0`	2^{12}	c_1, c_3	1	1
6	`000000000000fff0`	2^{12}	c_2	2	2

match the published strengthening of the known integral, rather than extend its round count.

The separate affine structures in Table 32 demonstrate finite divisibility even when linear trails exist. Their masks are specified at the ciphertext, after the last permutation. These large-data affine structures differ from the small coordinate cubes of Table 29.

Table 32: PRESENT linear-to-integral bounds for arbitrary equation values. Parentheses denote trail-parity enumeration as in Subsection E.4.

Rounds	Affine input structure	Data	Output parity	Linear t
5	$x_{28} \oplus x_{31} = b_0,\ x_{29} \oplus x_{30} = b_1$	2^{62}	$c_7 \oplus c_{39} \oplus c_{55}$	26
5	$x_{29} \oplus x_{30} = b_0$	2^{63}	$c_7 \oplus c_{39} \oplus c_{55}$	33 (**34**)
6	$x_{20} \oplus x_{23} = b_0,\ x_{21} \oplus x_{22} = b_1$	2^{62}	$c_{11} \oplus c_{27} \oplus c_{43} \oplus c_{59}$	7
6	$x_{21} \oplus x_{22} = b_0$	2^{63}	$c_{11} \oplus c_{27} \oplus c_{43} \oplus c_{59}$	17 (**18**)

A ZC mask family illustrates a different use of the conversion. Fix input mask e_0 and let the output mask range over

$$\beta \in \{v \in \mathbb{F}_2^{64} : v \preceq \texttt{eeeeeeeeeeeeeeee}\} \setminus \{0\}.$$

One support-SAT query, with the 48 permitted mask bits free, is UNSAT for 5-round PRESENT. Thus $L_{e_0,\beta} = 0$ for all $2^{48} - 1$ such masks simultaneously. Under decryption, fixing those 48 ciphertext bits and varying the other sixteen gives $N_1 = 2^{15}$ for plaintext bit x_0. This gives an exact-balance distinguisher with 2^{16} chosen-ciphertext queries. Families of this size were also studied in [BV24, Section 7.3].

For Table 33, fix $\delta = e_0$, write $c = E_K(x)$ and $c' = E_K(x + \delta)$, and count the inputs in $\mathcal{X}$ for which $c_0 \neq c'_0$. The exact mix-basis DL MILP bounds every required coefficient, including the one with input mask zero, before conversion. Because $\delta \notin A$, these are cross-coset tests and the number of distinct queries is twice the cube size. In particular, a half-codebook base structure requires a full codebook of queries. Table 10 bounds the corresponding random-permutation probabilities.

Table 33: PRESENT DL one-count bounds for $\delta = e_0$ and compared bits c_0, c_0'. All keys and affine offsets are covered.

Rounds	Active mask u	Base inputs	Distinct queries	DL t
5	`fffffffffffffff0`	2^{60}	2^{61}	**41**
5	`fffffffffffffffc`	2^{62}	2^{63}	**43**
5	`fffffffffffffffe`	2^{63}	2^{64}	**44**
6	`fffffffffffffffe`	2^{63}	2^{64}	**9**

F.5 Ascon

Specification. We use the 320-bit Ascon permutation [DEMS21]. The results concern neither the authenticated-encryption mode nor a later serialization convention. Let $S_j = \sum_{i=0}^{63} x_{64j+i} 2^i$ for $0 \le j < 5$. The packed state is $\sum_{j=0}^{4} S_j 2^{64j}$. A round first adds its constant to the low byte of S_2, then applies the Ascon five-bit S-box independently to the 64 columns, and finally applies

$$S_j \longmapsto S_j \oplus (S_j \ggg a_j) \oplus (S_j \ggg b_j),$$

where

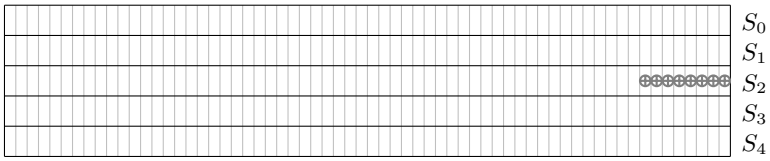
$$(a_j, b_j)_{j=0}^{4} = ((19, 28), (61, 39), (1, 6), (10, 17), (7, 41)).$$

In a column, the S-box input is ordered as $(x_i, x_{64+i}, x_{128+i}, x_{192+i}, x_{256+i})$, with the first component the most significant bit of the conventional five-bit S-box encoding. The 5-round applications use constants `87`, `78`, `69`, `5a`, `4b` in this order. All reported output bits are after the final diffusion layer. These statements concern a fixed public permutation, so no secret-key quantifier is needed. Figure 11 illustrates the three round operations using this state convention.

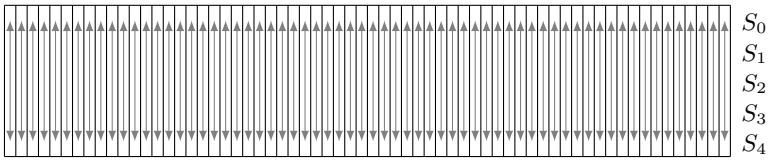
Results. Let P denote the five-round Ascon permutation. In Table 34, the linear MILP bounds all nonzero masks in the space orthogonal to each input structure. For the first row, the all-Walsh specialization of Joint gives an equivalent linear model and reaches its optimum. The remaining entries are certified lower bounds. All unmentioned input coordinates are free and every stated offset is covered.

Table 34: 5-round Ascon linear-to-integral one-count bounds. Fixing more input bits reduces the data complexity and the proved bound for the same output bit.

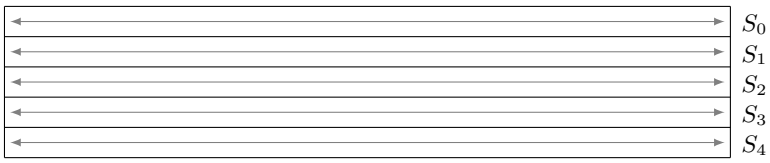
Rounds	Affine input structure	Data	Output	Linear t
5	$x_{256} = b_0$	2^{319}	c_{256}	**60**
5	$x_{256+j} = b_j,\ 0 \le j < 2$	2^{318}	c_{256}	53
5	$x_{256+j} = b_j,\ 0 \le j < 3$	2^{317}	c_{256}	50
5	$x_{256+j} = b_j,\ 0 \le j < 4$	2^{316}	c_{256}	47



(a) Round-constant addition p_C.



(b) Bit-sliced S-box layer p_S.



(c) Word-wise linear diffusion layer p_L.

Fig. 11: Ascon round operations. Row S_j consists of global bits $x_{64j}, \ldots, x_{64j+63}$, drawn from most to least significant (left to right). The constant affects $x_{128}, \ldots, x_{135}$, the rightmost eight cells of S_2. Vertical columns are S-box inputs.

