Theoretical Open Problems in Symmetric Cryptography: Verifiable LLM-Guided Analysis

Yufei Yuan[1,2(✉)], Yaoda Hu[2,3], Yixin Zhang[1,2], Lei Zhang[1,4], and Wenling Wu[1,3(✉)]

[1] Trusted Computing and Information Assurance Laboratory, Institute of Software Chinese Academy of Sciences, Beijing, 100190, China
yufei2021, wenling@iscas.ac.cn
[2] University of Chinese Academy of Sciences, Beijing, 100049, China
[3] Zhongguancun Laboratory, Beijing, 100194, China
[4] State Key Laboratory of Cryptology, P.O. Box 5159, Beijing 100878, China

Abstract. We present the Pilot–Sailor Framework, an LLM-guided system for studying theoretical open problems in symmetric cryptography. Pilot proposes intermediate statements and proof plans. Sailor attempts formal proofs, and the proof assistant admits only checked declarations to the verified context. We apply this methodology to Boolean-function theory and symmetric cryptanalysis through thirteen mathematical case studies, comprising complete resolutions, corrected formulations, counterexamples, and scoped quantitative advances. In particular, we prove the original pointwise Tu–Deng conjecture for all word lengths and admissible residues. We further characterize equality in this bound: if t has z zero bits, equality holds exactly when every cyclic gap between consecutive zeros is at least z. This criterion also gives a closed formula for the number of equality cases for each z. We also prove that, for $n = 2k \geq 6$ and $k < m < 2k$, every mapping $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ satisfies $\mathrm{NL}(F) \leq 2^{n-1} - 2^{n/2-1} - 2$. This improves both the covering-radius estimate and the bound obtained from Nyberg's obstruction and integrality. Using an exact computer-assisted spectral classification, we also prove that the maximum nonlinearity of a balanced Boolean function in eight variables is 116, resolving whether the value 118 can occur. Beyond the well-known long-standing problems highlighted above, we also establish new results for ten further research questions in symmetric cryptography.

Keywords: LLM-guided theorem proving · formal verification · Boolean functions · symmetric cryptanalysis

1 Introduction

The pursuit of fully rigorous and mechanically verifiable mathematics can be traced back to Hilbert's program, which sought to place mathematics on explicit formal foundations and to establish the consistency of these foundations through finitary reasoning [28]. This ambitious objective was fundamentally constrained by Gödel's incompleteness theorems and by the undecidability results of Church

and Turing [15, 24, 59]. These results showed that no sufficiently expressive formal system can simultaneously provide a complete and effective procedure for deciding all mathematical statements. Rather than ending the formalization of mathematics, these limitations redirected the field toward a more attainable objective: the construction of trustworthy systems that can verify individual proofs within precisely specified logical foundations. This line of research led from de Bruijn's Automath and the LCF architecture to modern interactive theorem provers such as Lean [9, 27, 40]. The successful formalization of major mathematical results, including the four-color theorem and the odd order theorem, has demonstrated that machine-checked reasoning can support proofs of substantial depth and complexity [25, 26].

Recent advances in large language models have introduced a new stage in this development by combining flexible mathematical reasoning with formal verification. FunSearch demonstrated that language models can be embedded into an evaluator-guided search process to discover improved mathematical constructions, including new solutions for the cap-set problem [49]. AlphaProof further showed that reinforcement learning and Lean-based verification can support formal reasoning at the level of the International Mathematical Olympiad [29]. More recent systems have extended these capabilities toward research-level mathematics. AlphaProof Nexus formally resolved several problems from the Erdős problem collection through AI-driven proof search [57], while the Rethlas and Archon framework combined conjecture resolution with machine-checked formal verification [30]. In parallel, a general-purpose reasoning model produced a construction that disproved a conjectured upper bound in the planar unit-distance problem, and the resulting argument was subsequently examined by independent mathematicians [1]. These developments indicate that reliable automated mathematics increasingly depends on a division of labor between heuristic reasoning and formal verification: language models explore proof strategies and generate candidate arguments, whereas proof assistants certify their logical correctness.

Formal verification has also become an established component of modern cryptographic research. Existing techniques have been successfully applied to mechanize game-based security proofs, analyze cryptographic protocols, verify implementations, and certify low-level arithmetic routines [3–5, 22, 65, 66]. These developments have substantially strengthened the reliability of cryptographic arguments and implementations. Nevertheless, their use has remained concentrated on the formal verification of explicitly specified claims. In most cases, human experts must first identify the theorem to be proved, select an appropriate model, and provide the essential proof structure before mechanized tools can complete or certify the argument. Formal methods in cryptography have therefore been used primarily to verify known constructions and security claims, rather than to investigate the unresolved theoretical questions that shape the field. This limitation is particularly pronounced in areas such as Boolean function theory and symmetric cryptanalysis, where many open problems and conjectures require the discovery of new mathematical arguments rather than the verification of predetermined proof obligations. The recent progress of LLM-based reasoning

systems naturally raises a broader question: can language-model agents, coupled with formal verification, be developed into a scalable methodology for obtaining substantive advances on collections of fundamental and frontier open problems in cryptography?

Turning this question into a practical research methodology is more difficult than applying a theorem prover to a fixed benchmark. One obstacle is the limited availability of reusable formal infrastructure for symmetric cryptanalysis. The mathlib library provides general theories for finite types, finite fields, matrices, and finite sums [55, 56]. These theories supply the underlying mathematics but not the domain objects used in cryptanalysis. One basic example is the difference distribution table (DDT) of a vectorial Boolean function, together with its differential spectrum and associated equivalence relations [8, 10, 21]. A reusable definition must support arbitrary mappings $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, rather than a single fixed S-box. It must also provide general lemmas for derivatives, fiber cardinalities, and transformations of the table. Otherwise, each new conjecture would require a separate encoding of the same cryptographic concepts. We therefore develop a common domain layer on top of mathlib for expressing individual cryptographic problems in Lean.

A further obstacle is semantic alignment. A formal statement may type-check while altering a quantifier, omitting a nondegeneracy condition, or encoding an unintended notion of equivalence. Existing autoformalization systems remain far from eliminating such errors [46, 62]. Proof generation must then cope with premise retrieval, domain shift, and multi-stage argument decomposition, particularly when the required definitions and intermediate lemmas are absent from the underlying library [51, 64]. Finally, LLMs may generate plausible but false intermediate claims or refer to nonexistent formal dependencies when uncertain [31, 61]. A proof assistant can reject an invalid derivation, but it cannot by itself determine whether the encoded statement faithfully represents the intended conjecture or whether a valid proof constitutes a novel and cryptographically meaningful advance. Addressing these obstacles requires an end-to-end methodology that combines domain-library construction, semantic validation, formal proof search, and independent mathematical assessment.

Our Contributions We introduce Pilot–Sailor, a proof workflow that separates intermediate-statement selection from formal proof construction and admits a statement to the reusable context only after proof-assistant acceptance. A mathematical audit checks the source formulation, hypotheses, semantics, and claim scope. We apply the workflow to thirteen questions in Boolean-function theory and symmetric cryptanalysis. Table 1 summarizes each question and result in ordinary language; full statements and proofs appear in Sections 3 and 4.

Table 1: Questions studied in this work and the corresponding contributions.

Problem	Description
Tu–Deng conjecture (Section 3.1)	The conjecture asks whether, for every binary word length and every admissible residue, no more than half of the associated pairs have total Hamming weight below the word length. We prove this pointwise statement for all parameters by translating circular carries into a monotone Boolean family and integrating its pivotal-edge profiles. We also characterize every equality case by a cyclic zero-gap condition and enumerate them according to the number of zero bits. [1]
Maximum vectorial nonlinearity (Section 3.2)	The problem asks how nonlinear a vectorial Boolean mapping can be when the number of output coordinates is greater than half, but less than all, of an even number of input coordinates. For at least six inputs, we improve the universal covering-radius and Nyberg-integrality bounds by two; we also determine the exact four-input, three-output value, without claiming tightness for the remaining parameters.
Balanced eight-variable nonlinearity (Section 3.3)	The open problem asks whether a balanced Boolean function in eight variables can have nonlinearity 118, while 116 was the largest known value. We rule out 118 through an exact computer-assisted spectral classification and inverse-Walsh inequalities. An explicit balanced truth table attains 116, so the maximum is exactly 116.
Affine derivatives of bent functions (Section 3.4)	The question asks for the largest possible number of distinct affine derivatives of a nonquadratic bent Boolean function. We prove that the maximum is one eighth of all directions, namely 2^{n-3}, in every even dimension $n \geq 6$.
Degree-drop hyperplanes (Section 3.5)	The first degree-stability problem asks whether a Boolean function having no affine hyperplane on which its degree drops can be put into the prescribed external private-edge normal form, up to an affine change of variables and lower-degree terms. We prove that such a representative exists throughout the stated range $n > r \geq 2$.
Functions without fast points (Section 3.6)	The second degree-stability problem asks for the analogous internal private-edge normal form when no nonzero direction lowers the derivative degree too far. We prove existence in the full range $n > r \geq 2$, while also ensuring that every ambient variable remains active.

Continued on the next page

[1] Our proof was obtained independently and contemporaneously with a very recent preprint work [34]; their cyclic-carry factorization and deletion argument are distinct from our monotone-family pivotal-edge proof.

Problem	Description
Differential factors (Section 3.7)	The conjecture asks whether a nonzero differential factor of a 4-bit S-box is equivalent to the corresponding difference-distribution-table column containing only zeros and fours. We disprove the literal unrestricted-function reading with a non-bijective counterexample and prove the equivalence for 4-bit permutation S-boxes of differential uniformity at most four.
Kaleyski's identities (Section 3.8)	Kaleyski conjectured that two aggregate DDT sums of every APN power function equal the corresponding cubic benchmark. We prove the second identity for every even dimension $n \geq 4$; the first identity, and therefore the conjunction in the original conjecture, remains open there.
Balanced quadratic functions (Section 3.9)	The problem asks for closed formulae counting balanced quadratic Boolean functions and for the weight distribution of the dual of the second-order Reed–Muller code. We give formulae for both the degree-at-most-two and exact-degree-two conventions and derive the full dual weight distribution. [2]
Hamming-layer annihilators (Section 3.10)	The problem asks for the degree threshold at which a dimension argument forces a restricted annihilator when a Boolean function is observed only on a fixed Hamming-weight layer. We bound the deficit from the symmetric layer distance by less than $\sqrt{n+1}$ and determine the threshold's first-order asymptotic behaviour.
dCUFN-SP active S-boxes (Section 4.1)	Bogdanov conjecture asserts a short-window lower bound on active S-boxes in contracting unbalanced Feistel networks under a two-matrix optimal diffusion hypothesis. We construct a valid period-six trail with only five active S-boxes in a four-round window for $d = 3$ and $m = 2$, refuting that stronger claim without affecting the established longer-window result.
38- and 40-branch type-II GFN diffusion (Section 4.2)	The open challenge asks for even–odd branch permutations that achieve full structural diffusion, in both the forward and inverse directions, after ten rounds on 38 and 40 branches. We give explicit witnesses meeting the known lower bound for both sizes; the 42-branch case and classification of all optimal permutations remain open.
Expanded equivalence of type-II GFNs (Section 4.3)	The problem asks for the sizes and number of expanded-equivalence classes when all type-II GFN branch permutations are allowed, rather than only even–odd ones. We prove that every class on $2k$ branches has $k!$ elements and that there are $(2k)!/k!$ classes.

[2] Although Carlet lists this as open, both answers follow from Sloane and Berlekamp's 1970 weight enumerator [50], via its middle coefficient and the MacWilliams transform. We rederive both using symplectic rank strata and Walsh-sum factorization.

The mathematical outcomes therefore have several distinct forms. The Tu–Deng, balanced-eight-variable nonlinearity, affine-derivative, and degree-stability results settle their stated ranges; our Tu–Deng analysis also determines all equality cases and counts them by the number of zero bits. The differential-factor analysis separates and resolves the two possible readings of the source statement, while the Feistel results combine an explicit counterexample with optimal finite witnesses. The remaining entries provide exact counts or deliberately scoped advances whose unresolved portions are stated in the table and in the corresponding sections.

The cryptographic formalization infrastructure developed for this work, the implementation details of our automated proving framework, and the associated Lean formalizations and proof source code will be made public in a subsequent version of this paper. Every proof in this manuscript has been checked by experts in cryptography.

2 An LLM-Based Automatic Proving System

This section presents the Pilot–Sailor Framework (PSF), an LLM-based framework for automated theorem proving. We use PSF to study the theoretical cryptographic problems treated in Sections 3 and 4. PSF consists of two subsystems, Pilot and Sailor. At iteration t, Pilot receives a fixed target theorem $s^\star$ and the current set of verified propositions $\mathcal{S}_t = \{s_i\}_{i=0}^{n_t}$, where each s_i is a formal proposition supported by a machine-checked proof. Pilot then proposes an intermediate lemma s_{next} that may advance the proof of $s^\star$. Given s_{next} and the propositions in $\mathcal{S}_t$, Sailor attempts to construct a formal proof $\hat{\pi}_{\text{next}}$. The proof assistant checks the candidate proof. If the proof is accepted, PSF updates the verified proposition set according to $\mathcal{S}_{t+1} = \mathcal{S}_t \cup \{s_{\text{next}}\}$. If the proof is rejected, s_{next} is not added to $\mathcal{S}_t$.

Both Pilot and Sailor use an LLM. We represent the LLM as an autoregressive conditional distribution p_θ, where θ denotes the model parameters. Let $\mathbf{x} = (x_1, \ldots, x_m)$ be an ordered tuple of task inputs, and let τ be a prompt template. A prompted generation is written as $y \sim p_\theta(\cdot \mid \text{Prompt}_\tau(\mathbf{x}))$, where $\text{Prompt}_\tau(\mathbf{x})$ denotes the token sequence obtained by inserting $\mathbf{x}$ into τ. At the token level, the conditional distribution factorizes as

$$p_\theta(y \mid u) = \prod_{j=1}^{|y|} p_\theta(y_j \mid u, y_{<j}), \qquad u = \text{Prompt}_\tau(\mathbf{x}).$$

Under this notation, the Pilot invocation at iteration t is expressed as

$$s_{\text{next}} \sim p_{\theta_{\text{P}}}(\cdot \mid \text{Prompt}_{\text{P}}(s^\star, \mathcal{S}_t)).$$

The corresponding Sailor invocation is expressed as

$$\hat{\pi}_{\text{next}} \sim p_{\theta_{\text{S}}}(\cdot \mid \text{Prompt}_{\text{S}}(s_{\text{next}}, \mathcal{S}_t)).$$

Here, $\mathrm{Prompt}_{\mathrm{P}}$ and $\mathrm{Prompt}_{\mathrm{S}}$ denote the prompt templates used by Pilot and Sailor, respectively. The parameters θ_{P} and θ_{S} may correspond to the same underlying LLM or to different models.

Models and per-problem budgets. The model runs reported in this work used OpenAI's GPT-5.5 and the more recently released GPT-5.6 Sol [43, 44]. We did not impose a uniform budget across all problems. For each problem, the budget governed the iteration bound T and the permitted model calls for lemma selection, proof construction, and repair. We assigned larger budgets to the first three problems in Table 1, which required more extensive proof search. The case studies should therefore be read as applications of PSF to problems of differing difficulty, rather than as a controlled comparison of model success rates under a fixed compute budget.

Algorithm 1 gives the abstract PSF loop. The predicate $\mathsf{Check}(\mathcal{S}, s, \hat{\pi})$ returns accept exactly when the proof assistant accepts $\hat{\pi}$ as a proof of s from the verified context $\mathcal{S}$. The iteration bound T makes the search procedure explicit without prescribing a subsystem-specific stopping rule.

Algorithm 1 Abstract Pilot–Sailor proof-search loop

Require: Target theorem $s^\star$, verified proposition set $\mathcal{S}_0$, and iteration bound T
Ensure: Status $b \in \{\mathsf{proved}, \mathsf{unresolved}\}$ and a verified proposition set $\mathcal{S}$

$t \leftarrow 0$, $\mathcal{S} \leftarrow \mathcal{S}_0$
while $s^\star \notin \mathcal{S} \wedge t < T$ **do**
  $s_{\mathrm{next}} \leftarrow \textsc{Pilot}(s^\star, \mathcal{S})$ ▷ Select the next formal proof obligation
  $\hat{\pi}_{\mathrm{next}} \leftarrow \textsc{Sailor}(s_{\mathrm{next}}, \mathcal{S})$ ▷ Synthesize a candidate derivation
  if $\mathsf{Check}(\mathcal{S}, s_{\mathrm{next}}, \hat{\pi}_{\mathrm{next}}) = \mathsf{accept}$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \cup \{s_{\mathrm{next}}\}$ ▷ Extend the trusted state only after checking
  else
    $\mathcal{S} \leftarrow \mathcal{S}$ ▷ A rejected proof leaves the trusted state unchanged
  end if
  $t \leftarrow t + 1$
end while
if $s^\star \in \mathcal{S}$ **then**
  return $(\mathsf{proved}, \mathcal{S})$
else
  return $(\mathsf{unresolved}, \mathcal{S})$
end if

3 Applications in Boolean Function Analysis

This section gives the full statements and proofs of the Boolean-function results. Each subsection states the source question before establishing the corresponding result.

The following notation remains in force throughout this section. Let $V_n = \mathbb{F}_2^n$, with addition written additively, and let $f : V_n \to \mathbb{F}_2$ be a Boolean function. For

a mapping $F : V_n \to V_m$,

$$\Delta_a F(x) = F(x+a) + F(x), \qquad \mathrm{DDT}_F(a,b) = \big|\{x \in V_n : \Delta_a F(x) = b\}\big|.$$

Its differential uniformity is

$$\mathrm{DU}(F) = \max_{a \in V_n \setminus \{0\},\, b \in V_m} \mathrm{DDT}_F(a,b).$$

The symbols $\deg(f)$, $\mathrm{wt}(f)$, and $\mathrm{supp}(f)$ denote algebraic degree, Hamming weight, and support. We use $\mathrm{RM}(r,n)$ for the binary Reed–Muller code of order r. The Walsh transform and nonlinearity of f are

$$\mathcal{W}_f(a) = \sum_{x \in V_n} (-1)^{f(x)+a\cdot x}, \qquad \mathrm{NL}(f) = 2^{n-1} - \frac{1}{2} \max_{a \in V_n} |\mathcal{W}_f(a)|.$$

For a vectorial function, $F_v = v \cdot F$ and

$$\mathrm{NL}(F) = \min_{0 \neq v \in V_m} \mathrm{NL}(F_v).$$

Finite-field notation is introduced locally when it is needed.

3.1 The Tu–Deng conjecture

Problem and prior work For a nonnegative integer $z < 2^k$, let $\mathrm{wt}_k(z)$ be the Hamming weight of its k-bit expansion. Tu and Deng introduced the following conjecture while constructing bent and balanced Boolean functions with optimal algebraic immunity [58]. The conjecture is a uniform statement in both the word length and the residue being added.

Open Question 3.1.1 (Tu–Deng conjecture). *Let $k \geq 2$, $M = 2^k - 1$, and $1 \leq t < M$. Define*

$$S(t,k) = \left\{ (a,b) \in \{0,\dots,M-1\}^2 : \begin{array}{l} a + b \equiv t \pmod{M}, \\ \mathrm{wt}_k(a) + \mathrm{wt}_k(b) < k \end{array} \right\}.$$

Then

$$|S(t,k)| \leq 2^{k-1}. \tag{3.1.1}$$

Carry recurrences and structured families were studied by Flori et al. [23]. Several later papers proved the conjecture for further classes of parameters [12–14, 16, 17, 48]. These classes include mirror-symmetric parameters and parameters of bounded Hamming weight. Spiegelhofer and Wallner established an exact circular-addition formulation and proved that the conjecture holds for almost all t as k tends to infinity [52]. An almost-sure result does not give the pointwise inequality in Conjecture 3.1.1. The first part in our proof translates the original pair count into a Boolean recurrence for cyclic carries. The second part bounds the corresponding increasing family by integrating the profiles of its pivotal edges. The same profiles also determine exactly when the pointwise bound is attained.

Circular addition and cyclic carries Write $\langle z \rangle_M$ for the representative of $z \bmod M$ in $\{0, \ldots, M-1\}$. We first record the precise relation between the pair count and circular addition. The endpoint $2^k - 1 = M$ is retained on the circular side.

Proposition 3.1.1 (Pair-count reduction). *For every $k \geq 2$ and $1 \leq t < M = 2^k - 1$,*

$$|S(t,k)| = |\{x \in \{0, \ldots, M\} : \mathrm{wt}_k(\langle x + t \rangle_M) < \mathrm{wt}_k(x)\}| \,. \tag{3.1.2}$$

Proof. For each $a \in \{0, \ldots, M-1\}$, the congruence in the definition of $S(t,k)$ determines

$$b = \langle t - a \rangle_M .$$

Set $x = M - a$. This is a bijection from the allowed values of a to $\{1, \ldots, M\}$, and

$$b = \langle x + t \rangle_M .$$

The k-bit word for $M - a$ is the bitwise complement of the word for a. Hence

$$\mathrm{wt}_k(x) = k - \mathrm{wt}_k(a).$$

It follows that

$$\mathrm{wt}_k(a) + \mathrm{wt}_k(b) < k \quad \Longleftrightarrow \quad \mathrm{wt}_k(\langle x + t \rangle_M) < \mathrm{wt}_k(x).$$

The missing value $x = 0$ never satisfies the strict inequality, so adjoining it does not change the count.

Let $d = (d_0, \ldots, d_{k-1}) \in \{0,1\}^k$ be the binary word for t, and put

$$m = |d| = \mathrm{wt}_k(t).$$

We identify each $x \in \{0, \ldots, M\}$ with its k-bit word $(x_0, \ldots, x_{k-1})$. Indices on all words are read modulo k. A cyclic carry word $c \in \{0,1\}^k$ obeys

$$c_{j+1} = \mathrm{Maj}(c_j, x_j, d_j), \tag{3.1.3}$$

where Maj is the majority function on three bits. For fixed x and d, the local map in (3.1.3) is the identity when $x_j \neq d_j$, and it is constant when $x_j = d_j$. Thus the cyclic recurrence has a unique closed carry word unless $x = 1 - d$. At this exceptional input both constant carry words are closed. We select the all-one carry word. Equivalently, $c(x)$ is the greatest closed carry word.

Lemma 3.1.1 (Carry-count equivalence). *With the preceding convention,*

$$|S(t,k)| = \big|\{x \in \{0,1\}^k : |c(x)| > m\}\big|. \tag{3.1.4}$$

Proof. Let $y = \langle x + t \rangle_M$. The digit and carry equations for circular addition are

$$x_j + d_j + c_j = y_j + 2c_{j+1}. \tag{3.1.5}$$

Their carry recurrence is (3.1.3). If $x \neq 1 - d$, the unique closed carry produces the canonical representative y. For $x = 1 - d$, the zero carry produces the all-one word, while the one carry produces the all-zero word. Selecting the one carry therefore agrees with the canonical representative of zero modulo M. Summing (3.1.5) over the cycle and using $\sum_j c_j = \sum_j c_{j+1}$ gives

$$|y| = |x| + m - |c(x)|. \tag{3.1.6}$$

Consequently $|y| < |x|$ holds exactly when $|c(x)| > m$. The result now follows from Proposition 3.1.1.

The biased measure of the bad family Define

$$\mathcal{B}_d = \{x \in \{0,1\}^k : |c(x)| > m\}. \tag{3.1.7}$$

The family $\mathcal{B}_d$ is increasing under the coordinatewise order. Indeed, every majority gate is increasing in its state and input bit. If $x \leq z$, the corresponding one-turn maps satisfy $\Phi_x \leq \Phi_z$. If the greatest fixed initial state of Φ_x is one, then one is also fixed by Φ_z. Otherwise the order of the two greatest fixed initial states is automatic. Propagation through the gates then gives

$$c(x) \leq c(z). \tag{3.1.8}$$

Let $F(p)$ be the p-biased measure of $\mathcal{B}_d$:

$$F(p) = \sum_{x \in \mathcal{B}_d} p^{|x|}(1-p)^{k-|x|}. \tag{3.1.9}$$

The word d is neither constant zero nor constant one. The all-zero input therefore has selected carry 0^k, and the all-one input has selected carry 1^k. Hence

$$F(0) = 0, \qquad F(1) = 1. \tag{3.1.10}$$

For $x_i = 0$, write $x^+ = x + e_i$. The directed edge (x, x^+) is *pivotal* if $x \notin \mathcal{B}_d$ and $x^+ \in \mathcal{B}_d$. Differentiating (3.1.9) coordinate by coordinate gives

$$F'(p) = \sum_{(x,x^+)\ \text{pivotal}} p^{|x|}(1-p)^{k-1-|x|}. \tag{3.1.11}$$

We next partition this sum into classes whose bias profiles can be compared term by term.

Classification of pivotal edges Fix a pivotal edge at coordinate i, and let

$$u = x^+, \qquad c = c(u), \qquad h = |c| - m > 0, \qquad U = \{j : c_j \neq d_j\}.$$

Since $1 \le m \le k-1$, we have $1 \le h \le k-1$. Call an input u compatible with c when c satisfies (3.1.3) for that input.

Lemma 3.1.2 (Pivotal interval). *Let u be compatible with the bad carry word c, and suppose $u_i = 1$. The edge obtained by changing u_i to zero is pivotal if and only if*

$$i \in U, \quad c_{i+r} = 1 \qquad (1 \le r \le h), \tag{3.1.12}$$

$$u_{i+r} = 0 \qquad \text{whenever } 1 \le r < h \text{ and } d_{i+r} = 1. \tag{3.1.13}$$

Proof. Compare the selected carry words below and above the edge. They are ordered by (3.1.8). At a gate other than i, both recurrences use the same unary map on the carry state. Such a map is either the identity or a constant. A difference between the two carry states can therefore end at one of these gates, but it cannot begin there. The difference set is a single forward cyclic interval starting after i, unless it fills the entire cycle. On this interval the upper carry is one and the lower carry is zero. Assume first that the interval does not fill the cycle. The two carry states agree on entering gate i. Flipping the input at that gate starts a difference precisely when $c_i \neq d_i$ and $c_{i+1} = 1$. The latter condition is included in (3.1.12). Once the upper and lower carries are one and zero, respectively, a zero digit preserves their difference whenever the upper carry remains one. A one digit preserves it exactly when the common input is zero. Conditions (3.1.12) and (3.1.13) are therefore equivalent to an interval of length at least h. If the interval has length ℓ, the carry weights differ by ℓ. The upper weight is $m + h$. Thus the lower endpoint is outside $\mathcal{B}_d$ exactly when $\ell \ge h$. This proves both directions in the noncyclic case. An interval that fills the cycle has upper carry 1^k and lower carry 0^k. If $d_i = 1$, the lower input is $1 - d$, whose selected carry is 1^k, so this does not give a pivotal edge. Hence a genuine full-cycle edge has $d_i = 0$. Its upper input is $1 - d$, and it satisfies (3.1.12) and (3.1.13). Conversely, if propagation under these conditions never merges the two carry states, the same full-cycle edge results. This also accounts for the exceptional input.

We group pivotal edges by i and their upper carry word c. Every input compatible with a nonconstant c selects it because its closed carry is unique. The only constant bad carry is 1^k, which is selected as the greatest closed carry. Thus the parametrization by compatible inputs does not add or remove an edge.

Lemma 3.1.3 (Pivotal subcube balance). *Fix a nonempty pivotal group (i, c). Set*

$$P = \{i + r : 1 \le r < h,\ d_{i+r} = 1\}, \qquad p_0 = |P|,$$
$$R = |\{j \in U : c_{j+1} = 1\}|.$$

On the $k-1$ coordinates other than i, the lower endpoints form a subcube with

$$\alpha = R - 1 \quad \textit{fixed ones}, \qquad \beta = |U| - R + p_0 \quad \textit{fixed zeros}. \tag{3.1.14}$$

Moreover,

$$\alpha - \beta = h - 1 - p_0 \geq 0. \tag{3.1.15}$$

Proof. If $j \notin U$, then $c_j = d_j$. Compatibility forces $c_{j+1} = c_j$ and leaves u_j free. If $j \in U$, compatibility fixes

$$u_j = c_{j+1}. \tag{3.1.16}$$

The coordinate i is one of the mismatch positions fixed to one at the upper endpoint, and it is omitted from the lower-endpoint profile. The other mismatch positions contribute $R-1$ fixed ones and $|U|-R$ fixed zeros. For every $j \in P$, Lemma 3.1.2 gives $c_j = 1 = d_j$. These are matching positions, and (3.1.13) fixes their input bits to zero. All remaining matching positions are free. This proves (3.1.14). It remains to compare the two fixed-bit counts. At every $j \notin U$, support gives $c_{j+1} = c_j$. Cyclic telescoping therefore gives

$$\sum_{j \in U} (c_{j+1} - c_j) = 0.$$

Since $c_j \neq d_j$ for $j \in U$,

$$\begin{aligned} 2R - |U| &= \sum_{j \in U} (2c_{j+1} - 1) \\ &= \sum_{j \in U} (2c_j - 1) \\ &= \sum_{j=0}^{k-1} (c_j - d_j) = h. \end{aligned} \tag{3.1.17}$$

Substitution into (3.1.14) gives

$$\alpha - \beta = 2R - |U| - 1 - p_0 = h - 1 - p_0.$$

The set P occupies at most the displayed $h-1$ positions, so the last quantity is nonnegative.

Integration of the pivotal profiles

Theorem 3.1.1 (Tu–Deng inequality). *For every $k \geq 2$ and every $1 \leq t \leq 2^k - 2$,*

$$|S(t,k)| \leq 2^{k-1}.$$

Proof. Consider one pivotal group with fixed-bit counts α and β. When the free coordinates are summed in (3.1.11), their factors sum to one. The contribution of this group to

$$\int_0^{1/2} \big(F'(1-p) - F'(p)\big)\, dp$$

is therefore

$$\int_0^{1/2} \left((1-p)^\alpha p^\beta - p^\alpha (1-p)^\beta\right)\, dp. \tag{3.1.18}$$

For $0 < p \le 1/2$, Lemma 3.1.3 gives

$$\frac{(1-p)^\alpha p^\beta}{p^\alpha (1-p)^\beta} = \left(\frac{1-p}{p}\right)^{\alpha-\beta} \ge 1.$$

The endpoints follow by continuity. Thus (3.1.18) is nonnegative. The groups partition all pivotal edges. Summing their contributions and using (3.1.10), we obtain

$$\begin{aligned} 1 - 2F(1/2) &= F(1) - F(1/2) - \big(F(1/2) - F(0)\big) \\ &= \int_0^{1/2} \big(F'(1-p) - F'(p)\big)\, dp \\ &\ge 0. \end{aligned}$$

Hence $F(1/2) \le 1/2$. By (3.1.4),

$$|S(t,k)| = 2^k F(1/2) \le 2^{k-1}.$$

Equality cases Let $z = k - \mathrm{wt}_k(t)$ be the number of zero bits in d. Write their positions as

$$0 \le r_1 < \cdots < r_z < k, \qquad r_{z+1} = r_1 + k,$$

and define the cyclic zero gaps

$$g_j = r_{j+1} - r_j \qquad (1 \le j \le z). \tag{3.1.19}$$

Theorem 3.1.2 (Equality characterization). *For every $k \ge 2$ and $1 \le t \le 2^k - 2$,*

$$|S(t,k)| = 2^{k-1} \quad \Longleftrightarrow \quad g_j \ge z \ \text{ for every } j. \tag{3.1.20}$$

Equivalently, each cyclic run of ones between consecutive zero bits has length at least $z-1$. In particular, the zero bits are isolated.

Proof. For a nonempty pivotal group G, let α_G and β_G be its numbers of fixed ones and fixed zeros. The proof of Theorem 3.1.1 gives the exact identity

$$1 - \frac{|S(t,k)|}{2^{k-1}} = \sum_G J_G, \qquad J_G = \int_0^{1/2} \left((1-p)^{\alpha_G} p^{\beta_G} - p^{\alpha_G}(1-p)^{\beta_G}\right)\, dp. \tag{3.1.21}$$

The sum ranges over the nonempty pivotal groups. Each J_G is nonnegative. It vanishes exactly when $\alpha_G = \beta_G$. In the notation of Lemma 3.1.3, this condition is

$$p_0 = h - 1. \tag{3.1.22}$$

Thus equality in the Tu–Deng bound holds exactly when every nonempty pivotal group satisfies (3.1.22). Suppose first that two cyclically consecutive zeros occur at positions i and $i+g$, where $g < z$. Take the upper carry word $c = 1^k$ and the upper input $u = 1-d$. Our carry convention selects c for this input. The pair (i, c) is a nonempty pivotal group with $h = z$, by Lemma 3.1.2. Among the positions $i+1, \ldots, i+z-1$, at least one is zero, namely $i+g$. Hence $p_0 \le z-2 < h-1$. This group has $J_G > 0$, so $|S(t,k)| < 2^{k-1}$. The gap condition is therefore necessary. Conversely, assume that every cyclic zero gap is at least z, and fix a nonempty pivotal group (i, c). Set

$$A = \{j : d_j = 0,\ c_j = 1\}, \qquad B = \{j : d_j = 1,\ c_j = 0\}.$$

Since $h = |c| - |d|$,

$$h = |A| - |B|, \qquad |B| \le z - h, \qquad h \le z. \tag{3.1.23}$$

Suppose that $p_0 < h - 1$, and choose the least r with $1 \le r < h$ and $d_{i+r} = 0$. If $d_i = 0$, then i and $i + r$ are cyclically consecutive zero positions. Their gap satisfies $r < h \le z$, contrary to the assumption. It remains to consider $d_i = 1$. Pivotality gives $c_i \ne d_i$, hence $c_i = 0$. Let q be the nearest zero position before i, in cyclic order, and let s be the cyclic distance from q to i. Every position from $q + 1$ through i has digit one. Compatibility of c with an upper input implies

$$c_j = d_j \quad \Longrightarrow \quad c_{j+1} = c_j.$$

If one of these positions had $c_j = 1 = d_j$, this implication would force $c_{j+1} = 1$. Starting from $c_i = 0$ and moving backwards therefore gives $c_j = 0$ at all s positions. They all belong to B, so $s \le |B|$. By the minimal choice of r, the positions q and $i + r$ are cyclically consecutive zeros. Their cyclic distance is

$$s + r \le |B| + h - 1 \le z - 1,$$

where the last inequality follows from (3.1.23). This again contradicts the assumed gap condition. We conclude that every nonempty pivotal group has $p_0 = h - 1$. All terms in (3.1.21) vanish, which proves equality.

Corollary 3.1.1 (Enumeration of equality cases). *Equality requires $z^2 \le k$. For a fixed number z of zero bits, the number of residues t satisfying $|S(t,k)| = 2^{k-1}$ is*

$$N_{k,z} = \frac{k}{z}\binom{k - z^2 + z - 1}{z - 1}, \qquad 1 \le z \le \lfloor \sqrt{k} \rfloor. \tag{3.1.24}$$

Proof. The gaps in (3.1.19) sum to k. Write $g_j = z + e_j$, where $e_j \ge 0$. Then

$$\sum_{j=1}^{z} e_j = k - z^2.$$

This proves the necessary condition $z^2 \le k$. When it holds, the number of ordered gap vectors with a distinguished initial zero is

$$\binom{k - z^2 + z - 1}{z - 1}.$$

There are k choices for the position of the distinguished zero. Every binary word is counted once for each of its z zero positions. Division by z gives (3.1.24).

As an independent finite check, we evaluated the original pair definition of $S(t, k)$ for every admissible t and every $2 \le k \le 14$. The gap criterion and the counts in Corollary 3.1.1 agree in all 32,738 cases. This computation is not used in either proof.

3.2 Bounding vectorial nonlinearity

Problem and prior bounds We use the Walsh transform, the components F_v, and scalar and vectorial nonlinearity with the meanings fixed at the beginning of this section. In particular, those conventions give

$$\mathrm{NL}(F) = 2^{n-1} - \frac{1}{2} \max_{\substack{0 \neq v \in V_m \\ a \in V_n}} |\mathcal{W}_{F_v}(a)|.$$

This is the convention used for S-boxes and vectorial Boolean functions throughout the Boolean-function literature [10].

Open Question 3.2.1 (Maximum vectorial nonlinearity). *Let $n = 2k$ and $k < m < n$. Determine*

$$\mathcal{N}(n, m) = \max_{F : V_n \to V_m} \mathrm{NL}(F).$$

In particular, determine the largest uniform improvement over the scalar bent bound in the output range $n/2 < m < n$.

Set $q = 2^k$, $N = 2^{2k} = q^2$, and

$$B = 2^{2k-1} - 2^{k-1} = \frac{N - q}{2}.$$

The covering radius of the first-order Reed–Muller code gives $\mathcal{N}(2k, m) \le B$. Equality for a given F requires every nonzero component F_v to be bent. Nyberg proved that such a vectorial bent function exists only when $m \le k$ [42]. Hence, for $m > k$, integrality yields $\mathcal{N}(2k, m) \le B - 1$. This is the first quantitative consequence of the Nyberg bound for the present problem. Wadayama et al. formulated $\mathcal{N}(n, m)$ in coding-theoretic terms and compared the covering-radius bound with the Chabaud–Vaudenay bound and nonexistence bounds for binary linear codes [11, 60]. The Chabaud–Vaudenay estimate applies in the restricted range $m > n - 2$, while the code bounds depend on the available tables

for individual parameters. Dib later obtained asymptotic estimates for almost all vectorial Boolean functions and retained the improvement of the covering-radius bound outside those parameter regimes as an open problem [20]. Xu, Carlet, Mesnager, and Wu revisited the range $n/2 < m < n$, recording that no tight general upper bound was known there and deriving upper bounds for selected classes of vectorial functions [63]. Another line of work determines how many bent components a vectorial function may have after crossing the Nyberg threshold [2, 37, 47]. Those results count bent components. They do not exclude the possibility that all remaining components have Walsh magnitude at most $q+2$, which is precisely the case that must be ruled out to improve $B-1$. Nagy studied a different nonlinearity parameter, namely the distance of a vectorial function from the set of affine vectorial functions, using thin and Sidon sets [41]. That parameter is distinct from the component nonlinearity used here, so the resulting bounds are complementary to the present statement.

