Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spectral Analysis and Strict Lower Bounds for Iterative Differentials of Nibble-Based SPNs

Wanli Dai

3242052051377@glut.edu.cn

Research Article

Keywords: differential cryptanalysis, Markov chain, provable lower bound, iterative differential, lightweight SPN, CRAFT, PRESENT, GIFT

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11560234/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Cryptography and Communications manuscript No.
(will be inserted by the editor)

Spectral Analysis and Strict Lower Bounds for Iterative Differentials of Nibble-Based SPNs

Wanli Dai

the date of receipt and acceptance should be inserted later

Abstract Heuristic clustering and single-trail searches are standard for evaluating differential attacks on nibble-based SPNs, yet they lack formal mathematical guarantees and can severely misestimate the true differential probability. In this paper, we establish a rigorous mathematical framework to address this. First, on the spectral side, we prove a structural obstruction: for any parallel SPN with a bijective linear layer, the second singular value of the round transition matrix is exactly 1 ($\sigma_2(M) = 1$). This proves that the classical operator-norm bound is structurally vacuous and isolates the second eigenvalue $\lambda_2(M)$ as the correct convergence index. Second, on the constructive side, we introduce an activity decomposition method that yields strict, entrywise-justified lower bounds. By restricting the transition matrix to differences with at most K active nibbles, we compute a strict lower bound for closed-loop iterative differentials; the resulting algorithm is fixed-parameter tractable in K, with a state count $\sum_{j\le K}\binom{N}{j}15^j$ independent of the full block size. To demonstrate its effectiveness, we apply our framework to CRAFT. We rigorously prove that the 11-round and 12-round differential probabilities are bounded below by $2^{-51.8503}$ and $2^{-55.9438}$, respectively, exceeding the best single-trail baseline by a factor of over 2^{36}. The framework is mathematically generic and transfers seamlessly to PRESENT-64 and GIFT-64. This work bridges the gap between discrete matrix theory and practical cryptanalysis, providing a provable baseline for future security evaluations.

Keywords differential cryptanalysis, Markov chain, provable lower bound, iterative differential, lightweight SPN, CRAFT, PRESENT, GIFT.

Corresponding author. Email: 3242052051377@glut.edu.cn.

Address(es) of author(s) should be given

Mathematics Subject Classification (2020) 94A60 · 05C50 · 15A18 · 60J10

1 Introduction

Estimating the probability of a differential has always been the bottleneck of differential cryptanalysis. The canonical answer, multiplying the S-box probabilities along the single best trail, is provable but can be exponentially weak, because a genuine differential is the sum of exponentially many trails. The alternative, assembling a cluster of trails as in the CRAFT analysis of Hadipour et al. [2], produces numbers that are far more accurate, yet nothing in the method certifies them: the cluster is a heuristic collection, and nobody can say how much probability still lies outside it. What is missing, then, is a lower bound that is at once non-trivial and mathematically guaranteed; that is the object of this paper.

The methods we rely on are combinatorial and algebraic rather than cryptographic folklore. The lower bound rests on an entrywise partial order of nonnegative matrices; the spectral obstruction rests on the Kronecker product and on the doubly-stochastic structure of the S-box kernel; and the long-round behaviour rests on the second eigenvalue of a nonnegative matrix. These are objects of discrete mathematics, and the instantiated ciphers, CRAFT, PRESENT, and GIFT, serve only as test beds for a method that applies to any nibble-based SPN.

In this paper we take the Markov view [4] $DP_r(\Delta) = (M^r)[\Delta, \Delta]$. The matrix M is astronomically large ($2^n \times 2^n$), so direct powering is impractical. We therefore analyze the chain through an *activity decomposition*: the state space is stratified by the number of active nibbles, and the transition operator is restricted to the low-activity subspace $W_K = \{A : \mathrm{act}(A) \le K\}$ of differences with at most K active nibbles, obtaining a substochastic matrix $M_K \le M$. Because the inequality is entrywise, every number we report is a *provable bound*, not a heuristic. A provable lower bound is the operative quantity: it states that the differential probability is at least the reported value—a certified statement rather than an unverifiable estimate. The residual mass $\bar{P}_{\mathrm{drop}}$ is computed exactly from the cipher structure and guarantees the lower bound: the truncation is entrywise, so $P_{\mathrm{comp}} \le DP_r(\Delta)$ always holds, and the reported lower bound never overestimates the true probability.