We also apply the coset-indicator identity to an affine family of input masks. Set

$$q = \texttt{b1d0d73df0153225}, \qquad \alpha_0 = e_1 + e_{193} + e_{257},$$
$$U = \langle e_{64j+k} : 0 \le j < 5,\ q_k = 1 \rangle.$$

Writing the output mask as five 64-bit words in increasing bit-index order (bits 0–63 first), we have $\beta = (\texttt{25b76d24c9125b69}, \texttt{3dd4b48f9974dc44}, \texttt{0000000000000000}, \texttt{0000000000000000}, \texttt{79b8971fd852f98c})$. One whole-family support query proves $L_{\alpha,\beta} = 0$ for every $\alpha \in \alpha_0 + U$, a family of size 2^{155}.

For $A = U^{\perp}$, the conclusion is

$$\sum_{x \in a + A} (-1)^{\alpha_0 \cdot x + \beta \cdot P(x)} = 0.$$

The input-dependent sign gives a weighted integral with data complexity 2^{165}, rather than ordinary balance of the output parity. This illustrates the conversion for affine ZC families, whose size and round coverage were already reported in [HGSE24].

F.6 GIFT-64

Specification. GIFT-64 [BPP+17] is a 64-bit SPN with a 128-bit key. Each round applies sixteen parallel GIFT S-boxes, PermBits, and AddRoundKey. The permutation sends input bit i to

$$P(i) = 4\left\lfloor \frac{i}{16} \right\rfloor + 16\left(\left(3\left\lfloor \frac{i \bmod 16}{4} \right\rfloor + (i \bmod 4)\right) \bmod 4\right) + (i \bmod 4).$$

The key is added at positions $4i$ and $4i+1$, for $0 \le i < 16$. The affine round constant is injected at positions $3, 7, 11, 15, 19, 23, 63$. The round is shown in Figure 12. Our output masks are at ciphertext, after all three operations. A mask specified immediately after the last S-box layer is transported through PermBits before it appears in a result table. The input-side extension uses the inverse of the public SubCells–PermBits map; there is no input whitening to invert. For constant-sensitive models, the stated extension uses the corresponding shifted sequence of round constants.

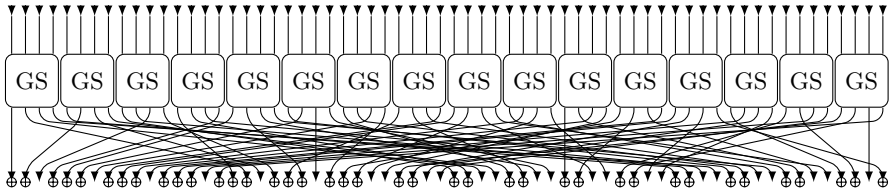


Fig. 12: A round of GIFT-64.

Results. Define $q(c) = c_{12} \oplus c_{29} \oplus c_{46} \oplus c_{55} \oplus c_{63}$. The 5-round linear result in Table 35 uses this multi-bit output, not individual output bits. By summing the local input contributions before taking valuations, the linear model improves the exponent from 52 to 53; Joint gives the same bound. The 6-round affine row is a separate linear-subspace result.

Table 35: GIFT-64 linear one-count bounds. Parenthesized rounds include the input extension at unchanged data complexity. The refined linear model incorporates exact local input propagation.

Rounds	Input structure	Data	Output	Linear t	Refined linear t	Joint t
5 (6)	$u =$ `ffffffff1ffeffff`	2^{60}	$q(c)$	52	**53**	**53**
6 (7)	$x_{10} \oplus x_{11} \oplus x_{40} \oplus x_{42} = b_0$	2^{63}	c_{45}	**7**	–	–

Table 36: Longer GIFT-64 coordinate-cube count bounds. Joint and Spatial Joint combine both representations. Ultrametric uses no Walsh transitions. These are our implementations, not published GIFT bounds from [BV24].

Rounds	u	Data	Output	Joint t	Spatial Joint t	Ultra-metric t
7 (8)	`fffff0fffffffff`	2^{60}	c_{58}	**4**	–	5
8 (9)	`fffff0fffffffff`	2^{60}	c_{44}	1	**2**	2
9 (10)	`ffffffffffffff8`	2^{61}	$c_4, c_{20}, c_{24}, c_{28}, c_{32}, c_{36}, c_{48}, c_{56}$	–	**1**	1

Table 36 gives one-count divisibility bounds for 7–9 rounds. Its 9-round row proves zero-sum separately for each of the 8 listed bits. The designers also give 9-round integrals with a 1-round input extension [BPP+17]; our parenthesized 10 rounds use the same extension principle. At 8 and 9 rounds, Spatial Joint uses a single Walsh coordinate near the output and matches the purely ultrametric bound; at 7 rounds, the ultrametric bound is 1 bit higher. Thus these examples demonstrate Joint's applicability, while the benefit of mixing representations is established by the SIMON and PRESENT comparisons. Table 37 records exact balance from support proofs on the stated structures.

Table 37: GIFT-64 exact counts from support proofs and conversion. The input extension preserves these all-round-key statements.

Rounds	u	Data	Output	Exact N_1
5 (6)	`ffffffff5ffeffff`	2^{61}	$q(c)$	2^{60}
6 (7)	`fffff0fffffffff`	2^{60}	c_{45}	2^{59}

For Table 38, we set $\delta = e_0$ and count inputs for which the indicated bits of $c = E_K(x)$ and $c' = E_K(x + \delta)$ differ. Here $\delta \in A$, so both encryptions remain inside the same structure and the base-input count equals the distinct-query count. Table 10 gives random-permutation bounds for the 5-round rows.

Table 38: GIFT-64 DL one-count bounds from exact mix-basis DL models and conversion, with $\delta = e_0$.

Rounds	u	Inputs / queries	Compared bit	DL t
4	`ffffbfffffffffff`	2^{63}	c_{63}, c'_{63}	**45**
5	`fffff0fffffffff`	2^{60}	c_{59}, c'_{59}	**15**
5	`fffff3fffffffff`	2^{62}	c_{59}, c'_{59}	**21**
5	`fffffbfffffffff`	2^{63}	c_{59}, c'_{59}	**26**

F.7 SKINNY-64

Specification. We use the single-tweakey cipher SKINNY-64-64 [BJK+16]. Its 4×4 state of nibbles is serialized row-wise from the most significant hexadecimal digit. Thus cell X_j, for $0 \le j < 16$, contains global bits $60-4j, \ldots, 63-4j$. Each round consists of SubCells, AddConstants, AddRoundTweakey, ShiftRows, and MixColumns. Only the first two tweakey rows are added to the state. The tweakey remains fixed across queries and follows the standard single-tweakey schedule. ShiftRows rotates row j right by j cells, and MixColumns multiplies each column by

$$M_{\mathrm{S}} = \begin{pmatrix} 1\,0\,1\,1 \\ 1\,0\,0\,0 \\ 0\,1\,1\,0 \\ 1\,0\,1\,0 \end{pmatrix} \quad \text{over } \mathbb{F}_2.$$

Figure 13 shows this round structure. Every reported output parity is evaluated after the complete final round, including ShiftRows and MixColumns. In particular, the cell ordering does not change the global least-significant-bit convention used for c_i.