Main result

Theorem 3.2.1 (Uniform bound beyond the Nyberg threshold). *Let $n = 2k \geq 6$ and $k < m < 2k$. Every $F : V_{2k} \to V_m$ satisfies*

$$\mathrm{NL}(F) \leq B - 2 = 2^{n-1} - 2^{n/2-1} - 2. \tag{3.2.1}$$

Equivalently,

$$\max_{\substack{0 \neq v \in V_m \\ a \in V_{2k}}} |\mathcal{W}_{F_v}(a)| \geq q + 4. \tag{3.2.2}$$

The comparison is $\mathcal{N}(2k, m) \leq B - 2$, whereas the covering-radius argument gives B and Nyberg's bound, together with integrality, gives $B-1$. In correlation terms, some component admits a linear approximation with correlation magnitude at least $(q+4)/q^2$. The two-unit improvement in nonlinearity is therefore equivalent to increasing the unavoidable Walsh magnitude from q at the scalar bent benchmark, or $q+2$ after the Nyberg obstruction, to $q+4$. The proof has two parts. A parity argument first shows that a hypothetical function with nonlinearity at least $B-1$ contains a large vectorial bent subfunction. This excludes $m \geq k+2$. In the critical case $m = k+1$, a congruence normalization forces the remaining component coset to have a two-level Walsh spectrum. The odd fibres of a maximal vectorial bent subfunction then contradict that spectrum.

Even-weight component space

Lemma 3.2.1 (Bent hyperplane of components). *Let $k \geq 3$, $m > k$, and $F : V_{2k} \to V_m$. Suppose that*

$$|\mathcal{W}_{F_v}(a)| \leq q + 2 \quad \text{for all } 0 \neq v \in V_m \text{ and } a \in V_{2k}. \tag{3.2.3}$$

Define $s = \sum_{x \in V_{2k}} F(x)$ and $K = s^{\perp}$. Then $s \neq 0$, $\dim K = m - 1$, and every component F_v with $0 \neq v \in K$ is bent. Consequently, $m - 1 \leq k$.

Proof. For every $v \in V_m$, summation in $\mathbb{F}_2$ gives

$$\mathrm{wt}(F_v) \equiv \sum_{x \in V_{2k}} v \cdot F(x) = v \cdot s \pmod 2. \tag{3.2.4}$$

Thus K is exactly the set of indices of the even-weight components. Fix $0 \neq v \in K$. For $a = 0$, the function $F_v + a \cdot x$ has even weight by (3.2.4). For $a \neq 0$, the linear function $a \cdot x$ has weight $N/2$, which is even, so $\mathrm{wt}(F_v + a \cdot x)$ is again even. It follows that

$$\mathcal{W}_{F_v}(a) = N - 2\,\mathrm{wt}(F_v + a \cdot x) \equiv 0 \pmod 4 \qquad (a \in V_{2k}). \tag{3.2.5}$$

Walsh Parseval gives

$$\sum_{a \in V_{2k}} \mathcal{W}_{F_v}(a)^2 = N^2. \tag{3.2.6}$$

If every coefficient had magnitude at most q, then the left-hand side would be at most $Nq^2 = N^2$. Equality is possible only when every coefficient has magnitude q, which is the bent case. Hence a non-bent F_v has a coefficient of magnitude strictly greater than q. Both q and the coefficients in (3.2.5) are divisible by 4, so this coefficient has magnitude at least $q + 4$. This contradicts (3.2.3). Therefore every nonzero component indexed by K is bent. If $s = 0$, then $K = V_m$, and all nonzero components of F are bent. This would make F vectorial bent, contrary to Nyberg's bound because $m > k$. Hence $s \neq 0$ and $\dim K = m - 1$. Choosing a basis of K produces a vectorial bent function from V_{2k} to V_{m-1}. Nyberg's bound applied once more gives $m - 1 \leq k$.

Corollary 3.2.1 (Noncritical output dimensions). *Let $k \geq 3$ and $k + 2 \leq m < 2k$. Then every $F : V_{2k} \to V_m$ satisfies $\mathrm{NL}(F) \leq B - 2$.*

Proof. If $\mathrm{NL}(F) \geq B - 1$, the Walsh expression for nonlinearity implies (3.2.3). Lemma 3.2.1 would then give $m - 1 \leq k$, contrary to $m \geq k + 2$.

Critical output dimension It remains to treat $m = k+1$. Assume throughout this part that $\mathrm{NL}(F) \geq B - 1$. Lemma 3.2.1 gives a k-dimensional component space $K = s^{\perp}$ whose nonzero members are bent. Choose a basis $v_1, \ldots, v_k$ of K and an index $v_{k+1} \notin K$. An invertible linear change of output coordinates preserves the set of nonzero components up to a bijective reindexing and therefore preserves $\mathrm{NL}(F)$. After applying this change and renaming the resulting function F, we may write

$$F = (G, h), \qquad G = (F_{v_1}, \ldots, F_{v_k}) : V_{2k} \to V_k, \qquad h = F_{v_{k+1}}.$$

The function G is vectorial bent with the maximal possible output dimension k. Since $v_{k+1} \cdot s = 1$, every function

$$f_u = h + u \cdot G, \qquad u \in V_k,$$

has odd weight by (3.2.4). Consequently, $\mathcal{W}_{f_u}(a) \equiv 2 \pmod 4$. Define

$$z_u(a) = \frac{1}{2}\mathcal{W}_{f_u}(a). \tag{3.2.7}$$

Each $z_u(a)$ is an odd integer, and the assumed lower bound on $\mathrm{NL}(F)$ gives

$$|z_u(a)| \le \frac{q}{2} + 1. \tag{3.2.8}$$

Lemma 3.2.2 (Simultaneous congruence normalization). *Let G, h, and $z_u(a)$ be as above. Put*

$$w = \mathrm{wt}(h), \qquad P = \sum_{x \in V_{2k}} h(x)G(x) \in V_k, \qquad S = \sum_{x \in V_{2k}} h(x)x \in V_{2k}.$$

There is a sign $\gamma \in \{\pm 1\}$ such that

$$z_u(a) \equiv \gamma(-1)^{u\cdot P + a\cdot S} \pmod 4 \qquad (u \in V_k,\ a \in V_{2k}). \tag{3.2.9}$$

In particular,

$$\widetilde{z}_u(a) = \gamma(-1)^{u\cdot P + a\cdot S} z_u(a) = 1 + 4c_u(a)$$

for integers $c_u(a)$ satisfying

$$-\frac{q}{8} \le c_u(a) \le \frac{q}{8}. \tag{3.2.10}$$

Proof. For $u \in V_k$ and $a \in V_{2k}$, let $r_{u,a}(x) = u \cdot G(x) + a \cdot x$. If $u \neq 0$, then $r_{u,a}$ is bent and has weight $N/2 \pm q/2$. Both terms are divisible by 4 because $k \ge 3$. If $u = 0$, then $r_{0,a}$ is zero or a nonzero linear function, with weight 0 or $N/2$. Hence

$$\mathrm{wt}(r_{u,a}) \equiv 0 \pmod 4 \qquad (u \in V_k,\ a \in V_{2k}). \tag{3.2.11}$$

The parity of the pointwise product $h\, r_{u,a}$ is

$$\mathrm{wt}(h\, r_{u,a}) \equiv \sum_x h(x)\big(u \cdot G(x) + a \cdot x\big) = u \cdot P + a \cdot S \pmod 2. \tag{3.2.12}$$

Using $\mathrm{wt}(h + r) = \mathrm{wt}(h) + \mathrm{wt}(r) - 2\,\mathrm{wt}(hr)$, equations (3.2.11) and (3.2.12) give

$$\mathrm{wt}(f_u + a \cdot x) \equiv w - 2(u \cdot P + a \cdot S) \pmod 4. \tag{3.2.13}$$

Since $N/2 \equiv 0 \pmod 4$, the identity $z_u(a) = N/2 - \mathrm{wt}(f_u + a \cdot x)$ now yields

$$z_u(a) \equiv -w + 2(u \cdot P + a \cdot S) \pmod 4. \tag{3.2.14}$$

The integer w is odd. Choose $\gamma \in \{\pm 1\}$ with $\gamma \equiv -w \pmod 4$. For $\beta \in \mathbb{F}_2$, one has $\gamma + 2\beta \equiv \gamma(-1)^\beta \pmod 4$. Taking $\beta = u \cdot P + a \cdot S$ in (3.2.14) proves (3.2.9). Thus $\widetilde{z}_u(a) \equiv 1 \pmod 4$, which defines the integers $c_u(a)$. Finally, (3.2.8) gives $-q/2 - 1 \le 1 + 4c_u(a) \le q/2 + 1$. Since $q/8$ is an integer, this is equivalent to (3.2.10).

Lemma 3.2.3 (Two-level spectrum forced by row moments). *Under the assumptions of Lemma 3.2.2,*

$$|\mathcal{W}_{f_u}(a)| \in \{q-2, q+2\} \qquad (u \in V_k,\ a \in V_{2k}). \tag{3.2.15}$$

Proof. Fix $u \in V_k$ and set $R_u = \sum_a \widetilde{z}_u(a)$. Walsh inversion at S gives

$$\begin{aligned} R_u &= \frac{\gamma(-1)^{u\cdot P}}{2} \sum_{a \in V_{2k}} \mathcal{W}_{f_u}(a)(-1)^{a\cdot S} \\ &= \frac{N}{2}\,\gamma(-1)^{h(S)+u\cdot(P+G(S))} \in \left\{-\frac{N}{2}, \frac{N}{2}\right\}. \end{aligned} \tag{3.2.16}$$

On the other hand, $\widetilde{z}_u(a) = 1 + 4c_u(a)$ and there are N frequencies, so

$$R_u = N + 4\sum_a c_u(a). \tag{3.2.17}$$

The two possible signs in (3.2.16) therefore give

$$\sum_a c_u(a) = -\frac{N}{8} \quad \text{or} \quad \sum_a c_u(a) = -\frac{3N}{8},$$

respectively. Walsh Parseval and $|\widetilde{z}_u(a)| = |z_u(a)|$ give

$$\sum_a \widetilde{z}_u(a)^2 = \frac{1}{4}\sum_a \mathcal{W}_{f_u}(a)^2 = \frac{N^2}{4}.$$

Expanding $\widetilde{z}_u(a) = 1 + 4c_u(a)$ yields

$$N + 8\sum_a c_u(a) + 16\sum_a c_u(a)^2 = \frac{N^2}{4}. \tag{3.2.18}$$

If $R_u = N/2$, substitution in (3.2.18) gives

$$\sum_a c_u(a)^2 = \frac{N^2}{64}. \tag{3.2.19}$$

Equation (3.2.10) gives $c_u(a)^2 \le q^2/64 = N/64$. There are N summands, so (3.2.19) attains the termwise upper bound. It follows that $|c_u(a)| = q/8$ for every a. If $R_u = -N/2$, the same substitution gives

$$\sum_a c_u(a)^2 = \frac{N^2}{64} + \frac{N}{8},$$

which is strictly larger than the termwise upper bound $N^2/64$. This case is impossible. Hence $\widetilde{z}_u(a) = 1 \pm q/2$ for every u, a. Since $|z_u(a)| = |\widetilde{z}_u(a)|$, equation (3.2.15) follows from $\mathcal{W}_{f_u}(a) = 2z_u(a)$.

Lemma 3.2.4 (Odd fibres exclude the two-level spectrum). *Let $G : V_{2k} \to V_k$ be maximal vectorial bent with $k \geq 3$, and let $h : V_{2k} \to \mathbb{F}_2$. The family $f_u = h + u \cdot G$ cannot satisfy* (3.2.15) *for every $u \in V_k$ and $a \in V_{2k}$.*

Proof. Products of Boolean functions in this proof are pointwise. First observe that every fibre of G has odd cardinality. Indeed, Fourier inversion for the indicator of $G^{-1}(y)$ gives

$$\begin{aligned} |G^{-1}(y)| &= \frac{1}{q} \sum_{v \in V_k} (-1)^{v \cdot y} \mathcal{W}_{v \cdot G}(0) \\ &= q + \sum_{0 \neq v \in V_k} \varepsilon_v (-1)^{v \cdot y}, \qquad \varepsilon_v \in \{\pm 1\}. \end{aligned} \tag{3.2.20}$$

The second sum in (3.2.20) contains $q - 1$ odd terms. It is odd, whereas q is even. Thus $|G^{-1}(y)|$ is odd. Fix $a \in V_{2k}$ and define

$$T_y(a) = \sum_{\substack{x \in V_{2k} \\ G(x) = y}} (-1)^{h(x) + a \cdot x} \qquad (y \in V_k). \tag{3.2.21}$$

Each $T_y(a)$ is a sum of an odd number of signs, hence is odd. Grouping the Walsh sum by the fibres of G gives the Hadamard relation

$$\mathcal{W}_{f_u}(a) = \sum_{y \in V_k} (-1)^{u \cdot y} T_y(a). \tag{3.2.22}$$

Parseval for this $q \times q$ Hadamard transform gives

$$\sum_{u \in V_k} \mathcal{W}_{f_u}(a)^2 = q \sum_{y \in V_k} T_y(a)^2. \tag{3.2.23}$$

Every odd square is congruent to 1 (mod 8), so $\sum_y T_y(a)^2 = q + 8L$ for some integer L. Dividing (3.2.23) by 4 and using (3.2.7), we obtain

$$\sum_{u \in V_k} z_u(a)^2 = \frac{q^2}{4} + 2qL \equiv 0 \pmod{2q}, \tag{3.2.24}$$

where $2q \mid q^2/4$ because $k \geq 3$. Suppose now that (3.2.15) holds. Then

$$z_u(a)^2 = \left(\frac{q}{2} \pm 1\right)^2 \equiv q + 1 \pmod{2q},$$

because $2q \mid q^2/4$ and $-q \equiv q \pmod{2q}$. Summing over the q choices of u gives

$$\sum_{u \in V_k} z_u(a)^2 \equiv q(q+1) \equiv q \pmod{2q}, \tag{3.2.25}$$

contradicting (3.2.24).

Proof (Proof of Theorem 3.2.1). Suppose that $\mathrm{NL}(F) \geq B-1$. This is equivalent to the spectral ceiling (3.2.3). If $m \geq k+2$, Corollary 3.2.1 gives an immediate contradiction. Thus $m = k+1$. Use the output decomposition $F = (G, h)$ constructed at the beginning of Section 3.2. Lemma 3.2.2 normalizes the half-Walsh spectrum modulo 4, and Lemma 3.2.3 forces $|\mathcal{W}_{f_u}(a)| \in \{q-2, q+2\}$ for every u, a. Lemma 3.2.4 rules out this spectrum. Hence $\mathrm{NL}(F) \geq B-1$ is impossible, which proves $\mathrm{NL}(F) \leq B-2$.

Four-variable verification The proof of Theorem 3.2.1 requires $k \geq 3$. In particular, the divisibilities $4 \mid q/2$ and $2q \mid q^2/4$ fail for $k = 2$. The only remaining parameter pair is $(n, m) = (4, 3)$. We determine it by a complete search in the first-order Reed–Muller quotient. The search is included here because the exact value is a computational claim rather than a consequence of the preceding congruence argument.

Proposition 3.2.1 (Exact four-variable value). *The maximum component nonlinearity of a function $F : V_4 \to V_3$ is*

$$\mathcal{N}(4,3) = 4 = 2^3 - 2^1 - 2. \tag{3.2.26}$$

Proof. Put $Q = \mathbb{F}_2^{16}/\mathrm{RM}(1,4)$. For a coset $C \in Q$, define its quotient weight by $d(C) = \min_{f \in C} \mathrm{wt}(f)$. Since the nonlinearity of a Boolean function is its distance from the affine functions, $\mathrm{NL}(f) = d([f])$. For $F : V_4 \to V_3$, the map $L_F : V_3 \to Q$ given by $L_F(v) = [F_v]$ is linear. If $\mathrm{NL}(F) > 0$, then L_F is injective. Indeed, a nonzero element of its kernel would index an affine component and would force $\mathrm{NL}(F) = 0$. Thus the eight component cosets form a three-dimensional subspace of Q. Conversely, choose representatives of a basis of any three-dimensional subspace of Q and use their truth tables as the three coordinates of F. This gives

$$\begin{aligned} \mathcal{N}(4,3) \geq t \quad \Longleftrightarrow \quad & \text{there is a three-space } U \leq Q \text{ such that} \\ & d(C) \geq t \quad \text{for every } 0 \neq C \in U \end{aligned} \tag{3.2.27}$$

for every positive integer t. Algorithm 2 tests this condition exhaustively. Once C_1, C_2, and $C_1 + C_2$ pass the threshold, the inner loop tests every C_3 outside their span and the three remaining sums. These are exactly the seven nonzero elements of $\langle C_1, C_2, C_3 \rangle$. Every three-space has such a basis, so a return value of $\perp$ is a complete nonexistence certificate. The quotient enumeration gives the distance distribution $(1, 16, 120, 560, 875, 448, 28)$ at distances $0, 1, \ldots, 6$, respectively. The call $\textsc{Search}(5)$ returns $\perp$. The call $\textsc{Search}(4)$ returns the basis $\texttt{000F}, \texttt{0033}, \texttt{0055}$. Here a 16-bit word encodes a truth table and its least significant bit is indexed by $0 \in V_4$. The seven nonzero representatives are $\texttt{000F}, \texttt{0033}, \texttt{0055}, \texttt{003C}, \texttt{005A}, \texttt{0066}, \texttt{0069}$, and each has quotient weight 4. Equation (3.2.27) now gives both inequalities in (3.2.26).

Algorithm 2 Complete three-space search in $\mathbb{F}_2^{16}/\operatorname{RM}(1,4)$

Require: The affine code $\mathcal{A} = \operatorname{RM}(1,4) \subseteq \mathbb{F}_2^{16}$
Ensure: A three-space of minimum quotient weight 4, and exhaustive failure at threshold 5

Enumerate the cosets $Q = \mathbb{F}_2^{16}/\mathcal{A}$ ▷ Each of the 2^{11} cosets contains 2^5 words.
for all $C \in Q$ **do**
    $r(C) \leftarrow \min C$ and $d(C) \leftarrow \min_{f \in C} \operatorname{wt}(f)$ ▷ The minimum for $r(C)$ uses the integer word ordering.
end for
function SEARCH(t)
    $E_t \leftarrow \{C \in Q \setminus \{0\} : d(C) \geq t\}$
    for all $C_1, C_2 \in E_t$ with $C_1 < C_2$ **do**
        if $C_1 + C_2 \notin E_t$ **then**
            continue
        end if
        $P \leftarrow \{0, C_1, C_2, C_1 + C_2\}$ ▷ P is the plane generated by C_1, C_2.
        for all $C_3 \in E_t \setminus P$ **do**
            if $C_1 + C_3, C_2 + C_3, C_1 + C_2 + C_3 \in E_t$ **then**
                return (C_1, C_2, C_3) ▷ All seven nonzero elements have distance at least t.
            end if
        end for
    end for
    return $\perp$ ▷ The search includes a basis for every three-space.
end function
return SEARCH(5) and SEARCH(4)

3.3 Maximum nonlinearity of balanced eight-variable functions

Problem statement and proof structure For a balanced Boolean function on eight variables, the largest previously known nonlinearity was 116, whereas the general spectral bound allowed 118. The possible value 118 was studied through its relation to resilient functions and was later stated as an open search target [36, 45]. We settle this finite extremal problem.

Open Question 3.3.1 (Maximum balanced nonlinearity in eight variables). *Determine*

$$\max\{\operatorname{NL}(f) : f : V_8 \to \mathbb{F}_2 \text{ is balanced}\}.$$

In particular, does a balanced eight-variable Boolean function of nonlinearity 118 exist?

Theorem 3.3.1 (Eight-variable case). *Every balanced Boolean function $f : V_8 \to \mathbb{F}_2$ satisfies*

$$\operatorname{NL}(f) \leq 116.$$

Equality is attained. Hence the maximum balanced nonlinearity in eight variables is exactly 116.

The proof is computer-assisted only in the finite classification of local spectral tables and in the verification of exact Farkas certificates. The mathematical reduction to that computation and the terminal exclusion of the remaining profiles are proved below. The local enumeration uses arbitrary-precision integers. The Farkas verifier uses exact rational arithmetic. The terminal certificates use integers only. Thus the finite search is specified mathematically and its output is checked by deterministic exact-arithmetic programs. This argument is not formalized in a proof assistant, and the artifact has not undergone an independent IACR audit. Appendix B gives the deterministic enumeration procedure and all global infeasibility certificates.

The proof has five steps. First, a hypothetical function of nonlinearity 118 has maximum Walsh magnitude 20 and a parity split of its spectrum. Second, each frequency in the odd half defines two seven-variable slices whose normalized spectra belong to nine local magnitude types. Third, exact arithmetic conditions give a finite enumeration of all local type tables. Fourth, a double count and exact Farkas certificates leave thirteen global spectral profiles. Finally, a common quadratic Boolean function and explicit inverse-Walsh inequalities exclude those thirteen profiles.

Notation and standard identities We use V_m, $\mathcal{W}_u$, $\mathrm{NL}(u)$, $\mathrm{wt}(u)$, and $\mathrm{supp}(u)$ with the meanings fixed at the beginning of this section. All vector spaces and dot products in this proof are over $\mathbb{F}_2$. We write $\mathbf{e}_j$ for a standard basis vector when its ambient space is clear. In particular, $\mathbf{e}_6 = (0, \dots, 0, 1) \in V_7$ in the odd-autocorrelation lemma and in the terminal certificates. Walsh inversion and Parseval are

$$\sum_{a \in V_m} (-1)^{a \cdot x} \mathcal{W}_u(a) = 2^m (-1)^{u(x)}, \tag{3.3.1}$$

$$\sum_{a \in V_m} \mathcal{W}_u(a)^2 = 2^{2m}. \tag{3.3.2}$$

For an integer-valued function $\varphi : V_7 \to \mathbb{Z}$, let

$$(\mathcal{H}\varphi)(t) = \sum_{a \in V_7} (-1)^{a \cdot t} \varphi(a)$$

be its unnormalized Hadamard transform. For a Boolean function $u : V_7 \to \mathbb{F}_2$, define its autocorrelation by

$$\mathrm{AC}_u(t) = \sum_{x \in V_7} (-1)^{u(x)+u(x+t)}.$$

For two Boolean functions $u_1, u_2 : V_7 \to \mathbb{F}_2$, define their cross-correlation by

$$\mathrm{CC}_{u_1,u_2}(t) = \sum_{x \in V_7} (-1)^{u_1(x)+u_2(x+t)}.$$

Character orthogonality gives the Wiener–Khinchin identities

$$\mathcal{H}(\mathcal{W}_u^2)(t) = 128\,\mathrm{AC}_u(t), \tag{3.3.3}$$
$$\mathcal{H}(\mathcal{W}_{u_1}\mathcal{W}_{u_2})(t) = 128\,\mathrm{CC}_{u_1,u_2}(t). \tag{3.3.4}$$

Reduction to maximum Walsh magnitude 20 Let $f : V_8 \to \mathbb{F}_2$ be balanced, and let $S = \mathrm{supp}(f)$. Then $|S| = 128$ and $\mathcal{W}_f(0) = 0$. For $\xi \neq 0$, define the support intersection number

$$n_S(\xi) = |\{x \in S : \xi \cdot x = 1\}|.$$

The complete character sum on V_8 vanishes. Therefore

$$\mathcal{W}_f(\xi) = -2\sum_{x\in S}(-1)^{\xi\cdot x} = -2(|S| - 2n_S(\xi)) = -256 + 4n_S(\xi). \tag{3.3.5}$$

Thus every Walsh coefficient is divisible by four.

Parseval and balancedness give

$$\sum_{\xi\neq 0}\mathcal{W}_f(\xi)^2 = 2^{16}.$$

If all 255 nonzero-frequency coefficients had magnitude at most 16, the left side would be at most $255 \cdot 16^2 = 65280 < 65536$, a contradiction. Hence $\max_\xi |\mathcal{W}_f(\xi)| \geq 20$. The relation between Walsh magnitude and nonlinearity fixed at the beginning of this section then gives $\mathrm{NL}(f) \leq 118$. If $\mathrm{NL}(f) > 116$, divisibility by four forces

$$\max_{\xi\in V_8}|\mathcal{W}_f(\xi)| = 20, \qquad \mathrm{NL}(f) = 118. \tag{3.3.6}$$

We assume (3.3.6) for the remainder of the exclusion argument.

Spectral parity and the affine split Define the support XOR and normalized Walsh spectrum by

$$c = \bigoplus_{x\in S} x \in V_8, \qquad \zeta_f(\xi) = \frac{\mathcal{W}_f(\xi)}{4}.$$

For $\xi \neq 0$, equation (3.3.5) gives $\zeta_f(\xi) = n_S(\xi) - 64$. Modulo two,

$$n_S(\xi) \equiv \sum_{x\in S}\xi\cdot x = \xi\cdot\left(\bigoplus_{x\in S} x\right) = \xi\cdot c.$$

Both sides are zero at $\xi = 0$. Hence

$$\zeta_f(\xi) \equiv \xi\cdot c \pmod 2 \qquad (\xi \in V_8). \tag{3.3.7}$$

The same vector is also the degree-seven coefficient vector in the algebraic normal form. Put $[8] = \{1, \dots, 8\}$, and write

$$f(x) = \bigoplus_{J \subseteq [8]} \alpha_J x^J, \qquad x^J = \prod_{i \in J} x_i,$$

where $\alpha_J \in \mathbb{F}_2$. All sums in this paragraph are taken in $\mathbb{F}_2$. The coefficient of the full monomial is

$$\alpha_{[8]} = \bigoplus_{x \in V_8} f(x) = 0,$$

because f is balanced. Fix $i \in [8]$, put $J_i = [8] \setminus \{i\}$, and let $H_i^0 = \{x \in V_8 : x_i = 0\}$. For $J \subseteq J_i$, the monomial x^J is one on $2^{7-|J|}$ points of H_i^0. Hence

$$\bigoplus_{x \in H_i^0} x^J = \begin{cases} 1, & J = J_i, \\ 0, & J \subsetneq J_i, \end{cases} \qquad \text{in } \mathbb{F}_2.$$

By linearity of the XOR,

$$\alpha_{J_i} = \bigoplus_{x \in H_i^0} f(x).$$

Since f is balanced, its weight is even. Thus

$$\alpha_{J_i} = \bigoplus_{x \in V_8 \colon x_i = 1} f(x) = \bigoplus_{x \in S} x_i = c_i.$$

Consequently, if $\mathbf{a}^{(7)} = (\alpha_{J_1}, \dots, \alpha_{J_8})$ denotes the degree-seven ANF coefficient vector, then $\mathbf{a}^{(7)} = c$. We use c for the support XOR in the spectral argument; the distinct symbol $\mathbf{a}^{(7)}$ denotes its ANF interpretation.

If $c = 0$, equation (3.3.7) makes every Walsh coefficient divisible by eight. Under (3.3.6), all Walsh magnitudes would then be at most 16, a contradiction. Hence $c \neq 0$. Therefore at least one degree-seven coefficient is nonzero, whereas $\alpha_{[8]} = 0$; consequently $\deg(f) = 7$. For $A \in \mathrm{GL}(V_8)$, put $f_A(x) = f(Ax)$. A change of variables gives

$$\bigoplus_{x \in \mathrm{supp}(f_A)} x = A^{-1}c, \qquad \mathcal{W}_{f_A}(\xi) = \mathcal{W}_f(A^{-\mathsf{T}}\xi).$$

Thus an invertible linear input change preserves balancedness and permutes the Walsh spectrum. Choose A such that $A\mathbf{e}_8 = c$, and replace f by f_A. We may therefore assume

$$c = \mathbf{e}_8.$$

Set

$$K = c^{\perp} = \{\xi \in V_8 : \xi_8 = 0\}.$$

Equations (3.3.7) and (3.3.6) yield

$$\begin{aligned} \xi \in K &\Longrightarrow |\mathcal{W}_f(\xi)| \in \{0, 8, 16\}, \\ \xi \notin K &\Longrightarrow |\mathcal{W}_f(\xi)| \in \{4, 12, 20\}. \end{aligned} \tag{3.3.8}$$

We call K the even half and $V_8 \setminus K$ the odd half of the frequency space.

Paired seven-variable slices Fix once and for all a linear isomorphism $\iota : V_7 \to K$. For each $L \in V_8 \setminus K$, define

$$\Phi_L : V_7 \times \mathbb{F}_2 \longrightarrow V_8, \qquad \Phi_L(a, \beta) = \iota(a) + \beta L.$$

The map Φ_L is invertible. Let $A_L : V_8 \to V_8$ be the linear map determined by $A_L^{-\mathsf{T}} = \Phi_L$, and put

$$f_L(y) = f(A_L y), \qquad y \in V_7 \times \mathbb{F}_2.$$

Then

$$\mathcal{W}_{f_L}(a, \beta) = \mathcal{W}_f(\Phi_L(a, \beta)).$$

Define the two restrictions

$$g_L(x) = f_L(x, 0), \qquad h_L(x) = f_L(x, 1).$$

Direct summation over the last input coordinate gives

$$\mathcal{W}_f(\iota(a) + \beta L) = \mathcal{W}_{g_L}(a) + (-1)^\beta \mathcal{W}_{h_L}(a). \tag{3.3.9}$$

At $a = 0$, equation (3.3.9) becomes

$$\mathcal{W}_{g_L}(0) + \mathcal{W}_{h_L}(0) = 0, \qquad \mathcal{W}_{g_L}(0) - \mathcal{W}_{h_L}(0) = \mathcal{W}_f(L).$$

The input translation $(x, \beta) \mapsto (x, \beta + 1)$ swaps g_L, h_L and negates the odd-half Walsh coefficients. It preserves all Walsh magnitudes. We use this equivalent representative and assume $\mathcal{W}_f(L) > 0$.

For the local analysis, suppress the subscript L from g_L, h_L, and define

$$p(a) = \frac{\mathcal{W}_g(a)}{2}, \qquad q(a) = \frac{\mathcal{W}_h(a)}{2}.$$

According as $|\mathcal{W}_f(L)|$ is $20, 12$, or 4, the ordered slice weights are

$$(\mathrm{wt}(g), \mathrm{wt}(h)) \in \{(59, 69), (61, 67), (63, 65)\}.$$

Since $\mathcal{W}_g(0) = 128 - 2\,\mathrm{wt}(g)$, these three cases give

$$p(0) = d, \qquad q(0) = -d, \qquad d \in \{5, 3, 1\}. \tag{3.3.10}$$

Both slice weights are odd, so all values of p and q are odd. Indeed, for any Boolean function $u : V_7 \to \mathbb{F}_2$,

$$\frac{\mathcal{W}_u(a)}{2} = 64 - \mathrm{wt}\big(u + (a \cdot x)\big).$$

The linear function $x \mapsto a \cdot x$ has even weight, including weight zero when $a = 0$. Therefore $\mathrm{wt}(u + (a \cdot x)) \equiv \mathrm{wt}(u) \pmod 2$, which proves the assertion.

Dividing (3.3.9) by four gives

$$\frac{\mathcal{W}_f(\iota(a))}{4} = \frac{p(a)+q(a)}{2}, \qquad \frac{\mathcal{W}_f(\iota(a)+L)}{4} = \frac{p(a)-q(a)}{2}.$$

For real numbers r, s, $\max\{|r+s|, |r-s|\} = |r| + |s|$. Therefore the maximum Walsh magnitude 20 is equivalent to

$$|p(a)| + |q(a)| \le 10 \qquad (a \in V_7). \tag{3.3.11}$$

The even-half normalized coefficient $(p(a)+q(a))/2$ is even by (3.3.7). Hence

$$p(a)+q(a) \equiv 0 \pmod 4, \qquad p(a) \equiv -q(a) \pmod 4. \tag{3.3.12}$$

For an odd-weight Boolean function $u : V_7 \to \mathbb{F}_2$, define its support XOR by

$$C_u = \bigoplus_{u(x)=1} x \in V_7.$$

For $a \neq 0$, let $M_u(a) = |\{x : u(x) = 1,\ a \cdot x = 1\}|$. Then

$$\frac{\mathcal{W}_u(a)}{2} = -\operatorname{wt}(u) + 2M_u(a), \qquad M_u(a) \equiv a \cdot C_u \pmod 2.$$

Since $\mathcal{W}_u(0)/2 = 64 - \operatorname{wt}(u)$, this proves

$$\frac{\mathcal{W}_u(a)}{2} \equiv \frac{\mathcal{W}_u(0)}{2} + 2a \cdot C_u \pmod 4. \tag{3.3.13}$$

Applying (3.3.13) to g, h and using (3.3.12) gives $C_g = C_h$.

Put $t = C_g = C_h$, and translate both slices by setting $g_t(x) = g(x+t)$ and $h_t(x) = h(x+t)$. For any odd-weight Boolean function u on V_7, a change of variables gives

$$C_{u(\cdot+t)} = C_u + t, \qquad \mathcal{W}_{u(\cdot+t)}(a) = (-1)^{a\cdot t}\mathcal{W}_u(a).$$

The first identity uses $\operatorname{wt}(u) \equiv 1 \pmod 2$. Thus both translated slices have support XOR zero. The second identity shows that their Walsh spectra are multiplied by the same character. It preserves their magnitudes, their pointwise products, and every local type count. We replace g, h by these translates and consequently assume

$$p(a) \equiv d \pmod 4, \qquad q(a) \equiv -d \pmod 4 \qquad (a \in V_7). \tag{3.3.14}$$

For each odd magnitude, exactly one sign has the prescribed residue modulo four. Thus the signs of $p(a)$ and $q(a)$ are fixed by their magnitudes. Equation (3.3.12) also gives $p(a)q(a) \equiv -1 \pmod 4$. Hence the function

$$\kappa(a) = \frac{p(a)q(a)+1}{4} \tag{3.3.15}$$

is integer-valued.

The nine local magnitude types The positive odd magnitude pairs satisfying (3.3.11), up to exchange, give nine types. For each local frequency a, equation (3.3.9) gives

$$|\mathcal{W}_f(\iota(a)+L)| = 2|p(a)-q(a)|, \qquad |\mathcal{W}_f(\iota(a))| = 2|p(a)+q(a)|.$$

The last two columns below record these odd-half and even-half magnitudes.

Type	$\{\lvert p\rvert, \lvert q\rvert\}$	odd	even	κ	$\kappa \bmod 2$
A	$\{9,1\}$	20	16	-2	0
B	$\{7,3\}$	20	8	-5	1
C	$\{7,1\}$	12	16	2	0
D	$\{5,5\}$	20	0	-6	0
E	$\{5,3\}$	4	16	4	0
F	$\{5,1\}$	12	8	-1	1
G	$\{3,3\}$	12	0	-2	0
H	$\{3,1\}$	4	8	1	1
I	$\{1,1\}$	4	0	0	0

For the fixed direction L, let $N_{\mathsf{T}}(L)$ be the number of local frequencies of type T. When no ambiguity is possible, we write N_{T}. With rows indexed by the odd-half magnitudes $20, 12, 4$ and columns indexed by the even-half magnitudes $16, 8, 0$, define the local contingency table

$$\mathbf{N}_L = \begin{pmatrix} N_{\mathsf{A}} & N_{\mathsf{B}} & N_{\mathsf{D}} \\ N_{\mathsf{C}} & N_{\mathsf{F}} & N_{\mathsf{G}} \\ N_{\mathsf{E}} & N_{\mathsf{H}} & N_{\mathsf{I}} \end{pmatrix}. \tag{3.3.16}$$

At local frequency $a = 0$, the even endpoint is the global zero frequency. The distinguished type is therefore $\mathsf{D}, \mathsf{G}, \mathsf{I}$ for slice weights $59, 61, 63$, respectively.

Necessary conditions for one local table This section derives every condition used by the local enumerator. Parseval for the two seven-variable slices gives

$$\sum_{a\in V_7} p(a)^2 = \sum_{a\in V_7} q(a)^2 = 4096.$$

The combined squared contributions $r^2 + s^2$ for types $\mathsf{A}, \dots, \mathsf{I}$ are

$$82, 58, 50, 50, 34, 26, 18, 10, 2.$$

Consequently every local table satisfies

$$\sum_{\mathsf{T}} N_{\mathsf{T}} = 128, \tag{3.3.17}$$

$$82N_{\mathsf{A}} + 58N_{\mathsf{B}} + 50N_{\mathsf{C}} + 50N_{\mathsf{D}} + 34N_{\mathsf{E}} + 26N_{\mathsf{F}} + 18N_{\mathsf{G}} + 10N_{\mathsf{H}} + 2N_{\mathsf{I}} = 8192. \tag{3.3.18}$$

Fourth-moment congruence We first prove the congruence used for the fourth moments of the two slices.

Lemma 3.3.1. *If $u : V_7 \to \mathbb{F}_2$ has odd weight, then $\mathrm{AC}_u(t)/4$ is odd for every nonzero $t \in V_7$.*

Proof. Choose $T \in \mathrm{GL}(V_7)$ with $T\mathbf{e}_6 = t$, and define $\widetilde{u} = u \circ T$. The change of variables preserves the weight, so $\widetilde{u}$ has odd weight. It also preserves the coefficient of the degree-seven monomial. Indeed, the coefficient of $x_0x_1\cdots x_6$ in the ANF of a Boolean function is its weight modulo two, so this coefficient is one for u. After composition with T, the top homogeneous component is multiplied by $\det(T) = 1$ in $\mathbb{F}_2$, so it is also one for $\widetilde{u}$. Moreover,

$$\mathrm{AC}_{\widetilde{u}}(\mathbf{e}_6) = \mathrm{AC}_u(t).$$

The derivative

$$\Delta_{\mathbf{e}_6}\widetilde{u}(x) = \widetilde{u}(x) + \widetilde{u}(x + \mathbf{e}_6)$$

is invariant under $x \mapsto x + \mathbf{e}_6$, and is therefore a function of $x_0, \ldots, x_5$ alone. Since the coefficient of $x_0\cdots x_6$ in $\widetilde{u}$ is one, the coefficient of $x_0\cdots x_5$ in this derivative is one; all lower-degree monomials in $\widetilde{u}$ contribute degree at most five. The induced six-variable function consequently has odd weight, say r_∂. Indeed, the coefficient of the full monomial of a Boolean function on six variables is the XOR of its truth-table values, hence its weight modulo two. The derivative on V_7 has weight $2r_\partial$, and hence

$$\mathrm{AC}_u(t) = \mathrm{AC}_{\widetilde{u}}(\mathbf{e}_6) = 128 - 2(2r_\partial) = 4(32 - r_\partial).$$

Since r_∂ is odd, $\mathrm{AC}_u(t)/4$ is odd.