To state the logical status of these numbers in one place: P_{comp} is not an approximation of $DP_r(\Delta)$ but the exactly computed probability of all Δ-to-Δ walks that stay inside W_K, and it is monotone in K (Theorem 3). Theorem 4 wraps it into the two-sided containment $P_{\mathrm{comp}} \le DP_r(\Delta) \le P_{\mathrm{comp}} + \bar{P}_{\mathrm{drop}}$, where $\bar{P}_{\mathrm{drop}}$ is the exactly computed mass of every discarded prefix. As K grows the lower endpoint is pushed upward and the upper endpoint downward, and both coincide with $DP_r(\Delta)$ at $K = N$; this is the branch-and-bound structure of the method (Remark 2), not a single-trail proxy. On full-size

ciphers the upper endpoint stays near 1 at every feasible K, so the operative provable quantity reported in the experiments is the lower endpoint.

This Markov formulation is not an additional modeling assumption. The standard differential-cryptanalysis toolchain—single-trail search by MILP/SAT [5–7] and probability estimation from the DDT, including multi-trail clustering [8,9]—already evaluates each round by the same one-round DDT transition and composes the per-round weights by multiplication, i.e. it implicitly adopts the round-independent model. The matrix M is simply the exact, entrywise statement of that default model; all MILP-, SAT-, and DDT-based estimates cited below operate under the same premise, with dependency-aware refinements as the exception that addresses the same caveat.

This is exact in the key-alternating nibble-based SPN setting studied here: the only non-linearity is the S-box, and the round key—including any round constant, which acts by XOR—cancels in the difference and is absorbed by averaging over independent uniform round keys (Section 3). The failure modes of the Markov cipher assumption exhibited for ARX ciphers by Xu et al. [16]—two consecutive or parallel modular additions, and the XOR of a branch with a round constant—concern the modular-addition non-linearity of ARX designs and do not arise in the S-box / linear-layer / key-alternating structure of CRAFT, PRESENT, and GIFT. Recent work has begun to verify individual differential trails against the algebraic constraints that the Markov cipher assumption neglects; Trail-Estimator [19] automatically detects such violations. In the key-alternating setting studied here, the per-round DDT model is exact, so no correction is needed.

These tools, however, answer a different question from the one studied here. MILP/SAT search [5–7] bounds the best single trail and the active-S-box count, the clustering analysis of [8] shows that the differential probability is a sum of trail contributions that can far exceed the best single trail, and multiple-differential techniques [9] aggregate several differentials statistically. None of these yields a rigorous lower bound that accounts for multi-path clustering: single-trail search gives only the weak bound edp^r, and the clustering and multiple-differential estimates carry no guarantee. Our K-active truncation provides a non-trivial, entrywise-justified lower bound together with an exact residual-mass certificate, so the reported bound is never an overestimate.

This separates our contribution from classical truncated differentials [10]: truncated differentials only assert the *existence* of a path through an activity pattern, without a probability, an error bound, or a provable lower bound. Our framework quantifies all kept paths, controls all discarded paths, and is generic across bit-permutation and MixColumns-based nibble permutations.

Complementing the constructive lower bound, one may ask whether a traditional spectral upper bound can provide a matching ceiling. A long-standing intuition holds that the per-round contraction of the S-box layer should propagate to the whole cipher, so that the differential probability approaches its uniform floor at a rate governed by the singular values of the round matrix. We show that this intuition fails in parallel. The round matrix is a Kronecker

product, and a Kronecker product multiplies the unit singular values of the S-box kernel outward, so the second singular value of the full round matrix is exactly $\sigma_2(M) = 1$ for any $N \geq 2$. The operator-norm route therefore collapses by algebra rather than by approximation, and the only quantity left to govern convergence is the second eigenvalue $\lambda_2(M)$. This negative result delineates the theory: it tells the community precisely where a spectral upper bound can exist and where it cannot. We stress that this concerns only the analytic spectral route and does not preclude combinatorial tools such as MILP/SAT from bounding the probability of individual trails.

We instantiate the framework for *iterative* differentials $\Delta \to \Delta$, whose input and output difference coincide. The activity decomposition is in fact generic for any differential pair (Δ, Δ'); we specialize to the closed-loop case because it carries the structure that makes the theory productive: a ψ-graph support closure that enumerates candidate iterative differences, a first-return decomposition and a doubly-stochastic long-round floor, and the single-trail baseline edp^r against which the provable multi-path lower bound can be measured. This specialization is a structural choice, not a restriction of the decomposition mechanism itself.