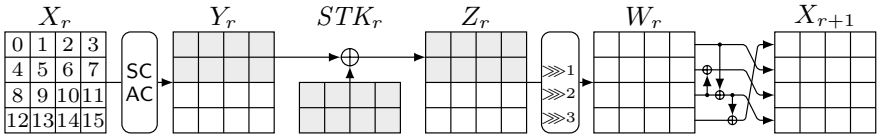


Fig. 13: A round of SKINNY-64. Cell j contains global bits $60-4j, \ldots, 63-4j$.

Results. The support proofs establish exact counts in Table 39, not just zero-sum properties. For the 11-round row, the input structure and output parity coincide with the specified relation in [ZYS+26]; its bit numbering is reversed relative to ours. Exact balance is a stronger constraint at the same data complexity, with its random-permutation probability given in Table 9. Table 40 gives the separate low-data zero-sum result, while Table 41 lists all four 11-round cube placements with exact balance.

On different input structures, Derbez and Fouque [DF20] report integral data complexities 2^{15}, 2^{47}, and 2^{63} at 8, 10, and 11 rounds. The exact-balance structures in Table 39 use 2^{12}, 2^{48}, and 2^{60}, respectively: less data at 8 and 11 rounds, but more at 10 rounds. For the separate 8-round 2^8 structure in Table 40, the contraction in Subsection E.7 proves zero-sum. Zero tweakey and zero offset give $N_1 = 124$, so exact balance does not hold here. Tweakey `becccfc892529f2f` and inactive offset `0063bb62d4ca9100` give $N_1 = 126$. Thus the proved exponent 1 is also the strongest possible uniform bound on this configuration.

Table 39: SKINNY-64 exact one-counts from linear support and conversion. ZYS denotes the published zero-sum bound of [ZYS+26] for the same 11-round structure and output parity.

Rounds	u	Data	Output parity	ZYS t	Exact N_1
8	`000000f00f00f000`	2^{12}	$c_0 \oplus c_{16} \oplus c_{32}$	–	2^{11}
10	`f0fffffffff00ff0`	2^{48}	$c_{12} \oplus c_{60}$	–	2^{47}
11	`0fffffffffffffff`	2^{60}	$c_{12} \oplus c_{28} \oplus c_{44}$	1	2^{59}

Table 40: The 8-round SKINNY case with data complexity 2^8. The parenthesized exponent uses Joint with exact contraction of the final layers.

Rounds	u	Data	Output parity	ZYS t	Joint t
8	`00000000000000ff`	2^8	$c_3 \oplus c_{19} \oplus c_{35}$	1	0 (1)

Table 41: All four 11-round SKINNY-64 cube placements with exact balance for $c_{12} \oplus c_{28} \oplus c_{44}$. All offsets and tweakeys are covered.

Rounds	u	Data	Exact N_1
11	`0fffffffffffffff`	2^{60}	2^{59}
11	`f0ffffffffffffff`	2^{60}	2^{59}
11	`ff0fffffffffffff`	2^{60}	2^{59}
11	`fff0ffffffffffff`	2^{60}	2^{59}

F.8 RC5-16

Specification. RC5 [Riv95] uses modular addition, XOR, and data-dependent rotations. Our 32-bit applications use RC5 with 16-bit words and a 128-bit master key. Let $s = \sum_{i=0}^{15} x_i 2^i$ and $t = \sum_{i=0}^{15} x_{16+i} 2^i$. Initial whitening adds the first two expanded subkeys to s and t modulo 2^{16}. A full round $r \geq 1$ then performs

$$s' = ((s \oplus t) \lll (t \bmod 16)) + k_{2r} \pmod{2^{16}},$$
$$t' = ((t \oplus s') \lll (s' \bmod 16)) + k_{2r+1} \pmod{2^{16}}.$$

The rotations depend on intermediate data, and consequently also on the key; they are not fixed key-selected rotations. Our 2.5-round experiments use the expanded-key table of a 3-round parent cipher, retain initial whitening, execute 2 full rounds, and then execute only the first update of round 3. After this update, we pack the output as $s' + 2^{16}t$. Full-round screening experiments use the stated number of complete rounds and its corresponding standard key expansion.

Results. The observations in Table 42 use 1,000 distinct keys in five disjoint batches of 200 keys. Each differential count evaluates all 2^{32} base inputs. Zero counts are included when taking valuations. Each row attains its displayed minimum on one key, and each is zero on at least 800 keys. These finite-divisibility

candidates were selected from the same sample, without independent holdout validation. Table 43 lists selected ID candidates from the same sample.

Table 42: RC5-16, 2.5-round ordered DDT-count divisibility. The candidates were selected from 1,000 sampled keys.

δ	γ	Sample minimum v_2(DDT)	Keys
80008000	04100000	**11**	1000
80008000	0c100000	**10**	1000

Table 43: Selected RC5-16, 2.5-round ID candidates. These cells are zero on all 1,000 sampled keys.

δ	Output differences γ
80008000	10300000, 11300000, 11500000, 11900000, 12300000, 13100000

The five batches share at least 89 zero cells and 63 finite-divisibility candidates with sampled minimum exponents 5–11. These are lower bounds on the number of candidates because each batch retains only a limited number of results. Hu et al. [HLJ+20] already give 2.5-round IDs for the same block size. Their reported search fixes paired input bits and one output bit, whereas the displayed output differences have Hamming weight 3 or 4. Our search therefore broadens the tested output differences at the same round count.

An exact SAT model of two encryptions with a shared symbolic master key could certify these ID candidates, including the data-dependent rotations. The results reported here remain empirical. A pair-existence test alone does not establish integral or finite count divisibility, which concern aggregate counts. The integral parity observations in Table 44 illustrate this additional aggregate information.

Table 44: RC5-16, 2.5-round integral parity observations. The data complexity is 2^{31}. Both affine offsets are covered on 1,000 holdout keys.

Input condition	β	Sample t	Keys
$x_{23} = b$	00020000	1	1000
$x_{24} = b$	00030000	1	1000

Toggling the most significant bit of both words preserves the rotation controls through whitening and cancels in the first word XOR, explaining the role of input difference `80008000`. Completing the third round removes surviving zero cells and extra divisibility in the scanned family in a separate 200-key experiment. No claim is made about all differentials.

F.9 ChiLow-32

Specification. Our ChiLow-32 applications concern its 32-bit data-only *decryption* permutation with a 128-bit key and a 64-bit tweak [BDG+25]. They do not include the parallel tag-generation path. Each analyzed round consists of the nonlinear ChiChi-32 layer, the fixed linear map

$$y_i = x_{(11i+5) \bmod 32} \oplus x_{(11i+9) \bmod 32} \oplus x_{(11i+12) \bmod 32}, \qquad 0 \le i < 32,$$

and the round injection derived from the key, tweak, and constant. The decryption input also has the specified initial key XOR.