Apply (3.3.3) to $\mathcal{W}_u/2$. Since

$$\mathcal{H}\left((\mathcal{W}_u/2)^2\right)(t) = 32\,\mathrm{AC}_u(t),$$

Hadamard Parseval gives

$$\begin{aligned}\sum_a \left(\frac{\mathcal{W}_u(a)}{2}\right)^4 &= \frac{1}{128}\sum_t \left[\mathcal{H}\left((\mathcal{W}_u/2)^2\right)(t)\right]^2 \\ &= \frac{1}{128}\sum_t \big(32\,\mathrm{AC}_u(t)\big)^2 \\ &= 128\sum_t \left(\frac{\mathrm{AC}_u(t)}{4}\right)^2.\end{aligned}$$

At $t = 0$, the term in parentheses is 32, whose square is zero modulo eight. By Lemma 3.3.1, each of the 127 nonzero shifts contributes one modulo eight. Multiplication by 128 gives $128 \cdot 127 \equiv 896 \pmod{1024}$. Thus

$$\sum_a \left(\frac{\mathcal{W}_u(a)}{2}\right)^4 \equiv 896 \pmod{1024}. \tag{3.3.19}$$

Adding (3.3.19) for g and h gives $2 \cdot 896 = 1792 \equiv 768 \pmod{1024}$. For each type T, let $\{r_{\mathsf{T}}, s_{\mathsf{T}}\}$ denote its magnitude pair in the nine-type table. We obtain

$$\sum_{\mathsf{T}} N_{\mathsf{T}} \big(r_{\mathsf{T}}^4 + s_{\mathsf{T}}^4\big) \equiv 768 \pmod{1024}, \tag{3.3.20}$$

A Hadamard-divisibility lemma

Lemma 3.3.2. *Let $\varphi : V_7 \to \mathbb{Z}$. If $(\mathcal{H}\varphi)(t) \in 32\mathbb{Z}$ for every $t \in V_7$, then $\varphi \bmod 2 \in \mathrm{RM}(2,7)$.*

Proof. Let $I_7 = \{0, \dots, 6\}$. For $U \subseteq I_7$, let $\boldsymbol{\chi}_U \in V_7$ be its incidence vector. Thus $(\boldsymbol{\chi}_U)_j = 1$ exactly when $j \in U$. For $J \subseteq I_7$, expand

$$\prod_{j \in J} x_j = 2^{-|J|} \prod_{j \in J} (1 - (-1)^{x_j}).$$

After summing against φ, this gives

$$\sum_{x} \varphi(x) \prod_{j \in J} x_j = 2^{-|J|} \sum_{U \subseteq J} (-1)^{|U|} (\mathcal{H}\varphi)(\boldsymbol{\chi}_U).$$

Every term in the last sum is divisible by 32. After division by $2^{|J|}$, the right side is divisible by $2^{5-|J|}$. It is therefore even when $|J| \le 4$. Hence $\varphi \bmod 2$ is orthogonal to every degree-at-most-four monomial. The Reed–Muller duality $\mathrm{RM}(4,7)^{\perp} = \mathrm{RM}(2,7)$ proves the claim.

The possible weights of a Boolean function in $\mathrm{RM}(2,7)$ are

$$\mathcal{Q} = \{0, 32, 48, 56, 64, 72, 80, 96, 128\}. \tag{3.3.21}$$

For completeness, let the alternating polar form have rank $2s$, $0 \le s \le 3$. If the linear part is nonzero on the radical, summation over a radical coordinate shows that the function is balanced. Otherwise the weights are $64 \pm 2^{6-s}$. These values give exactly (3.3.21).

The product first bit Equation (3.3.4), with $p = \mathcal{W}_g/2$ and $q = \mathcal{W}_h/2$, gives

$$\mathcal{H}(pq)(t) = 32\,\mathrm{CC}_{g,h}(t). \tag{3.3.22}$$

Let $D_t = \mathrm{wt}(g + h(\cdot + t))$. Both slice weights are odd, so D_t is even. Therefore

$$\mathrm{CC}_{g,h}(t) = 128 - 2D_t \equiv 0 \pmod 4.$$

Let δ_0 denote the point mass at the zero frequency: $\delta_0(a) = 1$ for $a = 0$ and $\delta_0(a) = 0$ otherwise. Then $\mathcal{H}\mathbf{1} = 128\delta_0$. Using (3.3.22), we obtain

$$(\mathcal{H}\kappa)(t) = 8\,\mathrm{CC}_{g,h}(t) + 32\delta_0(t).$$

This is divisible by 32. Lemma 3.3.2 shows $\kappa \bmod 2 \in \mathrm{RM}(2,7)$.

The κ column is obtained from $\kappa = (pq+1)/4$. It is unchanged when the two magnitudes are exchanged, and the displayed values are valid for all $d \in \{1,3,5\}$, because the residue normalization fixes the signs of p and q. The odd support of κ is therefore the union of types $\mathsf{B}, \mathsf{F}, \mathsf{H}$. Two necessary conditions follow:

$$N_{\mathsf{B}} + N_{\mathsf{F}} + N_{\mathsf{H}} \in \mathcal{Q}, \tag{3.3.23}$$

$$-2N_{\mathsf{A}} - 5N_{\mathsf{B}} + 2N_{\mathsf{C}} - 6N_{\mathsf{D}} + 4N_{\mathsf{E}} - N_{\mathsf{F}} - 2N_{\mathsf{G}} + N_{\mathsf{H}} \equiv 0 \pmod{32}. \tag{3.3.24}$$

The second condition is $(\mathcal{H}\kappa)(0) \equiv 0 \pmod{32}$.

The individual-slice first bit Set

$$u_p(a) = \frac{p(a)^2 - 1}{8}.$$

This is integral because $p(a)$ is odd. From (3.3.3),

$$(\mathcal{H}u_p)(t) = 4\,\mathrm{AC}_g(t) - 16\delta_0(t).$$

Let $v_p = u_p - 16\delta_0$. Since $\mathcal{H}\delta_0 = \mathbf{1}$, where $\mathbf{1}$ is the constant-one function, we have

$$\mathcal{H}v_p = 4\,\mathrm{AC}_g - 16\delta_0 - 16\mathbf{1}.$$

For $t = 0$,

$$(\mathcal{H}v_p)(0) = 4 \cdot 128 - 16 - 16 = 480.$$

For $t \neq 0$, Lemma 3.3.1 gives

$$(\mathcal{H}v_p)(t) = 4\,\mathrm{AC}_g(t) - 16 \equiv 0 \pmod{32}.$$

Hence $v_p \bmod 2 = u_p \bmod 2$ belongs to $\mathrm{RM}(2,7)$ by Lemma 3.3.2. The residue table

$\lvert p\rvert$	1	3	5	7	9
$(p^2-1)/8$	0	1	3	6	10

shows that the support of $u_p \bmod 2$ is $\{a : |p(a)| \in \{3,5\}\}$. Its cardinality lies in $\mathcal{Q}$. The same conclusion holds for q.

Unordered counts and orientations Eliminating N_{B} from (3.3.17) and (3.3.18) gives

$$N_{\mathsf{C}} + N_{\mathsf{D}} + 3N_{\mathsf{E}} + 4N_{\mathsf{F}} + 5N_{\mathsf{G}} + 6N_{\mathsf{H}} + 7N_{\mathsf{I}} = 3N_{\mathsf{A}} - 96. \tag{3.3.25}$$

For slice weights $59, 61, 63$, respectively, the count of the distinguished type $\mathsf{D}, \mathsf{G}, \mathsf{I}$ must be positive. The local enumerator generates every nonnegative solution of (3.3.17) and (3.3.25), and retains only those satisfying (3.3.20), (3.3.23), and (3.3.24).

The unordered type does not specify which slice receives the larger magnitude. For $\mathsf{T} \in \{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{E}, \mathsf{F}, \mathsf{H}\}$, define $N_{\mathsf{T}}^{p>q}$ and $N_{\mathsf{T}}^{q>p}$ as the numbers of copies for which the larger magnitude occurs in p and in q, respectively. Thus

$$\delta_{\mathsf{T}} = N_{\mathsf{T}}^{p>q} - N_{\mathsf{T}}^{q>p}, \qquad N_{\mathsf{T}}^{p>q} + N_{\mathsf{T}}^{q>p} = N_{\mathsf{T}}. \tag{3.3.26}$$

Each orientation determines one value in the set

$$\delta_{\mathsf{T}} \in \{-N_{\mathsf{T}}, -N_{\mathsf{T}}+2, \dots, N_{\mathsf{T}}\}.$$

Conversely, every value in this set is attained by an orientation of the N_{T} copies.

For every type T, let $(r_{\mathsf{T}}, s_{\mathsf{T}})$ be its magnitude pair ordered so that $r_{\mathsf{T}} \geq s_{\mathsf{T}}$. Define the residue-one signed magnitude function

$$\tau(m) = (-1)^{(m-1)/2} m \quad \text{for } m \in \{1,3,5,7,9\}. \tag{3.3.27}$$

Thus $\tau(1), \tau(3), \tau(5), \tau(7), \tau(9)$ are $1, -3, 5, -7, 9$, respectively. The elementary differences $r^2 - s^2$, $r^4 - s^4$, and $\tau(r) - \tau(s)$, in the order $\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{E}, \mathsf{F}, \mathsf{H}$, are

T	A	B	C	E	F	H
$r^2 - s^2$	80	40	48	16	24	8
$r^4 - s^4$	6560	2320	2400	544	624	80
$\tau(r) - \tau(s)$	8	-4	-8	8	4	-4

The first row gives the difference of the two Parseval identities:

$$80\delta_{\mathsf{A}} + 40\delta_{\mathsf{B}} + 48\delta_{\mathsf{C}} + 16\delta_{\mathsf{E}} + 24\delta_{\mathsf{F}} + 8\delta_{\mathsf{H}} = 0. \tag{3.3.28}$$

Let

$$F_4 = \sum_{\mathsf{T}} N_{\mathsf{T}} (r_{\mathsf{T}}^4 + s_{\mathsf{T}}^4).$$

Define

$$D_p = \sum_{a \in V_7} p(a)^4, \qquad D_q = \sum_{a \in V_7} q(a)^4, \qquad \Delta_4 = D_p - D_q.$$

Then $D_p + D_q = F_4$, while the second row of the table gives Δ_4 as the left-hand side below. The individual congruences $D_p \equiv D_q \equiv 896 \pmod{1024}$ mean that there are integers r_p, r_q such that

$$D_p = 896 + 1024 r_p, \qquad D_q = 896 + 1024 r_q.$$

Consequently,

$$F_4 = 1792 + 1024(r_p + r_q), \qquad \Delta_4 = 1024(r_p - r_q).$$

Because $r_p - r_q \equiv r_p + r_q \pmod 2$, we obtain

$$\Delta_4 \equiv F_4 - 1792 \pmod{2048}.$$

The combined condition (3.3.20) gives $F_4 \equiv 768 \pmod{1024}$, so $F_4 - 1792$ is a multiple of 1024. Therefore $F_4 - 1792 \equiv 1792 - F_4 \pmod{2048}$. Together with the combined condition, the displayed congruence is equivalent to the two separate fourth-moment congruences.

$$6560\delta_{\mathsf{A}} + 2320\delta_{\mathsf{B}} + 2400\delta_{\mathsf{C}} + 544\delta_{\mathsf{E}} + 624\delta_{\mathsf{F}} + 80\delta_{\mathsf{H}} \equiv 1792 - F_4 \pmod{2048}. \tag{3.3.29}$$

Put

$$S_0 = \sum_{\mathsf{T}} N_{\mathsf{T}}\big(\tau(r_{\mathsf{T}}) + \tau(s_{\mathsf{T}})\big)$$

and

$$\Delta_\tau = 8\delta_{\mathsf{A}} - 4\delta_{\mathsf{B}} - 8\delta_{\mathsf{C}} + 8\delta_{\mathsf{E}} + 4\delta_{\mathsf{F}} - 4\delta_{\mathsf{H}}.$$

Let

$$\eta_d = (-1)^{(d-1)/2} \in \{1, -1\}.$$

The fixed residues in (3.3.14) then give, for every local frequency a,

$$p(a) = \eta_d\, \tau(|p(a)|), \qquad q(a) = -\eta_d\, \tau(|q(a)|).$$

Define the two full slice sums

$$P = \sum_{a \in V_7} p(a), \qquad Q = -\sum_{a \in V_7} q(a). \tag{3.3.30}$$

At $x = 0$, Walsh inversion gives

$$P = 64(-1)^{g(0)}, \qquad Q = -64(-1)^{h(0)},$$

so $P, Q \in \{-64, 64\}$. The fixed residues determine how these sums depend on the orientations. Summing the larger and smaller entries according to the definition of δ_{T} gives

$$\sum_a \tau(|p(a)|) = \frac{S_0 + \Delta_\tau}{2}, \qquad \sum_a \tau(|q(a)|) = \frac{S_0 - \Delta_\tau}{2}.$$

Therefore

$$P = \eta_d \frac{S_0 + \Delta_\tau}{2}, \qquad Q = \eta_d \frac{S_0 - \Delta_\tau}{2}.$$

For $\eta_d = 1$, the four possible pairs (P, Q) give

(P, Q)	(S_0, Δ_τ)
$(64, 64)$	$(128, 0)$
$(64, -64)$	$(0, 128)$
$(-64, 64)$	$(0, -128)$
$(-64, -64)$	$(-128, 0)$

For $\eta_d = -1$, the signs in the second column are reversed. In both cases, Walsh inversion is therefore equivalent to

$$(S_0, \Delta_\tau) \in \{(-128, 0), (0, -128), (0, 128), (128, 0)\}. \tag{3.3.31}$$

Finally, define

$$U = N_{\mathsf{B}} + 2N_{\mathsf{D}} + 2N_{\mathsf{E}} + N_{\mathsf{F}} + 2N_{\mathsf{G}} + N_{\mathsf{H}}, \qquad \Delta_U = -\delta_{\mathsf{B}} + \delta_{\mathsf{F}} + \delta_{\mathsf{H}}.$$

The numbers of magnitude-3-or-5 coefficients in the two slices are

$$\frac{U + \Delta_U}{2} \in \mathcal{Q}, \qquad \frac{U - \Delta_U}{2} \in \mathcal{Q}. \tag{3.3.32}$$

Equations (3.3.28), (3.3.29), (3.3.31), and (3.3.32) are the exact orientation tests used in Algorithm 3.

Finite cutoff and local output Fix one of the two slices and denote its normalized Walsh spectrum by $\gamma : V_7 \to \mathbb{Z}$; thus $\gamma = p$ or $\gamma = q$. After the residue normalization, its values belong to $\{1, -3, 5, -7, 9\}$ or to the negative of this set. On the first set,

$$192\mathbf{1}_{\{\gamma=9\}} \le \gamma^2 + 10\gamma + 21,$$

and on the negative set,

$$192\mathbf{1}_{\{\gamma=-9\}} \le \gamma^2 - 10\gamma + 21.$$

For one slice, Parseval gives $\sum_\alpha \gamma(\alpha)^2 = 4096$, while Walsh inversion at zero gives $|\sum_\alpha \gamma(\alpha)| = 64$. Let r_9 denote the number of coefficients of this slice whose absolute value is nine. Summing the applicable inequality over 128 frequencies yields

$$192r_9 \le 4096 + 10 \cdot 64 + 21 \cdot 128 = 7424.$$

Thus each slice has at most 38 coefficients of magnitude nine. Type A contains exactly one such coefficient across p, q, and no other type contains one. Therefore

$$N_{\mathsf{A}} \le 76. \tag{3.3.33}$$

The exact enumeration checks every N_{A} from 0 through 76. The first feasible values are 40, 42, and 45 for slice weights 59, 61, and 63. At those first layers, the surviving unordered count vectors, in the order

$$(N_{\mathsf{A}}, N_{\mathsf{B}}, N_{\mathsf{C}}, N_{\mathsf{D}}, N_{\mathsf{E}}, N_{\mathsf{F}}, N_{\mathsf{G}}, N_{\mathsf{H}}, N_{\mathsf{I}}),$$

are

$$\begin{aligned}
w = 59: &\quad (40, 64, 0, 24, 0, 0, 0, 0, 0), \quad (40, 64, 8, 16, 0, 0, 0, 0, 0),\\
w = 61: &\quad (42, 64, 6, 14, 0, 0, 2, 0, 0),\\
w = 63: &\quad (45, 56, 9, 16, 0, 0, 0, 0, 2), \quad (45, 61, 10, 7, 1, 3, 0, 0, 1).
\end{aligned}$$

The complete local families are reconstructed by Algorithm 3.

Global contingency constraints The local table for a direction $L \notin K$ records the joint magnitudes of the even and odd endpoints of the pairs $\{\iota(a), \iota(a) + L\}$, where $\iota : V_7 \to K$ is the fixed isomorphism used above. We now aggregate these local tables over all 128 odd directions.

Define the global magnitude classes

$$O_i = \{b \in V_8 \setminus K : |\mathcal{W}_f(b)| = i\}, \qquad i \in \{20, 12, 4\},$$

and

$$E_j = \{b \in K : |\mathcal{W}_f(b)| = j\}, \qquad j \in \{16, 8, 0\}.$$

For $L \notin K$, define the cell count

$$X_{ij}(L) = |\{e \in E_j : e + L \in O_i\}|. \tag{3.3.34}$$

Under the coordinates $\iota : V_7 \to K$, the nine values $X_{ij}(L)$ are the entries of the local table $\mathbf{N}_L$ in (3.3.16).

Lemma 3.3.3 (Global cell double count). *For every $i \in \{20, 12, 4\}$ and $j \in \{16, 8, 0\}$,*

$$\sum_{L \notin K} X_{ij}(L) = |O_i|\,|E_j|. \tag{3.3.35}$$

Proof. Count ordered pairs $(e, o) \in E_j \times O_i$. The sum $L = e+o$ is nonzero modulo K, hence $L \notin K$, and $e+L = o$. Thus each pair contributes once to the left-hand side of (3.3.35). Conversely, every element counted by $X_{ij}(L)$ determines exactly one ordered pair $(e, e+L) \in E_j \times O_i$. This proves the equality.

Write a global profile as

$$\pi = (o_{20}, o_{12}, o_4; e_{16}, e_8, e_0), \tag{3.3.36}$$

where

$$o_i = |O_i| \quad (i \in \{20, 12, 4\}), \qquad e_j = |E_j| \quad (j \in \{16, 8, 0\}).$$

The distinct letters o and e record the odd and even halves, respectively. Let $\mathcal{T}_w(\pi)$ be the finite set of admissible local type-count vectors for slice class w, with complete row and column margins equal to π. For $\boldsymbol{\nu} \in \mathcal{T}_w(\pi)$, write

$$\boldsymbol{\nu} = \begin{pmatrix} \nu_{\mathsf{A}} & \nu_{\mathsf{B}} & \nu_{\mathsf{D}} \\ \nu_{\mathsf{C}} & \nu_{\mathsf{F}} & \nu_{\mathsf{G}} \\ \nu_{\mathsf{E}} & \nu_{\mathsf{H}} & \nu_{\mathsf{I}} \end{pmatrix}.$$

Introduce the nonnegative integer $\lambda_{w,\boldsymbol{\nu}}$ for the number of odd directions of slice class w whose local vector is $\boldsymbol{\nu}$. Let $\boldsymbol{\lambda}$ be the column vector formed by all these variables. The three direction counts are therefore

$$\sum_{\boldsymbol{\nu} \in \mathcal{T}_{59}(\pi)} \lambda_{59,\boldsymbol{\nu}} = o_{20}, \qquad \sum_{\boldsymbol{\nu} \in \mathcal{T}_{61}(\pi)} \lambda_{61,\boldsymbol{\nu}} = o_{12}, \qquad \sum_{\boldsymbol{\nu} \in \mathcal{T}_{63}(\pi)} \lambda_{63,\boldsymbol{\nu}} = o_4. \tag{3.3.37}$$

For each cell (i, j), the double count in Lemma 3.3.3 gives

$$\sum_{w \in \{59,61,63\}} \sum_{\boldsymbol{\nu} \in \mathcal{T}_w(\pi)} \lambda_{w,\boldsymbol{\nu}}\, \nu_{ij} = o_i e_j, \qquad i \in \{20, 12, 4\}, \quad j \in \{16, 8, 0\}, \tag{3.3.38}$$

where ν_{ij} denotes the entry in the row labelled by i and the column labelled by j. In the matrix below the nine cells are flattened in the row-major order

$$(20, 16), (20, 8), (20, 0), (12, 16), (12, 8), (12, 0), (4, 16), (4, 8), (4, 0).$$

Equivalently, when a local table is written as a vector, its nine entries are

$$(\nu_{\mathsf{A}}, \nu_{\mathsf{B}}, \nu_{\mathsf{D}}; \nu_{\mathsf{C}}, \nu_{\mathsf{F}}, \nu_{\mathsf{G}}; \nu_{\mathsf{E}}, \nu_{\mathsf{H}}, \nu_{\mathsf{I}})^{\mathsf{T}}.$$

The distinguished type at local frequency zero is already included in every $\boldsymbol{\nu}$, so no additional correction is made in (3.3.38).

Equations (3.3.37) and (3.3.38), together with $\lambda_{w,\boldsymbol{\nu}} \geq 0$, form a finite linear Diophantine system. This is a necessary-condition model. A local vector is retained when at least one orientation passes the local tests, and the variables do not impose orientation compatibility between different directions. Every genuine spectrum nevertheless gives a feasible solution of this relaxed system, so its infeasibility is sufficient for exclusion. In matrix notation write it as

$$\mathbf{M}_\pi \boldsymbol{\lambda} = \boldsymbol{b}_\pi, \qquad \boldsymbol{\lambda} \geq 0. \tag{3.3.39}$$

The first three rows of $\mathbf{M}_\pi$ are the direction indicators in (3.3.37); its remaining nine rows are the entries of the local vector $\boldsymbol{\nu}$. In the same order, the right-hand side is

$$\boldsymbol{b}_\pi = (o_{20}, o_{12}, o_4; o_{20}e_{16}, o_{20}e_8, o_{20}e_0, o_{12}e_{16}, o_{12}e_8, o_{12}e_0, o_4e_{16}, o_4e_8, o_4e_0)^{\mathsf{T}}. \tag{3.3.40}$$

Lemma 3.3.4 (Farkas exclusion certificate). *If $\boldsymbol{y} \in \mathbb{Q}^{12}$ satisfies*

$$\mathbf{M}_\pi^{\mathsf{T}} \boldsymbol{y} \geq 0 \quad \textit{componentwise}, \qquad \boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{y} = -1, \tag{3.3.41}$$

then the system (3.3.39) *has no nonnegative real solution.*

Proof. If $\boldsymbol{\lambda} \geq 0$ solved (3.3.39), then

$$0 \leq \boldsymbol{\lambda}^{\mathsf{T}} \mathbf{M}_\pi^{\mathsf{T}} \boldsymbol{y} = (\mathbf{M}_\pi \boldsymbol{\lambda})^{\mathsf{T}} \boldsymbol{y} = \boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{y} = -1,$$

a contradiction. Therefore no nonnegative integer solution exists either.

Proposition 3.3.1 (Exact global survivors). *The local enumeration has 96 distinct weight-59 margin profiles. Fourteen of them have a positive weight-61 or weight-63 margin but no local family in that required class. Of the remaining 82 profiles, 69 are excluded by certificates satisfying* (3.3.41)*. The exact thirteen survivors are listed in Table 2.*

Proof (Exact finite verification). Equation (3.3.6) guarantees a Walsh coefficient of magnitude 20, and (3.3.8) places every such coefficient in the odd half. Hence $o_{20} > 0$, so every global profile must occur as a weight-59 margin profile. The deterministic local dump has 96 such profiles. A direct lookup removes the fourteen profiles for which one of the required nonzero classes has no local family. For each of the remaining 82 profiles, the artifact records either a rational vector $\boldsymbol{y}$ satisfying (3.3.41) or no certificate. The verifier reconstructs every column of $\mathbf{M}_\pi$ from the local dump, reconstructs $\boldsymbol{b}_\pi$ from (3.3.40), and checks all inequalities and the exact equality using rational arithmetic. All 69 recorded certificates pass. The set-theoretic complement consists of exactly the thirteen profiles in Table 2. Thus the accounting is

$$96 = 14 + 69 + 13.$$

Table 2. The thirteen global profiles left by the exact contingency model.

ID	o_{20}	o_{12}	o_4	e_{16}	e_8	e_0
U1	105	13	10	70	56	2
U2	106	10	12	70	56	2
B1	112	8	8	60	64	4
B2	112	16	0	56	64	8
B3	113	5	10	60	64	4
B4	114	2	12	60	64	4
B5	114	10	4	56	64	8
B6	115	7	6	56	64	8
B7	116	4	8	56	64	8
B8	116	12	0	52	64	12
B9	117	1	10	56	64	8
B10	120	8	0	48	64	16
B11	128	0	0	40	64	24

The common quadratic and its orbits The parity of $\kappa = (pq+1)/4$ is supported on types $\mathsf{B}, \mathsf{F}, \mathsf{H}$, exactly the column of the local table whose even-half magnitude is eight. For the fixed isomorphism $\iota : V_7 \to K$, the even endpoint at local index a is $\iota(a)$. Hence the global function

$$R(a) = 1 \quad \Longleftrightarrow \quad |\mathcal{W}_f(\iota(a))| = 8 \tag{3.3.42}$$

is independent of the odd direction. The common translation used to impose the residues (3.3.14) multiplies p and q by the same character, so it leaves pq, κ, and R unchanged. Also $R(0) = 0$, because the local zero frequency has type D, G, or I. The type table and (3.3.15) show that $R = \kappa \bmod 2$. The product-first-bit argument therefore gives

$$R \in \mathrm{RM}(2,7), \qquad \mathrm{wt}(R) = |E_8| = e_8,$$

where the weight identity uses the bijection $\iota : V_7 \to K$.

For a quadratic function $R : V_7 \to \mathbb{F}_2$ with $R(0) = 0$, define its polar form by

$$B_R(x,y) = R(x+y) + R(x) + R(y).$$

Its rank is even; write it as $2s$.

Lemma 3.3.5 (Quadratic orbit classification). *If* $\mathrm{wt}(R) = 64$*, then* R *is linearly equivalent to exactly one of*

$$R_s(x) = Q_s(x) + x_{2s}, \qquad Q_s(x) = \sum_{i=0}^{s-1} x_{2i}x_{2i+1}, \qquad s \in \{0,1,2,3\}. \tag{3.3.43}$$

If $\mathrm{wt}(R) = 56$*, then* R *is linearly equivalent to*

$$Q_3(x) = x_0x_1 + x_2x_3 + x_4x_5, \tag{3.3.44}$$

independent of x_6*.*

Proof. If $R' = R \circ T$ for $T \in \mathrm{GL}(V_7)$, then

$$B_{R'}(x, y) = B_R(Tx, Ty).$$

Thus a linear equivalence preserves the polar rank. On the nondegenerate quotient, the induced map is symplectic. When R vanishes on the radical, it induces a nonsingular quadratic form on this quotient, whose Arf invariant is also preserved.

Choose a decomposition $V_7 = \mathcal{U} \oplus \mathcal{Z}$ such that B_R restricts to a direct sum of s hyperbolic planes on $\mathcal{U}$, while $\mathcal{Z}$ is the radical and has dimension $7 - 2s$. Write $x_{\mathcal{U}} \in \mathcal{U}$ and $x_{\mathcal{Z}} \in \mathcal{Z}$. In suitable coordinates,

$$R(x_{\mathcal{U}}, x_{\mathcal{Z}}) = Q_s(x_{\mathcal{U}}) + \ell_{\mathcal{U}}(x_{\mathcal{U}}) + \ell_{\mathcal{Z}}(x_{\mathcal{Z}}), \tag{3.3.45}$$

where $\ell_{\mathcal{U}} : \mathcal{U} \to \mathbb{F}_2$ and $\ell_{\mathcal{Z}} : \mathcal{Z} \to \mathbb{F}_2$ are linear forms. If $\ell_{\mathcal{Z}} \neq 0$, summing $(-1)^{R(x_{\mathcal{U}}, x_{\mathcal{Z}})}$ over any radical coordinate on which $\ell_{\mathcal{Z}}$ is nonzero gives zero. Hence R is balanced. Choose radical coordinates $z_0, z_1, \ldots$ with $\ell_{\mathcal{Z}}(x_{\mathcal{Z}}) = z_0$. The invertible linear shear

$$(x_{\mathcal{U}}, z_0, z_1, \ldots) \longmapsto (x_{\mathcal{U}}, z_0 + \ell_{\mathcal{U}}(x_{\mathcal{U}}), z_1, \ldots)$$

transforms (3.3.45) into $Q_s(x_{\mathcal{U}}) + z_0$. This proves the four balanced normal forms. Their polar ranks $2s$ are different, so the forms are pairwise inequivalent.

If $\ell_{\mathcal{Z}} = 0$, the character sum on the nondegenerate quotient has magnitude 2^s, repeated 2^{7-2s} times over the radical. Hence

$$\sum_x (-1)^{R(x)} = \pm 2^{7-s}, \qquad \mathrm{wt}(R) = 64 \mp 2^{6-s}.$$

Weight 56 forces $s = 3$ and the positive sign. The two equivalence classes on the nondegenerate quotient are distinguished by the Arf invariant. The positive-sign class is the all-hyperbolic form Q_3, because

$$\sum_{u,v \in \mathbb{F}_2} (-1)^{uv} = 2 \quad \Longrightarrow \quad \sum_{x \in V_6} (-1)^{Q_3(x)} = 8.$$

The remaining coordinate is radical and does not occur in (3.3.44).

Profiles B1–B11 have $e_8 = 64$. They therefore use one of the four forms in (3.3.43). Profiles U1 and U2 have $e_8 = 56$ and use (3.3.44). In each terminal case, we apply the linear equivalence to all local frequency coordinates before using the displayed normal form.

For the terminal calculations, the dual coordinate change must be made explicit. Suppose $T \in \mathrm{GL}(V_7)$ satisfies

$$R(Ta') = R_s(a').$$

Define

$$p'(a') = p(Ta'), \quad q'(a') = q(Ta'), \quad \iota' = \iota \circ T, \quad \Omega_i' = T^{-1}\Omega_i.$$

The corresponding primal coordinate is

$$x' = T^{\mathsf{T}}x, \qquad x = T^{-\mathsf{T}}x'.$$

Thus $a \cdot x = a' \cdot x'$ whenever $a = Ta'$, and

$$(\mathcal{H}p')(x') = (\mathcal{H}p)(T^{-\mathsf{T}}x').$$

Consequently the inverse-Walsh box inequalities and all character-sum certificates are preserved by relabeling. A terminal set is a set of primal coordinates, so it must be specified in the new coordinate x'; an old primal set $\mathcal{X}$ is represented by $T^{\mathsf{T}}\mathcal{X}$. We omit primes after this relabeling.

Closed inverse-Walsh certificates We now exclude the thirteen profiles without a solver. Fix an odd direction whose smaller slice has weight $w \in \{59, 61\}$, and set $d = 64 - w$. Thus $d = 5$ for $w = 59$ and $d = 3$ for $w = 61$. In the normalized spectra, $p(0) = d$ and $q(0) = -d$. Walsh inversion gives

$$(\mathcal{H}p)(x) = 64(-1)^{g(x)}, \qquad (\mathcal{H}q)(x) = 64(-1)^{h(x)}. \tag{3.3.46}$$

For this fixed direction, pull each global odd magnitude class back to the local frequency space:

$$\Omega_i^{(L)} = \{a \in V_7 : \iota(a) + L \in O_i\}, \qquad i \in \{20, 12, 4\}. \tag{3.3.47}$$

The map $a \mapsto \iota(a) + L$ is a bijection from V_7 to $V_8 \setminus K$. Hence $|\Omega_i^{(L)}| = |O_i| = o_i$. We suppress the superscript L below, because only one direction is used at a time. Define the sums with the zero frequency removed:

$$P_x = \sum_{a \neq 0} (-1)^{a \cdot x} p(a), \qquad Q_x = \sum_{a \neq 0} (-1)^{a \cdot x} q(a).$$

From (3.3.46) we obtain the exact intervals

$$-64 - d \leq P_x \leq 64 - d, \tag{3.3.48}$$

$$-64 + d \leq Q_x \leq 64 + d. \tag{3.3.49}$$

Let $\mathcal{X} \subseteq V_7$ and let $\epsilon_x \in \{\pm 1\}$ for each $x \in \mathcal{X}$. Put

$$\Theta(a) = \sum_{x \in \mathcal{X}} \epsilon_x (-1)^{a \cdot x}.$$

Choosing the upper endpoint of the P_x interval when $\epsilon_x = 1$, the lower endpoint when $\epsilon_x = -1$, and the opposite endpoints for Q_x, gives

$$\sum_{a \neq 0} \Theta(a)(p(a) - q(a)) \leq 128|\mathcal{X}| - 2d\Theta(0). \tag{3.3.50}$$

Indeed, the left side equals $\sum_{x \in \mathcal{X}} \epsilon_x (P_x - Q_x)$, and each summand is at most $128 - 2d\epsilon_x$. Summing these bounds proves (3.3.50).

Signed category identity Recall the signed-magnitude function τ from (3.3.27); it satisfies $\tau(m) \equiv 1 \pmod 4$. For a fixed local frequency a, put $r_a = |p(a)|$ and $s_a = |q(a)|$. The fixed residues in (3.3.14) imply

$$\begin{array}{r|cc} & p(a) & q(a) \\ \hline w = 59 & \tau(r_a) & -\tau(s_a) \\ w = 61 & -\tau(r_a) & \tau(s_a) \end{array}$$

Write $\tau(r_a) = 1 + 4j$ and $\tau(s_a) = 1 + 4k$, where $j, k \in \mathbb{Z}$. In both slice classes,

$$\kappa(a) = \frac{1 - \tau(r_a)\tau(s_a)}{4} \equiv j + k \pmod 2,$$

and

$$\frac{\tau(r_a) + \tau(s_a)}{2} = 1 + 2(j + k).$$

Therefore define

$$\sigma(a) = \begin{cases} 5, & a \in \Omega_{20}, \\ -3, & a \in \Omega_{12}, \\ 1, & a \in \Omega_4. \end{cases} \tag{3.3.51}$$

Since $R = \kappa \bmod 2$, direct substitution of the fifteen possible ordered magnitude pairs gives

$$\frac{1}{2}(-1)^{R(a)}(p(a) - q(a)) = \sigma(a) \qquad (w = 59), \tag{3.3.52}$$

$$\frac{1}{2}(-1)^{R(a)}(q(a) - p(a)) = \sigma(a) \qquad (w = 61). \tag{3.3.53}$$

The calculation is displayed in Appendix B.3; hence the sign in the middle row is fixed as -3, not $+3$.

Balanced quadratic profiles Suppose $R = R_s = Q_s + x_{2s}$. Define the terminal set

$$\mathcal{X}_s = \{x : x_{2s} = 1,\ x_j = 0 \text{ for } j > 2s\}. \tag{3.3.54}$$

It has size 2^{2s}. The identity

$$\sum_{u,v \in \mathbb{F}_2} (-1)^{uv + \alpha u + \beta v} = 2(-1)^{\alpha\beta}$$

factors over the s hyperbolic planes and gives

$$\Theta_s(a) := \sum_{x \in \mathcal{X}_s} (-1)^{R_s(x) + a \cdot x} = -2^s (-1)^{R_s(a)}. \tag{3.3.55}$$

For $w = 59$ choose $\epsilon = -1$, and for $w = 61$ choose $\epsilon = 1$. In the notation of (3.3.50), set $\epsilon_x = \epsilon(-1)^{R_s(x)}$ for $x \in \mathcal{X}_s$. Its character sum is

$$\Theta(a) = \sum_{x \in \mathcal{X}_s} \epsilon_x (-1)^{a \cdot x} = \epsilon \Theta_s(a).$$

Equations (3.3.52)–(3.3.53) and (3.3.55) give

$$\epsilon\Theta_s(a)(p(a)-q(a)) = 2^{s+1}\sigma(a). \tag{3.3.56}$$

Use (3.3.50) with this choice of $\mathcal{X}_s$ and ϵ_x. Since $\Theta(0) = -\epsilon 2^s$, its right-hand side is

$$128 \cdot 2^{2s} + \epsilon 2^{s+1} d.$$

At the removed zero frequency, the local type is D for $w = 59$ and G for $w = 61$. Thus the left-hand side of the summed pointwise identity is, respectively,

$$2^{s+1}(5o_{20} - 5 - 3o_{12} + o_4), \qquad 2^{s+1}(5o_{20} - 3o_{12} + 3 + o_4).$$

Inserting $d = 5$ and $d = 3$ cancels the zero-frequency correction in both cases and yields the necessary inequality

$$5o_{20} - 3o_{12} + o_4 \le 64 \cdot 2^s. \tag{3.3.57}$$

The eleven balanced profiles have $o_{20} > 0$, so a weight-59 direction is available. Their scores $5o_{20} - 3o_{12} + o_4$ are

ID	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11
score	544	512	560	576	544	560	576	544	592	576	640

For $s \le 2$, the right-hand side of (3.3.57) is at most 256, so all eleven profiles are excluded. For $s = 3$, it is 512, so only B2 remains.