2 Contributions

1. **A structural spectral theorem for parallel SPNs.** We prove that the round matrix of an N-nibble SPN factors as a Kronecker product of identical S-box kernels, so that its second singular value satisfies $\sigma_2(M) = 1$ for every $N \geq 2$. The classical operator-norm (singular-value) bound μ^r is therefore structurally vacuous, and the correct convergence rate of the chain is the second eigenvalue $\lambda_2(M)$; this is an algebraic property of nonnegative matrices, independent of any approximation.
2. **A strict lower-bound framework via entrywise truncation.** Decomposing the nonnegative transition matrix by activity and keeping the low-activity part gives an entrywise inequality $M_K \leq M$, hence a strict lower bound $P_{\mathrm{comp}} = (M_K^r)[\Delta, \Delta] \leq DP_r(\Delta)$ with an exactly computed residual-mass certificate. We prove the ψ-graph support-closure characterization of one-round iterative differentials, the first-return decomposition, monotonicity of the lower bound in K, and a reverse-reachability tightening; the bound is realized by a fixed-parameter-tractable enumerator whose state count $|W_K| = \sum_{j \leq K} \binom{N}{j} 15^j$ is independent of the full block size 2^n.
3. **Numerical results.** On CRAFT the method proves lower bounds $2^{-51.8503}$ and $2^{-55.9438}$ at rounds 11 and 12, exceeding the best single trail by factors of about 2^{36} and 2^{40}; a $K = 3, 4, 5, 6$ ablation identifies $K = 4$ as the Pareto-optimal trade-off, and the bounds transfer unchanged to PRESENT-64 and GIFT-64.

3 Preliminaries

Let the block be $n = 4N$ bits and write a difference as $A = (a_0, \ldots, a_{N-1})$, $a_m \in \mathbb{F}_2^4$. A round applies a bijective 4-bit S-box S to every nibble and an $\mathbb{F}_2$-linear layer L: a bit permutation $L = P$ for PRESENT/GIFT, applied after the S-boxes, and $L = \mathrm{PN} \circ \mathrm{MC}$ for CRAFT, applied before the S-boxes (MixColumns, then PermuteNibbles, then the S-boxes). Keys and constants are omitted because they act by XOR. Accordingly, $DP_r(\Delta) = (M^r)[\Delta, \Delta]$ is the expected differential probability over the round keys under the standard round-key-independence assumption (hypothesis of stochastic equivalence), the same model that MILP/SAT- and DDT-based differential search adopts by default.

Define the difference distribution table $\mathrm{DDT}_S[a][d] = \#\{x : S(x) \oplus S(x \oplus a) = d\}$. Let z and y denote the S-box input and output differences of the round: $z = A$ and $y = L^{-1}(B)$ (write $d(B) = L^{-1}(B)$) when the S-boxes come first (PRESENT/GIFT), and $z = L(A)$ and $y = B$ when the linear layer comes first (CRAFT). The one-round transition matrix is

$$M[A][B] = \prod_{m=0}^{N-1} \frac{\mathrm{DDT}_S[z_m, y_m]}{16}. \tag{1}$$

Let $\mathrm{supp}(A) = \{m : a_m \neq 0\}$, $\mathrm{act}(A) = |\mathrm{supp}(A)|$, and $W_K = \{A : \mathrm{act}(A) \leq K\}$.

Lemma 1 *M is row-stochastic. If S is bijective, M is column-stochastic, hence doubly stochastic. If $A \neq 0$ then $M[A][B] > 0$ implies $B \neq 0$.*

Proof For fixed A the S-box input z is fixed and $y \mapsto B$ is a bijection, so $\sum_B M[A][B] = \prod_m \big(\sum_{y_m} \mathrm{DDT}_S[z_m][y_m]/16\big) = 1$. For fixed B the S-box output y is fixed and $z \mapsto A$ is a bijection, so $\sum_A M[A][B] = \prod_m \big(\sum_{z_m} \mathrm{DDT}_S[z_m][y_m]/16\big) = 1$, because $\sum_a \mathrm{DDT}_S[a][d] = 16$ for every d when S is bijective. Finally $\mathrm{DDT}_S[a][0] = 0$ for $a \neq 0$.

Throughout, the linear-algebraic background—nonnegative matrices, stochastic matrices, Kronecker products, and the singular-value/eigenvalue theory of matrices—is standard; we follow [17] for nonnegative matrices and Markov chains, and [18] for matrix analysis and spectral theory.