For completeness, write $\chi_s(v)_i = v_i \oplus (1 \oplus v_{i+1})v_{i+2}$, with indices modulo s. ChiChi-32 first applies χ_{15} to bits $0, \ldots, 14$ and χ_{17} to bits $15, \ldots, 31$. It then XORs a linear function of the layer input into output positions $13, \ldots, 16$:

$$\ell_{13} = x_{13} \oplus x_{16}, \qquad \ell_{14} = x_{14} \oplus x_{15},$$
$$\ell_{15} = x_{13} \oplus x_{15} \oplus x_{16}, \ell_{16} = x_{14} \oplus x_{16}.$$

The key and tweak are fixed during every evaluation. Our reduced-round permutation includes the last diffusion layer and round injection; it is not the variant that omits final diffusion. Thus x and c in these tables denote the input and output of this decryption permutation, respectively.

Results. Global input and output indices refer to this permutation. All halfspaces in Table 45 cover both offsets through the exact linear conversion. Table 46 gives a separate smaller-data parity observation on a zero-offset cube.

Table 45: ChiLow-32 decryption count observations. The last row has constant parity 1 on every tested key/tweak pair, despite having $t = 0$.

Rounds	Input structure	Data	β	Sample t	Key/tweak pairs
4	$x_0 = b$	2^{31}	`25c72ab9`	**3**	1000
4	$x_0 = b$	2^{31}	`1b97a34d`	**3**	1000
5	$x_0 = b$	2^{31}	`15265ab9`	1	1000
5	$x_0 = b$	2^{31}	`269142a8`	1	1000
5	$x_2 = b$	2^{31}	`50da630c`	1	1000
5	$x_1 = b$	2^{31}	`4dafb20f`	0	1000

Table 46: A smaller-data ChiLow-32 decryption parity observation. The coordinate cube has offset $a = 0$.

Rounds	Active mask u	Data	β	Sample t	Key/tweak pairs
4	`0000aaaa`	2^8	`372f469a`	1	1000

The designers [BDG+25] gave a 5-round integral under a fixed key and tweak. Peng et al. [PKA+26] give low-data single-tweak integrals and a 6-round multiple-tweak integral, which changes the input model. Our results add count-divisibility information for combinations of output bits, together with a constant-one parity; the round and data complexities differ from those earlier integrals.

F.10 KATAN32

Specification. KATAN32 [DDK09] has a 32-bit block, an 80-bit key, and 254 rounds. The two feedback registers occupy global bits $0, \ldots, 18$ and $19, \ldots, 31$. With x_i denoting a round-input bit, define

$$f_r = x_{31} \oplus x_{26} \oplus x_{27}x_{24} \oplus \mathrm{IR}_r x_{22} \oplus k_{2r},$$
$$g_r = x_{18} \oplus x_7 \oplus x_{12}x_{10} \oplus x_8 x_3 \oplus k_{2r+1},$$

where IR_r is the specified public irregular sequence. Each round sets $x'_0 = f_r$, $x'_{19} = g_r$, and $x'_i = x_{i-1}$ for the remaining positions. The key sequence begins with the master-key bits and satisfies

$$k_i = k_{i-80} \oplus k_{i-61} \oplus k_{i-50} \oplus k_{i-13} \qquad (i \geq 80).$$

All experiments start at round 0 and use this key schedule. The output follows the same register packing as the input.

Results. The finite-count rows in Table 47 cover both affine offsets. For the first row of Table 47, the observed valuation is exactly 3 on every one of the 1,000 keys. The varying counts are divisible by 8 but not 16, rather than exactly balanced. The two rows using $x_3 = b$ are related by an exact register shift and are not independent evidence.

The long-round parity observations in Table 48 use zero-offset cubes and one common 1,000-key sample. Sun et al. [SWLW16] report 90- and 99-round zero-sum properties at data 2^{30} and 2^{31}. Hu and Wang [HW19] subsequently report 100-round zero-sum and 101-round constant parity. Our 100- and 101-round observations reproduce that round coverage. At 101 rounds the surviving parity is one, not zero.

The dimensions of the sampled constant-parity spaces at rounds 99, 100, and 101 are respectively 3, 2, and 1 on these 1,000 keys. At 102 rounds, a separate 100-key sample gives dimension 0: the output-XOR differences span $\mathbb{F}_2^{32}$, excluding constant parity for every fixed nonzero output mask on this cube. This exclusion concerns that cube only; the other 31 coordinate halfspaces were not tested.

Table 47: KATAN32 finite one-count observations. Each row uses 1,000 validation keys.

Rounds	Input structure	Data	β	Sample t	Keys
84	$x_0 = b$	2^{31}	00040000	3	1000
84	$x_3 = b$	2^{31}	00020000	3	1000
85	$x_3 = b$	2^{31}	00040000	3	1000
84	$x_0 \oplus x_2 = b$	2^{31}	00020000	3	1000

Table 48: KATAN32 constant-parity observations on zero-offset cubes. All direct rows use the same 1,000 keys. Only the 91-round row is shift-derived.

Rounds	u	Data	Output	$N_1 \bmod 2$	Evaluation
90	fffffffc	2^{30}	c_{17}, c_{18}	0	Direct
91	fffffffc	2^{30}	c_{18}	0	Shift-derived
99	fffffffe	2^{31}	c_{17}, c_{18}	0	Direct
99	fffffffe	2^{31}	c_{16}	1	Direct
100	fffffffe	2^{31}	c_{18}	0	Direct
100	fffffffe	2^{31}	c_{17}	1	Direct
101	fffffffe	2^{31}	c_{18}	**1**	Direct

F.11 KeeLoq

Specification. KeeLoq is a 32-bit feedback-shift-register block cipher with a 64-bit key and 528 rounds [IKD+08]. Let $q(z)$ be bit z of the 32-bit truth table 3a5c742e. A round r computes

$$z = x_1 + 2x_9 + 4x_{20} + 8x_{26} + 16x_{31}, \qquad f_r = x_0 \oplus x_{16} \oplus q(z) \oplus k_{r \bmod 64},$$

and updates $x'_i = x_{i+1}$ for $0 \le i < 31$ and $x'_{31} = f_r$. Our positive results concern the first 64 rounds under the cipher's cyclic key schedule. They are properties of this reduced permutation, not of full 528-round KeeLoq.

Results. For the 64-round reduced permutation, each input mask in Table 49 and Table 50 defines $\mathcal{X} = \{x : \alpha \cdot x = b\}$, with data complexity 2^{31} and either offset. The master key is fixed during each full-domain evaluation.

Table 49: KeeLoq finite count valuations, exact on each of 1,000 tested keys.

Rounds	α	Data	β	Exact $v_2(N_1)$	Keys
64	31200000	2^{31}	00000003	**14**	1000
64	06000000	2^{31}	00000009	**10**	1000

For the two rows of Table 49, the observed count sets are respectively $\{2^{30} - 2^{14}, 2^{30} + 2^{14}\}$ and $\{2^{30} - 2^{10}, 2^{30} + 2^{10}\}$. These two-value restrictions give more information than divisibility alone.