Strict certificate for B2 For B2, R is linearly equivalent to $R_3 = Q_3 + x_6$. Use the dual coordinate change specified above, with $R(Ta') = R_3(a')$, and then omit primes. This is a bijective relabeling, so it preserves the inverse-Walsh box inequalities and all cardinalities. Since $o_{12} = 16$, choose a weight-61 direction and set

$$\mathcal{X}_{B_2} = \{x : x_6 = 1,\ Q_3(x_0, \dots, x_5) = 1\}. \tag{3.3.58}$$

The six-variable Gauss sum gives

$$\Theta_{B_2}(a) := \sum_{x \in \mathcal{X}_{B_2}} (-1)^{a \cdot x} = 32(-1)^{a_6}\mathbf{1}_{\{a_0 = \cdots = a_5 = 0\}} - 4(-1)^{R_3(a)}. \tag{3.3.59}$$

Consequently $\Theta_{B_2}(0) = 28$, $\Theta_{B_2}(\mathbf{e}_6) = -28$, and outside $\{0, \mathbf{e}_6\}$ it equals $-4(-1)^{R_3(a)}$. For $a \neq 0$, define

$$\mathcal{C}_{B_2}(a) = 320\mathbf{1}_{\{a=\mathbf{e}_6\}} - 40\mathbf{1}_{\Omega_{20}}(a) + 24\mathbf{1}_{\Omega_{12}}(a) + \Theta_{B_2}(a)(p(a) - q(a)). \tag{3.3.60}$$

For $a \notin \{0, \mathbf{e}_6\}$, (3.3.53) gives $\Theta_{B_2}(a)(p(a) - q(a)) = 8\sigma(a)$. Substitution into (3.3.60) gives $0, 0, 8$ on the odd-half magnitude classes $20, 12, 4$, respectively. At $a = \mathbf{e}_6$, we have $R_3(\mathbf{e}_6) = 1$. Equation (3.3.53) then gives $p(\mathbf{e}_6) - q(\mathbf{e}_6) = 2\sigma(\mathbf{e}_6)$. Since $\Theta_{B_2}(\mathbf{e}_6) = -28$, the three possible values are

$\lvert\mathcal{W}_f(\iota(\mathbf{e}_6) + L)\rvert$	20	12	4
$\mathcal{C}_{B_2}(\mathbf{e}_6)$	$320 - 40 - 56 \cdot 5$	$320 + 24 - 56(-3)$	$320 - 56 = 264$

and hence are $0, 512, 264$. Thus $\mathcal{C}_{B_2}$ is pointwise nonnegative. The box inequality (3.3.50), used without signs, gives

$$\sum_{a\neq 0} \Theta_{B_2}(a)(p(a)-q(a)) \le 128\cdot 28 - 2\cdot 3\cdot 28 = 28\cdot 122.$$

Removing the local index $a=0$ removes the odd endpoint L, which is one of the 16 entries in Ω_{12}. Thus summing (3.3.60) and using the B2 profile gives

$$0 \le \sum_{a\neq 0} \mathcal{C}_{B_2}(a) \le 320 - 40\cdot 112 + 24(16-1) + 28\cdot 122 = -384,$$

a contradiction.

The two weight-56 profiles For U1 and U2, R is linearly equivalent to Q_3, independent of x_6. Use the dual coordinate change specified above, with $R(Ta') = Q_3(a')$, and then omit primes. The box inequalities and profile counts are preserved. Since both have $o_{12} > 0$, choose a weight-61 direction and set

$$\mathcal{X}_U = \{x : x_6 = 0,\ Q_3(x_0,\dots,x_5) = 1\}. \tag{3.3.61}$$

Then $|\mathcal{X}_U| = 28$ and

$$\Theta_U(a) := \sum_{x\in\mathcal{X}_U} (-1)^{a\cdot x} = 32\mathbf{1}_{\{a_0=\cdots=a_5=0\}} - 4(-1)^{Q_3(a_0,\dots,a_5)}. \tag{3.3.62}$$

Thus $\Theta_U(0) = \Theta_U(\mathbf{e}_6) = 28$, and outside $\{0,\mathbf{e}_6\}$ it equals $-4(-1)^{R(a)}$. For $a\neq 0$, put

$$\mathcal{C}_U(a) = 320\mathbf{1}_{\{a=\mathbf{e}_6\}} - 40\mathbf{1}_{\Omega_{20}}(a) + 24\mathbf{1}_{\Omega_{12}}(a) - 8\mathbf{1}_{\Omega_4}(a) + \Theta_U(a)(p(a)-q(a)). \tag{3.3.63}$$

For $a\notin\{0,\mathbf{e}_6\}$, equation (3.3.53) gives $\Theta_U(a)(p(a)-q(a)) = 8\sigma(a)$, which cancels the three category terms in (3.3.63). At $a=\mathbf{e}_6$, we have $R(\mathbf{e}_6)=0$, so $p(\mathbf{e}_6)-q(\mathbf{e}_6) = -2\sigma(\mathbf{e}_6)$. Together with $\Theta_U(\mathbf{e}_6)=28$, this gives

$\lvert\mathcal{W}_f(\iota(\mathbf{e}_6)+L)\rvert$	20	12	4
$\mathcal{C}_U(\mathbf{e}_6)$	$320-40-56\cdot 5$	$320+24-56(-3)$	$320-8-56$

and hence the values $0, 512, 256$. Thus $\mathcal{C}_U$ is pointwise nonnegative. The same box bound and the omission of the zero-frequency O_{12} entry give

$$\begin{aligned} 0 &\le \sum_{a\neq 0} \mathcal{C}_U(a) \\ &\le 320 - 40o_{20} + 24(o_{12}-1) - 8o_4 + 28\cdot 122 \\ &= 8\big(464 - (5o_{20} - 3o_{12} + o_4)\big). \end{aligned} \tag{3.3.64}$$

For U1 and U2 the scores are 496 and 512, so the upper bounds are -256 and -384, respectively. This excludes both profiles.

Sharpness witness. Identify V_8 with $\{0, \dots, 255\}$ by the standard binary expansion. Read the following hexadecimal string from least significant bit to most significant bit within each digit:

```
67942f45a35eb17a9ce59c58697830ed
24ef1c648575ebc72203d46c888b74ee
```

The resulting truth table has weight 128. Its Walsh spectrum has maximum magnitude 24, so its nonlinearity is $128 - 24/2 = 116$. The complete Walsh value distribution is recorded in Appendix B.5.

Proof (Proof of Theorem 3.3.1). If a balanced eight-variable function had nonlinearity greater than 116, equations (3.3.5)–(3.3.6) would put it in the maximum-magnitude-20 case. The local enumeration and the global contingency equations reduce this case to the thirteen profiles in Table 2, by Proposition 3.3.1. The quadratic orbit classification and the certificates above exclude all thirteen. The displayed truth table attains 116, so the upper bound is sharp.

3.4 Affine derivatives of nonquadratic bent functions

Problem Surface Let $f : V_n \to \mathbb{F}_2$ be a Boolean function. We use $\Delta_a f$ for the derivative fixed at the beginning of this section. Carlet's list of open problems on bent and plateaued functions asks for the maximum possible number of distinct affine derivatives of a nonquadratic bent function in even dimension. The same question asks, equivalently, for the maximum dimension of the vector space of directions a for which $\Delta_a f$ is affine [10]. Carlet also records the standard lower-bound construction: a quadratic bent function in $n - 6$ variables direct-summed with a six-variable cubic Maiorana–McFarland bent function gives at least 2^{n-3} affine derivatives.

Open Question 3.4.1 (Carlet 13.3(14)). *For even n, determine the largest possible number of distinct affine derivatives of a nonquadratic bent n-variable Boolean function. Equivalently, determine the maximal dimension of the vector space of directions a such that $\Delta_a f$ is affine.*

We resolve this question for the full nontrivial bent range $n \geq 6$. The restriction is natural: bent Boolean functions exist only in even dimension, and nonquadratic bent functions first occur in dimension six. The result is not a classification of the extremal functions and does not address second-order-bent or all-balanced second-derivative questions.

Definitions All vector spaces are over $\mathbb{F}_2$, and all sums below are in $\mathbb{F}_2$. A Boolean function $f : V_n \to \mathbb{F}_2$ is *affine* if

$$f(x) = \ell(x) + c$$

for some linear form $\ell : V_n \to \mathbb{F}_2$ and some $c \in \mathbb{F}_2$. It is *quadratic* in this note if its algebraic normal form has degree at most two, equivalently if $f \in \mathrm{RM}(2, n)$.

Thus *nonquadratic* means that $f \notin \mathrm{RM}(2,n)$. The affine-derivative direction set is $\mathcal{A}(f) = \{a \in V_n : \Delta_a f \text{ is affine}\}$. It is a linear subspace of V_n. Indeed, $0 \in \mathcal{A}(f)$, and if $\Delta_a f$ and $\Delta_b f$ are affine, then

$$\Delta_{a+b} f(x) = \Delta_a f(x+b) + \Delta_b f(x)$$

is affine; scalar closure is immediate over $\mathbb{F}_2$. For even n, f is *bent* if its Walsh transform, under the convention fixed at the beginning of this section, satisfies

$$\mathcal{W}_f(u)^2 = 2^n \qquad \text{for all } u \in V_n.$$

We use the standard derivative characterization: f is bent if and only if $\Delta_a f$ is balanced for every $a \neq 0$.

The Upper Bound

Lemma 3.4.1 (Large affine-derivative space forces quadraticity). *Let $f : V_n \to \mathbb{F}_2$. If*

$$\dim \mathcal{A}(f) > n - 3,$$

then $f \in \mathrm{RM}(2,n)$.

Proof. Put $U = \mathcal{A}(f)$. The assumption says that U has codimension at most two. Let $e_1, \ldots, e_n$ be the coordinate basis of V_n, and fix three distinct coordinates i, j, k. The coordinate subspace

$$E_{ijk} = \langle e_i, e_j, e_k \rangle$$

has dimension three, so $E_{ijk} \cap U \neq \{0\}$. Choose $0 \neq a \in E_{ijk} \cap U$. Since $a \in U$, the derivative $\Delta_a f$ is affine, and hence every second derivative of $\Delta_a f$ is zero. Choose $b, c \in E_{ijk}$ such that a, b, c form a basis of E_{ijk}. Then

$$\Delta_b \Delta_c \Delta_a f = 0.$$

We next compare this derivative with the coordinate third derivative on the same three-dimensional subspace. For every $x \in V_n$, since (a, b, c) is a basis of E_{ijk},

$$\Delta_b \Delta_c \Delta_a f(x) = \sum_{\lambda \in \mathbb{F}_2^3} f(x + \lambda_1 a + \lambda_2 b + \lambda_3 c) = \sum_{y \in E_{ijk}} f(x+y).$$

The coordinate vectors e_i, e_j, e_k are another basis of E_{ijk}, and hence

$$\Delta_{e_i} \Delta_{e_j} \Delta_{e_k} f(x) = \sum_{y \in E_{ijk}} f(x+y) = \Delta_b \Delta_c \Delta_a f(x).$$

Therefore the coordinate third derivative

$$\Delta_{e_i} \Delta_{e_j} \Delta_{e_k} f$$

vanishes identically. Since this holds for every triple i, j, k, all coordinate third derivatives of f vanish. Writing the ANF of f as

$$f(x) = \sum_{S \subseteq [n]} c_S \prod_{\ell \in S} x_\ell,$$

we have, for distinct i, j, k,

$$\Delta_{e_i} \Delta_{e_j} \Delta_{e_k} f(x) = \sum_{S \supseteq \{i,j,k\}} c_S \prod_{\ell \in S \setminus \{i,j,k\}} x_\ell.$$

Thus no ANF coefficient c_S with $|S| \geq 3$ can be nonzero. Hence $f \in \mathrm{RM}(2, n)$.

Theorem 3.4.1 (Upper bound). *If $f : V_n \to \mathbb{F}_2$ is nonquadratic, then*

$$\dim \mathcal{A}(f) \leq n - 3.$$

Proof. This is the contrapositive of Lemma 3.4.1.

Sharpness Let $n \geq 6$ be even, and write $n = 2r + 6$. Split variables as

$$(x_1, \ldots, x_r, y_1, \ldots, y_r, u_1, u_2, u_3, v_1, v_2, v_3).$$

On the first $2r$ variables set

$$q(x, y) = \sum_{i=1}^{r} x_i y_i.$$

This is the standard quadratic bent form. On the last six variables set

$$h(u, v) = u_1 v_1 + u_2 v_2 + (u_1 u_2 + u_3) v_3.$$

This is Maiorana–McFarland: $h(u, v) = v \cdot \pi(u)$, where

$$\pi(u_1, u_2, u_3) = (u_1, u_2, u_3 + u_1 u_2)$$

is a permutation of $\mathbb{F}_2^3$. Hence h is bent, and the direct sum

$$f(x, y, u, v) = q(x, y) + h(u, v)$$

is bent.

Lemma 3.4.2 (The six-variable tail). *For*

$$h(u, v) = u_1 v_1 + u_2 v_2 + (u_1 u_2 + u_3) v_3$$

on $\mathbb{F}_2^6$, the space $\mathcal{A}(h)$ has dimension three. Moreover h is not quadratic.

Proof. Let a direction be written as $(s,t) \in \mathbb{F}_2^3 \times \mathbb{F}_2^3$. Since

$$h(u,v) = v \cdot \pi(u), \qquad \pi(u) = (u_1, u_2, u_3 + u_1u_2),$$

we have

$$\Delta_{(s,t)}h(u,v) = v \cdot (\pi(u+s) + \pi(u)) + t \cdot \pi(u+s).$$

The vector $\pi(u+s) + \pi(u)$ has first two coordinates s_1, s_2 and third coordinate

$$s_3 + s_1u_2 + s_2u_1 + s_1s_2.$$

If $s_1 \neq 0$ or $s_2 \neq 0$, then the coefficient of v_3 depends on u, producing a quadratic term in $\Delta_{(s,t)}h$. Thus an affine derivative requires $s_1 = s_2 = 0$. Under this condition the v-part is affine, and the remaining term $t \cdot \pi(u+s)$ is affine if and only if $t_3 = 0$, because the only nonlinear coordinate of π is $u_3 + u_1u_2$. Hence

$$\mathcal{A}(h) = \{(0,0,s_3,t_1,t_2,0) : s_3,t_1,t_2 \in \mathbb{F}_2\},$$

which has dimension three. The term $u_1u_2v_3$ shows that h has algebraic degree three, so it is not quadratic.

Lemma 3.4.3 (Direct-sum direction space). *For $f = q \oplus h$ as above,*

$$\dim \mathcal{A}(f) = n - 3.$$

Proof. Every derivative of the quadratic form q is affine. Since q and h are supported on disjoint variables,

$$\Delta_{(a,b)}f = \Delta_a q + \Delta_b h.$$

The first summand is affine for every a, while the second is affine exactly when $b \in \mathcal{A}(h)$. Therefore

$$\mathcal{A}(f) = \mathbb{F}_2^{n-6} \times \mathcal{A}(h),$$

and Lemma 3.4.2 gives

$$\dim \mathcal{A}(f) = (n-6) + 3 = n - 3.$$

Theorem 3.4.2 (Exact direction-space maximum). *For every even $n \geq 6$, the maximum of $\dim \mathcal{A}(f)$, over all nonquadratic bent functions $f : V_n \to \mathbb{F}_2$, is exactly $n-3$.*

Proof. The upper bound is Theorem 3.4.1. The direct-sum function constructed above is bent because both summands are bent and Walsh transforms factor under disjoint direct sums. It is nonquadratic because the six-variable tail has degree three. Lemma 3.4.3 gives $\dim \mathcal{A}(f) = n - 3$.

Distinct Affine Derivatives

Lemma 3.4.4 (Injectivity for bent functions). *Let $f : V_n \to \mathbb{F}_2$ be bent. If $a, b \in V_n$ and $\Delta_a f = \Delta_b f$, then $a = b$. Consequently, the number of distinct affine derivatives of f is*

$$|\mathcal{A}(f)| = 2^{\dim \mathcal{A}(f)}.$$

Proof. If $\Delta_a f = \Delta_b f$, then

$$f(x + a) + f(x) = f(x + b) + f(x)$$

for all x, so $f(y) = f(y + a + b)$ for all y. Thus $\Delta_{a+b} f$ is the zero function. If $a+b \neq 0$, this contradicts the derivative characterization of bentness, since every nonzero derivative of a bent function is balanced. Hence $a = b$. The displayed count follows because $\mathcal{A}(f)$ is a finite $\mathbb{F}_2$-vector space.

Theorem 3.4.3 (Resolution of Carlet Open Problem 13.3(14)). *For every even $n \geq 6$, the largest possible number of distinct affine derivatives of a nonquadratic bent n-variable Boolean function is*

$$2^{n-3}.$$

Proof. For every nonquadratic bent f, Theorem 3.4.1 gives $\dim \mathcal{A}(f) \leq n - 3$, and Lemma 3.4.4 converts this to at most 2^{n-3} distinct affine derivatives. The direct-sum construction in Theorem 3.4.2 is nonquadratic and bent and has $\dim \mathcal{A}(f) = n - 3$, so Lemma 3.4.4 gives exactly 2^{n-3} distinct affine derivatives.

3.5 Degree-drop hyperplanes

Top Forms and the Exact Questions Let $V = V_n$, with coordinate basis $e_1, \ldots, e_n$, and let

$$\mathcal{B}_n = \mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$$

be the Boolean polynomial ring. Every Boolean function has a unique algebraic normal form (ANF) [10]. If f has degree r, write

$$f = f^{[r]} + f^{[<r]}, \qquad f^{[r]} = \sum_{S \in \binom{[n]}{r}} c_S x_S, \qquad x_S = \prod_{i \in S} x_i, \tag{3.5.1}$$

where $c_S \in \mathbb{F}_2$ and $\deg(f^{[<r]}) < r$. Its top support is $A(f) = \{S \in \binom{[n]}{r} : c_S = 1\}$. For a nonempty support A, define

$$\mathrm{Var}(A) = \bigcup_{S \in A} S, \qquad \mathrm{Core}(A) = \bigcap_{S \in A} S.$$

Thus $i \in \mathrm{Var}(A)$ means that coordinate i is active, while $i \not\in \mathrm{Core}(A)$ means that at least one supported edge avoids i. The top component is naturally identified with the alternating form

$$\omega_f = \sum_{S=\{i_1<\cdots<i_r\}\in\binom{[n]}{r}} c_S e^*_{i_1} \wedge \cdots \wedge e^*_{i_r} \in \bigwedge^r V^*. \tag{3.5.2}$$

Although signs disappear over $\mathbb{F}_2$, the relation $u \wedge u = 0$ remains. Consequently, the standard exterior basis is indexed by the same squarefree monomials as the degree-r ANF. For an invertible affine map $\phi(x) = Mx + a$, the translation contributes only terms of degree below r, and

$$\omega_{(f\circ\phi)^{[r]}} = \bigwedge^r (M^*)\omega_f. \tag{3.5.3}$$

Indeed, the raw coefficient of a squarefree degree-r monomial is the permanent of the corresponding minor of M. Over $\mathbb{F}_2$, permanent and determinant coincide. Terms selecting a translation constant, or selecting a repeated Boolean variable, have degree below r. Equation (3.5.3) also shows that the top-form action is invertible. We write $f \sim_{r-1} g$ if some affine automorphism ϕ satisfies $\deg(f \circ \phi - g) < r$. This is the affine action on $\mathrm{RM}(r,n)/\mathrm{RM}(r-1,n)$ used in the source paper. Working in this quotient is essential: even a linear substitution can create lower-degree Boolean terms, as

$$x_1x_2\big|_{x_1\leftarrow x_1+x_2} = x_1x_2 + x_2.$$

For distinct coordinates i, t and an r-set S containing i and avoiding t, put $S - i + t = (S \setminus \{i\}) \cup \{t\}$. The two source conditions can now be written as follows:

$$\mathsf{P}_-(A): \quad \forall i \in \mathrm{Var}(A),\ \exists S_i \in A: \quad i \not\in S_i \quad \text{and} \quad \forall t \in S_i,\ S_i - t + i \not\in A; \tag{3.5.4}$$

$$\mathsf{P}_+(A): \quad \forall i \in \mathrm{Var}(A),\ \exists T_i \in A: \quad i \in T_i \quad \text{and} \quad \forall t \not\in T_i,\ T_i - i + t \not\in A. \tag{3.5.5}$$

Condition P_- is precisely published Theorem 5, and P_+ is the private-edge condition in published Corollary 8. An affine hyperplane $H \subseteq V$ is degree dropping for a degree-r function f if $\deg(f|_H) < r$. A nonzero $a \in V$ is a fast point if $\deg(\Delta_a f) < r - 1$. Both predicates depend only on $f^{[r]}$: lower-degree terms cannot create degree r after restriction or degree $r-1$ after differentiation.

Open Question 3.5.1 (Carlet–Feukoua–Sălăgean, Open Problem 1). *Is every Boolean function with no degree-drop hyperplane affine equivalent, modulo its lower-degree part, to a function whose top support satisfies* P_-*?*

Restriction and Derivative Kernels For $\omega \in \bigwedge^r V^*$, define the wedge and contraction maps

$$W_\omega : V^* \longrightarrow \bigwedge^{r+1} V^*, \qquad W_\omega(\ell) = \ell \wedge \omega, \tag{3.5.6}$$

$$C_\omega : V \longrightarrow \bigwedge^{r-1} V^*, \qquad C_\omega(a) = \iota_a \omega. \tag{3.5.7}$$

Lemma 3.5.1 (Restriction–wedge bridge). *Let $1 \le r < n$, let f have degree r, and let $0 \neq \ell \in V^*$. The following are equivalent:*

1. *one, equivalently both, of the parallel hyperplanes $\ell(x) = 0$ and $\ell(x) = 1$ is degree dropping for f;*
2. *$\ell \wedge \omega_f = 0$;*
3. *$\omega_f = \ell \wedge \eta$ for some $\eta \in \bigwedge^{r-1} V^*$.*

Consequently, f has no degree-drop affine hyperplane if and only if

$$\ker W_{\omega_f} = \{0\}. \tag{3.5.8}$$

Proof. Extend ℓ to a basis $y_1 = \ell, y_2, \ldots, y_n$ of V^*. There is a unique decomposition

$$\omega_f = y_1 \wedge \eta + \theta,$$

where $\eta \in \bigwedge^{r-1} \operatorname{span}(y_2, \ldots, y_n)$ and $\theta \in \bigwedge^r \operatorname{span}(y_2, \ldots, y_n)$. In the corresponding Boolean coordinates, restriction to $y_1 = b$ turns the first summand into a polynomial of degree at most $r-1$, for either $b \in \mathbb{F}_2$. The degree-r component on either parallel hyperplane is therefore exactly θ. Hence the restriction drops degree exactly when $\theta = 0$, which is equivalent to the third assertion. Moreover, $y_1 \wedge \omega_f = y_1 \wedge \theta$. Wedge multiplication by y_1 is injective on $\bigwedge^r \operatorname{span}(y_2, \ldots, y_n)$, since it sends distinct standard basis monomials to distinct basis monomials containing y_1. Thus the second assertion is also equivalent to $\theta = 0$. Varying over all nonzero ℓ proves (3.5.8).

Lemma 3.5.2 (Derivative–contraction bridge). *Let f have degree $r \ge 2$. For every $a \in V$, the degree-$(r-1)$ component of $\Delta_a f$ is the alternating form $\iota_a \omega_f$. Consequently, f has no fast point if and only if*

$$\ker C_{\omega_f} = \{0\}. \tag{3.5.9}$$

Proof. For a degree-r monomial x_S, direct expansion gives

$$\prod_{j \in S} (x_j + a_j) + \prod_{j \in S} x_j = \sum_{i \in S} a_i x_{S \setminus \{i\}} + \text{terms of degree at most } r-2.$$

The two degree-r terms cancel, and selecting at least two entries from a lowers the degree by at least two. Therefore, for every $(r-1)$-set T, the coefficient of x_T in the top component of $\Delta_a f$ is

$$\sum_{i \notin T} a_i c_{T \cup \{i\}}. \tag{3.5.10}$$

This is exactly the coefficient of e_T^* in $\iota_a\omega_f$. Signs are immaterial over $\mathbb{F}_2$. The lower-degree part of f contributes only degrees at most $r-2$ to the derivative. Hence $\deg(\Delta_a f) < r-1$ precisely when $\iota_a\omega_f = 0$. Excluding every nonzero a is equivalent to (3.5.9).

Lemma 3.5.3 (Covariance of the two kernels). *Let $M \in \mathrm{GL}(V)$ and $\omega' = \bigwedge^r(M^*)\omega$. Then*

$$\ker W_{\omega'} = M^*(\ker W_\omega), \tag{3.5.11}$$

$$\ker C_{\omega'} = M^{-1}(\ker C_\omega). \tag{3.5.12}$$

In particular, injectivity of either map is invariant under every invertible affine change of variables.

Proof. For $\ell \in V^*$, naturality of exterior powers gives

$$\ell \wedge \omega' = \bigwedge^{r+1}(M^*)\left(((M^*)^{-1}\ell) \wedge \omega\right).$$

The exterior power on the right is invertible, proving (3.5.11). For $a \in V$, direct evaluation on $r-1$ vectors gives

$$\iota_a\omega' = \bigwedge^{r-1}(M^*)\big(\iota_{Ma}\omega\big).$$

Again the exterior power is invertible, so the left side vanishes exactly when $Ma \in \ker C_\omega$, which proves (3.5.12). Translations do not change the top form by (3.5.3).

The two kernel conditions have elementary support consequences that will be used after normalization.

Lemma 3.5.4 (Kernel injectivity removes the support obstructions). *Let $A = \mathrm{supp}(\omega)$.*

1. *If $\ker W_\omega = \{0\}$, then $\mathrm{Core}(A) = \varnothing$; equivalently, every coordinate is avoided by some edge of A.*
2. *If $\ker C_\omega = \{0\}$, then $\mathrm{Var}(A) = [n]$; equivalently, every coordinate is active.*

Proof. If every supported edge contained i, then $e_i^* \wedge \omega = 0$, contradicting wedge injectivity. This proves the first assertion. If no supported edge contained i, then $\iota_{e_i}\omega = 0$, contradicting contraction injectivity. This proves the second.

The coefficient formulas also recover the two published sufficient directions, with the ambient guard made explicit on the fast-point side.

Lemma 3.5.5 (Private-edge certificates). *Let $0 \neq \omega \in \bigwedge^r V^*$, with support A.*

1. *If $\mathsf{P}_-(A)$ holds, then $\ker W_\omega = \{0\}$.*

2. If $\mathsf{P}_+(A)$ *holds and* $\mathrm{Var}(A) = [n]$*, then* $\ker C_\omega = \{0\}$.

Proof. First observe that $\mathsf{P}_-(A)$ supplies a private avoiding edge for every ambient coordinate. This is part of its definition for active coordinates. If $i \notin \mathrm{Var}(A)$, choose any $S_i \in A$. Then $i \notin S_i$, and every replacement $S_i - t + i$ contains the inactive coordinate i, so it is absent from A. Now let $0 \neq \ell = \sum_i \lambda_i e_i^*$, and choose i with $\lambda_i = 1$. Let S_i be the private avoiding edge just described and put $U = S_i \cup \{i\}$. The coefficient of e_U^* in $\ell \wedge \omega$ is

$$\sum_{j \in U} \lambda_j c_{U \setminus \{j\}} = \lambda_i c_{S_i} = 1,$$

because every term indexed by $j \in S_i$ contains the forbidden replacement $S_i - j + i$. Hence $W_\omega(\ell) \neq 0$, proving the first assertion. For the second assertion, let $0 \neq a \in V$ and choose i with $a_i = 1$. The ambient guard makes i active, so $\mathsf{P}_+(A)$ supplies a private edge T_i containing i. Put $T = T_i \setminus \{i\}$. Equation (3.5.10) gives

$$[e_T^*](\iota_a \omega) = \sum_{j \notin T} a_j c_{T \cup \{j\}} = a_i c_{T_i} = 1,$$

because every term with $j \notin T_i$ contains the forbidden replacement $T_i - i + j$. Thus $C_\omega(a) \neq 0$, proving the second assertion.

A Two-Sided Lexicographic Normal Form Put $\mathcal{E}_{n,r} = \binom{[n]}{r}$. Order its edges by their characteristic vectors

$$(\mathbf{1}_{1 \in S}, \ldots, \mathbf{1}_{n \in S})$$

lexicographically, with $0 < 1$. Enumerate the edges in this order and compare supports $A \subseteq \mathcal{E}_{n,r}$ by their incidence vectors, again lexicographically with $0 < 1$. Thus, at the first edge where two supports differ, the support omitting that edge is the smaller one. For distinct $a, b \in [n]$, define

$$\mathcal{D}_{a,b} = \{S \in \mathcal{E}_{n,r} : a \in S,\ b \notin S\}, \tag{3.5.13}$$

$$q_{a,b}(S) = S - a + b \qquad (S \in \mathcal{D}_{a,b}). \tag{3.5.14}$$

The map $q_{a,b}$ is a bijection from $\mathcal{D}_{a,b}$ to $\mathcal{D}_{b,a}$, with inverse $q_{b,a}$.

Lemma 3.5.6 (Replacement preserves edge order). *For* $S, T \in \mathcal{D}_{a,b}$*,*

$$S < T \quad \Longleftrightarrow \quad q_{a,b}(S) < q_{a,b}(T).$$

Proof. The characteristic vectors of S and T have the same bits at both a and b, namely 1 and 0. Their first differing coordinate is therefore outside $\{a, b\}$. Replacement changes the two common bits in both vectors and leaves all other bits fixed, so the first difference and its orientation are unchanged.

Let $\tau_{a,b}$ be the elementary transvection

$$x_a \longleftarrow x_a + x_b, \qquad x_j \longleftarrow x_j \quad (j \neq a). \tag{3.5.15}$$

It is an involution and hence belongs to $\mathrm{GL}(n,2)$. Its action on top forms will be written explicitly as

$$\tau_{a,b} \cdot \omega := \bigwedge^{r} (\tau_{a,b}^{*})\omega. \tag{3.5.16}$$

Lemma 3.5.7 (Top coefficients under a transvection). *Let*

$$\omega = \sum_{S \in \mathcal{E}_{n,r}} c_S e_S^*, \qquad \tau_{a,b} \cdot \omega = \sum_{T \in \mathcal{E}_{n,r}} c'_T e_T^*.$$

Then

$$c'_T = \begin{cases} c_T + c_{T-b+a}, & b \in T \text{ and } a \notin T, \\ c_T, & \text{otherwise.} \end{cases} \tag{3.5.17}$$

If $A = \mathrm{supp}(\omega)$, the positions whose incidence bits change are exactly

$$q_{a,b}(A \cap \mathcal{D}_{a,b}). \tag{3.5.18}$$

Proof. Consider a source monomial x_S. If $a \notin S$, it is unchanged. If $a \in S$ and $b \notin S$, then

$$x_S \longmapsto x_S + x_{S-a+b}.$$

If $a, b \in S$, write $S = \{a, b\} \cup R$. Boolean reduction gives

$$x_a x_b x_R \longmapsto (x_a + x_b) x_b x_R = x_a x_b x_R + x_b x_R,$$

and the additional term has degree $r-1$. Thus a target coefficient can receive an additional degree-r contribution only when $T \in \mathcal{D}_{b,a}$, and then its unique source is $T - b + a$. This is (3.5.17). Over $\mathbb{F}_2$, such a target toggles exactly when its source coefficient is one, proving (3.5.18).

Call a support transvection-minimal if no elementary transvection makes it lexicographically smaller. A least support in a full $\mathrm{GL}(n,2)$-orbit is transvection-minimal.

Lemma 3.5.8 (Pivot lemma). *Let A be transvection-minimal. If $A \cap \mathcal{D}_{a,b} \neq \varnothing$ and S is its least member, then*

$$q_{a,b}(S) \notin A. \tag{3.5.19}$$

Proof. By Lemma 3.5.7, the changed positions are $q_{a,b}(A \cap \mathcal{D}_{a,b})$. Lemma 3.5.6 shows that their least member is $q_{a,b}(S)$, so every earlier incidence bit is unchanged. If $q_{a,b}(S) \in A$, its bit changes from 1 to 0. The transformed support is lexicographically smaller. This contradicts transvection minimality, so (3.5.19) holds.

Theorem 3.5.1 (Two-sided private-edge normal form). *Let $1 \le r \le n$, and let $0 \neq \omega \in \bigwedge^r (\mathbb{F}_2^n)^*$. There is an $M \in \mathrm{GL}(n,2)$ such that, for*

$$A = \mathrm{supp}\big(\bigwedge^r (M^*)\omega\big),$$

both of the following hold:

1. *for every $i \in \mathrm{Var}(A)$, there is an $S_i^+ \in A$ with $i \in S_i^+$ and*

$$S_i^+ - i + t \not\in A \qquad \text{for every } t \not\in S_i^+;$$

2. *for every $i \not\in \mathrm{Core}(A)$, there is an $S_i^- \in A$ with $i \not\in S_i^-$ and*

$$S_i^- - t + i \not\in A \qquad \text{for every } t \in S_i^-.$$

In particular, every nonzero top form has an affine representative satisfying the printed private-edge condition P_+, without any kernel hypothesis.

Proof. The $\mathrm{GL}(n,2)$-orbit of ω is finite. Choose an orbit element ω' whose support A is lexicographically least. This support is transvection-minimal, and membership of ω' in the orbit supplies the required M. Fix $i \in \mathrm{Var}(A)$, and let S_i^+ be the least supported edge containing i. For every $t \not\in S_i^+$, this edge is the least member of $A \cap \mathcal{D}_{i,t}$: every member of that family contains i, and S_i^+ was least among all supported edges containing i. Applying Lemma 3.5.8 with $(a,b) = (i,t)$ gives

$$S_i^+ - i + t = q_{i,t}(S_i^+) \not\in A.$$

This proves the first assertion. Now fix $i \not\in \mathrm{Core}(A)$, and let S_i^- be the least supported edge avoiding i. For every $t \in S_i^-$, this edge is the least member of $A \cap \mathcal{D}_{t,i}$: every member of that family avoids i, and S_i^- was least among all supported edges avoiding i. Applying Lemma 3.5.8 with $(a,b) = (t,i)$ gives

$$S_i^- - t + i = q_{t,i}(S_i^-) \not\in A.$$

This proves the second assertion.

The global orbit minimum is convenient for existence, but orbit enumeration is unnecessary.

Corollary 3.5.1 (Descent). *Starting from any nonzero degree-r top support, repeatedly apply any elementary transvection that strictly decreases its support incidence vector. The process terminates at a support satisfying both conclusions of Theorem 3.5.1. The product of the transvections is an explicit linear witness.*

Proof. There are only $2^{\binom{n}{r}}$ supports, and each step strictly decreases a total order, so the process terminates. Its terminal support is transvection-minimal. The proof of Theorem 3.5.1 used only this local minimality, so it applies verbatim.

The resulting bound of $2^{\binom{n}{r}} - 1$ steps is deliberately crude. It is a termination bound, not a polynomial-time normalization claim.

Resolution of the Two Open Problems We first apply the external half of the normal form.

Theorem 3.5.2 (Resolution of Open Problem 1). *Let $n > r \geq 2$, and let $f : V \to \mathbb{F}_2$ have degree r. If f has no degree-drop affine hyperplane, then there exist $M \in \mathrm{GL}(n, 2)$ and a homogeneous degree-r function g such that*

$$\deg(f \circ M - g) < r \qquad \textit{and} \qquad \mathsf{P}_-(A(g)).$$

In fact the witness in (3.5.4) *exists for every coordinate $i \in [n]$, not only for $i \in \mathrm{Var}(A(g))$.*

Proof. Let $\omega = \omega_f$. By Lemma 3.5.1, the no-drop hypothesis is equivalent to $\ker W_\omega = \{0\}$. Apply Theorem 3.5.1, and write $\omega' = \bigwedge^r(M^*)\omega$ and $A = \mathrm{supp}(\omega')$. Kernel covariance, Lemma 3.5.3, gives $\ker W_{\omega'} = \{0\}$. Lemma 3.5.4 therefore gives $\mathrm{Core}(A) = \varnothing$. Every coordinate satisfies the premise of the external half of Theorem 3.5.1, so for each $i \in [n]$ there is an edge $S_i^- \in A$ avoiding i whose every replacement $S_i^- - t + i$ is absent. In particular, $\mathsf{P}_-(A)$ holds. Define $g = \sum_{S \in A} x_S$. By affine top-form covariance (3.5.3), g is the degree-r component of $f \circ M$. Hence $\deg(f \circ M - g) < r$, so $f \sim_{r-1} g$. Moreover, $f \circ M$ itself has degree-r support $A(g)$, so the actual affine transform, not only its homogeneous quotient representative, satisfies P_-. This is exactly the affine quotient representative asked for in Open Problem 1. The affine witness has zero translation.

We next apply the internal half. The proof is direct and does not invoke complement duality or Theorem 3.5.2.

3.6 Functions without fast points

The definitions and normal-form lemmas required for this problem appear in Section 3.5. The ambient-variable condition is retained because it is necessary when the support is viewed in a space larger than its active variable set.

Open Question 3.6.1 (Carlet–Feukoua–Sălăgean, Open Problem 2). *Is every Boolean function with no fast point affine equivalent, modulo its lower-degree part, to a function whose top support satisfies P_+?*

We prove affirmative answers in the full stated range $n > r \geq 2$. For Open Problem 2, the representative will additionally satisfy $\mathrm{Var}(A) = [n]$, repairing the ambient-variable issue described in the introduction.

Theorem 3.6.1 (Resolution of Open Problem 2). *Let $n > r \geq 2$, and let $f : V \to \mathbb{F}_2$ have degree r. If f has no fast point, then there exist $M \in \mathrm{GL}(n, 2)$ and a homogeneous degree-r function g such that*

$$\deg(f \circ M - g) < r, \qquad \mathsf{P}_+(A(g)), \qquad \mathrm{Var}(A(g)) = [n].$$

Thus g satisfies both the private-edge condition printed in published Corollary 8 and the ambient-variable guard needed for that condition to imply the absence of fast points.