4 Theory

4.1 Markov Differential Probability

Definition 1 Let X_t be the difference after t rounds with $X_0 = \Delta$. Define $DP_r(\Delta) = \Pr[X_r = \Delta \mid X_0 = \Delta]$.

Theorem 1 $DP_r(\Delta) = (M^r)[\Delta, \Delta]$.

Proof Given $X_t = A_t$, the distribution of X_{t+1} is the row $M[A_t, \cdot]$; the Markov property gives conditional independence across rounds. Expanding over intermediate differences $(A_1, \ldots, A_{r-1})$,

$$DP_r(\Delta) = \sum_{A_1,\ldots,A_{r-1}} \prod_{t=0}^{r-1} M[A_t, A_{t+1}] = (M^r)[\Delta, \Delta], \tag{2}$$

with $A_0 = A_r = \Delta$.

4.2 Support Closure of Iterative Differentials

For a bit permutation $L = P$, write $\psi(m, b) = \lfloor P^{-1}[4m + b]/4 \rfloor$ for the nibble that receives the b-th output bit of nibble m.

Theorem 2 *Let $\mathcal{S} = \mathrm{supp}(\Delta)$. Then Δ is a one-round iterative differential iff (C1) for every $m \in \mathcal{S}$ and every set bit b of a_m, $\psi(m, b) \in \mathcal{S}$; and (C2) for every $m \in \mathcal{S}$, $\mathrm{DDT}_S[a_m, d_m(\Delta)] > 0$.*

Proof If $a_m[b] = 1$, the corresponding output bit of $d(\Delta)$ is 1, and a bijective S-box forces the source nibble to be active, giving (C1); (C2) is feasibility. Conversely, (C1) forces $d_m(\Delta) = 0$ for $m \notin \mathcal{S}$, and (C2) makes every active nibble feasible.

4.3 Activity Decomposition and the Interval Bound

Definition 2 Define $M_K[A][B] = M[A][B]$ if $B \in W_K$ and 0 otherwise. Let $q_0 = e_\Delta$ (a row vector) and $q_t = q_{t-1} M_K$, and $P_{\mathrm{comp}} = (M_K^r)[\Delta, \Delta] = q_r(\Delta)$.

Theorem 3 (Strict lower bound) *$P_{\mathrm{comp}} \le DP_r(\Delta)$, and P_{comp} is nondecreasing in K.*

Proof Entrywise $0 \le M_K \le M$ and $W_K \subseteq W_{K+1}$; matrix products preserve the entrywise order.

Definition 3 The mass discarded at step t is

$$\delta_t = \sum_A q_{t-1}(A) \sum_{B \notin W_K} M[A][B], \tag{3}$$

and the total dropped mass over r rounds is

$$\bar{P}_{\mathrm{drop}} = \sum_{t=1}^{r} \delta_t. \tag{4}$$

Remark 1 $\bar{P}_{\mathrm{drop}}$ is computed exactly from M and W_K; it equals the probability that a true r-round trajectory leaves W_K at least once. It is a structural bound, not an empirical estimate.

Theorem 4 (Provable interval) $P_{\text{comp}} \le DP_r(\Delta) \le P_{\text{comp}} + \bar{P}_{\text{drop}}$.

Proof Partition $\{X_r = \Delta\}$ by the first time $\tau \le r$ at which the trajectory leaves W_K. Trajectories with $\tau > r$ evolve exactly as M_K and contribute P_{comp}. The remaining trajectories have probability at most $\Pr[\tau \le r] = \bar{P}_{\text{drop}}$.

Remark 2 (Branch-and-bound refinement) Theorems 3 and 4 together give a branch-and-bound reading. The lower bound P_{comp} is the exact mass of the Δ-to-Δ walks retained by M_K and is nondecreasing in K; the upper bound $P_{\text{comp}} + \bar{P}_{\text{drop}}$ equals $1 - \sum_{B \ne \Delta}(M_K^r)[\Delta, B]$ by row stochasticity of M, so it is nonincreasing in K. Thus the lower bound is pushed upward and the upper bound is pushed downward as K grows, and at $K = N$ no state is truncated, $\bar{P}_{\text{drop}} = 0$, and both bounds coincide with $DP_r(\Delta)$ exactly. On full-size ciphers the upper bound stays close to 1 until K approaches N, so the operative quantity is the exact lower bound P_{comp}.