Table 50: KeeLoq exact halfspace-count observations for selected ZC candidates. Each row gives the corresponding halfspace count on 1,000 validation keys.

Rounds	α	Data	β	Exact N_1	Keys
64	`00040000`	2^{31}	`00000001`	2^{30}	1000
64	`40000000`	2^{31}	`000013e1`	2^{30}	1000
64	`80000000`	2^{31}	`0000266c`	2^{30}	1000
64	`24000000`	2^{31}	`00000064`	2^{30}	1000
64	`60000000`	2^{31}	`00000a33`	2^{30}	1000
64	`c0000000`	2^{31}	`00001f97`	2^{30}	1000

Together, Table 49 and Table 50 illustrate how all-output screening distinguishes finite count restrictions from exact balance, including for multi-bit masks.

F.12 Excluding Universal Coordinate-Cube Properties

For a coordinate halfspace $\mathcal{X}_i = \{x : x_i = 0\}$, let $Q_{i,j} = \bigoplus_{x \in \mathcal{X}_i} E_{K_j}(x)$ be its output-XOR vector. A mask β has constant parity across the sampled keys if and only if $\beta \cdot (Q_{i,j} + Q_{i,0}) = 0$ for every j. If their binary span has rank n, only $\beta = 0$ remains. This is an exact all-output exclusion from finitely many counterexamples, not a probabilistic inference that no property is likely to exist. The counterexample key may differ between masks.

Table 51 covers all 32 coordinate halfspaces for each listed permutation. For each halfspace, the output-XOR differences from 50 keys or key/tweak pairs have rank 32. This excludes an all-key constant parity for every fixed nonzero output mask on these halfspaces. For one-count divisibility by 2^t, $t \geq 1$, Subsection E.13 also excludes smaller coordinate cubes satisfying the bound at every offset. These exclusions concern coordinate cubes at the plaintext input, not the input-extended structures $R_0^{-1}(a + A)$ used for our 15-round SIMON32 and 8-round SPECK32 results. The conclusion does not extend to probabilistic distinguishers, isolated smaller cubes, weak-key classes, nonlinear outputs, or other affine orientations. For ChiLow it is an exclusion over the allowed key/tweak configurations, not every restricted fixed-tweak family.

The 102-round KATAN32 observation in Subsection F.10 gives the same rank conclusion for one cube only. We do not claim complete coordinate-halfspace exclusions for BEANIE, KeeLoq, or ToySPN.

F.13 Probabilistic Divisibility

Our search for deterministic divisibility also revealed conditions that hold for only some keys and input offsets, but more often than for a random permuta-

Table 51: Exclusion of universal constant-parity properties on coordinate halfspaces. All cubes have size 2^{31}, and every difference matrix has rank 32.

Permutation	Rounds	Coordinate cubes	Keys/configurations per cube
SIMON32	15	32	50
SPECK32	8	32	50
RC5-16	3	32	50
ChiLow-32 decryption	6	32	50

tion. We report selected observations to motivate the analytical study of probabilistic integral and divisibility properties as future work. These are empirical observations, not additional all-key claims.

Probabilities over Input Offsets and Keys. Fix an active mask u, nonzero output mask β, and threshold t, and let $A = u \wedge \mathbb{F}_2^n$. For GIFT and SPECK, we use $\mathcal{X}_a = R_0^{-1}(a + A)$, where R_0 is the public first-round map without key addition or round constants. Their reported round counts include this extension, with unchanged data complexity, but the structures need not be affine. For the other ciphers, $\mathcal{X}_a = a + A$. For each key K, we consider the probability over uniformly chosen offsets

$$q_t(K) = \Pr_{a \in \mathcal{R}_A} \left[2^t \mid N_1(\mathcal{X}_a, \beta; E_K) \right].$$

We estimated $q_t(K)$ by drawing independent uniform offsets for each sampled key. For 7-round SPECK32/64, we instead enumerated all 1,024 distinct offsets, obtaining exact $q_t(K)$ for each tested key. We then compared the per-key results across uniformly sampled keys. Table 52 reports their mean, which estimates $\Pr_{K,a}[2^t \mid N_1]$. For Ascon, which is unkeyed, only the offset is sampled. We enumerated every plaintext in each structure, so the one-counts themselves are exact. We fixed the active masks and output masks before validation on fresh samples. The table thresholds for GIFT, SPECK, and 7-round PRESENT were also fixed before validation. We evaluated the random-permutation percentages using Proposition 2, summing over counts j with $2^t \mid j$. This law depends only on the block size and $M = |\mathcal{X}_a|$, so it also applies to the input-extended structures. These are probabilities for fixed events, not for the entire search. We do not treat different structures under one permutation as independent ideal instances.

PRESENT. For 6-round PRESENT-80, we tested $c_0 \oplus c_{16} \oplus c_{48}$ on $\mathcal{X}_a = a + A$ with active mask $u =$ `0000000f00000f00`. The event $32 \mid N_1$ with $N_1 \notin \{0, 2^7, 2^8\}$ occurs in 1.623% of all samples, compared with 0.003169% for a random permutation. The exclusion removes constant and balanced outputs, so this excess concerns finite divisibility.

Larger structures give two 7-round examples without an input-side extension (Table 52). With data complexity 2^{18}, we use $u =$ `0000111100005fff` and output $c_{16} \oplus c_{48}$ to observe an excess probability of evenness. With data complexity 2^{20}, we use $u =$ `000011110000ffff` and output $c_0 \oplus c_{16} \oplus c_{48}$ for the stronger

Table 52: Selected observations over keys and input offsets. The event is $2^t \mid N_1$, and data complexity is per structure. GIFT and SPECK round counts include the input-side extension. The last column gives keys × offsets per key. Ascon uses offsets only.

				Probability (%)		
Cipher	Rounds	Data	t	Observed	Random	Samples
PRESENT-80	6	2^8	5	19.850	4.985	10,000 × 2,048
PRESENT-80	7	2^{18}	1	67.067	50.000	2,000 × 2,048
PRESENT-80	7	2^{20}	4	40.362	6.250	2,000 × 2,048
GIFT-64	7	2^8	2	33.218	25.000	10,000 × 2,048
GIFT-128	7	2^8	6	31.024	4.982	10,000 × 2,048
GIFT-128	8	2^8	2	37.902	25.000	10,000 × 2,048
GIFT-128	9	2^{16}	2	46.522	25.000	10,000 × 2,048
SPECK32/64	6	2^9	6	17.233	3.524	10,000 × 2,048
SPECK32/64	7	2^{22}	2	60.110	25.000	10,000 × 1,024
Ascon	5	2^{13}	2	34.020	25.000	100,000 offsets

condition $16 \mid N_1$. Each example uses a separate fresh sample of 2,000 keys and 2,048 offsets per key. For the latter structure, all sampled one-counts are divisible by 4. The event $16 \mid N_1$ with $N_1 \notin \{0, 2^{19}, 2^{20}\}$ occurs in 40.201% of all samples, compared with 6.172% for a random permutation, so the excess is not due to exact balance.