Proof. Let $\omega = \omega_f$. By Lemma 3.5.2, the no-fast hypothesis is equivalent to $\ker C_\omega = \{0\}$. Apply Theorem 3.5.1, and put $\omega' = \bigwedge^r(M^*)\omega$, with support A. Contraction-kernel covariance gives $\ker C_{\omega'} = \{0\}$. By Lemma 3.5.4, every coordinate is active: $\mathrm{Var}(A) = [n]$. The internal half of Theorem 3.5.1 supplies, for each active coordinate i, a supported edge S_i^+ containing i such that $S_i^+ - i + t \not\in A$ for every $t \not\in S_i^+$. This is exactly $\mathsf{P}_+(A)$. As in the proof of Theorem 3.5.2, define $g = \sum_{S \in A} x_S$. Equation (3.5.3) gives $\deg(f \circ M - g) < r$. Moreover, $f \circ M$ itself has degree-r support $A(g)$, so the actual affine transform satisfies both P_+ and $\mathrm{Var}(A(g)) = [n]$. Thus g is the required affine quotient representative, again with zero translation.

Let $\mathcal{N}_W$ and $\mathcal{N}_C$ denote the top forms with trivial wedge and contraction kernels, respectively. Let $\mathcal{O}_-$ denote the union of affine top-form orbits meeting P_-, and let $\mathcal{O}_+^{\mathrm{all}}$ denote the corresponding union for P_+ with $\mathrm{Var} = [n]$. The two proofs give the concise inclusions

$$\mathcal{N}_W \subseteq \mathcal{O}_-, \qquad \mathcal{N}_C \subseteq \mathcal{O}_+^{\mathrm{all}}. \tag{3.6.1}$$

By Lemmas 3.5.1 and 3.5.2, these are precisely the semantic source classes of Open Problems 1 and 2. If a top form satisfies both kernel conditions, one lexicographically minimal representative simultaneously satisfies both private-edge systems.

3.7 Differential factors and zero-four DDT columns

Problem Surface Let $S : V_n \to V_m$. Tezcan and Özbudak introduced a *differential factor* as an input translation that preserves all pairs with a prescribed output difference [54]. More precisely, a nonzero $\lambda \in V_n$ is a differential factor for $\mu \in V_m$ when

$$S(x) + S(y) = \mu \quad \Longrightarrow \quad S(x + \lambda) + S(y + \lambda) = \mu \tag{3.7.1}$$

for every $x, y \in V_n$. Addition denotes XOR. The symbols $\mathrm{DDT}_S(a, b)$ and $\mathrm{DU}(S)$ have the meanings fixed at the beginning of this section. In their study of differentially 4-uniform 4-bit S-boxes, Tezcan and Özbudak posed the following statement.

Open Question 3.7.1 (Tezcan–Özbudak, Conjecture 1). *Let S be a differentially 4-uniform 4-by-4 S-box and let $\mu \neq 0$. Then S has a differential factor for μ if and only if the μ-column of* DDT_S *consists only of zeros and fours.*

The source defines an S-box as a function and does not put "bijective" in the displayed conjecture. Its surrounding theory repeatedly specializes to bijective S-boxes, and the cryptanalytic application uses permutation S-boxes. This distinction matters: the two readings have different answers. The corrected-attack paper of Tezcan revisited the attack accounting but did not prove Conjecture 3.7.1 [53]. Later work studied when a differential factor is practically useful, a question logically separate from its existence [33]. The result below concerns the existence equivalence only.

An Unrestricted Counterexample Let $K = \mathbb{F}_2[t]/(t^4 + t + 1) \cong \mathbb{F}_{16}$ and let $K_0 \cong \mathbb{F}_4$ be its unique subfield of order four. We use the additive groups of these fields as binary vector spaces.

Lemma 3.7.1 (The fourth-power linear map). *The $\mathbb{F}_2$-linear map $L : K \to K$, $L(z) = z^4 + z$, has $\ker L = \operatorname{im} L = K_0$. Every element of K_0 therefore has exactly four preimages under L.*

Proof. The equation $z^4 = z$ characterizes the elements of K_0, so $\ker L = K_0$ and $|\ker L| = 4$. Hence $|\operatorname{im} L| = 16/4 = 4$. Moreover,

$$L(z)^4 = z^{16} + z^4 = z + z^4 = L(z),$$

which shows $\operatorname{im} L \subseteq K_0$. The two sets have the same cardinality. The fiber statement follows from linearity.

Theorem 3.7.1 (Counterexample to the unrestricted reading). *Define $S : K \to K$ by $S(x) = x^5$, and set $\mu = 1$. Then:*

1. *S is non-bijective and $\mathrm{DU}(S) = 4$;*
2. *$\mathrm{DDT}_S(0, 1) = 0$ and $\mathrm{DDT}_S(a, 1) = 4$ for every $a \neq 0$;*
3. *S has no nonzero differential factor for $\mu = 1$.*

Consequently, Conjecture 3.7.1 is false when "S-box" is read as an arbitrary function.

Proof. The multiplicative group K^* has order 15. The fifth-power map on K^* has image of order $15/\gcd(5, 15) = 3$, namely K_0^*. Thus $S(K) = K_0$, and S is not bijective. Fix $a \neq 0$ and write $x = az$. In characteristic two,

$$\begin{aligned} S(x + a) + S(x) &= (x + a)^5 + x^5 \\ &= ax^4 + a^4x + a^5 \\ &= a^5(z^4 + z + 1). \end{aligned} \tag{3.7.2}$$

Since $(a^5)^4 = a^{20} = a^5$, we have $a^5 \in K_0^*$. By Lemma 3.7.1, every value attained by the derivative in (3.7.2) has four preimages, and no other value has a preimage. Hence $\mathrm{DU}(S) = 4$. In particular, the equation

$$a^5(z^4 + z + 1) = 1$$

is equivalent to

$$L(z) = a^{-5} + 1 \in K_0.$$

It has four solutions for every $a \neq 0$. The row $a = 0$ contains zero in the $\mu = 1$ column, proving the asserted column pattern. Suppose that $\lambda \neq 0$ were a differential factor for $\mu = 1$. For every $a \neq 0$, the fiber just computed is nonempty; choose an x with $S(x + a) + S(x) = 1$. The factor property gives $S(x + a + \lambda) + S(x + \lambda) = 1$. Subtracting the two derivative values using (3.7.2) yields

$$a\lambda^4 + a^4\lambda = 0.$$

After division by $a\lambda$, this says

$$(a/\lambda)^3 = 1 \qquad \text{for every } a \in K^*.$$

Only three elements of the cyclic group K^* satisfy $u^3 = 1$, whereas a ranges over all fifteen nonzero elements. This contradiction excludes every nonzero factor.

The proof is field-theoretic and does not depend on a table enumeration. In the polynomial basis determined by $t^4 + t + 1$, the value table of x^5 is

$$(0, 1, 6, 6, 7, 7, 7, 6, 1, 7, 1, 6, 1, 6, 7, 1),$$

which provides a convenient independent implementation check.

The Permutation Case From now on let $V = \mathbb{F}_2^4$, let $S : V \to V$ be a permutation, and fix $\mu \neq 0$. Define the *partner involution*

$$T(x) = S^{-1}(S(x) + \mu). \tag{3.7.3}$$

For $a \in V$, define its direction fiber

$$P_a = \{x \in V : T(x) = x + a\}. \tag{3.7.4}$$

Lemma 3.7.2 (Partner and DDT fibers). *The map T is a fixed-point-free involution. The sets $(P_a)_{a \in V}$ partition V, $P_0 = \varnothing$, and*

$$|P_a| = \mathrm{DDT}_S(a, \mu) \qquad (a \in V). \tag{3.7.5}$$

Moreover, a nonzero vector λ is a differential factor for μ if and only if translation by λ commutes with T:

$$T(x + \lambda) = T(x) + \lambda \qquad (x \in V). \tag{3.7.6}$$

Proof. Applying S twice in (3.7.3) and using characteristic two gives $T(T(x)) = x$. A fixed point would imply $S(x) = S(x) + \mu$, contrary to $\mu \neq 0$. Every x lies in the unique fiber with direction $a = T(x) + x$, which proves the partition and $P_0 = \varnothing$. Now

$$T(x) = x + a \quad \Longleftrightarrow \quad S(x + a) + S(x) = \mu,$$

so (3.7.5) follows. Finally, if $y = T(x)$, then $S(x) + S(y) = \mu$. Condition (3.7.1) says exactly that the partner of $x + \lambda$ is $y + \lambda$, which is (3.7.6). The converse is the same implication in reverse.

The reverse implication of the conjecture uses one finite-geometric fact.

Lemma 3.7.3 (Four-plane partition). *Let an affine four-space over $\mathbb{F}_2$ be partitioned into four affine 2-flats*

$$V = (x_1 + U_1) \,\dot\cup\, (x_2 + U_2) \,\dot\cup\, (x_3 + U_3) \,\dot\cup\, (x_4 + U_4),$$

where each U_i is two-dimensional. Then

$$U_1 \cap U_2 \cap U_3 \cap U_4$$

contains a nonzero vector.

Proof. First observe that $U_i \cap U_j \neq \{0\}$ whenever $i \neq j$. Otherwise $U_i + U_j = V$, and any coset of U_i meets every coset of U_j, contrary to the disjointness of the partition. Suppose for a contradiction that the intersection of all four direction spaces is zero. Put

$$L = U_1 \cap U_2, \qquad W = U_1 + U_2.$$

Then L is a line and W is three-dimensional. After relabelling, we may assume that U_3 does not contain L. Its intersections with U_1 and U_2 are two distinct lines, so they span U_3; hence $U_3 \subset W$. If U_4 does not contain L, the same argument puts U_4 in W. If U_4 contains L, choose a nonzero vector in $U_4 \cap U_3$. It cannot lie in L, because U_3 does not contain L. Together with L, it spans U_4, again showing $U_4 \subset W$. Thus all four direction spaces lie in W. Every affine plane $x_i + U_i$ is now contained in one of the two cosets of W in V. Each coset has eight points and is covered by exactly two of the four planes. Two disjoint affine planes in an affine 3-space must be parallel: if their direction spaces were different, their sum would be the whole 3-space and the two cosets would intersect. Therefore the two planes in one W-coset have a common direction space, say A, and the two in the other coset have a common direction space, say B. But two two-dimensional subspaces $A, B \leq W$ intersect in a line. This line lies in all four U_i, a contradiction.

A 0/4 column produces a factor

Proposition 3.7.1. *Let $S : V \to V$ be a permutation and $\mu \neq 0$. If*

$$\mathrm{DDT}_S(a, \mu) \in \{0, 4\} \qquad \textit{for every } a \in V,$$

then S has a nonzero differential factor for μ.

Proof. By Lemma 3.7.2, the nonempty direction fibers P_a partition the 16-element space V. Exactly four of them are nonempty, each of size four. If $x \in P_a$, then $T(x) = x + a$. Involutivity gives $T(x + a) = x$, so P_a is stable under translation by a. A four-point set stable under a nonzero translation is the union of two cosets of the line $\{0, a\}$. Choosing representatives x, y from the two cosets gives

$$P_a = \{x, x + a, y, y + a\} = x + \mathrm{span}\{a, x + y\},$$

an affine 2-flat. The four nonempty fibers therefore satisfy Lemma 3.7.3. Choose a nonzero vector λ in the intersection of their four direction spaces. Translation by λ preserves every nonempty P_a. On such a fiber, $T(x) = x + a$, and hence

$$T(x + \lambda) = x + \lambda + a = T(x) + \lambda.$$

The empty fibers cause no condition, so (3.7.6) holds on all of V. Lemma 3.7.2 shows that λ is a differential factor.

A factor forces a 0/4 column

Proposition 3.7.2. *Let $S : V \to V$ be a permutation with $\mathrm{DU}(S) \leq 4$, let $\mu \neq 0$, and suppose that S has a differential factor $\lambda \neq 0$ for μ. Then*

$$\mathrm{DDT}_S(a, \mu) \in \{0, 4\} \qquad \textit{for every } a \in V.$$

Proof. Each direction fiber P_a is stable under translation by a, as in the proof of Proposition 3.7.1. It is also stable under translation by λ, by Lemma 3.7.2. If $a \notin \{0, \lambda\}$, then a and λ are linearly independent over $\mathbb{F}_2$. The translation group $\mathrm{span}\{a, \lambda\}$ acts freely on P_a, so $|P_a|$ is divisible by four. We have $P_0 = \varnothing$, and the fibers partition V. Therefore

$$|P_\lambda| = 16 - \sum_{a \neq \lambda} |P_a| \equiv 0 \pmod 4.$$

For every nonzero a, differential uniformity gives $|P_a| \leq 4$. Thus all the fiber sizes are zero or four. Equation (3.7.5) converts this conclusion into the claimed DDT-column statement.

Combining the two propositions gives the corrected form of the conjecture.

Theorem 3.7.2 (Permutation-S-box resolution). *Let $S : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ be a permutation with $\mathrm{DU}(S) \leq 4$, and let $\mu \neq 0$. The following are equivalent:*

1. *S has a nonzero differential factor for μ;*
2. *the μ-column of DDT_S consists only of zeros and fours.*

Proof. Proposition 3.7.2 proves the forward implication, and Proposition 3.7.1 proves the reverse implication.

The hypothesis $\mathrm{DU}(S) \leq 4$ is slightly more general than the source wording $\mathrm{DU}(S) = 4$. No classification of 4-bit permutations is used. The only dimension-specific input is Lemma 3.7.3.

3.8 The second identity in Kaleyski's conjecture

Introduction Let $K = \mathbb{F}_{2^n}$. For a function $F : K \to K$, we apply the DDT convention fixed at the beginning of this section to the additive group of K. The function F is almost perfect nonlinear (APN) if $\mathrm{DDT}_F(a, b) \leq 2$ for every $a \neq 0$

and every b. Kaleyski's two-point modification problem led to two distinguished sums of these entries [32]. For a power function $F_d(x) = x^d$, put

$$S_1(d) = \sum_{y \in K} \mathrm{DDT}_{F_d}(y, F_d(y) + 1), \quad S_2(d) = \sum_{y \in K} \mathrm{DDT}_{F_d}(y + 1, F_d(y)), \tag{3.8.1}$$

$$B_n = \sum_{y \in K} \mathrm{DDT}_{x^3}(y, y^3 + 1). \tag{3.8.2}$$

Open Question 3.8.1 (Kaleyski, Conjecture 21). *If F_d is an APN power function on $\mathbb{F}_{2^n}$, then*

$$S_1(d) = B_n \qquad \textit{and} \qquad S_2(d) = B_n.$$

Kaleyski evaluated the cubic benchmark explicitly [32, Eq. (34)]:

$$B_n = \begin{cases} 2(2^{n-1} - 1), & n \text{ odd}, \\ 2(2^{n-1} + 2^{n/2} - 1), & n \text{ even and } n/2 \text{ odd}, \\ 2(2^{n-1} - 2^{n/2} - 1), & n \text{ even and } n/2 \text{ even}. \end{cases} \tag{3.8.3}$$

Mihaila and Thornburgh subsequently proved Conjecture 3.8.1 for odd $n \geq 3$ [39, Prop. 5.3]. Their proof expresses the sums through multiplicities in an associated graph. They also give a fractional-map preimage formula for a support-indicator sum [39, Prop. 5.2]; passing from that formula to a DDT sum uses the APN fact that every nonempty fiber at a nonzero direction has two elements. Immediately after Proposition 5.3, they state that even dimension remains open. Since their range begins at $n = 3$, the remaining even-dimensional range begins at $n = 4$. Our result closes the second equality in that range. It is a family theorem for every exponent satisfying the APN hypothesis, not a verification for a list of known Gold, Kasami, or Dobbertin exponents. In particular, F_d is not assumed to be a permutation. The first equality remains open in even dimension and is outside the claim proved below.

An Incidence Normal Form Fix a positive exponent e. Define the derivative at one and the image of the power map by

$$D_e(z) = (z + 1)^e + z^e, \qquad I_e = \{x^e : x \in K\}. \tag{3.8.4}$$

We also use

$$H_e = \{t \in K^* : t^e = 1\}, \qquad \kappa_e = |H_e|, \tag{3.8.5}$$

$$P_e = \{z \in K : D_e(z) \neq 0 \text{ and } D_e(z) \in I_e\}, \qquad p_e = |P_e|. \tag{3.8.6}$$

The set P_e is the source profile of the unit derivative relative to the power image.

Lemma 3.8.1 (Second-sum incidence). *Suppose $F_e(x) = x^e$ is APN on a binary field K with $|K| > 2$. Then $e > 0$ and*

$$S_2(e) + 2 = \sum_{z \in K} \big|\{u \in K : u^e = D_e(z)\}\big|. \tag{3.8.7}$$

Proof. If $e = 0$, then F_e is constant and $\mathrm{DDT}_{F_e}(1, 0) = |K| > 2$, contrary to APN. Thus $e > 0$. For $y \neq 1$, set $a = y + 1$. Scaling $x = az$ in the equation counted by $\mathrm{DDT}_{F_e}(y + 1, y^e)$ gives $D_e(z) = (y/(y + 1))^e$. The fractional map

$$\phi : K \setminus \{1\} \longrightarrow K \setminus \{1\}, \qquad \phi(y) = \frac{y}{y+1},$$

is an involution in characteristic two. At every nonzero direction, a nonempty derivative fiber contains the distinct pair $x, x + a$; APN makes this the entire fiber. The $y = 1$ summand has direction zero and is zero. Consequently,

$$S_2(e) = 2\big|\{u \in K : u \neq 1,\ u^e \in \mathrm{im}(D_e)\}\big|. \tag{3.8.8}$$

Since $1^e = D_e(0) = 1$, adjoining the omitted point $u = 1$ gives $S_2(e) + 2 = 2|\{u \in K : u^e \in \mathrm{im}(D_e)\}|$. Double-count the incidence set $\mathcal{E}_e = \{(u, z) \in K^2 : u^e = D_e(z)\}$. For each u with $u^e \in \mathrm{im}(D_e)$, the APN derivative D_e has exactly two preimages. Counting first by u and then by z proves (3.8.7).

Lemma 3.8.2 (Power fibers and derivative zeros). *For $e > 0$,*

$$\sum_{z \in K} \big|\{u \in K : u^e = D_e(z)\}\big| = Z_e + \kappa_e p_e, \tag{3.8.9}$$

where $Z_e = |\{z \in K : D_e(z) = 0\}|$. Moreover,

$$Z_e + 1 = \kappa_e. \tag{3.8.10}$$

Proof. The fiber of $u \mapsto u^e$ above zero is the singleton $\{0\}$. For $b \in I_e \setminus \{0\}$, choose $a \neq 0$ with $a^e = b$. Division by a gives a bijection

$$\{u \in K : u^e = b\} \longleftrightarrow H_e, \qquad u \longmapsto u/a.$$

A nonzero value outside I_e has no preimage. Splitting the sum according to these three cases gives (3.8.9). If $D_e(z) = 0$, then $z \neq 0$ and

$$t = \frac{z+1}{z}$$

lies in $H_e \setminus \{1\}$. Conversely, every $t \in H_e \setminus \{1\}$ gives a derivative zero at $z = (t + 1)^{-1}$. These maps are inverse, proving $Z_e = \kappa_e - 1$.

Combining the two lemmas gives the normal form that drives the proof.

Proposition 3.8.1 (Kernel–profile normal form). *If $F_e(x) = x^e$ is APN and $|K| > 2$, then*

$$\boxed{S_2(e) + 3 = \kappa_e(p_e + 1).} \tag{3.8.11}$$

Proof. Equations (3.8.7), (3.8.9), and (3.8.10) give

$$S_2(e) + 3 = (Z_e + 1) + \kappa_e p_e = \kappa_e(p_e + 1).$$

The constants in (3.8.11) are structural. The $+2$ repairs the omitted direction $y = 1$, while the additional $+1$ comes from the zero fiber of the power map and the identity $Z_e + 1 = \kappa_e$.

The Kernel Has Three Elements

Lemma 3.8.3 (Even-dimensional APN kernel). *Let $n \geq 4$ be even and let $F_d(x) = x^d$ be APN on $K = \mathbb{F}_{2^n}$. Then*

$$\kappa_d = 3. \tag{3.8.12}$$

The cube map also satisfies $\kappa_3 = 3$.

Proof. By (3.8.10) and APN,

$$\kappa_d - 1 = Z_d = \mathrm{DDT}_{F_d}(1, 0) \leq 2.$$

The set H_d is a subgroup of K^*, whose order $2^n - 1$ is odd. Thus κ_d is odd, and hence $\kappa_d \in \{1, 3\}$. Because n is even, $3 \mid 2^n - 1$. Choose a nontrivial element $\omega \in K^*$ of order three. It satisfies

$$\omega^2 + \omega = 1, \qquad \omega + 1 = \omega^2. \tag{3.8.13}$$

Suppose that $\kappa_d = 1$. Then $\theta = \omega^d \neq 1$; otherwise $\omega \in H_d$. Since $\theta^3 = 1$, the nontrivial cubic root θ satisfies $\theta^2 + \theta = 1$. Therefore

$$D_d(\omega) = (\omega + 1)^d + \omega^d = (\omega^d)^2 + \omega^d = 1.$$

We also have $D_d(0) = D_d(1) = 1$. The three distinct inputs $0, 1, \omega$ in the fiber $D_d^{-1}(1)$ contradict $\mathrm{DDT}_{F_d}(1, 1) \leq 2$. Hence $\kappa_d = 3$. Finally, the equation $t^3 = 1$ has exactly three solutions in the cyclic group K^*, because $3 \mid |K^*|$. Thus $\kappa_3 = 3$.

This argument is independent of any classification of APN monomials. It also shows why a permutation hypothesis would be inappropriate here: the conclusion $\kappa_d = 3$ means that the nonzero d-th-power map is three-to-one.

Equality of the Source Profiles Since K^* is cyclic, it has a unique subgroup of each order dividing $|K^*|$. A group endomorphism of K^* with kernel of size three has image of size $(2^n - 1)/3$. Lemma 3.8.3 therefore gives

$$I_d = I_3. \tag{3.8.14}$$

Here the equality includes zero; the nonzero parts are the common index-three subgroup of K^*.

Lemma 3.8.4 (Ratio profile). *For every positive exponent e, define*

$$T_e = \{t \in K^* : t \neq 1,\ t^e + 1 \neq 0,\ t^e + 1 \in I_e\}.$$

Then

$$p_e = |T_e| + 2. \tag{3.8.15}$$

If n is even and F_d is APN, then

$$p_d = p_3. \tag{3.8.16}$$

Proof. The exceptional sources $z = 0, 1$ both belong to P_e, since $D_e(0) = D_e(1) = 1$. On $K \setminus \{0, 1\}$, the map

$$z \longmapsto t = \frac{z+1}{z}$$

is a bijection onto $K^* \setminus \{1\}$, with inverse $z = (t+1)^{-1}$. It satisfies

$$D_e(z) = z^e(t^e + 1). \tag{3.8.17}$$

The nonzero part of I_e is a subgroup, and z^e lies in that subgroup. Thus the two conditions defining P_e become $t^e + 1 \neq 0$ and $t^e + 1 \in I_e$. This proves (3.8.15). For $e = d$ and $e = 3$, the maps $t \mapsto t^e$ have the same image $I_d \setminus \{0\} = I_3 \setminus \{0\}$ and every fiber has size three. Partition T_e by $b = t^e$. The fiber $b = 1$ contributes nothing, because $b + 1 = 0$; for every other b in the common image, either all three preimages contribute or none do, according as $b + 1 \in I_d = I_3$. Hence

$$|T_d| = 3\big|\{b \in I_d \setminus \{0, 1\} : b + 1 \in I_d\}\big| = |T_3|.$$

Equation (3.8.16) now follows from (3.8.15).

Connecting the Cubic Benchmark The benchmark B_n in (3.8.2) is the cubic *first* sum. It is therefore not legitimate simply to substitute $e = 3$ into Proposition 3.8.1, which concerns the second sum. The next lemma supplies the required bridge and tracks the constant term in the cubic derivative.

Lemma 3.8.5 (Cubic benchmark normal form). *If $n \geq 4$ is even, then*

$$B_n + 3 = 3(p_3 + 1). \tag{3.8.18}$$

Proof. The cube map is APN in every dimension. Its derivative at one is

$$D_3(z) = (z+1)^3 + z^3 = z^2 + z + 1. \tag{3.8.19}$$

We first normalize the first sum. The $y = 0$ summand in B_n vanishes. For $y \neq 0$, scale $x = yz$ and then put $u = y^{-1}$. As in Lemma 3.8.1, nonempty APN fibers have size two. Adding the omitted $u = 0$, which contributes because $D_3(0) = 1$, and double-counting the resulting incidences gives

$$B_n + 2 = \sum_{z \in K} \big|\{u \in K : u^3 = z^2 + z\}\big|. \tag{3.8.20}$$

Choose ω as in (3.8.13). Translation by ω is a bijection of K, and

$$(z+\omega)^2 + (z+\omega) = z^2 + z + 1 = D_3(z).$$

Reindexing (3.8.20) therefore yields

$$B_n + 2 = \sum_{z \in K} \big|\{u \in K : u^3 = D_3(z)\}\big|. \tag{3.8.21}$$

Applying Lemma 3.8.2 with $e = 3$, and using $\kappa_3 = 3$, gives

$$B_n + 3 = \kappa_3(p_3 + 1) = 3(p_3 + 1).$$

The $+1$ in (3.8.19) is essential. It is precisely the Artin–Schreier translation above that turns $z^2 + z$ into the unit cubic derivative.

Main Result

Theorem 3.8.1 (Second identity in even dimension). *Let $n \geq 4$ be even, let $K = \mathbb{F}_{2^n}$, and let $F_d(x) = x^d$ be APN. Then*

$$\sum_{y \in K} \mathrm{DDT}_{F_d}(y+1, F_d(y)) = \sum_{y \in K} \mathrm{DDT}_{x^3}(y, y^3 + 1). \tag{3.8.22}$$

Equivalently,

$$S_2(d) = B_n = \begin{cases} 2^n + 2^{n/2+1} - 2, & n \equiv 2 \pmod 4, \\ 2^n - 2^{n/2+1} - 2, & n \equiv 0 \pmod 4. \end{cases} \tag{3.8.23}$$

Proof. Proposition 3.8.1, Lemma 3.8.3, and Lemma 3.8.4 give

$$S_2(d) + 3 = \kappa_d(p_d + 1) = 3(p_3 + 1).$$

By Lemma 3.8.5, this equals $B_n + 3$. Cancelling three proves (3.8.22). The closed form (3.8.23) is the even-dimensional specialization of (3.8.3).

3.9 Balanced functions of degree at most two and second-order Reed–Muller codes

Problem Surface Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. Carlet's list of general open questions on Boolean and vectorial functions contains the following item.

Open Question 3.9.1 (Carlet 13.2(3)). *Find simple formulae for the number of balanced quadratic functions in n variables and for the weight distribution of the dual $\mathrm{RM}(n-3, n)$ of the second-order Reed–Muller code.*

We spell out both components of Open Question 3.9.1. The first component is a scalar count of balanced quadratic Boolean functions. The second component is a full code distribution: for every $n \geq 3$ and every weight w, it gives the number of words of weight w in $\mathrm{RM}(n-3, n)$. The connection between the two components is structural. The middle-weight words of $\mathrm{RM}(2, n)$ are precisely the balanced degree-at-most two Boolean functions, while the full dual distribution is obtained from the entire $\mathrm{RM}(2, n)$ distribution by the MacWilliams identity.

Related Work Sloane and Berlekamp determined the weight enumerator of the second-order binary Reed–Muller code in 1970 [50]. In the notation of the present note, their theorem gives the following distribution for $\mathrm{RM}(2, n)$: the only nonzero weights are

$$0, \qquad 2^n, \qquad 2^{n-1}, \qquad 2^{n-1} \pm 2^{n-r-1} \quad (1 \le r \le \lfloor n/2 \rfloor),$$

with the non-middle multiplicities

$$A_{n,r} 2^{2r}.$$

The balanced degree-at-most-two count is the coefficient of the middle weight 2^{n-1} in this distribution. The exact-degree-two count is obtained from the same coefficient by subtracting the balanced affine words $2(2^n - 1)$. The proof route below is different from the 1970 derivation. Sloane and Berlekamp use matrix notation for binary quadratic forms. They reduce these forms by elementary affine transformations, organize the reduction by a flow chart of terminal matrix types, and derive recurrences for the number of forms reducing to each type. The present derivation uses the alternating polar form of a quadratic Boolean function as the primary invariant. It counts rank-$2r$ polar forms by choosing the radical and a nondegenerate alternating form on the quotient, and then counts affine perturbations by separating the linear part on the radical. Thus the invariant controlling the answer is the same rank of a binary quadratic form, but the enumeration is expressed through symplectic rank strata and Walsh-sum factorization rather than through the Sloane–Berlekamp reduction recurrences. The standard Reed–Muller duality relation

$$\mathrm{RM}(r, n)^{\perp} = \mathrm{RM}(n - r - 1, n)$$

gives $\mathrm{RM}(n-3, n) = \mathrm{RM}(2, n)^{\perp}$ for $n \ge 3$, and the MacWilliams identity transforms the known $\mathrm{RM}(2, n)$ weight enumerator into the full distribution of its dual [35]. The closed formula in Theorem 3.9.3 is this MacWilliams transform written coefficient-by-coefficient with an explicit Krawtchouk convention. The present note rewrites these coding-theoretic facts in the notation of quadratic Boolean functions and identifies the middle-weight coefficient with the balanced-function clause of Carlet 13.2(3). Balanced quadratic functions also appear in IACR work on S-box classification. De Meyer and Bilgin classified balanced quadratic vectorial Boolean functions up to affine equivalence in small dimensions and gave efficient classification procedures [38]. That line is orthogonal to the present note: it concerns affine-equivalence classes of vectorial functions, whereas the formulae below count scalar functions and codewords.

Definitions and Conventions All vector spaces are over $\mathbb{F}_2$. Let $n \ge 0$. A Boolean function $f : V_n \to \mathbb{F}_2$ is *balanced* if $\sum_{x \in V_n} (-1)^{f(x)} = 0$. Equivalently, for $n \ge 1$, f is balanced if and only if $\mathrm{wt}(f) = 2^{n-1}$, because

$$\sum_{x \in V_n} (-1)^{f(x)} = |\{x : f(x) = 0\}| - |\{x : f(x) = 1\}| = 2^n - 2\,\mathrm{wt}(f).$$

Every Boolean function of algebraic degree at most two has a unique algebraic normal form

$$f(x) = \sum_{0 \le i < j < n} q_{ij} x_i x_j + \sum_{i=0}^{n-1} \ell_i x_i + c, \qquad q_{ij}, \ell_i, c \in \mathbb{F}_2.$$

We call $q(x) = \sum_{0 \le i < j < n} q_{ij} x_i x_j$ the *quadratic part*. Thus "degree at most two" includes affine functions. Under the exact-degree convention, one also requires $q \neq 0$. We state both counts. The *polar form* of a quadratic part q is $\beta_q(x, y) = q(x + y) + q(x) + q(y)$. It is an alternating bilinear form: $\beta_q(x, x) = 0$ for all x. In the coordinate basis $e_0, \ldots, e_{n-1}$,

$$\beta_q(e_i, e_j) = q_{ij} \qquad (i < j).$$

Hence $q \mapsto \beta_q$ is a bijection from quadratic coefficient vectors $(q_{ij})_{i<j}$ to alternating bilinear forms on V_n. The rank of q is the rank of β_q. Its radical is

$$\mathrm{Rad}(\beta_q) = \{x \in V_n : \beta_q(x, y) = 0 \text{ for all } y \in V_n\}.$$

An alternating form has even rank; we write this rank as $2r$. For a binary linear code $\mathcal{C} \le \mathbb{F}_2^N$, its weight distribution is the sequence

$$W_j(\mathcal{C}) = |\{c \in \mathcal{C} : \mathrm{wt}(c) = j\}|, \qquad 0 \le j \le N.$$

The dual code is

$$\mathcal{C}^\perp = \{x \in \mathbb{F}_2^N : \langle x, c \rangle = 0 \text{ for all } c \in \mathcal{C}\},$$

where

$$\langle x, c \rangle = \sum_{i=1}^{N} x_i c_i \in \mathbb{F}_2.$$

Following the convention at the beginning of this section, we realize $\mathrm{RM}(r, n)$ as the evaluation code of Boolean polynomials in n variables of algebraic degree at most r. Thus $\mathrm{RM}(2, n)$ is exactly the evaluation code of the functions counted in the first component. We use the following Krawtchouk convention. For $0 \le w, j \le N$,

$$K_w^{(N)}(j) = \sum_{t=0}^{w} (-1)^t \binom{j}{t} \binom{N-j}{w-t}.$$

Throughout, $\binom{a}{b} = 0$ when $b < 0$ or $b > a$.

Counting Alternating Forms of Given Rank

Lemma 3.9.1 (Rank strata of alternating forms). *Let $0 \le r \le \lfloor n/2 \rfloor$. The number of alternating bilinear forms on V_n of rank $2r$ is*

$$A_{n,r} = \frac{2^{r(r-1)} \prod_{i=0}^{2r-1} (2^{n-i} - 1)}{\prod_{i=1}^{r} (2^{2i} - 1)}.$$

Proof. Let $k = n - 2r$. An alternating form β of rank $2r$ has radical $R = \mathrm{Rad}(\beta)$ of dimension k. Conversely, if $R \leq V_n$ has dimension k, then alternating forms on V_n with radical exactly R are in bijection with nondegenerate alternating forms on the quotient V_n/R. Indeed, a form with radical R factors through V_n/R, and the fact that the radical is exactly R says precisely that the induced form on V_n/R is nondegenerate. Therefore the desired number is

$$\begin{bmatrix} n \\ 2r \end{bmatrix}_2 N_{2r},$$

where $\begin{bmatrix} n \\ 2r \end{bmatrix}_2$ is the number of $2r$-dimensional subspaces of V_n and N_{2r} is the number of nondegenerate alternating forms on a fixed $2r$-dimensional $\mathbb{F}_2$-space. Counting ordered bases of a $2r$-dimensional subspace in two ways gives

$$\begin{bmatrix} n \\ 2r \end{bmatrix}_2 = \frac{\prod_{i=0}^{2r-1}(2^n - 2^i)}{\prod_{i=0}^{2r-1}(2^{2r} - 2^i)}.$$

It remains to compute N_{2r}. Fix one nondegenerate alternating form ω on a $2r$-dimensional space W. The group $\mathrm{GL}(W)$ acts transitively on nondegenerate alternating forms by pullback, and the stabilizer of ω is $\mathrm{Sp}_{2r}(2)$. Thus

$$N_{2r} = \frac{|\,\mathrm{GL}_{2r}(2)|}{|\,\mathrm{Sp}_{2r}(2)|}.$$

The general linear group has order

$$|\,\mathrm{GL}_{2r}(2)| = \prod_{i=0}^{2r-1}(2^{2r} - 2^i).$$

For the symplectic group, count ordered symplectic bases $(u_1, v_1, \ldots, u_r, v_r)$. After j pairs have been chosen, their span is nondegenerate of dimension $2j$, and its orthogonal complement has dimension $2r - 2j$. The next vector u_{j+1} can be any nonzero vector in that complement, giving $2^{2r-2j} - 1$ choices. Once u_{j+1} is fixed, the condition $\omega(u_{j+1}, v_{j+1}) = 1$ cuts out an affine hyperplane of size $2^{2r-2j-1}$ in the same complement. Hence

$$|\,\mathrm{Sp}_{2r}(2)| = \prod_{j=0}^{r-1}(2^{2r-2j} - 1)2^{2r-2j-1} = 2^{r^2}\prod_{i=1}^{r}(2^{2i} - 1).$$

Combining the displayed formulae gives

$$\begin{bmatrix} n \\ 2r \end{bmatrix}_2 N_{2r} = \frac{\prod_{i=0}^{2r-1}(2^n - 2^i)}{2^{r^2}\prod_{i=1}^{r}(2^{2i} - 1)}.$$

Finally,

$$\prod_{i=0}^{2r-1}(2^n - 2^i) = 2^{0+1+\cdots+(2r-1)}\prod_{i=0}^{2r-1}(2^{n-i} - 1),$$

and $0 + 1 + \cdots + (2r-1) - r^2 = r(r-1)$. This is the claimed expression.

Weights of Affine Perturbations

Lemma 3.9.2 (Signed sums at fixed polar rank). *Let $q : V_n \to \mathbb{F}_2$ be a quadratic part whose polar form has rank $2r$, and let $R = \mathrm{Rad}(\beta_q)$. Write a linear form ℓ as*

$$\ell = \ell_H + \ell_R$$

with respect to a direct sum $V_n = H \oplus R$, where H is any complement on which β_q is nondegenerate. The restriction of q to R is a linear form; denote it by

$$\lambda_R = q|_R \in \mathrm{Lin}(R, \mathbb{F}_2).$$

Then the signed sum

$$S(q + \ell + c) = \sum_{x \in V_n} (-1)^{q(x)+\ell(x)+c}$$

is zero if and only if $\ell_R + \lambda_R \neq 0$. If $\ell_R + \lambda_R = 0$, then

$$S(q + \ell + c) = \pm 2^{n-r}.$$

Proof. For $z, z' \in R$, the radical condition gives

$$q(z + z') + q(z) + q(z') = \beta_q(z, z') = 0.$$

Thus $q|_R = \lambda_R$ is linear. Moreover, for $h \in H$ and $z \in R$, $\beta_q(h, z) = 0$, so

$$q(h + z) = q(h) + \lambda_R(z).$$

Choose a symplectic basis $u_1, v_1, \ldots, u_r, v_r$ of H. In the corresponding coordinates, the canonical quadratic part on H is $q_0(a, b) = \sum_{i=1}^{r} a_i b_i$. The restrictions $q|_H$ and q_0 have the same polar form. Their difference is therefore a linear form $\lambda_H \in \mathrm{Lin}(H, \mathbb{F}_2)$. Hence

$$q(h + z) + \ell(h + z) + c = q_0(h) + (\ell_H + \lambda_H)(h) + (\ell_R + \lambda_R)(z) + c.$$

The signed sum factors as

$$\begin{aligned}\sum_{h \in H,\, z \in R} &(-1)^{q_0(h)+(\ell_H+\lambda_H)(h)+(\ell_R+\lambda_R)(z)+c} \\ &= (-1)^c \left(\sum_{h \in H} (-1)^{q_0(h)+(\ell_H+\lambda_H)(h)} \right) \left(\sum_{z \in R} (-1)^{(\ell_R+\lambda_R)(z)} \right).\end{aligned}$$

The radical factor is explicit. If $\ell_R + \lambda_R = 0$, then every term is 1, so the factor is $|R| = 2^{n-2r}$. Otherwise, choose $z_0 \in R$ with $(\ell_R + \lambda_R)(z_0) = 1$. Translation by z_0 pairs the terms in the sum and changes their sign, so the factor is 0. It remains only to see that the H-factor never vanishes. Write

$$(\ell_H + \lambda_H)(a, b) = \sum_{i=1}^{r} (\alpha_i a_i + \gamma_i b_i).$$

Then

$$\sum_{a,b\in\mathbb{F}_2^r}(-1)^{q_0(a,b)+(\ell_H+\lambda_H)(a,b)}=\prod_{i=1}^{r}\sum_{s,t\in\mathbb{F}_2}(-1)^{st+\alpha_i s+\gamma_i t}.$$

For one pair,

$$\begin{aligned}\sum_{s,t\in\mathbb{F}_2}(-1)^{st+\alpha s+\gamma t}&=\sum_{t\in\mathbb{F}_2}(-1)^{\gamma t}+(-1)^{\alpha}\sum_{t\in\mathbb{F}_2}(-1)^{(1+\gamma)t}\\&=2(-1)^{\alpha\gamma}.\end{aligned}$$

Thus the H-factor is $2^r(-1)^{\sum_i \alpha_i\gamma_i}$. Combining the two factors proves that the total signed sum is 0 exactly when $\ell_R+\lambda_R\neq 0$, and has magnitude $2^r\cdot 2^{n-2r}=2^{n-r}$ when $\ell_R+\lambda_R=0$.

Lemma 3.9.3 (Balanced perturbations). *Let $q:V_n\to\mathbb{F}_2$ be a quadratic part whose polar form has rank $2r$. The number of affine perturbations $q+\ell+c$, with $\ell\in\mathrm{Lin}(V_n,\mathbb{F}_2)$ and $c\in\mathbb{F}_2$, that are balanced is*

$$2^{2r+1}(2^{n-2r}-1).$$

Proof. Keep the notation of Lemma 3.9.2. A perturbation is balanced if and only if its signed sum is zero, and the lemma shows that this is equivalent to $\ell_R+\lambda_R\neq 0$. Translation by the fixed form λ_R permutes $\mathrm{Lin}(R,\mathbb{F}_2)$. Thus there are 2^{2r} choices for ℓ_H, $2^{n-2r}-1$ choices for ℓ_R that satisfy this condition, and two choices for the constant c. Therefore the number of balanced perturbations is $2^{2r}(2^{n-2r}-1)2=2^{2r+1}(2^{n-2r}-1)$.

Lemma 3.9.4 (Unbalanced weights at fixed nonzero polar rank). *Let $q:V_n\to\mathbb{F}_2$ have polar rank $2r$ with $r\geq 1$. Among the affine perturbations $q+\ell+c$, exactly 2^{2r} have weight*

$$2^{n-1}-2^{n-r-1},$$

and exactly 2^{2r} have weight

$$2^{n-1}+2^{n-r-1}.$$

Proof. By Lemma 3.9.2, the unbalanced perturbations are exactly those with $\ell_R=\lambda_R$. This fixes ℓ_R, while ℓ_H has 2^{2r} possible values. For each fixed ℓ_H, the two constants $c=0$ and $c=1$ multiply the signed sum by opposite signs. Hence one of the two constants gives signed sum $+2^{n-r}$, and the other gives signed sum -2^{n-r}. For any Boolean function f, $S(f)=\sum_x(-1)^{f(x)}=2^n-2\,\mathrm{wt}(f)$. Hence $\mathrm{wt}(f)=2^{n-1}-S(f)/2$. Thus $S(f)=+2^{n-r}$ and $S(f)=-2^{n-r}$ give weights $2^{n-1}-2^{n-r-1}$ and $2^{n-1}+2^{n-r-1}$, respectively. Since each of the 2^{2r} choices of ℓ_H contributes one perturbation of each sign, the two multiplicities are both 2^{2r}.

The Balanced-Quadratic Count

Theorem 3.9.1 (Degree-at-most-two balanced count). *For every $n \geq 0$, the number M_n of balanced Boolean functions $V_n \to \mathbb{F}_2$ of algebraic degree at most two is*

$$M_n = \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{2^{r(r-1)} \prod_{i=0}^{2r-1} (2^{n-i} - 1)}{\prod_{i=1}^{r} (2^{2i} - 1)} \, 2^{2r+1} (2^{n-2r} - 1).$$

Proof. Every degree-at-most-two Boolean function has a unique decomposition $f = q + \ell + c$ into its quadratic part q, linear part ℓ, and constant term c. Group the quadratic parts by the rank $2r$ of their polar forms. By Lemma 3.9.1, the number of possible quadratic parts of rank $2r$ is $A_{n,r}$. For each such part, Lemma 3.9.3 gives exactly

$$2^{2r+1}(2^{n-2r} - 1)$$

balanced choices of (ℓ, c). Summing over $0 \leq r \leq \lfloor n/2 \rfloor$ gives the formula.

Corollary 3.9.1 (Exact-degree-two balanced count). *For every $n \geq 0$, the number of balanced Boolean functions $V_n \to \mathbb{F}_2$ of algebraic degree exactly two is*

$$M_n - 2(2^n - 1).$$

Equivalently, it is

$$\sum_{r=1}^{\lfloor n/2 \rfloor} \frac{2^{r(r-1)} \prod_{i=0}^{2r-1} (2^{n-i} - 1)}{\prod_{i=1}^{r} (2^{2i} - 1)} \, 2^{2r+1} (2^{n-2r} - 1).$$

Proof. The affine functions are exactly the degree-at-most-two functions with quadratic part $q = 0$, i.e. the $r = 0$ stratum. An affine function $\ell + c$ is balanced if and only if $\ell \neq 0$, and then the constant c is arbitrary. Hence the balanced affine contribution is $2(2^n - 1)$. Removing this contribution from Theorem 3.9.1 gives the exact-degree-two count. The displayed summation form is the same statement with the $r = 0$ term omitted.

The Weight Distribution of RM$(2, n)$

Lemma 3.9.5 (Dimension of the second-order code). *Let $n \geq 0$, and put $k = 1 + n + \binom{n}{2}$. Then* RM$(2, n)$ *has dimension k and cardinality 2^k.*

Proof. The coefficient space of degree-at-most-two ANFs has one constant coefficient, n linear coefficients, and $\binom{n}{2}$ quadratic coefficients, hence dimension k. It remains to check that the evaluation map from this coefficient space to $\mathbb{F}_2^{V_n}$ is injective. Suppose that

$$f(x) = \sum_{i<j} q_{ij} x_i x_j + \sum_i \ell_i x_i + c$$

evaluates to zero for every $x \in V_n$. Evaluating at 0 gives

$$c = f(0) = 0.$$

Evaluating at e_i gives

$$\ell_i = f(e_i) + f(0) = 0 \qquad (0 \le i < n).$$

Finally, for $i < j$, evaluating at $e_i + e_j$ gives

$$q_{ij} = f(e_i + e_j) + f(e_i) + f(e_j) + f(0) = 0.$$

All coefficients are zero, so the evaluation map is injective. Therefore the dimension and cardinality of $\mathrm{RM}(2, n)$ are k and 2^k.

Theorem 3.9.2 (Second-order Reed–Muller weight distribution). *Let $n \ge 1$, $N = 2^n$, and let*

$$M_n = \sum_{r=0}^{\lfloor n/2 \rfloor} A_{n,r}\, 2^{2r+1}(2^{n-2r} - 1).$$

The number $W_j(\mathrm{RM}(2, n))$ of words of weight j in $\mathrm{RM}(2, n)$ is as follows:

$$W_j(\mathrm{RM}(2,n)) = \begin{cases} 1, & j = 0, \\ 1, & j = N, \\ M_n, & j = 2^{n-1}, \\ A_{n,r}2^{2r}, & j = 2^{n-1} - 2^{n-r-1} \text{ for some } 1 \le r \le \lfloor n/2 \rfloor, \\ A_{n,r}2^{2r}, & j = 2^{n-1} + 2^{n-r-1} \text{ for some } 1 \le r \le \lfloor n/2 \rfloor, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. The code $\mathrm{RM}(2, n)$ consists of the evaluation words of all functions

$$f = q + \ell + c$$

of algebraic degree at most two. By Lemma 3.9.5, different coefficient triples give different codewords, so we may count coefficient triples. First consider the rank-zero stratum $q = 0$. If $\ell = 0$, the two constants give the all-zero word and the all-one word, of weights 0 and N. If $\ell \neq 0$, then $\ell + c$ is balanced for both constants c, so these words have weight 2^{n-1}. This contributes one word of weight 0, one word of weight N, and $2(2^n - 1)$ middle-weight words. This is exactly the $r = 0$ term of M_n, since $A_{n,0} = 1$. Now fix $r \ge 1$. Lemma 3.9.1 gives $A_{n,r}$ choices for the quadratic part of polar rank $2r$. For each such part, Lemma 3.9.3 contributes

$$2^{2r+1}(2^{n-2r} - 1)$$

middle-weight words, and Lemma 3.9.4 contributes 2^{2r} words of weight

$$2^{n-1} - 2^{n-r-1}$$

and 2^{2r} words of weight

$$2^{n-1} + 2^{n-r-1}.$$

Multiplying by the number $A_{n,r}$ of quadratic parts gives the displayed non-middle multiplicities. Finally, summing the middle-weight contributions over all ranks gives exactly

$$M_n = \sum_{r=0}^{\lfloor n/2 \rfloor} A_{n,r}\, 2^{2r+1}(2^{n-2r} - 1).$$

There are no other possibilities: every degree-at-most-two word has exactly one quadratic rank $2r$, and Lemma 3.9.2 shows that its signed sum is either 0, $+2^{n-r}$, or -2^{n-r} according to the cases already counted.

The MacWilliams Step

Lemma 3.9.6 (Binary MacWilliams formula in Krawtchouk form). *Let $\mathcal{C} \leq \mathbb{F}_2^N$ be a binary linear code. For $0 \leq w \leq N$,*

$$W_w(\mathcal{C}^\perp) = \frac{1}{|\mathcal{C}|} \sum_{j=0}^{N} W_j(\mathcal{C})\, K_w^{(N)}(j).$$

Proof. For $x, y \in \mathbb{F}_2^N$, write $\chi_y(x) = (-1)^{\langle x,y \rangle}$, where the inner product is computed in $\mathbb{F}_2$ and then read as the exponent 0 or 1. First fix a word y of weight j. We compute the character sum over all words x of weight w. Let $S = \operatorname{supp}(y)$, so $|S| = j$. If $|\operatorname{supp}(x) \cap S| = t$, then $\langle x, y \rangle \equiv t \pmod 2$, and the character value is $(-1)^t$. The number of such x is $\binom{j}{t}\binom{N-j}{w-t}$, because one chooses t support positions inside S and $w - t$ support positions outside S. Summing over all possible t gives

$$\sum_{\substack{x \in \mathbb{F}_2^N \\ \mathrm{wt}(x) = w}} \chi_y(x) = \sum_{t=0}^{w} (-1)^t \binom{j}{t} \binom{N-j}{w-t} = K_w^{(N)}(j).$$

Next fix $x \in \mathbb{F}_2^N$. The sum

$$\sum_{y \in \mathcal{C}} \chi_y(x)$$

equals $|\mathcal{C}|$ if $x \in \mathcal{C}^\perp$, because then every term is 1. If $x \notin \mathcal{C}^\perp$, choose $y_0 \in \mathcal{C}$ with $\langle x, y_0 \rangle = 1$. The translation $y \mapsto y + y_0$ is a bijection of $\mathcal{C}$, and it changes $\chi_y(x)$ to $-\chi_y(x)$. Hence the sum is equal to its negative and therefore is 0. Thus

$$\mathbf{1}_{\mathcal{C}^\perp}(x) = \frac{1}{|\mathcal{C}|} \sum_{y \in \mathcal{C}} \chi_y(x).$$

Now sum this identity over all x of weight w:

$$\begin{aligned} W_w(\mathcal{C}^\perp) &= \sum_{\substack{x\in\mathbb{F}_2^N \\ \mathrm{wt}(x)=w}} \mathbf{1}_{\mathcal{C}^\perp}(x) \\ &= \frac{1}{|\mathcal{C}|} \sum_{\substack{x\in\mathbb{F}_2^N \\ \mathrm{wt}(x)=w}} \sum_{y\in\mathcal{C}} \chi_y(x) \\ &= \frac{1}{|\mathcal{C}|} \sum_{y\in\mathcal{C}} \sum_{\substack{x\in\mathbb{F}_2^N \\ \mathrm{wt}(x)=w}} \chi_y(x) \\ &= \frac{1}{|\mathcal{C}|} \sum_{y\in\mathcal{C}} K_w^{(N)}(\mathrm{wt}(y)) \\ &= \frac{1}{|\mathcal{C}|} \sum_{j=0}^{N} W_j(\mathcal{C})\, K_w^{(N)}(j). \end{aligned}$$

This is the desired identity.

Lemma 3.9.7 (The relevant Reed–Muller duality). *For $n \geq 3$,*

$$\mathrm{RM}(n-3,n) = \mathrm{RM}(2,n)^\perp.$$

Proof. For $S \subseteq \{0,\dots,n-1\}$, write

$$m_S(x) = \prod_{i\in S} x_i$$

for the corresponding evaluation word. The words m_S with $|S| \leq r$ form the standard basis of $\mathrm{RM}(r,n)$. This basis has size $\sum_{i=0}^{r} \binom{n}{i}$; independence follows from uniqueness of the ANF, equivalently from the usual Mobius inversion on the Boolean cube. Let $|S| \leq 2$ and $|T| \leq n-3$. Their inner product is

$$\langle m_S, m_T \rangle = \sum_{x\in\mathbb{F}_2^n} \prod_{i\in S\cup T} x_i \in \mathbb{F}_2.$$

If $S \cup T \neq \{0,\dots,n-1\}$, then at least one coordinate is free, so the number of points for which all variables in $S\cup T$ are 1 is even, and the sum is 0 in $\mathbb{F}_2$. Since

$$|S\cup T| \leq |S| + |T| \leq 2 + (n-3) = n-1,$$

this is always the case. Hence every basis word of $\mathrm{RM}(n-3,n)$ is orthogonal to every basis word of $\mathrm{RM}(2,n)$, and therefore

$$\mathrm{RM}(n-3,n) \subseteq \mathrm{RM}(2,n)^\perp.$$

It remains to compare dimensions. The dimension of $\mathrm{RM}(2,n)$ is

$$1 + n + \binom{n}{2}.$$

The dimension of $\mathrm{RM}(n-3,n)$ is

$$\sum_{i=0}^{n-3}\binom{n}{i} = 2^n - \binom{n}{n-2} - \binom{n}{n-1} - \binom{n}{n} = 2^n - \binom{n}{2} - n - 1.$$

Thus

$$\dim \mathrm{RM}(n-3,n) = 2^n - \dim \mathrm{RM}(2,n) = \dim\bigl(\mathrm{RM}(2,n)^{\perp}\bigr).$$

The inclusion of subspaces with equal finite dimension is equality.

Closed Formula for the Dual Distribution

Theorem 3.9.3 (Weight distribution of $\mathrm{RM}(n-3,n)$). *Let $n \geq 3$, put $N = 2^n$ and $k = 1 + n + \binom{n}{2}$, and define*

$$M_n = \sum_{r=0}^{\lfloor n/2 \rfloor} A_{n,r}\, 2^{2r+1}(2^{n-2r}-1).$$

For every $0 \leq w \leq N$, the number of words of weight w in $\mathrm{RM}(n-3,n)$ is

$$\boxed{\begin{aligned}&\frac{1}{2^k}\Bigg(K_w^{(N)}(0) + K_w^{(N)}(N) + M_n K_w^{(N)}(2^{n-1}) \\ &\qquad + \sum_{r=1}^{\lfloor n/2 \rfloor} A_{n,r} 2^{2r}\Big(K_w^{(N)}(2^{n-1}-2^{n-r-1}) + K_w^{(N)}(2^{n-1}+2^{n-r-1})\Big)\Bigg).\end{aligned}}$$

Proof. Let $\mathcal{C} = \mathrm{RM}(2,n)$. By Lemma 3.9.5, $|\mathcal{C}| = 2^k$. By Lemma 3.9.7, the code whose distribution is requested in Carlet's question is $\mathrm{RM}(n-3,n) = \mathcal{C}^{\perp}$. Therefore Lemma 3.9.6 gives

$$W_w(\mathrm{RM}(n-3,n)) = \frac{1}{2^k}\sum_{j=0}^{N} W_j(\mathrm{RM}(2,n)) K_w^{(N)}(j).$$

Now substitute the complete distribution of $\mathrm{RM}(2,n)$ from Theorem 3.9.2. The weights 0 and N contribute $K_w^{(N)}(0)$ and $K_w^{(N)}(N)$, respectively, because each occurs once. The middle weight 2^{n-1} contributes $M_n K_w^{(N)}(2^{n-1})$. For each $1 \leq r \leq \lfloor n/2 \rfloor$, the weights $2^{n-1}-2^{n-r-1}$ and $2^{n-1}+2^{n-r-1}$ both have multiplicity $A_{n,r}2^{2r}$. Hence their combined contribution is

$$A_{n,r}2^{2r}\Big(K_w^{(N)}(2^{n-1}-2^{n-r-1}) + K_w^{(N)}(2^{n-1}+2^{n-r-1})\Big).$$

Summing these contributions and dividing by 2^k gives exactly the displayed formula. The quotient is an integer because it is the cardinality $W_w(\mathrm{RM}(n-3,n))$.

Theorem 3.9.4 (Resolution of Carlet 13.2(3)). *Open Question 3.9.1 has the following all-n resolution. The number of balanced Boolean functions in n variables of algebraic degree at most two is the formula of Theorem 3.9.1. If quadratic is understood as exact algebraic degree two, the number is the formula of Corollary 3.9.1. For $n \geq 3$, the full weight distribution of the dual Reed–Muller code $\mathrm{RM}(n-3,n)$ is the formula of Theorem 3.9.3.*

Proof. The first sentence is Theorem 3.9.1. The exact-degree convention is obtained by deleting precisely the balanced affine functions, as proved in Corollary 3.9.1. The second code-theoretic component is Theorem 3.9.3, which applies for the natural range $n \geq 3$ in which the code $\mathrm{RM}(n-3,n)$ is defined with nonnegative order.

3.10 Restricted annihilators on a Hamming layer

Problem and threshold Carlet's list of open problems on Boolean functions includes, in Problem 13.7(7) of *Boolean Functions for Cryptography and Coding Theory*, a restricted algebraic-immunity question in which inputs are confined to a fixed Hamming layer. The problem is motivated by the standard $\lceil n/2 \rceil$ dimension bound for algebraic immunity, but asks what the corresponding threshold becomes when the truth table is observed only on vectors of Hamming weight k. In this layer-restricted model the relevant linear space is not the full degree-d polynomial space on $\mathbb{F}_2^n$, but the span of monomial evaluations restricted to $\Omega_{n,k}$.

Open Question 3.10.1 (Carlet 13.7(7)). *For $1 \leq k \leq n$, determine the least integer threshold $e = e(n,k)$ such that a dimension argument forces a nonzero restricted annihilator of either a Boolean function on the Hamming layer $\Omega_{n,k} = \{S \subseteq [n] : |S| = k\}$ or its complement. Equivalently, determine and estimate*

$$e(n,k) = \min\left\{ d \in \mathbb{N} : 2\binom{n}{d} > \binom{n}{k} \right\},$$

and describe its asymptotic behavior relative to the standard algebraic-immunity scale $D(n) = \lceil n/2 \rceil$.

This question is useful because it separates a purely combinatorial obstruction from the construction of special Boolean functions. A sharp or asymptotically sharp understanding of $e(n,k)$ gives the universal degree scale at which annihilators must appear on a layer, independently of the field and independently of the support pattern of the function. The result below proves the universal upper bound and shows that the threshold has the same first-order behavior as the symmetric layer distance $m = \min(k, n-k)$: after normalization by $\lceil n/2 \rceil$, no extra first-order loss is introduced by the dimension argument. The proof is intentionally field-independent. The Boolean function is used only to partition the Hamming layer, and the linear algebra is carried out over an arbitrary field. Let $\mathbb{N} = \{0, 1, \ldots\}$, let $1 \leq k \leq n$, and write

$$[n] = \{1, \ldots, n\}, \qquad \Omega_{n,k} = \{S \subseteq [n] : |S| = k\}.$$

For $d \in \mathbb{N}$, put $\mathcal{U}_d = \{T \subseteq [n] : |T| \le d\}$. If F is a field and $a = (a_T)_{T \in \mathcal{U}_d} \in F^{\mathcal{U}_d}$, define the induced function on the layer by

$$p_a(S) = \sum_{\substack{T \subseteq S \\ |T| \le d}} a_T, \qquad S \in \Omega_{n,k}.$$

Definition 3.10.1 (Restricted layer annihilator). *Let $f : \Omega_{n,k} \to \{0,1\}$. A nonzero coefficient vector $a \in F^{\mathcal{U}_d}$ is a restricted annihilator of f of degree at most d if $f(S) = 1 \implies p_a(S) = 0$ for every $S \in \Omega_{n,k}$. It is a restricted annihilator of $1 - f$ of degree at most d if the same condition holds on the zero-set of f. We denote by $\mathrm{RAI}_F(f; n, k)$ the least d for which one of these two annihilators exists.*

Define the combinatorial threshold

$$E(n,k) = \min\{0 \le d \le n : 2\binom{n}{d} > \binom{n}{k}\}, \qquad m = \min(k, n-k).$$

Proposition 3.10.1 (Layer-counting upper bound). *For every field F and every function $f : \Omega_{n,k} \to \{0,1\}$,*

$$\mathrm{RAI}_F(f; n, k) \le E(n,k).$$

Moreover, $E(n,k)$ is well-defined and $E(n,k) \le m$.

Proof. Since $\binom{n}{m} = \binom{n}{k}$, the index $d = m$ satisfies $2\binom{n}{m} = 2\binom{n}{k} > \binom{n}{k}$. Hence the defining set for $E(n,k)$ is nonempty and $E(n,k) \le m$. In particular, $E(n,k) \le k$. Write $E = E(n,k)$, let $\mathcal{T}_E = \{T \subseteq [n] : |T| = E\}$, and put $V_E = F^{\mathcal{T}_E}$. Define an F-linear map $L_E : V_E \to F^{\Omega_{n,k}}$ by

$$(L_E c)(S) = \sum_{\substack{T \subseteq S \\ |T| = E}} c_T.$$

Let $A_1 = \{S \in \Omega_{n,k} : f(S) = 1\}$ and $A_0 = \{S \in \Omega_{n,k} : f(S) = 0\}$. For $i \in \{0,1\}$, let $R_i : F^{\Omega_{n,k}} \to F^{A_i}$ be restriction to A_i. Suppose, for contradiction, that both maps $R_1 L_E$ and $R_0 L_E$ have trivial kernel. They are then injective, and therefore

$$\binom{n}{E} = \dim_F V_E \le |A_1|, \qquad \binom{n}{E} = \dim_F V_E \le |A_0|.$$

Adding the two inequalities gives

$$2\binom{n}{E} \le |A_0| + |A_1| = \binom{n}{k},$$

contradicting the definition of E. Thus one of the two kernels is nonzero. Choose $0 \ne c \in V_E$ in such a kernel. If $R_1 L_E(c) = 0$, then $L_E(c)$ vanishes on the support

of f; if $R_0 L_E(c) = 0$, then it vanishes on the zero-set of f. Embed c into a degree-at-most-E coefficient vector by setting

$$a_T = \begin{cases} c_T, & |T| = E, \\ 0, & |T| < E. \end{cases}$$

Then $a \neq 0$ and $p_a(S) = (L_E c)(S)$ for every $S \in \Omega_{n,k}$. Hence a is a restricted annihilator of either f or $1 - f$, and $\mathrm{RAI}_F(f; n, k) \leq E$.

Lemma 3.10.1 (Uniform binomial-loss estimate). *Assume $0 \leq m \leq n/2$. For every integer s with $0 \leq s \leq m$,*

$$(n + 1 + s^2)\binom{n}{m - s} \leq (n + 1)\binom{n}{m}.$$

Proof. We induct on s. The case $s = 0$ is equality. Assume $s + 1 \leq m$ and set

$$A = n + 1 + s^2, \qquad B = m - s, \qquad Q = 2s + 1, \qquad D = n - m + s + 1.$$

Since $m \leq n - m$, we have $B + Q = m + s + 1 \leq D$. Also $B \leq m \leq n \leq A$. Therefore

$$(A + Q)B = AB + QB \leq AB + QA = A(B + Q) \leq AD.$$

Equivalently,

$$(n + 1 + (s + 1)^2)(m - s) \leq (n + 1 + s^2)(n - m + s + 1).$$

Using

$$\binom{n}{m - s}(m - s) = \binom{n}{m - s - 1}(n - m + s + 1),$$

we obtain the one-step inequality

$$(n + 1 + (s + 1)^2)\binom{n}{m - s - 1} \leq (n + 1 + s^2)\binom{n}{m - s}.$$

Combining this with the induction hypothesis proves the claim.

Theorem 3.10.1. *Let $E = E(n, k)$ and $m = \min(k, n - k)$. Then*

$$(m - E)^2 < n + 1,$$

and in particular

$$(m - E(n, k))^2 \leq n + 1.$$

Consequently, for any sequence $1 \leq k_t \leq n_t$, if

$$D_t = \left\lceil \frac{n_t}{2} \right\rceil, \qquad m_t = \min(k_t, n_t - k_t), \qquad D_t \to \infty,$$

and if $m_t / D_t \to L$, then

$$\frac{E(n_t, k_t)}{D_t} \to L.$$

Proof. Since $m \le k$ and $m \le n-k$, we have $2m \le n$. Let

$$r = m - E.$$

By Proposition 3.10.1, $0 \le r \le m$. Apply Lemma 3.10.1 with $s = r$. Since $m - r = E$, this gives

$$(n+1+r^2)\binom{n}{E} \le (n+1)\binom{n}{m}.$$

On the other hand, $\binom{n}{m} = \binom{n}{k}$ and the definition of E gives

$$2\binom{n}{E} > \binom{n}{k} = \binom{n}{m}.$$

Hence

$$(n+1+r^2)\binom{n}{E} \le (n+1)\binom{n}{m} < 2(n+1)\binom{n}{E}.$$

Dividing by the positive integer $\binom{n}{E}$ yields $r^2 < n+1$, as claimed. For the asymptotic statement, write $E_t = E(n_t, k_t)$. The deficit bound gives

$$0 \le \left(\frac{m_t - E_t}{D_t}\right)^2 \le \frac{n_t + 1}{D_t^2}.$$

Since $D_t = \lceil n_t/2 \rceil$, we have $n_t \le 2D_t$. Also $D_t \ge 1$, so $n_t + 1 \le 3D_t$. Therefore

$$0 \le \left(\frac{m_t - E_t}{D_t}\right)^2 \le \frac{3}{D_t} \to 0.$$

It follows that $(m_t - E_t)/D_t \to 0$. Finally,

$$\frac{E_t}{D_t} = \frac{m_t}{D_t} - \frac{m_t - E_t}{D_t} \to L.$$

4 Applications in Symmetric Cryptanalysis and Design

This section studies three problems concerning diffusion and equivalence in symmetric cryptography. The first result concerns a contracting unbalanced Feistel network. The second constructs type-II GFN diffusion permutations, and the third counts expanded-equivalence classes of type-II GFN branch permutations. These contributions are presented alongside the Boolean-function results.

Table 3 fixes the notation used across the three subsections. Construction-specific objects are defined where they first occur.

Table 3. Common notation for the symmetric-cryptography applications.

Notation	Meaning
$\mathbb{F}_{2^n}$	Binary extension field with 2^n elements.
$\mathrm{wt}(x)$, $\mathrm{wt}(A)$	Hamming weight of a vector and entrywise weight of a binary matrix.
$\mathrm{Sym}(\Omega)$	Symmetric group on the finite set Ω.
P, Q, π	Branch permutations; products are composed from right to left.
$\mathrm{DR}(\pi)$	First round of full structural diffusion for π.
$\mathrm{DR}_{\max}(\pi)$	Maximum of the forward and inverse diffusion rounds.
$\lvert X\rvert$	Cardinality of a finite set X.

The chapter begins with the active-S-box conjecture because it illustrates a necessary distinction between a proved long-window bound and a stronger short-window statement. The proof gives all matrices and verifies every round of the counterexample.

4.1 Refuting the short-window dCUFN-SP bound

Problem Surface Bogdanov introduced d-branch contracting unbalanced Feistel networks with SP-type F-functions and contracting MDS diffusion, denoted dCUFN-SP, in the course of studying provable resistance of generalized Feistel constructions to differential and linear cryptanalysis [7]. The F-function of round i is

$$F_i(x_i, \ldots, x_{i+d-2}, k_i) = \theta_i\big(\gamma_i((x_i, \ldots, x_{i+d-2}) \oplus k_i)\big),$$

where γ_i is a bricklayer of $(d-1)m$ bijective n-bit S-boxes and θ_i is a contracting $\mathbb{F}_{2^n}$-linear map represented by an $m \times (d-1)m$ matrix M_i. Here d is the number of Feistel branches, n is the cell width, and m is the number of n-bit cells in one branch. Equivalently, if the whole block has b bits, then one branch has $b/d = mn$ bits and

$$m = \frac{b}{nd}.$$

Thus one branch difference $X_i = \Delta x_i$ is an element of $(\mathbb{F}_{2^n})^m$, and its weight $D_i = \mathrm{wt}(X_i)$ is an integer between 0 and m. In the paper's notation $n > 2$ and $m \geq 2$. For a bundle $u = (u_1, \ldots, u_c) \in (\mathbb{F}_{2^n})^c$, write

$$\mathrm{wt}(u) = |\{j : u_j \neq 0\}|.$$

For a linear map $\theta : (\mathbb{F}_{2^n})^c \to (\mathbb{F}_{2^n})^m$, its branch number is

$$\mathrm{B}(\theta) = \min_{u \neq 0}\big(\mathrm{wt}(u) + \mathrm{wt}(\theta(u))\big).$$

Bogdanov calls θ a linear optimal diffusion mapping, or ODM, when $\mathrm{B}(\theta) = m + 1$. Proposition 4 of the DCC paper proves the following small-d, longer-window differential lower bound. For two $m \times (d-1)m$ diffusion matrices A and B, write $[A \mid B]$ for their horizontal concatenation. Thus $[A \mid B]$ is an $m \times 2(d-1)m$ matrix and defines the linear map

$$(u, v) \in (\mathbb{F}_{2^n})^{(d-1)m} \times (\mathbb{F}_{2^n})^{(d-1)m} \longmapsto Au + Bv \in (\mathbb{F}_{2^n})^m.$$

This is the "double-matrix" object used in Bogdanov's Proposition 3 and Proposition 4 hypotheses. In the present note, the phrase two-matrix block always refers to this horizontal concatenation; it is not a plaintext block or a Feistel state block.

Source Fact 1 (Bogdanov DCC 2011, Proposition 4). For $d = 3$ and $d = 4$, if every two-matrix block $[M_i \mid M_{i+d}]$ is an ODM, then every $2(d+1)$ consecutive rounds of dCUFN-SP contain at least $2(d-1)(m+1)$ differentially active S-boxes [7, Prop. 4].

The DCC article adds a footnote to this proposition. That footnote keeps the same $2(d+1)$-round window and conjectures only the extension in the branch parameter d.

Open Question 4.1.1 (Bogdanov DCC 2011 footnote, $2(d+1)$-round extension). *The statement of Source Fact 1 also holds for $d > 4$; equivalently, under the same two-matrix ODM hypothesis, every $2(d+1)$ consecutive rounds contain at least $2(d-1)(m+1)$ differentially active S-boxes for all $d > 4$ [7, footnote after Prop. 4].*

The doctoral thesis then records a different, stronger form immediately after Proposition 4 in Chapter 3, Section 3.4.3, on printed page 46 [6, p. 46]. This thesis statement changes the window length from $2(d+1)$ to $d+1$ and halves the right-hand side accordingly.

Open Question 4.1.2 (Bogdanov thesis, stronger $d+1$-round form). *Under the same two-matrix ODM hypothesis as in Source Fact 1, every $d+1$ consecutive rounds of dCUFN-SP contain at least $(d-1)(m+1)$ differentially active S-boxes.*

The rest of this note gives a counterexample to Conjecture 4.1.2. The notation is aligned with Bogdanov's difference-weight proof surface: X_i denotes the m-cell branch difference Δx_i, $D_i = \mathrm{wt}(X_i)$, and round i takes the two trailing branch differences (X_{i+1}, X_{i+2}) as the F-function input when $d = 3$. Thus the differential output of the F-function in round i must equal $X_i + X_{i+3}$.

A Useful ODM Criterion for $m = 2$ All arithmetic below is over $\mathbb{F} = \mathbb{F}_{16} = \mathbb{F}_2[\alpha]/(\alpha^4 + \alpha + 1)$. For $c \geq 3$ and pairwise distinct nonzero elements $s_1, \ldots, s_c \in \mathbb{F}^\times$, define

$$N(s_1, \ldots, s_c) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ s_1 & s_2 & \cdots & s_c \end{pmatrix}.$$

Lemma 4.1.1. *The linear map $N(s_1, \ldots, s_c) : \mathbb{F}^c \to \mathbb{F}^2$ is an ODM; equivalently,* $\mathrm{B}(N(s_1, \ldots, s_c)) = 3$.

Proof. Let $u \in \mathbb{F}^c$ be nonzero. If $\mathrm{wt}(u) = 1$, then

$$N(s_1, \ldots, s_c)u = u_j(1, s_j)^T$$

for the unique nonzero coordinate u_j. Since $u_j \neq 0$ and $s_j \neq 0$, the output has weight 2, so $\mathrm{wt}(u) + \mathrm{wt}(Nu) = 3$. If $\mathrm{wt}(u) = 2$, say the support is $\{a, b\}$, then the two active columns $(1, s_a)^T$ and $(1, s_b)^T$ are linearly independent because

$$\det \begin{pmatrix} 1 & 1 \\ s_a & s_b \end{pmatrix} = s_b + s_a \neq 0.$$

Hence $Nu \neq 0$, so $\mathrm{wt}(u) + \mathrm{wt}(Nu) \geq 3$. If $\mathrm{wt}(u) \geq 3$, the same inequality is immediate. Therefore $\mathrm{B}(N) \geq 3$. Conversely, $c \geq 3$ columns in the two-dimensional codomain are linearly dependent. There is a nonzero vector u supported on at most three coordinates with $Nu = 0$, so $\mathrm{B}(N) \leq 3$. Thus $\mathrm{B}(N) = 3 = m + 1$.

This criterion will be applied both to 2×4 single-round matrices M_i and to 2×8 two-matrix blocks $[M_i \mid M_{i+3}]$.

The Period-Six Trail Set $d = 3$, $m = 2$, and let the branch-difference sequence be 6-periodic, so that $X_{i+6} = X_i$, with

i	0	1	2	3	4	5
X_i	$(1,1)$	$(1,0)$	$(0,0)$	$(0,0)$	$(1,0)$	$(1,1)$
D_i	2	1	0	0	1	2.