GIFT. For GIFT, the following masks specify A in the extended structure $\mathcal{X}_a = R_0^{-1}(a + A)$. The 7-round GIFT-64 example uses $u =$ `0000000000000207f` and output c_{12}. For GIFT-128 at 7 and 8 rounds, $A = \langle e_0, \ldots, e_7 \rangle$, with outputs $c_{21} \oplus c_{116}$ and c_{20}, respectively. At 9 rounds, $A = \langle e_0, \ldots, e_{15} \rangle$ and the output is c_{36}. The PRESENT and GIFT experiments use the standard key schedules and the first r encryption rounds, including the final bit permutation and key addition (Subsection F.4, Subsection F.6). GIFT-128 uses the 128-bit variant specified in [BPP+17]. For 7-round GIFT-128, the event $64 \mid N_1$ with $N_1 \notin \{0, 2^7, 2^8\}$ occurs in 6.098% of all samples, compared with approximately 3.29×10^{-14}% for a random permutation. With c_{21} alone and the same input mask, the corresponding event occurs in 1.496% of a separate sample of 10,000 keys and 2,048 offsets per key. For 9-round GIFT-128, divisibility by 16 occurs in 11.625% of samples, compared with 6.25% for a random permutation. At this round count, the excess at higher divisibility thresholds largely reflects the increased probability that N_1 is even.

SPECK. For SPECK32/64, we used $\beta =$ `02050204` on $\mathcal{X}_a = R_0^{-1}(a + A)$ (Subsection F.2). The active masks defining A are $u =$ `e07b0000` at 6 rounds and $u =$ `ff7f039e` at 7 rounds. The 6-round example needs only 512 plaintexts per structure. The event $64 \mid N_1$ with $N_1 \notin \{0, 2^8, 2^9\}$ occurs in 9.508% of all samples, compared with approximately 6.90×10^{-7}% for a random permutation. At

7 rounds, each structure contains 2^{22} plaintexts, and every sampled one-count is even. Divisibility by 4 occurs in 60.110% of samples, exceeding the approximately 50% random-permutation probability conditioned on evenness. Thus the excess is not explained by evenness alone.

Ascon. We evaluated the last 5 rounds of the Ascon permutation's 12-round sequence, using the bit numbering in Subsection F.5. In $\mathcal{X}_a = a + A$, only bits $0, \ldots, 5, 7, \ldots, 13$ of the first word S_0 vary. The other input bits specify the offset. The output is c_{129}, bit 1 of word S_2, after the final diffusion layer. Over 100,000 fresh offsets, divisibility by 16 occurs in 8.341% of samples, compared with 6.25% for a random permutation. Requiring also $N_1 \notin \{0, 2^{12}, 2^{13}\}$ excludes constant and balanced outputs and leaves 7.143% of all samples, compared with 5.368% for the same random-permutation event.

Figure 14, Figure 15, and Figure 16 show how the observed probabilities decrease as the required divisor increases, using the samples of Table 52. Table 53 gives the observed and random-permutation percentages for each divisor. The divisibility events include exact balance. The PRESENT, GIFT-128 and SPECK32/64 examples above show that the excess persists when constant and balanced outputs are excluded.

Table 53: Divisibility frequencies for the samples in Table 52. Each cell gives the observed percentage above the random-permutation percentage, both rounded to two decimal places. Column subheadings give the number of rounds and data complexity M.

Divisor of N_1	PRESENT-80			GIFT-64	GIFT-128			SPECK32/64		Ascon
	6	7	7	7	7	8	9	6	7	5
	2^8	2^{18}	2^{20}	2^8	2^8	2^8	2^{16}	2^9	2^{22}	2^{13}
2	96.96	67.07	100.00	66.32	100.00	74.42	93.07	93.72	100.00	68.42
	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00
4	74.57	33.71	100.00	33.22	98.79	37.90	46.52	75.36	60.11	34.02
	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00
8	45.71	16.86	78.57	16.61	88.47	19.05	23.26	57.63	30.05	16.79
	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50	12.50
16	27.41	8.43	40.36	8.52	67.77	9.79	11.62	38.51	15.02	8.34
	6.34	6.25	6.25	6.34	6.34	6.34	6.25	6.25	6.25	6.25
32	19.85	4.21	20.24	6.55	45.71	7.01	5.81	27.23	7.53	4.10
	4.99	3.13	3.13	4.99	4.99	4.99	3.13	3.65	3.13	3.13
64	18.32	2.10	10.12	6.52	31.02	6.89	2.91	17.23	3.76	2.11
	4.98	1.56	1.56	4.98	4.98	4.98	1.56	3.52	1.56	1.56
128	18.23	1.05	5.06	6.52	24.93	6.89	1.45	10.68	1.88	1.24
	4.98	0.78	0.78	4.98	4.98	4.98	0.78	3.52	0.78	0.91

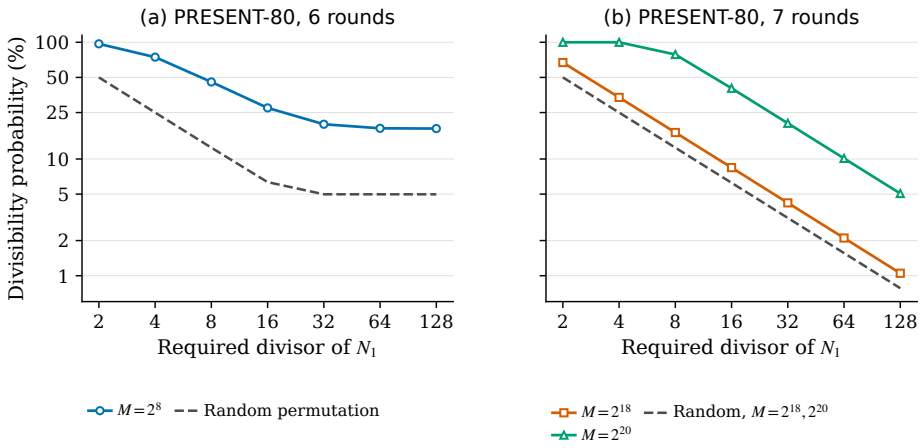


Fig. 14: PRESENT-80 divisibility probabilities at 6 and 7 rounds from Table 53, with logarithmic vertical axes. Solid marked curves show observations. Gray curves show random-permutation probabilities at the same data complexity M.