There are no three consecutive zero branch differences, so every full three-branch state window is nonzero. The choice $m = 2$ means that each branch X_i has two $\mathbb{F}_{16}$-cells. Since $d = 3$, the F-function input in each round consists of the two trailing branches and therefore has $(d-1)m = 4$ S-box positions. Take every S-box to be the inverse permutation on $\mathbb{F}$, with $S(0) = 0$. The zero input difference gives zero output difference. The active transition $1 \mapsto 1$ is also allowed: with $\omega = \alpha^5 = \alpha^2 + \alpha$, one has $\omega^2 + \omega + 1 = 0$, hence

$$S(\omega + 1) + S(\omega) = \omega + (\omega + 1) = 1.$$

We use output difference 1 at every active S-box. The coordinate permutation in the SP layer is the identity. Thus the vector entering the linear layer in round i is exactly

$$I_i = (X_{i+1}, X_{i+2}) \in \mathbb{F}^4.$$

The required F-output is

$$Y_i = X_i + X_{i+3} \in \mathbb{F}^2.$$

Define the six round matrices, extended periodically by $M_{i+6} = M_i$, as follows:

$$M_0 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \alpha & \alpha^2 & \alpha^3 \end{pmatrix}, \qquad M_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha+1 & \alpha^2+1 & \alpha^2+\alpha & \alpha^3 \end{pmatrix},$$

$$M_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha & \alpha^2 & 1 & \alpha^3 \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha+1 & \alpha^2+1 & \alpha^3+1 & \alpha^3+\alpha+1 \end{pmatrix},$$

$$M_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \alpha & \alpha^2 & \alpha^2+\alpha+1 \end{pmatrix}, \quad M_5 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \alpha+1 & \alpha^2+1 & \alpha^2+\alpha+1 & \alpha^2+\alpha \end{pmatrix}.$$

Lemma 4.1.2. *Each M_i is an ODM. Moreover, each two-matrix block $[M_i \mid M_{i+3}]$, with indices read modulo 6, is an ODM.*

Proof. Each displayed matrix has the form $N(s_1, s_2, s_3, s_4)$, and in each row of slopes the four elements are nonzero and pairwise distinct. Lemma 4.1.1 therefore proves that every M_i is an ODM. For the two-matrix block condition it is enough, by periodicity, to check the three pairs (M_0, M_3), (M_1, M_4), and (M_2, M_5). The reason is that the hypotheses use $[M_i \mid M_{i+3}]$, and the sequence of matrices is 6-periodic, so the remaining pairs repeat these three. For a matrix of the form

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ s_1 & s_2 & \cdots & s_c \end{pmatrix},$$

we call $\{s_1, \ldots, s_c\}$ its set of slopes. This terminology is only a compact way to refer to the second-row entries: the j-th column is $(1, s_j)^T$, so distinct slopes are exactly what make any two columns linearly independent. For a two-matrix block $[M_a \mid M_b]$, the slope set is obtained by concatenating the second row of M_a with the second row of M_b. Thus

$$\begin{aligned} S_{0,3} &= \{1, \alpha, \alpha^2, \alpha^3, \alpha+1, \alpha^2+1, \alpha^3+1, \alpha^3+\alpha+1\}, \\ S_{1,4} &= \{\alpha+1, \alpha^2+1, \alpha^2+\alpha, \alpha^3, 1, \alpha, \alpha^2, \alpha^2+\alpha+1\}, \\ S_{2,5} &= \{\alpha, \alpha^2, 1, \alpha^3, \alpha+1, \alpha^2+1, \alpha^2+\alpha+1, \alpha^2+\alpha\}. \end{aligned}$$

For example, $S_{0,3}$ is the second row of M_0,

$$(1, \alpha, \alpha^2, \alpha^3),$$

followed by the second row of M_3,

$$(\alpha+1, \alpha^2+1, \alpha^3+1, \alpha^3+\alpha+1).$$

The two other displayed sets are obtained in the same way from (M_1, M_4) and (M_2, M_5). In each line the eight field elements are nonzero and pairwise distinct. Hence each corresponding two-matrix block is of the form $N(s_1, \ldots, s_8)$ with pairwise distinct nonzero slopes. Lemma 4.1.1 applies again and gives branch number $3 = m + 1$.

Period-Six Feistel Verification

Lemma 4.1.3. *Extend the branch sequence and the round matrices by*

$$X_{i+6} = X_i, \qquad M_{i+6} = M_i \qquad (i \geq 0).$$

For every round $i \geq 0$, put

$$I_i = (X_{i+1}, X_{i+2}) \in \mathbb{F}^4, \qquad Y_i = M_i I_i^T \in \mathbb{F}^2.$$

Then

$$Y_i = X_i + X_{i+3} \qquad \text{for every } i \geq 0.$$

Consequently the $d = 3$ Feistel differential update maps

$$(X_i, X_{i+1}, X_{i+2}) \longmapsto (X_{i+1}, X_{i+2}, X_i + Y_i) = (X_{i+1}, X_{i+2}, X_{i+3})$$

in every round. Hence the listed branch-difference sequence is a valid period-six Feistel differential trail for the specified round matrices.

Proof. Since both X_i and M_i are 6-periodic, it is enough to verify the six residue classes $i = 0, \ldots, 5$. The required identities

$$M_i(X_{i+1}, X_{i+2})^T = X_i + X_{i+3}$$

are exactly the following:

i	$I_i = (X_{i+1}, X_{i+2})$	$X_i + X_{i+3}$	$M_i I_i^T$
0	$(1,0,0,0)$	$(1,1)$	$(1,1)$
1	$(0,0,0,0)$	$(0,0)$	$(0,0)$
2	$(0,0,1,0)$	$(1,1)$	$(1,1)$
3	$(1,0,1,1)$	$(1,1)$	$(1,1)$
4	$(1,1,1,1)$	$(0,0)$	$(0,0)$
5	$(1,1,1,0)$	$(1,1)$	$(1,1)$.

The first three rows follow directly from the first, zero, and third active columns, respectively. For round 3, the active columns are $1, 3, 4$; their first coordinates sum to $1 + 1 + 1 = 1$, and their second coordinates sum to

$$(\alpha + 1) + (\alpha^3 + 1) + (\alpha^3 + \alpha + 1) = 1.$$

For round 4, the full input is active and

$$1 + 1 + 1 + 1 = 0, \qquad 1 + \alpha + \alpha^2 + (\alpha^2 + \alpha + 1) = 0.$$

For round 5, the active columns are $1, 2, 3$, and

$$1 + 1 + 1 = 1, \qquad (\alpha + 1) + (\alpha^2 + 1) + (\alpha^2 + \alpha + 1) = 1.$$

All sums are in characteristic two. Therefore $M_i I_i^T = X_i + X_{i+3}$ for $i = 0, \ldots, 5$, and periodicity gives the same equality for every $i \geq 0$. In a 3-branch contracting Feistel round, the first two output branches are the shifted input branches

(X_{i+1}, X_{i+2}), and the last output branch is the old first branch plus the F-output difference Y_i. Thus the output state is

$$(X_{i+1}, X_{i+2}, X_i + Y_i) = (X_{i+1}, X_{i+2}, X_{i+3}).$$

After six rounds the state returns to

$$(X_{i+6}, X_{i+7}, X_{i+8}) = (X_i, X_{i+1}, X_{i+2}),$$

so the recurrence realizes the claimed period-six trail.

The two structurally important cases in the table are worth isolating. In round 1, the F-input is zero; the F-output is zero, and the Feistel equation forces $X_4 = X_1$. In round 4, the F-input is $(1, 1, 1, 1)$, but the chosen ODM has this full-weight vector in its kernel. This is compatible with branch number 3: ODM excludes kernel vectors of weight 1 and 2, not kernel vectors of weight 4.

Theorem 4.1.1. *The $d + 1$-round active-S-box bound in Conjecture 4.1.2 is false.*

Proof. The construction above is a valid period-six dCUFN-SP difference trail under the source propagation semantics. The S-box transitions are valid for the inverse S-box specified above, the linear layers are ODMs by Lemma 4.1.2, the two-matrix restrictions $[M_i \mid M_{i+3}]$ are ODMs by the same lemma, and the Feistel recurrence is verified in Lemma 4.1.3. For $d = 3$, round i has

$$A_i = \mathrm{wt}(X_{i+1}) + \mathrm{wt}(X_{i+2}) = D_{i+1} + D_{i+2}$$

active S-boxes. In the first four rounds,

i	0	1	2	3
A_i	1	0	1	3.

Therefore

$$\sum_{i=0}^{3} A_i = 1 + 0 + 1 + 3 = 5.$$

However, for $d = 3$ and $m = 2$, the conjectured lower bound is

$$(d - 1)(m + 1) = 2 \cdot 3 = 6.$$

The four-round window $0, 1, 2, 3$ is a $d + 1$-round window with fewer active S-boxes than the conjectured lower bound. This refutes the conjecture.

4.2 Finding ten-round 38- and 40-branch diffusion permutations

The Open Challenge A type-II Generalised Feistel Network (GFN) on $2k$ branches updates the left member of every adjacent pair and then permutes all branches. The structural diffusion round measures how many rounds are needed

before every output branch depends on every input branch. Following the two-sided convention of the source paper, the design criterion is the maximum of the diffusion rounds computed for the branch permutation and its inverse. Appendix A.1 gives the complete round transformation, explains the meaning of a $2k$-branch state, and illustrates the type-II update. Derbez, Fouque, Lambin, and Mollimard developed an efficient search for even–odd permutations and determined tight values for many branch sizes [19]. For $k = 19, 20, 21$, corresponding to 38, 40, and 42 branches, they proved that nine rounds are impossible and found permutations with diffusion round eleven. They stated the remaining task explicitly: find a permutation with diffusion round ten. We address the first two unresolved sizes.

Open Question 4.2.1. *For each $k \in \{19, 20\}$, does there exist an even–odd permutation π on $2k$ branches such that $\mathrm{DR}_{\max}(\pi) = 10$? Here $\mathrm{DR}_{\max}$ includes full diffusion in both the forward and inverse directions.*

The lower bound from [19] shows that a positive answer is optimal. It remains to provide one witness for each size and to verify both against exactly the same round semantics.

Diffusion Semantics Let $\Omega_k = \{0, 1, \ldots, 2k-1\}$. The positions $(2i, 2i+1)$ form the i-th Feistel pair. An even–odd permutation is determined by $p, q \in \mathrm{Sym}(\{0, \ldots, k-1\})$ through

$$\pi(2i) = 2p(i) + 1, \qquad \pi(2i+1) = 2q(i). \tag{4.2.1}$$

This is the convention of the source paper. For $\rho \in \mathrm{Sym}(\Omega_k)$, let $D_j^{(r)}(\rho) \subseteq \Omega_k$ be the set of input positions on which output position j depends after r rounds. Initially,

$$D_j^{(0)}(\rho) = \{j\}. \tag{4.2.2}$$

The type-II Feistel layer updates the left branch of each pair and leaves the right branch unchanged. If $\widetilde{D}^{(r)}$ denotes the dependency state immediately before the branch permutation, then

$$\widetilde{D}_{2i}^{(r)} = D_{2i}^{(r)}(\rho) \cup D_{2i+1}^{(r)}(\rho), \tag{4.2.3}$$

$$\widetilde{D}_{2i+1}^{(r)} = D_{2i+1}^{(r)}(\rho), \tag{4.2.4}$$

$$D_{\rho(s)}^{(r+1)}(\rho) = \widetilde{D}_s^{(r)}. \tag{4.2.5}$$

Equations (4.2.2)–(4.2.5) are a complete finite specification of the structural computation. They do not depend on the branch width or on the internal choice of round functions. We say that ρ is fully diffused after r rounds when

$$D_j^{(r)}(\rho) = \Omega_k \quad \text{for every } j \in \Omega_k.$$

Write $\mathrm{DR}(\rho)$ for the first such round. Following Derbez et al.,

$$\mathrm{DR}_{\max}(\rho) = \max\{\mathrm{DR}(\rho), \mathrm{DR}(\rho^{-1})\}. \tag{4.2.6}$$

Table 4. First ten entries of the witness permutations.

i	0	1	2	3	4	5	6	7	8	9
$p(i)$	9	6	17	13	1	3	16	7	18	4
$q(i)$	13	17	3	12	11	1	7	16	9	5

Table 5. Remaining entries of the witness permutations.

i	10	11	12	13	14	15	16	17	18
$p(i)$	14	15	5	8	0	10	11	12	2
$q(i)$	0	18	10	15	14	6	2	4	8

Once a dependency state is full, the union operation preserves fullness. Thus full diffusion at round r, together with non-full diffusion at round $r-1$, proves that the first full-diffusion round is exactly r.

The 38-Branch Witness Set $k = 19$. The two permutations below are written in image-list form, so the entry in column i is the image of i. Both rows contain every integer in $\{0, \dots, 18\}$ exactly once, hence they define permutations. Let $\pi \in \mathrm{Sym}(\Omega_{19})$ be obtained from them by (4.2.1). For a compact summary of the dependency computation, define

$$m_r(\rho) = \min_j |D_j^{(r)}(\rho)|, \quad M_r(\rho) = \max_j |D_j^{(r)}(\rho)|, \quad T_r(\rho) = \sum_j |D_j^{(r)}(\rho)|.$$

Applying (4.2.2)–(4.2.5) gives Table 6. Together with the displayed witness and the recurrence, the table is a reproducible finite certificate. The total $38^2 = 1444$ in the last column of each direction confirms that every one of the 38 dependency sets has size 38. The round-nine failure is also recorded by explicit missing dependencies:

$$\{26, 32\} \cap D_0^{(9)}(\pi) = \varnothing,$$
$$\{0, 4, 34\} \cap D_0^{(9)}(\pi^{-1}) = \varnothing.$$

In fact 17 forward outputs and 18 inverse outputs are not full at round nine; the two displayed rows are sufficient witnesses.

Theorem 4.2.1 (Resolution of the 38-branch challenge). *For the permutations $p, q \in \mathrm{Sym}(\{0, \dots, 18\})$ in Tables 4 and 5, let π be defined by* (4.2.1). *Then*

$$\mathrm{DR}(\pi) = \mathrm{DR}(\pi^{-1}) = \mathrm{DR}_{\max}(\pi) = 10.$$

Consequently, the open challenge in Question 4.2.1 has an affirmative answer for $k = 19$.

Table 6. Dependency-set certificate for the witness and its inverse.

	π			π^{-1}		
round r	m_r	M_r	T_r	m_r	M_r	T_r
0	1	1	38	1	1	38
1	1	2	57	1	2	57
2	2	3	95	2	3	95
3	3	5	152	3	5	152
4	5	8	247	5	8	247
5	8	13	398	8	13	398
6	12	21	625	12	21	625
7	18	32	924	18	32	924
8	25	38	1223	26	38	1223
9	33	38	1400	34	38	1400
10	38	38	1444	38	38	1444

Proof. Equations (4.2.2)–(4.2.5), evaluated with the displayed p and q, give Table 6. At round ten the minimum dependency-set cardinality is 38 in both directions. Since every dependency set is a subset of the 38-element set Ω_{19}, all of them equal Ω_{19}. Neither direction is full at round nine, as shown by the explicit missing dependencies above. Fullness is preserved by subsequent rounds, so the first full round is ten in both directions. Equation (4.2.6) now gives $\mathrm{DR}_{\max}(\pi) = 10$.

Theorem 4.2.1 is witness-level: its proof is the finite evaluation of a stated recurrence. Global optimality within the even–odd class follows by combining the theorem with the source paper's proof that no $k = 19$ even–odd permutation can reach full diffusion in nine rounds [19].

The 40-Branch Witness Set $k = 20$. As above, the entry in column i is the image of i.

Table 7. First ten entries of the 40-branch witness permutations.

i	0	1	2	3	4	5	6	7	8	9
$p(i)$	11	13	7	2	8	0	14	18	16	3
$q(i)$	13	11	1	9	18	6	10	8	4	5

Table 8. Remaining entries of the 40-branch witness permutations.

i	10	11	12	13	14	15	16	17	18	19
$p(i)$	17	6	1	15	5	19	10	4	9	12
$q(i)$	19	0	7	12	3	17	14	16	2	15

Each row contains every element of $\{0, \ldots, 19\}$ exactly once. Thus both rows define permutations. Let $\pi_{40} \in \mathrm{Sym}(\Omega_{20})$ be their even–odd permutation under (4.2.1). Applying the dependency recurrence from Section 4.2 gives Table 9. The forward and inverse summaries agree at every round, although the underlying dependency sets need not agree.

Table 9. Dependency-set certificate for π_{40} and π_{40}^{-1}.

round r	π_{40}			π_{40}^{-1}		
	m_r	M_r	T_r	m_r	M_r	T_r
0	1	1	40	1	1	40
1	1	2	60	1	2	60
2	2	3	100	2	3	100
3	3	5	160	3	5	160
4	5	8	260	5	8	260
5	8	13	420	8	13	420
6	13	21	676	13	21	676
7	20	31	1004	20	31	1004
8	27	38	1320	27	38	1320
9	35	40	1532	35	40	1532
10	40	40	1600	40	40	1600

The two directions are not fully diffused at round nine. More precisely, $\{2, 24\} \cap D_0^{(9)}(\pi_{40}) = \varnothing$ and $\{20, 32\} \cap D_0^{(9)}(\pi_{40}^{-1}) = \varnothing$.

Theorem 4.2.2 (40 branches). *Let p, q be the permutations displayed above, and let π_{40} be defined by* (4.2.1)*. Then*

$$\mathrm{DR}(\pi_{40}) = \mathrm{DR}(\pi_{40}^{-1}) = \mathrm{DR}_{\max}(\pi_{40}) = 10.$$

Consequently, Question 4.2.1 also has an affirmative answer for $k = 20$.

Proof. The displayed rows are permutations, so (4.2.1) defines $\pi_{40} \in \mathrm{Sym}(\Omega_{20})$. Applying (4.2.2)–(4.2.5) yields Table 9. At round 10, the minimum dependency-set cardinality is 40 for both π_{40} and π_{40}^{-1}. Every such set is contained in the 40-element set Ω_{20}, hence every output depends on all inputs in both directions after ten rounds. The two explicit empty intersections above show that output position zero fails to depend on all inputs at round nine in either direction. Therefore $\mathrm{DR}(\pi_{40}) = \mathrm{DR}(\pi_{40}^{-1}) = 10$, and (4.2.6) yields $\mathrm{DR}_{\max}(\pi_{40}) = 10$.

The source paper proves that nine rounds are impossible for $k = 20$ within the even–odd class [19]. Theorem 4.2.2 therefore attains the optimal diffusion round in that class.

4.3 Counting expanded-equivalence classes

Problem Surface Type-II GFNs are specified, among other choices, by a permutation of the $2k$ branches. Derbez and Euler write this permutation as $P \in S_{2k}$ and use the round layer $P \circ F_i^k$ for the Feistel step F_i^k [18]. Their expanded-equivalence relation identifies two branch permutations when the associated GFNs differ only by a round-compatible relabelling of ordered input/output pairs. For the even-odd subfamily, where even branch indices are sent to odd branch indices and odd branch indices are sent to even branch indices, they prove a sharp class-size result: the $(k!)^2$ even-odd permutations decompose into $k!$ classes of size $k!$. The conclusion of their paper asks for the corresponding class sizes when non-even-odd permutations are also considered.

Open Question 4.3.1 (Expanded-equivalence classes). *For type-II GFNs with $2k$ branches, let expanded equivalence be defined by the existence of a pair relabelling A such that*

$$Q^n A P^{-n}$$

is again a pair relabelling for every positive integer n. When P and Q range over all permutations of the $2k$ branches, not only over the even-odd permutations, determine the number of expanded-equivalence classes and the cardinality of each class.

We solve this class-size and quotient-count component for the unrestricted domain S_{2k}. The theorem below does not claim a new GFN with improved diffusion, a new active-S-box bound, or a security-preservation theorem for all attack models. It is the exact finite-group counting statement underlying the expanded-equivalence relation.

Definitions

Branch positions and products of permutations. For an integer $k \geq 1$, write $[k] = \{0, \ldots, k-1\}$ and $\Omega_k = \{0, \ldots, 2k-1\}$. For a finite set Y, $\mathrm{Sym}(Y)$ denotes the symmetric group of all bijections $Y \to Y$. We write products of permutations from right to left, so $(UV)(x) = U(V(x))$. Thus $Q^n A P^{-n}$ denotes the composed permutation

$$Q^n \circ A \circ (P^n)^{-1}$$

on branch positions. It is not a product of branch values.

Branch permutations. A $2k$-branch permutation is an element

$$P \in S_{2k} := \mathrm{Sym}(\Omega_k).$$

In a type-II GFN this is the public shuffle in the round layer $P \circ F_i^k$; the counting problem considered here depends only on this permutation layer and on the relabellings that commute with the Feistel step. The branch positions are grouped into ordered pairs

$$(2i, 2i+1), \qquad i \in [k].$$

Equivalently, one may identify Ω_k with $[k] \times \{0,1\}$ by $(i,b) \mapsto 2i+b$.

Even-odd and parity-preserving permutations. Following Derbez–Euler, a branch permutation $P \in S_{2k}$ is *even-odd* if it sends every even position to an odd position and every odd position to an even position:

$$P(2i) \equiv 1 \pmod 2, \qquad P(2i+1) \equiv 0 \pmod 2 \qquad (i \in [k]).$$

Equivalently, there are unique $L, R \in \mathrm{Sym}([k])$ such that

$$P = \Pi_{L,R}, \qquad \Pi_{L,R}(2i) = 2L(i)+1, \quad \Pi_{L,R}(2i+1) = 2R(i).$$

We denote the set of such permutations by EO_k; it has cardinality $(k!)^2$. A *parity-preserving* permutation, called an even permutation in [18], is one of the form

$$\Phi_{L,R}(2i) = 2L(i), \qquad \Phi_{L,R}(2i+1) = 2R(i)+1 \qquad (L, R \in \mathrm{Sym}([k])).$$

We write Ev_k for this parity-preserving subset of S_{2k}. The open question addressed here is not restricted to EO_k: P and Q range over all of S_{2k}, including permutations that are not even-odd.

Permutations of pairs. The subgroup of ordered pair relabellings is

$$S_k^{\mathrm{p}} = \{\Phi_{\tau,\tau} : \tau \in \mathrm{Sym}([k])\} \le S_{2k}.$$

Equivalently, $A \in S_k^{\mathrm{p}}$ if and only if there exists $\tau \in \mathrm{Sym}([k])$ such that

$$A(2i) = 2\tau(i), \qquad A(2i+1) = 2\tau(i)+1 \qquad (i \in [k]).$$

Thus A maps the ordered pair $(2i, 2i+1)$ to the ordered pair $(2\tau(i), 2\tau(i)+1)$. This is the subgroup denoted S_k^p in Derbez–Euler; equivalently, it is the set of branch permutations satisfying

$$A(2i+1) = A(2i)+1 \qquad \text{for every } i \in [k].$$

In particular $|S_k^{\mathrm{p}}| = k!$. Under the identification $\Omega_k \simeq [k] \times \{0,1\}$, the same subgroup is

$$H_k = \{\widehat{\tau} : \tau \in \mathrm{Sym}([k])\} \le \mathrm{Sym}([k] \times \{0,1\}), \qquad \widehat{\tau}(i,b) = (\tau(i), b).$$

This is the same ordered-pair relabelling subgroup in product coordinates.

Definition 4.3.1 (Branch expanded equivalence). *Let $P, Q \in S_{2k}$. We say that P and Q are* expanded-equivalent*, and write $P \sim_{\exp} Q$, if there exists $A \in S_k^{\mathrm{p}}$ such that*

$$Q^n A P^{-n} \in S_k^{\mathrm{p}} \qquad \text{for every positive integer } n.$$

The witness A is an initial relabelling of ordered pairs; the condition says that after each positive number of rounds the induced relabelling is still an ordered pair relabelling.

Definition 4.3.2 (Expanded equivalence). *Let G be a group and $H \leq G$. For $P, Q \in G$, write $P \sim_{\exp} Q$ if there exists $A \in H$ such that*

$$Q^n A P^{-n} \in H \qquad \text{for every } n \geq 1.$$

Definition 4.3.2 is the group-theoretic form of Definition 4.3.1. In the GFN specialization, one takes $G = S_{2k}$ and $H = S_k^{\mathrm{p}}$, or equivalently $G = \mathrm{Sym}([k] \times \{0,1\})$ and $H = H_k$.

The Counting Theorem The proof is slightly cleaner in a finite group G with an arbitrary subgroup H. The GFN statement is then obtained by substituting $G = S_{2k}$ and $H = S_k^{\mathrm{p}}$. Equivalently, one may use $G = \mathrm{Sym}([k] \times \{0,1\})$ and $H = H_k$.

Lemma 4.3.1 (Expanded relation as an intertwining trace). *Let G be a group, $H \leq G$, and $P, Q \in G$. The following are equivalent:*

$$\exists A \in H,\ \forall n \geq 1,\ Q^n A P^{-n} \in H,$$

and

$$\exists (S_n)_{n \geq 0} \subseteq H, \quad Q S_n = S_{n+1} P \quad \text{for every } n \geq 0.$$

Proof. If A satisfies the expanded condition, set $S_0 = A$ and $S_n = Q^n A P^{-n}$ for $n \geq 1$. Then $S_n \in H$ and

$$Q S_n = Q^{n+1} A P^{-n} = S_{n+1} P.$$

Conversely, suppose such a sequence S_n is given. Induction on n gives

$$S_n = Q^n S_0 P^{-n}$$

for all $n \geq 0$. Taking $A = S_0 \in H$ proves the expanded condition.

For a fixed $P \in G$, define the stable-increment set

$$\mathcal{T}_P = \{R \in H : P^n R P^{-n} \in H \text{ for all } n \geq 0\}.$$

It is nonempty because $1 \in \mathcal{T}_P$.

Lemma 4.3.2 (Stable-increment normal form). *Let G be a group, $H \leq G$, and $P, Q \in G$. The intertwining-trace condition of Lemma 4.3.1 holds if and only if there exist $S \in H$ and $R \in \mathcal{T}_P$ such that*

$$Q = SRPS^{-1}.$$

Proof. Assume first that $QS_n = S_{n+1}P$ with all $S_n \in H$. Put

$$S = S_0, \qquad R = S_0^{-1}S_1.$$

Then $R \in H$, and a direct calculation from the recurrence gives

$$S_n^{-1}S_{n+1} = P^n R P^{-n}.$$

Since $S_n, S_{n+1} \in H$, this shows $R \in \mathcal{T}_P$. At $n = 0$, the same recurrence gives

$$Q = S_1 P S_0^{-1} = S_0 R P S_0^{-1}.$$

Conversely, suppose $Q = SRPS^{-1}$, with $S \in H$ and $R \in \mathcal{T}_P$. Define

$$B_n = (RP)^n P^{-n}.$$

Since $B_{n+1} = B_n(P^n R P^{-n})$, induction gives $B_n \in H$. Put $S_n = SB_n$. Then $S_n \in H$, and

$$QS_n = SRPS^{-1}SB_n = SRPB_n = SB_{n+1}P = S_{n+1}P.$$

For fixed P, let

$$\mathcal{C}(P) = \{Q \in G : P \sim_{\exp} Q\}.$$

By Lemmas 4.3.1 and 4.3.2, $\mathcal{C}(P)$ is the image of

$$\Phi_P : H \times \mathcal{T}_P \longrightarrow G, \qquad \Phi_P(S, R) = SRPS^{-1}.$$

Lemma 4.3.3 (Constant fibers). *Assume G is finite. For every $P \in G$ and every $(S, R) \in H \times \mathcal{T}_P$, the fiber of Φ_P above $\Phi_P(S, R)$ has cardinality $|\mathcal{T}_P|$.*

Proof. We describe the fiber explicitly. Let $T \in \mathcal{T}_P$. Define

$$S' = ST, \qquad R' = T^{-1}RPTP^{-1}.$$

Then $S' \in H$, and $R' \in \mathcal{T}_P$ because

$$P^n R' P^{-n} = (P^n T^{-1} P^{-n})(P^n R P^{-n})(P^{n+1} T P^{-(n+1)})$$

is a product of elements of H. A direct cancellation gives

$$S'R'P(S')^{-1} = SRPS^{-1}.$$

Thus every $T \in \mathcal{T}_P$ gives an element of the fiber. Conversely, suppose

$$S'R'P(S')^{-1} = SRPS^{-1}$$

with $S' \in H$ and $R' \in \mathcal{T}_P$. Set $T = S^{-1}S' \in H$. Rearranging the displayed equality gives

$$R' = T^{-1}RPTP^{-1}.$$

It also gives $PTP^{-1} = R^{-1}TR'$. Since $R, R' \in \mathcal{T}_P$, induction on n yields $P^nTP^{-n} \in H$ for all $n \geq 0$; hence $T \in \mathcal{T}_P$. The two constructions are inverse to one another, so the fiber is in bijection with $\mathcal{T}_P$.

Theorem 4.3.1 (Class-size theorem for expanded equivalence). *Let G be a finite group and $H \leq G$. For the expanded relation of Definition 4.3.2, every equivalence class has cardinality $|H|$. Consequently, the number of classes is $|G|/|H|$.*

Proof. The relation is an equivalence: reflexivity uses $A = 1$; symmetry follows by taking inverses of Q^nAP^{-n}; transitivity follows by multiplying the two round-n witnesses. Fix $P \in G$. By Lemmas 4.3.1 and 4.3.2, $\mathcal{C}(P)$ is the image of $\Phi_P : H \times \mathcal{T}_P \to G$. Lemma 4.3.3 says that every fiber of this map has size $|\mathcal{T}_P|$. Hence

$$|H|\,|\mathcal{T}_P| = |H \times \mathcal{T}_P| = |\mathcal{C}(P)|\,|\mathcal{T}_P|.$$

Since $1 \in \mathcal{T}_P$, cancellation gives $|\mathcal{C}(P)| = |H|$. The quotient count follows because G is partitioned into equal-size classes.

Resolution of the Derbez–Euler Question

Theorem 4.3.2 (GFN class count). *For type-II GFNs with $2k$ branches, expanded equivalence on all branch permutations has classes of size $k!$, and the number of classes is*

$$\frac{(2k)!}{k!}.$$

Proof. Set $G = S_{2k} = \mathrm{Sym}(\Omega_k)$ and $H = S_k^{\mathrm{p}}$. The map

$$\mathrm{Sym}([k]) \longrightarrow S_k^{\mathrm{p}}, \qquad \tau \longmapsto \Phi_{\tau,\tau},$$

is a bijection, so $|H| = |S_k^{\mathrm{p}}| = |\,\mathrm{Sym}([k])| = k!$. Also

$$|G| = |S_{2k}| = |\,\mathrm{Sym}(\Omega_k)| = (2k)!.$$

Applying Theorem 4.3.1 gives the stated class size and class count.

The theorem extends the even-odd count from Derbez and Euler's ToSC 2024 paper: their restricted domain has $(k!)^2$ permutations and $k!$ classes, whereas the full branch-permutation domain has $(2k)!$ permutations and therefore $(2k)!/k!$ classes. The common class size is $k!$.

5 Conclusions

The case studies show that cryptographic conjectures require more than a type-correct formal statement. Their hypotheses, endpoint conventions, and equivalent formulations must be checked against the source before proof search begins. Several source questions admit only partial resolutions or require a corrected cryptographic interpretation; their exact hypotheses and resolved ranges are therefore stated in the contribution table and the corresponding theorems.

For the Tu–Deng conjecture, we give a complete proof by a route different from the recent preprint cited in Section 3.1. The pair count is first converted into a Hamming-weight decrease under circular addition. The cyclic carries form a monotone Boolean family. Grouping their pivotal edges by the upper carry word produces subcubes that fix at least as many ones as zeros. Integrating their biased profiles proves the pointwise half bound for every admissible parameter. The same local balance identifies all tight cases. If the binary expansion of t contains z zero bits, the bound is attained exactly when every cyclic gap between consecutive zeros is at least z. We also obtain a closed formula for the number of such residues with each value of z.

For vectorial Boolean functions with $n = 2k \geq 6$ input variables and $k < m < 2k$ output coordinates, we improve the covering-radius upper bound on component nonlinearity by two. The proof first extracts a maximal vectorial bent subfunction from a hypothetical extremal map. In the critical output dimension, Walsh congruences force a two-level spectrum, while the odd fibre sizes of the bent subfunction exclude that spectrum. The argument gives a uniform upper bound; except for $(n, m) = (4, 3)$, it does not establish tightness.

For balanced Boolean functions in eight variables, we prove that the maximum nonlinearity is 116. For a hypothetical function of nonlinearity 118, a parity condition partitions the Walsh spectrum into two affine halves. Exact slice constraints and a global double count leave thirteen spectral profiles. A common quadratic support and inverse-Walsh inequalities exclude each profile. An explicit truth table attains 116.

Acknowledgments The author sincerely thanks everyone who has offered valuable feedback and devoted their valuable time.

References

1. Alon, N., Bloom, T.F., Gowers, W.T., Litt, D., Sawin, W., Shankar, A., Tsimerman, J., Wang, V., Wood, M.M.: Remarks on the disproof of the unit distance conjecture (2026). https://doi.org/10.48550/arXiv.2605.20695, arXiv preprint arXiv:2605.20695
2. Bapić, A., Pasalic, E., Polujan, A., Pott, A.: Vectorial boolean functions with the maximum number of bent components beyond the nyberg's bound. Designs, Codes and Cryptography **92**(3), 531–552 (2024). https://doi.org/10.1007/s10623-022-01180-7
3. Barbosa, M., Barthe, G., Bhargavan, K., Blanchet, B., Cremers, C., Liao, K., Parno, B.: SoK: Computer-aided cryptography. In: 2021 IEEE Symposium on Security and Privacy. pp. 777–795. IEEE Computer Society (May 2021). https://doi.org/10.1109/SP40001.2021.00008
4. Barthe, G., Grégoire, B., Heraud, S., Zanella-Béguelin, S.: Computer-aided security proofs for the working cryptographer. In: Advances in Cryptology – CRYPTO 2011. Lecture Notes in Computer Science, vol. 6841, pp. 71–90. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22792-9_5
5. Blanchet, B.: A computationally sound mechanized prover for security protocols. In: 2006 IEEE Symposium on Security and Privacy. pp. 140–154. IEEE Computer Society, Oakland, California, USA (May 2006). https://doi.org/10.1109/SP.2006.1
6. Bogdanov, A.: Analysis and Design of Block Cipher Constructions. Ph.D. thesis, Ruhr University Bochum (2011), chapter 3
7. Bogdanov, A.: On unbalanced Feistel networks with contracting MDS diffusion. Designs, Codes and Cryptography **59**(1–3), 35–58 (2011). https://doi.org/10.1007/s10623-010-9462-0
8. Boura, C., Canteaut, A., Jean, J., Suder, V.: Two notions of differential equivalence on S-boxes. Designs, Codes and Cryptography **87**(2–3), 185–202 (2019). https://doi.org/10.1007/s10623-018-0496-z
9. de Bruijn, N.G.: The mathematical language AUTOMATH, its usage, and some of its extensions. In: Laudet, M., Lacombe, D., Nolin, L., Schützenberger, M.P. (eds.) Symposium on Automatic Demonstration. Lecture Notes in Mathematics, vol. 125, pp. 29–61. Springer, Berlin, Heidelberg (1970). https://doi.org/10.1007/BFb0060623
10. Carlet, C.: Boolean Functions for Cryptography and Coding Theory. Cambridge University Press, Cambridge (2020). https://doi.org/10.1017/9781108606806
11. Chabaud, F., Vaudenay, S.: Links between differential and linear cryptanalysis. In: Advances in Cryptology – EUROCRYPT '94. Lecture Notes in Computer Science, vol. 950, pp. 356–365. Springer (1995). https://doi.org/10.1007/BFb0053450
12. Chen, Y., Guo, F., Gong, Z., Cai, W.: One note about the Tu–Deng conjecture in case $\mathrm{wt}(t) = 5$. IEEE Access **7**, 13799–13802 (2019). https://doi.org/10.1109/ACCESS.2019.2894074
13. Chen, Y., Lin, L., Wei, C.: About the Tu–Deng conjecture for $\mathrm{wt}(t) \leq 10$. Cryptology ePrint Archive, Paper 2020/227 (2020), https://eprint.iacr.org/2020/227
14. Cheng, K., Hong, S., Zhong, Y.: A note on the Tu–Deng conjecture. Journal of Systems Science and Complexity **28**(3), 702–724 (2015). https://doi.org/10.1007/s11424-015-2240-3
15. Church, A.: An unsolvable problem of elementary number theory. American Journal of Mathematics **58**(2), 345–363 (1936). https://doi.org/10.2307/2371045

16. Cusick, T.W., Li, Y., Stănică, P.: On a combinatorial conjecture. Integers **11**(2), 185–203 (2011). https://doi.org/10.1515/integ.2011.017
17. Deng, G., Yuan, P.: On a combinatorial conjecture of Tu and Deng. Integers **12**, A48 (2012)
18. Derbez, P., Euler, M.: Equivalence of generalised Feistel networks. IACR Transactions on Symmetric Cryptology **2024**(1), 412–440 (2024). https://doi.org/10.46586/tosc.v2024.i1.412-440
19. Derbez, P., Fouque, P.A., Lambin, B., Mollimard, V.: Efficient search for optimal diffusion layers of generalized Feistel networks. IACR Transactions on Symmetric Cryptology **2019**(2), 218–240 (2019). https://doi.org/10.46586/tosc.v2019.i2.218-240
20. Dib, S.: Asymptotic nonlinearity of vectorial boolean functions. Cryptography and Communications **6**(2), 103–115 (2014). https://doi.org/10.1007/s12095-013-0090-1
21. Dunkelman, O., Huang, S.: Reconstructing an S-box from its difference distribution table. IACR Transactions on Symmetric Cryptology **2019**(2), 193–217 (2019). https://doi.org/10.46586/tosc.v2019.i2.193-217
22. Erbsen, A., Philipoom, J., Gross, J., Sloan, R., Chlipala, A.: Simple high-level code for cryptographic arithmetic: With proofs, without compromises. In: 2019 IEEE Symposium on Security and Privacy. pp. 1202–1219. IEEE (2019). https://doi.org/10.1109/SP.2019.00005
23. Flori, J.P., Randriambololona, H., Cohen, G., Mesnager, S.: On a conjecture about binary strings distribution. In: Sequences and Their Applications – SETA 2010. Lecture Notes in Computer Science, vol. 6338, pp. 346–358. Springer (2010). https://doi.org/10.1007/978-3-642-15874-2_30
24. Gödel, K.: Über formal unentscheidbare sätze der Principia Mathematica und verwandter systeme I. Monatshefte für Mathematik und Physik **38**, 173–198 (1931). https://doi.org/10.1007/BF01700692
25. Gonthier, G.: Formal proof—the four-color theorem. Notices of the American Mathematical Society **55**(11), 1382–1393 (2008)
26. Gonthier, G., Asperti, A., Avigad, J., Bertot, Y., Cohen, C., Garillot, F., Le Roux, S., Mahboubi, A., O'Connor, R., Ould Biha, S., Pasca, I., Rideau, L., Solovyev, A., Tassi, E., Théry, L.: A machine-checked proof of the odd order theorem. In: Blazy, S., Paulin-Mohring, C., Pichardie, D. (eds.) Interactive Theorem Proving. Lecture Notes in Computer Science, vol. 7998, pp. 163–179. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39634-2_14
27. Gordon, M.J.C., Milner, A.J., Wadsworth, C.P.: Edinburgh LCF: A Mechanised Logic of Computation, Lecture Notes in Computer Science, vol. 78. Springer, Berlin, Heidelberg (1979). https://doi.org/10.1007/3-540-09724-4
28. Hilbert, D.: Über das unendliche. Mathematische Annalen **95**, 161–190 (1926). https://doi.org/10.1007/BF01206605
29. Hubert, T., Mehta, R., Sartran, L., Horváth, M.Z., Žužić, G., Wieser, E., Huang, A., Schrittwieser, J., Schroecker, Y., Masoom, H., Bertolli, O., Zahavy, T., Mandhane, A., Yung, J., Beloshapka, I., Ibarz, B., Veeriah, V., Yu, L., Nash, O., Lezeau, P., Mercuri, S., Sönne, C., Mehta, B., Davies, A., Zheng, D., Pedregosa, F., Li, Y., von Glehn, I., Rowland, M., Albanie, S., Velingker, A., Schmitt, S., Lockhart, E., Hughes, E., Michalewski, H., Sonnerat, N., Hassabis, D., Kohli, P., Silver, D.: Olympiad-level formal mathematical reasoning with reinforcement learning. Nature **651**, 607–613 (2026). https://doi.org/10.1038/s41586-025-09833-y

30. Ju, H., Gao, G., Jiang, J., Wu, B., Sun, Z., Liu, S., Chen, L., Wang, Y., Wang, Y., Wang, Z., He, W., Wu, P., Xiao, L., Liu, R., Dai, B., Dong, B.: Automated conjecture resolution with formal verification (2026). https://doi.org/10.48550/arXiv.2604.03789, arXiv preprint arXiv:2604.03789
31. Kalai, A.T., Nachum, O., Vempala, S.S., Zhang, E.: Why language models hallucinate (2025). https://doi.org/10.48550/arXiv.2509.04664, arXiv preprint arXiv:2509.04664
32. Kaleyski, N.S.: Changing APN functions at two points. Cryptography and Communications **11**(6), 1165–1184 (2019). https://doi.org/10.1007/s12095-019-00366-6
33. Kim, S., Shin, M., Shin, H., Kim, I., Kim, S., Kwon, D., Hong, D., Sung, J., Hong, S.: Differential factors revisited: A sufficient condition for the practical use of differential factors. IEICE Transactions on Fundamentals **E107-A**(12), 1669–1675 (2024). https://doi.org/10.1587/transfun.2024EAP1058
34. Liu, R., Luo, H., Xie, T.: A complete proof for Tu–Deng conjecture. arXiv preprint arXiv:2608.05187 (2026). https://doi.org/10.48550/arXiv.2608.05187, https://arxiv.org/abs/2608.05187
35. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland (1977)
36. Maity, S., Maitra, S.: Minimum distance between bent and 1-resilient boolean functions. In: Fast Software Encryption. Lecture Notes in Computer Science, vol. 3017, pp. 143–160. Springer, Berlin, Heidelberg (2004). https://doi.org/10.1007/978-3-540-25937-4_10
37. Mesnager, S., Zhang, F., Tang, C., Zhou, Y.: Further study on the maximum number of bent components of vectorial functions. Designs, Codes and Cryptography **87**(11), 2597–2610 (2019). https://doi.org/10.1007/s10623-019-00639-4
38. Meyer, L.D., Bilgin, B.: Classification of balanced quadratic functions. IACR Transactions on Symmetric Cryptology **2019**(2), 169–192 (2019). https://doi.org/10.13154/tosc.v2019.i2.169-192
39. Mihaila, M., Thornburgh, D.: On lower bounds for the distances between APN functions. Cryptography and Communications (2026). https://doi.org/10.1007/s12095-026-00908-9, published online 6 June 2026
40. de Moura, L., Ullrich, S.: The Lean 4 theorem prover and programming language. In: Platzer, A., Sutcliffe, G. (eds.) Automated Deduction—CADE 28. Lecture Notes in Computer Science, vol. 12699, pp. 625–635. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-79876-5_37
41. Nagy, G.P.: Sidon sets, thin sets, and the nonlinearity of vectorial boolean functions. Journal of Combinatorial Theory, Series A **212**, 106001 (2025). https://doi.org/10.1016/j.jcta.2024.106001
42. Nyberg, K.: Perfect nonlinear S-boxes. In: Advances in Cryptology – EUROCRYPT '91. Lecture Notes in Computer Science, vol. 547, pp. 378–386. Springer (1991). https://doi.org/10.1007/3-540-46416-6_32
43. OpenAI: GPT-5.5 model (2026), https://developers.openai.com/api/docs/models/gpt-5.5, openAI API model documentation. Accessed August 12, 2026
44. OpenAI: GPT-5.6 Sol model (2026), https://developers.openai.com/api/docs/models/gpt-5.6-sol, openAI API model documentation. Accessed August 12, 2026
45. Picek, S., Marchiori, E., Batina, L., Jakobovic, D.: Combining evolutionary computation and algebraic constructions to find cryptography-relevant boolean functions. In: Parallel Problem Solving from Nature – PPSN XIII. Lecture Notes

in Computer Science, vol. 8672, pp. 822–831. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10762-2_81
46. Poiroux, A., Weiss, G., Kunčak, V., Bosselut, A.: Improving autoformalization using type checking (2024). https://doi.org/10.48550/arXiv.2406.07222, arXiv preprint arXiv:2406.07222
47. Pott, A., Pasalic, E., Muratovic-Ribic, A., Bajric, S.: On the maximum number of bent components of vectorial functions. IEEE Transactions on Information Theory **64**(1), 403–411 (2018). https://doi.org/10.1109/TIT.2017.2749421
48. Qarboua, S., Schrek, J., Fontaine, C.: New results about Tu–Deng's conjecture. In: 2016 IEEE International Symposium on Information Theory. pp. 485–489. IEEE (2016). https://doi.org/10.1109/ISIT.2016.7541346
49. Romera-Paredes, B., Barekatain, M., Novikov, A., Balog, M., Kumar, M.P., Dupont, E., Ruiz, F.J.R., Ellenberg, J.S., Wang, P., Fawzi, O., Kohli, P., Fawzi, A.: Mathematical discoveries from program search with large language models. Nature **625**, 468–475 (2024). https://doi.org/10.1038/s41586-023-06924-6
50. Sloane, N.J.A., Berlekamp, E.R.: Weight enumerator for second-order Reed–Muller codes. IEEE Transactions on Information Theory **16**(6), 745–751 (1970). https://doi.org/10.1109/TIT.1970.1054553
51. Song, P., Yang, K., Anandkumar, A.: Towards large language models as copilots for theorem proving in Lean (2024). https://doi.org/10.48550/arXiv.2404.12534, arXiv preprint arXiv:2404.12534
52. Spiegelhofer, L., Wallner, M.: The Tu–Deng conjecture holds almost surely. The Electronic Journal of Combinatorics **26**(1), P1.28 (2019). https://doi.org/10.37236/7178
53. Tezcan, C.: Differential factors revisited: Corrected attacks on present and serpent. In: Lightweight Cryptography for Security and Privacy. LNCS, vol. 9542, pp. 21–33. Springer (2016). https://doi.org/10.1007/978-3-319-29078-2_2
54. Tezcan, C., Özbudak, F.: Differential factors: Improved attacks on SERPENT. In: Lightweight Cryptography for Security and Privacy. Lecture Notes in Computer Science, vol. 8898, pp. 69–84. Springer (2015). https://doi.org/10.1007/978-3-319-16363-5_5
55. The mathlib Community: The Lean mathematical library. In: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs. pp. 367–381. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3372885.3373824
56. The mathlib Community: Mathlib 4 documentation (2026), https://leanprover-community.github.io/mathlib4_docs/Mathlib.html, accessed July 21, 2026
57. Tsoukalas, G., Kovsharov, A., Shirobokov, S., Surina, A., Firsching, M., Bérczi, G., Ruiz, F.J.R., Suggala, A., Wagner, A.Z., Wieser, E., Yu, L., Huang, A., Horváth, M.Z., Ferrauiolo, A., Michalewski, H., Grosu, C., Hubert, T., Balog, M., Kohli, P., Chaudhuri, S.: Advancing mathematics research with AI-driven formal proof search (2026). https://doi.org/10.48550/arXiv.2605.22763, arXiv preprint arXiv:2605.22763
58. Tu, Z., Deng, Y.: A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity. Designs, Codes and Cryptography **60**(1), 1–14 (2011). https://doi.org/10.1007/s10623-010-9413-9
59. Turing, A.M.: On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London Mathematical Society **s2-42**(1), 230–265 (1937). https://doi.org/10.1112/plms/s2-42.1.230

60. Wadayama, T., Hada, T., Wakasugi, K., Kasahara, M.: Upper and lower bounds on maximum nonlinearity of n-input m-output boolean function. Designs, Codes and Cryptography **23**(1), 23–34 (2001). https://doi.org/10.1023/A:1011207501748
61. Wang, S., Yu, L., Lou, S., Yan, F., Yang, C., Cui, Q., Zhou, J.: Improving autoformalization using direct dependency retrieval. In: Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 18023–18040. Association for Computational Linguistics (Jul 2026)
62. Wu, Y., Jiang, A.Q., Li, W., Rabe, M.N., Staats, C., Jamnik, M., Szegedy, C.: Autoformalization with large language models. In: Advances in Neural Information Processing Systems. vol. 35, pp. 32353–32368. Curran Associates, Inc. (2022)
63. Xu, Y., Carlet, C., Mesnager, S., Wu, C.: Classification of bent monomials, constructions of bent multinomials and upper bounds on the nonlinearity of vectorial functions. IEEE Transactions on Information Theory **64**(1), 367–383 (2018). https://doi.org/10.1109/TIT.2017.2750663
64. Yang, K., Swope, A.M., Gu, A., Chalamala, R., Song, P., Yu, S., Godil, S., Prenger, R.J., Anandkumar, A.: LeanDojo: Theorem proving with retrieval-augmented language models. In: Advances in Neural Information Processing Systems. vol. 36. Curran Associates, Inc. (2023)
65. Ye, K.Q., Green, M., Sanguansin, N., Beringer, L., Petcher, A., Appel, A.W.: Verified correctness and security of mbedTLS HMAC-DRBG. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. pp. 2007–2020. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3133956.3133974
66. Zinzindohoué, J.K., Bhargavan, K., Protzenko, J., Beurdouche, B.: HACL*: A verified modern cryptographic library. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. pp. 1789–1806. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3133956.3134043

Appendix

A Supplementary Background

A.1 Type-II Generalised Feistel Networks

This subsection fixes the network model used in Section 4.2, following the notation of Derbez et al. [19]. A Generalised Feistel Network (GFN) extends the two-branch Feistel construction by partitioning the state into more than two branches and applying several Feistel updates in parallel. The adjective *type-II* specifies the pairing and update pattern described below.

Branches and the round transformation Let the branch width be w bits. A $2k$-branch state is an ordered tuple

$$X^{(r)} = (X_0^{(r)}, X_1^{(r)}, \ldots, X_{2k-1}^{(r)}) \in (\{0,1\}^w)^{2k}$$

at the input to round r. Thus $2k$ is the number of state words, not the block size in bits. There are k adjacent Feistel pairs $(X_{2i}^{(r)}, X_{2i+1}^{(r)})$, for $0 \le i < k$, and

the complete state has $2kw$ bits. For example, the witnesses in Section 4.2 use $k = 19$ and $k = 20$, hence 38 and 40 branches, respectively. Their structural diffusion properties do not depend on w.

Let $F_{r,i} : \{0,1\}^w \to \{0,1\}^w$ be the round function for pair i in round r. The parallel type-II Feistel layer produces

$$\widetilde{X}_{2i}^{(r)} = X_{2i}^{(r)} \oplus F_{r,i}(X_{2i+1}^{(r)}), \qquad \widetilde{X}_{2i+1}^{(r)} = X_{2i+1}^{(r)}. \tag{A.1.1}$$

A branch permutation $\rho \in \mathrm{Sym}(\{0, \ldots, 2k-1\})$ is then applied to all positions:

$$X_{\rho(s)}^{(r+1)} = \widetilde{X}_s^{(r)} \qquad (0 \le s < 2k). \tag{A.1.2}$$

The round functions need not be permutations. As in an ordinary Feistel network, the unchanged right branch permits inversion of the pairwise update once the branch permutation has been undone. Figure 1 shows the parallel pair layer followed by the global shuffle.

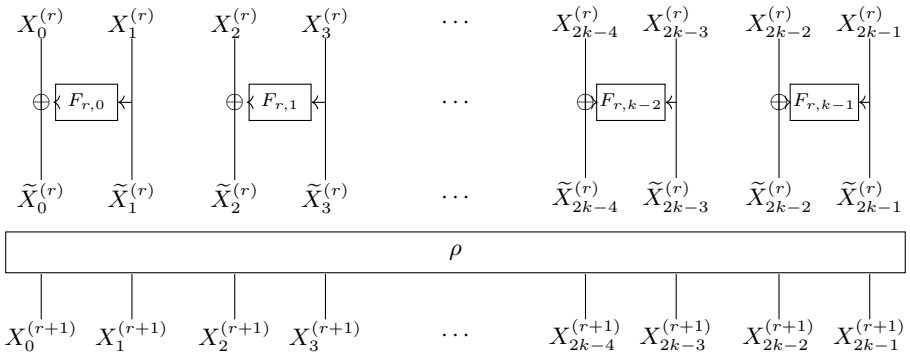


Fig. 1. Round structure of a $2k$-branch type-II GFN, adapted from [19, Fig. 1] with notation aligned to Eqs. (A.1.1)–(A.1.2). Each odd-indexed branch enters a round function whose output is XORed into the adjacent even-indexed branch. The permutation ρ then routes the intermediate branch at source position s to destination position $\rho(s)$.

Even–odd branch permutations The diffusion problem in Section 4.2 restricts ρ to the even–odd class. Two permutations $p, q \in \mathrm{Sym}(\{0, \ldots, k-1\})$ determine ρ by

$$\rho(2i) = 2p(i) + 1, \qquad \rho(2i+1) = 2q(i). \tag{A.1.3}$$

Hence every even source position is routed to an odd destination and every odd source position to an even destination. Because p and q are permutations, the two image sets are disjoint and together contain all $2k$ positions, so ρ is a permutation. This restriction describes the public branch shuffle. It imposes no algebraic condition on the round functions $F_{r,i}$.

The two rows listed for each witness in Section 4.2 are the image lists of p and q. Equation (A.1.3) converts those k-point permutations into the corresponding

$2k$-point branch permutation. For instance, an entry $p(i) = j$ sends the updated left branch of pair i, at source position $2i$, to destination position $2j + 1$.

Two-sided structural diffusion Here structural diffusion records which input branches can influence each output branch. The standard structural model treats each $F_{r,i}$ as a black box whose output carries the dependency of its input; it does not use a particular algebraic representation of the round function. Initially, output position j depends only on input position j. Under (A.1.1), the updated left branch of a pair inherits the dependencies of both members, while the right branch retains its previous dependencies. Equation (A.1.2) then relabels these sets. This gives exactly the recurrence in Equations (4.2.2)–(4.2.5).

Full structural diffusion after r rounds means that every output branch depends on all $2k$ input branches. The value $\mathrm{DR}(\rho)$ is the first such round in the forward direction. For inverse-direction diffusion, the convention of the source paper applies the same structural recurrence to ρ^{-1}. The resulting two-sided design criterion is $\mathrm{DR}_{\max}(\rho) = \max\{\mathrm{DR}(\rho), \mathrm{DR}(\rho^{-1})\}$. This is a structural criterion: it does not quantify differential probabilities, linear correlations, active S-boxes, or the security of a concrete instantiation.

B Exact Audit Data for the Eight-Variable Classification

This appendix supplies the finite data needed to audit the computer-assisted step in Section 3.3. It gives a deterministic enumeration procedure and every global infeasibility certificate. The full list of local tables is not printed: it contains 3,652 nine-entry vectors and is reproduced exactly by Algorithm 3. No floating-point calculation or solver verdict is used when checking the data below.

B.1 Reconstructing the local table families

Write a local table as the vector

$$\mathbf{N} = (N_{\mathsf{A}}, N_{\mathsf{B}}, N_{\mathsf{C}}, N_{\mathsf{D}}, N_{\mathsf{E}}, N_{\mathsf{F}}, N_{\mathsf{G}}, N_{\mathsf{H}}, N_{\mathsf{I}}).$$

Its row and column margins, in the notation of Section 3.3, are

$$\begin{aligned}\mathrm{marg}(\mathbf{N}) = (&N_{\mathsf{A}} + N_{\mathsf{B}} + N_{\mathsf{D}}, N_{\mathsf{C}} + N_{\mathsf{F}} + N_{\mathsf{G}}, N_{\mathsf{E}} + N_{\mathsf{H}} + N_{\mathsf{I}};\\ &N_{\mathsf{A}} + N_{\mathsf{C}} + N_{\mathsf{E}}, N_{\mathsf{B}} + N_{\mathsf{F}} + N_{\mathsf{H}}, N_{\mathsf{D}} + N_{\mathsf{G}} + N_{\mathsf{I}}).\end{aligned} \tag{B.1.1}$$

For slice weight $w = 59, 61, 63$, respectively, the distinguished zero-frequency type is $\mathsf{D}, \mathsf{G}, \mathsf{I}$. Algorithm 3 is an exhaustive mathematical specification of the local computation. The set $\mathcal{Q}$, the quantities $F_4, S_0, \Delta_\tau, U, \Delta_U$, and all cited conditions are defined in Section 3.3.

Algorithm 3 Exact reconstruction of all admissible local tables

Require: A slice weight $w \in \{59, 61, 63\}$
Ensure: The table families $\mathcal{T}_w(\pi)$, indexed by their margins
Set $Z_{59} = \mathsf{D}$, $Z_{61} = \mathsf{G}$, and $Z_{63} = \mathsf{I}$
Initialize every $\mathcal{T}_w(\pi)$ to the empty set
for $N_{\mathsf{A}} = 0, 1, \ldots, 76$ **do**
    for all nonnegative $N_{\mathsf{C}}, N_{\mathsf{D}}, N_{\mathsf{E}}, N_{\mathsf{F}}, N_{\mathsf{G}}, N_{\mathsf{H}}, N_{\mathsf{I}}$ satisfying $N_{\mathsf{C}}+N_{\mathsf{D}}+3N_{\mathsf{E}}+4N_{\mathsf{F}}+5N_{\mathsf{G}}+6N_{\mathsf{H}}+7N_{\mathsf{I}} = 3N_{\mathsf{A}} - 96$ **do**
        Set $N_{\mathsf{B}} = 128 - N_{\mathsf{A}} - N_{\mathsf{C}} - N_{\mathsf{D}} - N_{\mathsf{E}} - N_{\mathsf{F}} - N_{\mathsf{G}} - N_{\mathsf{H}} - N_{\mathsf{I}}$
        if $N_{\mathsf{B}} < 0$ or the count of type Z_w is zero **then**
            continue
        end if
        Discard $\mathbf{N}$ unless it satisfies (3.3.20), (3.3.23), and (3.3.24)
        Set feasible $\leftarrow$ false
        for all $\delta_{\mathsf{T}} \in \{-N_{\mathsf{T}}, -N_{\mathsf{T}}+2, \ldots, N_{\mathsf{T}}\}$, $\mathsf{T} \in \{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{E}, \mathsf{F}, \mathsf{H}\}$ **do**
            if (3.3.28), (3.3.29), (3.3.31), and (3.3.32) all hold **then**
                feasible $\leftarrow$ true; **break**
            end if
        end for
        if feasible **then**
            Insert $\mathbf{N}$ into $\mathcal{T}_w(\mathrm{marg}(\mathbf{N}))$
        end if
    end for
end for
Sort each family lexicographically and remove duplicate vectors
return the nonempty families $\mathcal{T}_w(\pi)$

The cutoff $N_{\mathsf{A}} \le 76$ is proved in Eq. (3.3.33); hence Algorithm 3 omits no table. Its output has the following sizes.

Slice weight w	First feasible N_{A}	Margin profiles	Local tables
59	40	96	1,461
61	42	85	1,254
63	45	77	937

The five tables at the first feasible layers are exactly those displayed after Eq. (3.3.33). As a further check on the grouping, the following 14 weight-59 margin profiles have no local family for at least one other slice weight that occurs with positive multiplicity. They are therefore excluded before the global linear systems are formed.

Profile π	Missing family	Profile π	Missing family
$(90, 26, 12; 92, 32, 4)$	$w = 63$	$(94, 30, 4; 84, 32, 12)$	$w = 63$
$(101, 25, 2; 70, 56, 2)$	$w = 63$	$(109, 17, 2; 62, 56, 10)$	$w = 63$
$(112, 16, 0; 60, 48, 20)$	$w = 61$	$(113, 13, 2; 54, 72, 2)$	$w = 63$
$(113, 13, 2; 60, 48, 20)$	$w = 61, 63$	$(114, 10, 4; 54, 72, 2)$	$w = 61, 63$
$(115, 7, 6; 54, 72, 2)$	$w = 61, 63$	$(116, 4, 8; 54, 72, 2)$	$w = 61, 63$
$(117, 1, 10; 54, 72, 2)$	$w = 61, 63$	$(121, 5, 2; 46, 72, 10)$	$w = 63$
$(122, 2, 4; 46, 72, 10)$	$w = 61, 63$	$(125, 1, 2; 46, 56, 26)$	$w = 63$

B.2 Complete global infeasibility certificates

Fix a profile $\pi = (o_{20}, o_{12}, o_4; e_{16}, e_8, e_0)$. For every $\boldsymbol{\nu} \in \mathcal{T}_w(\pi)$, define the column vector

$$\boldsymbol{m}_{w,\boldsymbol{\nu}} = (\boldsymbol{d}_w; \nu_{\mathsf{A}}, \nu_{\mathsf{B}}, \nu_{\mathsf{D}}, \nu_{\mathsf{C}}, \nu_{\mathsf{F}}, \nu_{\mathsf{G}}, \nu_{\mathsf{E}}, \nu_{\mathsf{H}}, \nu_{\mathsf{I}})^{\mathsf{T}} \in \mathbb{Z}^{12},$$

The nine local entries are in row-major order $(\mathsf{A}, \mathsf{B}, \mathsf{D}; \mathsf{C}, \mathsf{F}, \mathsf{G}; \mathsf{E}, \mathsf{H}, \mathsf{I})$, corresponding to

$$(20, 16), (20, 8), (20, 0), (12, 16), (12, 8), (12, 0),$$
$$(4, 16), (4, 8), (4, 0).$$

Here the direction-class indicator vectors are $\boldsymbol{d}_{59} = (1, 0, 0)$, $\boldsymbol{d}_{61} = (0, 1, 0)$, and $\boldsymbol{d}_{63} = (0, 0, 1)$. These symbols are distinct from the even-half counts e_{16}, e_8, e_0. The right-hand side vector is

$$\begin{aligned}\boldsymbol{b}_\pi = (&o_{20}, o_{12}, o_4; o_{20}e_{16}, o_{20}e_8, o_{20}e_0,\\ &o_{12}e_{16}, o_{12}e_8, o_{12}e_0, o_4e_{16}, o_4e_8, o_4e_0)^{\mathsf{T}}.\end{aligned} \tag{B.2.1}$$

Let $\mathbf{M}_\pi$ have all applicable columns $\boldsymbol{m}_{w,\boldsymbol{\nu}}$. Let $\boldsymbol{\lambda}$ be the vector of local-family multiplicities. The global conditions are $\mathbf{M}_\pi \boldsymbol{\lambda} = \boldsymbol{b}_\pi$ with $\boldsymbol{\lambda} \geq 0$.

For each of the 69 profiles in Table 10, the displayed primitive integer vector $\boldsymbol{z}_\pi$ satisfies

$$\mathbf{M}_\pi^{\mathsf{T}} \boldsymbol{z}_\pi \geq 0, \qquad \boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{z}_\pi < 0. \tag{B.2.2}$$

The first inequality is checked by taking the scalar product with every column reconstructed by Algorithm 3. If a nonnegative solution $\boldsymbol{\lambda}$ existed, then $0 \leq \boldsymbol{\lambda}^{\mathsf{T}} \mathbf{M}_\pi^{\mathsf{T}} \boldsymbol{z}_\pi = \boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{z}_\pi < 0$, a contradiction. Coordinates not displayed in the table are zero. The coordinate order is the order used in Eq. (B.2.1), indexed from 0 to 11. The main text states the same certificates in normalized rational form. For each displayed row, put

$$\boldsymbol{y}_\pi = \frac{\boldsymbol{z}_\pi}{-\boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{z}_\pi}.$$

Then $\mathbf{M}_\pi^{\mathsf{T}} \boldsymbol{y}_\pi \geq 0$ and $\boldsymbol{b}_\pi^{\mathsf{T}} \boldsymbol{y}_\pi = -1$, so the appendix vectors are positive integer rescalings of the certificates in Lemma 3.3.4.

Table 10: Integer Farkas certificates; omitted coordinates of $\boldsymbol{z}_\pi$ are zero.

Profile π	Nonzero coordinates of $\boldsymbol{z}_\pi$	$\boldsymbol{b}_\pi^{\mathsf{T}}\boldsymbol{z}_\pi$
$(92, 28, 8; 88, 32, 8)$	$z_7 = -1$	-896
$(96, 32, 0; 80, 32, 16)$	$z_7 = -1$	-1024
$(98, 26, 4; 76, 48, 4)$	$z_{10} = -1$	-192
$(99, 23, 6; 76, 48, 4)$	$z_2 = 24,\ z_9 = 1,\ z_{10} = -2,\ z_{11} = -5$	-96
$(100, 20, 8; 76, 48, 4)$	$z_0 = -4,\ z_1 = -4,\ z_2 = -4,\ z_{10} = 1,\ z_{11} = 2$	-64
$(100, 28, 0; 72, 48, 8)$	$z_5 = 1,\ z_7 = -1,\ z_8 = 1$	-320
$(101, 17, 10; 76, 48, 4)$	$z_0 = -6,\ z_1 = -6,\ z_2 = -6,\ z_{10} = 1,\ z_{11} = 2$	-208
$(101, 25, 2; 72, 48, 8)$	$z_9 = -1,\ z_{11} = 1$	-128
$(102, 14, 12; 76, 48, 4)$	$z_0 = 8,\ z_1 = 8,\ z_2 = 8,\ z_7 = -2,\ z_{10} = -1$	-896
$(102, 22, 4; 70, 56, 2)$	$z_9 = -1,\ z_{10} = 1,\ z_{11} = 3$	-32
$(102, 22, 4; 72, 48, 8)$	$z_0 = 4,\ z_1 = 4,\ z_2 = 4,\ z_9 = -2,\ z_{10} = -1$	-256
$(103, 11, 14; 76, 48, 4)$	$z_0 = -42,\ z_1 = -42,\ z_2 = -42,\ z_4 = 1,\ z_7 = -1$	-960
$(103, 19, 6; 70, 56, 2)$	$z_5 = -2,\ z_8 = -2,\ z_{10} = 1$	-152
$(103, 19, 6; 72, 48, 8)$	$z_0 = -6,\ z_1 = -6,\ z_2 = -6,\ z_{10} = 1,\ z_{11} = 2$	-384
$(104, 8, 16; 76, 48, 4)$	$z_7 = -8,\ z_9 = 1,\ z_{10} = -3,\ z_{11} = 1$	-4096
$(104, 16, 8; 70, 56, 2)$	$z_0 = -2,\ z_1 = -2,\ z_2 = -2,\ z_5 = -2,\ z_8 = -2,\ z_{10} = 1$	-288
$(104, 16, 8; 72, 48, 8)$	$z_7 = -2,\ z_9 = 1,\ z_{11} = 1$	-896
$(105, 5, 18; 76, 48, 4)$	$z_1 = -23,\ z_4 = 2,\ z_5 = -23,\ z_7 = -71$	-16735
$(105, 13, 10; 72, 48, 8)$	$z_0 = -90,\ z_1 = -90,\ z_2 = -90,\ z_4 = 2,\ z_{10} = 1$	-960
$(105, 21, 2; 66, 56, 6)$	$z_9 = -1$	-132
$(106, 2, 20; 76, 48, 4)$	$z_7 = -1$	-96
$(106, 10, 12; 72, 48, 8)$	$z_0 = 4,\ z_1 = 4,\ z_2 = 4,\ z_7 = -2,\ z_{10} = -1$	-1024
$(106, 18, 4; 66, 56, 6)$	$z_9 = -3,\ z_{11} = 1$	-768
$(106, 18, 4; 68, 48, 12)$	$z_{10} = -1$	-192
$(107, 7, 14; 70, 56, 2)$	$z_5 = -1,\ z_7 = -2,\ z_8 = -1,\ z_9 = 1$	-32
$(107, 7, 14; 72, 48, 8)$	$z_4 = 1,\ z_7 = -47$	-10656
$(107, 15, 6; 66, 56, 6)$	$z_{10} = -1,\ z_{11} = 2$	-264
$(107, 15, 6; 68, 48, 12)$	$z_0 = -90,\ z_1 = -90,\ z_2 = -90,\ z_4 = 2,\ z_{10} = 1$	-960
$(108, 4, 16; 70, 56, 2)$	$z_0 = 106,\ z_1 = 106,\ z_2 = 122,\ z_4 = -2,\ z_5 = 4,\ z_8 = 4,$ $z_{10} = -3$	-64
$(108, 4, 16; 72, 48, 8)$	$z_{10} = -1$	-768
$(108, 12, 8; 66, 56, 6)$	$z_0 = 2,\ z_1 = 2,\ z_2 = 2,\ z_{10} = -1,\ z_{11} = 2$	-96
$(108, 12, 8; 68, 48, 12)$	$z_0 = 4,\ z_1 = 4,\ z_2 = 4,\ z_7 = -2,\ z_{10} = -1$	-1024
$(108, 20, 0; 64, 48, 16)$	$z_0 = 4,\ z_1 = 4,\ z_7 = -1$	-448
$(109, 1, 18; 70, 56, 2)$	$z_7 = -1,\ z_{11} = 1$	-20
$(109, 9, 10; 66, 56, 6)$	$z_0 = 12,\ z_1 = 12,\ z_2 = 12,\ z_7 = -2,\ z_{10} = -1$	-32
$(109, 9, 10; 68, 48, 12)$	$z_4 = 1,\ z_7 = -47,\ z_{10} = -23$	-26112
$(109, 17, 2; 64, 48, 16)$	$z_9 = -1,\ z_{11} = 1$	-96
$(110, 6, 12; 66, 56, 6)$	$z_0 = 10,\ z_1 = 10,\ z_2 = 10,\ z_7 = -2,\ z_{10} = -1$	-64
$(110, 6, 12; 68, 48, 12)$	$z_{10} = -1$	-576
$(110, 14, 4; 60, 64, 4)$	$z_{10} = -1$	-256
$(110, 14, 4; 62, 56, 10)$	$z_9 = 1,\ z_{10} = -1,\ z_{11} = -3$	-96
$(110, 14, 4; 64, 48, 16)$	$z_7 = -2,\ z_9 = 1,\ z_{11} = 1$	-1024
$(111, 3, 14; 66, 56, 6)$	$z_0 = 8,\ z_1 = 8,\ z_2 = 8,\ z_7 = -2,\ z_{10} = -1$	-96
$(111, 11, 6; 60, 64, 4)$	$z_2 = 24,\ z_9 = 1,\ z_{10} = -2,\ z_{11} = -5$	-384
$(111, 11, 6; 62, 56, 10)$	$z_0 = 4,\ z_1 = 4,\ z_2 = 4,\ z_7 = -2,\ z_{10} = 1,\ z_{11} = 4$	-144
$(111, 11, 6; 64, 48, 16)$	$z_0 = -4,\ z_1 = -4,\ z_2 = -4,\ z_7 = -2,\ z_9 = 1,\ z_{11} = 1$	-1088
$(112, 0, 16; 66, 56, 6)$	$z_5 = -1,\ z_{11} = 2$	-480
$(112, 8, 8; 62, 56, 10)$	$z_0 = -102,\ z_1 = -102,\ z_2 = -102,\ z_4 = 2,\ z_{10} = 1$	-64
$(112, 8, 8; 64, 48, 16)$	$z_{10} = -1$	-384
$(113, 5, 10; 62, 56, 10)$	$z_0 = 8,\ z_1 = 8,\ z_2 = 8,\ z_7 = -2,\ z_{10} = -1$	-96
$(113, 13, 2; 56, 64, 8)$	$z_9 = -1,\ z_{11} = 1$	-96
$(113, 13, 2; 58, 56, 14)$	$z_{10} = -1$	-112
$(114, 2, 12; 62, 56, 10)$	$z_4 = 1,\ z_6 = 53,\ z_8 = 53,\ z_{10} = -26$	-3456
$(114, 10, 4; 58, 56, 14)$	$z_0 = 2,\ z_1 = 2,\ z_2 = 2,\ z_7 = -2,\ z_{10} = 3$	-192
$(114, 10, 4; 60, 48, 20)$	$z_{10} = -1$	-192
$(115, 7, 6; 58, 56, 14)$	$z_{10} = -1,\ z_{11} = 2$	-168
$(116, 4, 8; 58, 56, 14)$	$z_0 = 106,\ z_1 = 106,\ z_2 = 106,\ z_3 = -2,\ z_8 = 2,\ z_9 = -1,$ $z_{11} = 1$	-128
$(116, 12, 0; 56, 48, 24)$	$z_7 = -1$	-576
$(117, 1, 10; 58, 56, 14)$	$z_7 = 1,\ z_{11} = -1$	-84

Continued on the next page

Table 10 (continued)

Profile π	Nonzero coordinates of $\boldsymbol{z}_\pi$	$\boldsymbol{b}_\pi^{\mathsf{T}}\boldsymbol{z}_\pi$
$(117, 9, 2; 52, 64, 12)$	$z_9 = -1,\ z_{11} = 1$	-80
$(117, 9, 2; 54, 56, 18)$	$z_{10} = -1$	-112
$(118, 6, 4; 52, 64, 12)$	$z_7 = -2,\ z_{10} = 1,\ z_{11} = 4$	-320
$(118, 6, 4; 54, 56, 18)$	$z_4 = -1,\ z_9 = 26,\ z_{11} = -25$	-2792
$(119, 3, 6; 52, 64, 12)$	$z_0 = -49,\ z_1 = -49,\ z_2 = -49,\ z_3 = 1,\ z_8 = -1$	-120
$(119, 3, 6; 54, 56, 18)$	$z_7 = -6,\ z_9 = 2,\ z_{10} = -1,\ z_{11} = 2$	-480
$(120, 0, 8; 52, 64, 12)$	$z_9 = -2,\ z_{10} = 1$	-320
$(120, 0, 8; 54, 56, 18)$	$z_4 = -1,\ z_9 = 9,\ z_{11} = 9$	-1536
$(121, 5, 2; 50, 56, 22)$	$z_{10} = -1$	-112
$(122, 2, 4; 50, 56, 22)$	$z_6 = 1,\ z_8 = 1,\ z_{10} = -1$	-80

Together with the 14 missing-family profiles, these certificates leave exactly the 13 profiles displayed in Table 2. Thus the partition of the 96 weight-59 margin profiles is

$$96 = 14 + 69 + 13.$$

B.3 The fifteen signed categories

This table verifies the sign identity used in Eqs. (3.3.52)–(3.3.53). It suffices to display the weight-61 case. Here $p \equiv 3 \pmod 4$, $q \equiv 1 \pmod 4$, and

$$R = \frac{pq+1}{4} \bmod 2.$$

For each of the fifteen ordered magnitude pairs allowed by (3.3.11), the final column equals $5, -3, 1$ in the odd rows $20, 12, 4$, respectively.

$\lvert p\rvert$	$\lvert q\rvert$	p	q	R	odd row	$\frac{1}{2}(-1)^R(q-p)$
1	1	-1	1	0	4	1
1	3	-1	-3	1	4	1
1	5	-1	5	1	12	-3
1	7	-1	-7	0	12	-3
1	9	-1	9	0	20	5
3	1	3	1	1	4	1
3	3	3	-3	0	12	-3
3	5	3	5	0	4	1
3	7	3	-7	1	20	5
5	1	-5	1	1	12	-3
5	3	-5	-3	0	4	1
5	5	-5	5	0	20	5
7	1	7	1	0	12	-3
7	3	7	-3	1	20	5
9	1	-9	1	0	20	5

B.4 Checks for the terminal exclusions

For a residual profile, put $\mathrm{Score}(\pi) = 5o_{20} - 3o_{12} + o_4$. Table 11 records the integer substitutions used in Section 3.3. When $\mathrm{wt}(R) = 64$, the quadratic normal forms

with $s \leq 2$ require $\mathrm{Score}(\pi) \leq 256$, which fails for every B_i. For $s = 3$, the bound is 512 and only B_2 remains. Its special inverse-Walsh certificate has upper bound -384. When $\mathrm{wt}(R) = 56$, the two special certificates have upper bound $8(464 - \mathrm{Score}(\pi))$, giving -256 and -384.

Profile	o_{20}	o_{12}	o_4	$\mathrm{Score}(\pi)$
U_1	105	13	10	496
U_2	106	10	12	512
B_1	112	8	8	544
B_2	112	16	0	512
B_3	113	5	10	560
B_4	114	2	12	576
B_5	114	10	4	544
B_6	115	7	6	560
B_7	116	4	8	576
B_8	116	12	0	544
B_9	117	1	10	592
B_{10}	120	8	0	576
B_{11}	128	0	0	640

Table 11. Integer data for the terminal profile exclusions.

B.5 Walsh audit of the sharpness witness

The hexadecimal truth table in Section 3.3 decodes to 256 bits by reading each hexadecimal digit from its least significant bit to its most significant bit. Let $v_x = (-1)^{f(x)}$. Starting with the array $(v_x)_{x \in V_8}$, perform the butterfly update $(u, v) \longmapsto (u+v, u-v)$ at strides $1, 2, 4, \ldots, 128$. After the eighth pass, the array is the Walsh spectrum of f. The resulting multiplicities are

Walsh value	-24	-20	-16	-12	-8	-4	0	4	8	12	16	20	24
Multiplicity	5	40	31	16	19	8	2	8	23	16	39	40	9

The multiplicities sum to 256, and the squared values sum to 2^{16}, as required by Parseval. The zero-frequency coefficient is zero because the truth table has weight 128. Its largest Walsh magnitude is 24, so its nonlinearity is $128 - 24/2 = 116$.