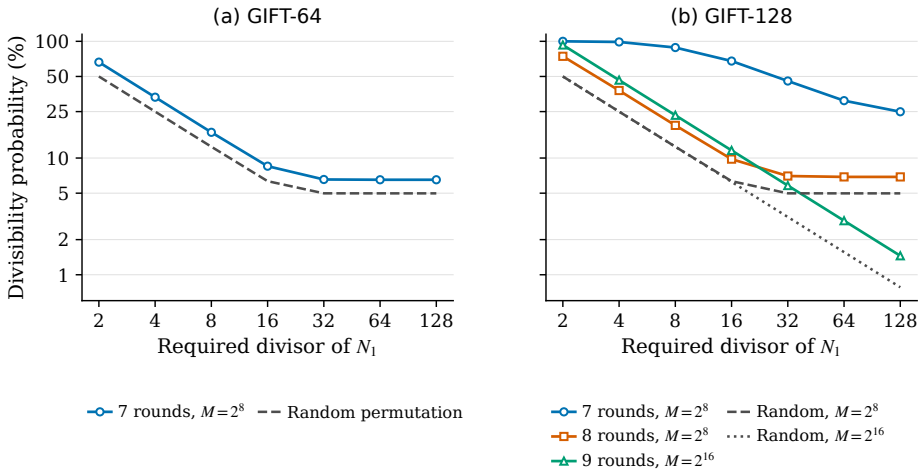


Fig. 15: GIFT-64 and GIFT-128 divisibility probabilities, following the conventions of Figure 14. Round counts include the input-side extension.

Variation across Keys. Figure 17 compares the per-key estimates of $q_t(K)$. The bars extend from the 1st to the 99th percentile of these estimates across sampled keys. For PRESENT-80 and the GIFT examples, the observed variation is close to

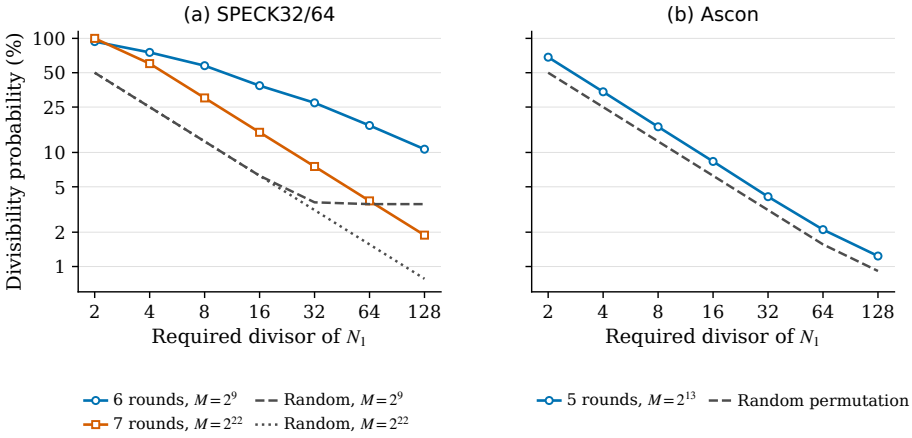


Fig. 16: SPECK32/64 and Ascon divisibility probabilities, following the conventions of Figure 14. SPECK round counts include the input-side extension.

that expected from sampling offsets alone. For example, the 1st–99th percentile range of the 7-round PRESENT-80 estimates for $16 \mid N_1$ is 37.99%–42.72%. Their standard deviation is 1.06 percentage points, compared with 1.08 points expected from offset sampling if all keys had the same probability. The samples therefore suggest similar probabilities across keys, but do not establish key independence. In contrast, the 1st–99th percentile range of the 6-round SPECK32/64 estimates is 8.40%–33.84%. Their standard deviation is 7.25 percentage points, compared with about 0.83 points expected from offset sampling if all keys had the same probability. At 7 rounds, the exact offset probabilities take five values between 53.125% and 65.625% across the 10,000 tested keys. This variation cannot be attributed to offset sampling, since all offsets were evaluated. Figure 18 contrasts the distributions for 7-round GIFT-128 and 6-round SPECK32/64 at the same threshold $64 \mid N_1$. Each key contributes one estimate $\widehat{q}_6(K)$ over the sampled offsets. The vertical lines mark the corresponding random-permutation means, not distributions across permutations.

SIMON: Probabilities at Fixed Offsets. We used SIMON32/64 with its standard key schedule and the first r encryption rounds (Subsection F.1). We first fixed the offset and measured the frequency across keys, rather than across offsets for each key. For fixed active mask u, output mask β and threshold t, let $\mathcal{X}_a = a + (u \wedge \mathbb{F}_2^{32})$ and consider

$$p_t(a) = \Pr_K\left[2^t \mid N_1(\mathcal{X}_a, \beta; E_K)\right],$$

where K is uniform over the 64-bit master-key space. We enumerated each cube completely, so only the keys were sampled. We fixed the masks and input structures before validation on fresh keys, independently of the discovery samples.

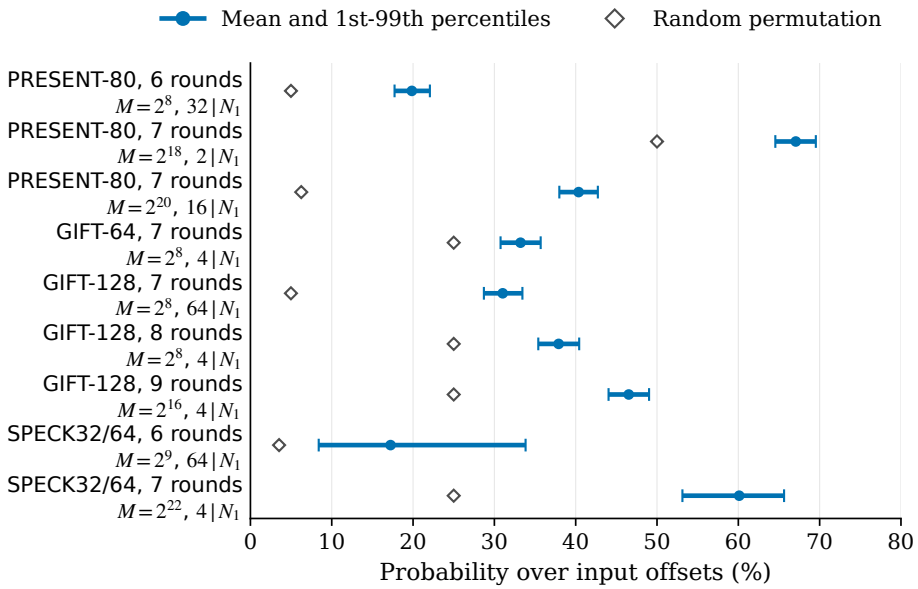


Fig. 17: Means and 1st–99th percentile ranges of per-key estimates of $q_t(K)$ for the events in Table 52. The ranges include offset-sampling variation except for 7-round SPECK32/64. They are not confidence intervals.

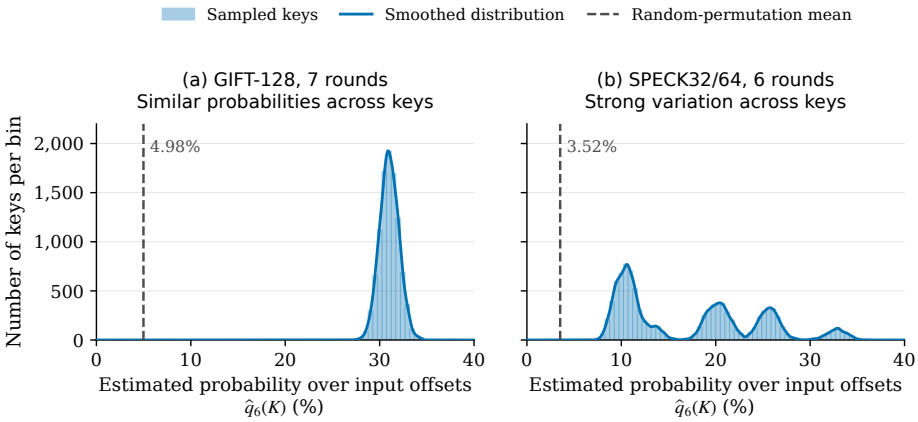


Fig. 18: Key dependence of divisibility by 64 on input-extended structures, using 10,000 keys and 2,048 offsets per key in each panel. Both distributions include offset-sampling variation.

For the main illustration, u = 55555555: every even-indexed input bit varies, giving data complexity 2^{16}. The offset a specifies the fixed odd-indexed bits.

Comparison with Random Permutations. For the same fixed cube and nonzero output mask, we evaluate the random-permutation probability numerically using Proposition 2. For data complexity 2^{16}, the probabilities of divisibility by 4, 8 and 16 are approximately 25%, 12.5% and 6.25%. Figure 19 compares this baseline with the observed frequencies and shows the effect of changing the offset. At zero offset, we tested c_7 at 8 rounds and c_6 at 9 and 10 rounds, using 10,000 fresh keys per curve. The offset comparison uses the same separate sample of 4,000 keys at each offset, not 16,000 independent trials.

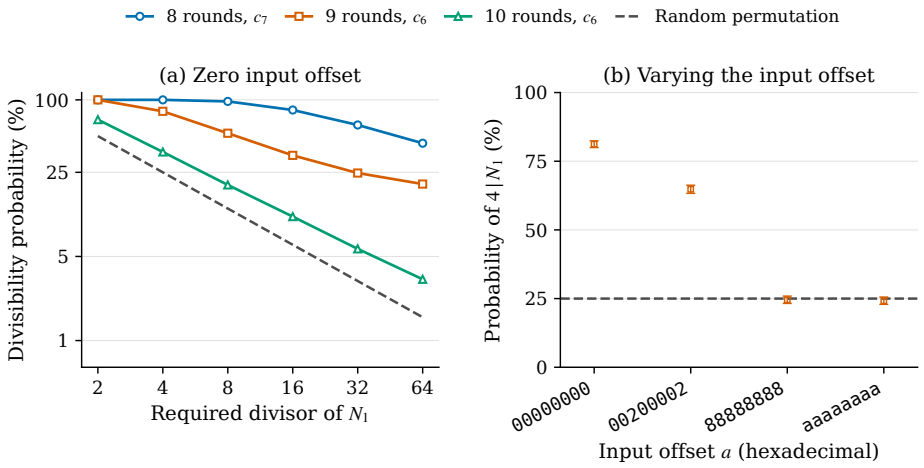


Fig. 19: Probabilistic one-count divisibility on SIMON32/64 with active mask `55555555` and data complexity 2^{16}. Panel (a): zero offset, 10,000 keys per curve, logarithmic vertical axis. Panel (b): 9 rounds, output c_6, the same 4,000 keys at each offset. Error bars in (b) are pointwise 95% Wilson intervals, not simultaneous bands. In (a), each interval extends less than one percentage point from its estimate and is omitted for clarity. Gray dashed lines give the numerically evaluated random-permutation probabilities.

At 8 rounds, $8 \mid N_1$ holds for 97.07% of the fresh keys, and at 9 rounds it holds for 52.81%. Exact balance $N_1 = 2^{15}$ occurs for 16.28%, 16.59% and 1.46% of the zero-offset samples at 8, 9 and 10 rounds, respectively, and is included in the plotted events. These effects are not explained solely by exact balance. Among samples without exact balance, $8 \mid N_1$ holds for 96.50% and 43.42% at 8 and 9 rounds, respectively. Changing the fixed input bits can substantially weaken the effect: at 9 rounds, $4 \mid N_1$ holds for 81.225% of the separate zero-offset sample, but for 24.55% and 24.20% at offsets `88888888` and `aaaaaaaa`. These fixed-offset frequencies therefore do not describe every offset.

Table 54 also gives examples with lower data complexity and parities of two output bits. At 7 rounds, we use $u =$ `000000ff` and output c_9. At 10 rounds,

we use $u =$ `55155555` with c_0, and $u =$ `55555555` with $c_0 \oplus c_7$ or $c_0 \oplus c_{14}$. All these examples use zero offset. The 10-round effects are weaker but remain above the corresponding baseline on fresh keys. Much of their excess comes from the increased probability of evenness, rather than an independent gain at every threshold.

Table 54: Selected SIMON32/64 frequencies at zero offset on s fresh keys. The three 10-round rows with $u =$ `55555555` share the same keys and are not independent replications.

					Observed frequency (%)		
Rounds	Active mask u	Data	Output	s	$4 \mid N_1$	$8 \mid N_1$	$16 \mid N_1$
7	`000000ff`	2^8	c_9	100,000	100.00	72.59	48.58
8	`55555555`	2^{16}	c_7	10,000	100.00	97.07	82.35
9	`55555555`	2^{16}	c_6	10,000	80.30	52.81	34.57
10	`55555555`	2^{16}	c_6	10,000	36.90	19.64	10.71
10	`55155555`	2^{15}	c_0	10,000	35.32	17.89	9.70
10	`55555555`	2^{16}	$c_0 \oplus c_7$	10,000	33.03	16.28	8.17
10	`55555555`	2^{16}	$c_0 \oplus c_{14}$	10,000	31.92	15.62	7.41

SIMON: Probabilities over All Offsets. To see whether the probability over offsets varies with the key, we also enumerated all offsets for 64 fresh keys. For the 7-round row of Table 54, the per-key probability of $8 \mid N_1$ ranges from 69.043% to 75.439%, with mean 72.652%, compared with approximately 12.5% for a random permutation. For the 9-round row, the corresponding range is 18.207%–19.394%, with mean 18.818%. These per-key probabilities are exact over all 2^{24} and 2^{16} offsets, respectively. Only the keys are sampled. In particular, 52.81% is the frequency across keys at zero offset, whereas 18.818% also averages over all offsets. Thus a property can depend strongly on the fixed input bits while its probability over offsets changes little across keys.

Open Question. Can we compute the probabilities of integral and divisibility properties over keys and input offsets analytically, without exhaustive evaluation? Can these methods guide tools that search for strong probabilistic divisibility distinguishers? A fuller description would give the distribution of integral one-counts modulo 2^t and explain when probabilities over offsets remain stable across keys. A useful explanation should distinguish information beyond parity from effects caused by exact balance, and identify which input structures preserve the excess probability. Our experiments provide concrete examples for this question, not an all-key property or a general law over offsets. The paper's main results concern deterministic constraints.