4.4 Reverse-Reachability Tightening

Theorem 5 (Reverse-reachability bound) *Let $\mathcal{R}_s(\Delta)$ be the set of support patterns from which* $\text{supp}(\Delta)$ *is reachable in at most s rounds, and partition the mass dropped at step t by the support P of the dropped state, $\delta_t = \sum_P \delta_t[P]$. Then*

$$DP_r(\Delta) \;\le\; P_{\text{comp}} + \sum_{t=1}^{r} \sum_{P \in \mathcal{R}_{r-t}(\Delta)} \delta_t[P] \;\le\; P_{\text{comp}} + \bar{P}_{\text{drop}}. \tag{5}$$

Proof A trajectory dropped at step t at a state with support P can still reach Δ at step r only if P is backward-reachable from $\text{supp}(\Delta)$ in the remaining $r - t$ rounds. Dropped mass outside $\mathcal{R}_{r-t}(\Delta)$ contributes zero to $(M^r)[\Delta, \Delta]$, so it can be removed from the error bound. Since $\mathcal{R}_{r-t}(\Delta)$ is a subset of all supports, the bound never exceeds the naive one.

Remark 3 (Sound over-approximation of $\mathcal{R}_s$) Because a bijective S-box preserves nibble activity, the output support equals the S-box input support. For CRAFT the linear layer is nibble-wise: $Z_i = \bigoplus_{j \in T_i} A_j$. The support transition is therefore characterized by the forced bits

$$\begin{aligned} \text{must0}(P) &= \{i : |T_i \cap P| = 0\}, \\ \text{must1}(P) &= \{i : |T_i \cap P| = 1\}. \end{aligned} \tag{6}$$

A support Q is reachable from P only if Q contains $\text{must1}(P)$ and avoids $\text{must0}(P)$. Using this superset of the true relation in the backward BFS yields a sound over-approximation of $\mathcal{R}_s$, hence a valid bound.

Remark 4 (Exact support reachability for column-wise layers) For CRAFT the linear layer is $L = \text{PN} \circ \text{MC}$, and MC acts independently on the four nibble columns. The exact one-step support relation therefore factorizes over columns:

each column reads four nibbles and emits four nibbles, so its local reachable-support set is computed by at most 15^4 enumerations and is stored as a subset of the 16 four-bit masks; the full relation is the product of the four column relations. This gives the exact closure $\mathcal{R}_s$, not an over-approximation, at negligible cost.

Remark 5 (Exact reachability establishes short-round saturation) For CRAFT's $\Delta^\star$ the exact closure grows from 12 patterns at $s = 2$ to 424 at $s = 4$, 5485 at $s = 5$, 30764 at $s = 6$, and 57159 at $s = 7$, saturating to all 65535 nonzero patterns only at $s = 11$; the forced-bit over-approximation saturates already at $s = 5$. The exact bound is therefore far stronger: for $r \leq 9$ the dropped mass is entirely outside $\mathcal{R}_{r-t}$ for every t, so the dropped mass vanishes (equivalently, $P_{\mathrm{comp}}(K{=}4)$ is exact for $r \leq 9$). It excludes 12.5% of the dropped mass at $r = 10$ and 1% at $r = 11$, and essentially nothing beyond $r = 12$. Exact reachability thus establishes saturation at short rounds, but because $\mathcal{R}_s$ saturates at $s = 11$ it cannot make $\bar{P}_{\mathrm{drop}}$ informative in the long-round regime, where the eigenvalue analysis below applies.

4.5 First-Return Decomposition and Long-Round Spectral Floor

Theorem 6 (First-return decomposition) *$DP_r(\Delta) = \sum_{j=1}^{r} f_j \, DP_{r-j}(\Delta)$, with $DP_0(\Delta) = 1$, where $f_j = \Pr[X_1 \neq \Delta, \ldots, X_{j-1} \neq \Delta, X_j = \Delta \mid X_0 = \Delta]$.*

Proof Partition $\{X_r = \Delta\}$ by the first return time j and apply the Markov property.

Lemma 2 (Double stochasticity of the S-box kernel) *For a bijective 4-bit S-box, the normalized DDT matrix $G = \mathrm{DDT}/16$ is doubly stochastic and is the direct sum $G = [1] \oplus G'$, where $G' = (\mathrm{DDT}_S[a][b]/16)_{a,b\neq 0}$ is a 15×15 doubly stochastic matrix.*

Proof Row sums are 16 because each input difference partitions the 16 plaintexts. Column sums are 16 because $\sum_a \mathrm{DDT}_S[a][b]$ counts the pairs (x, y) with $S(x) \oplus S(y) = b$; for fixed x the equation has exactly one solution y and $y \neq x$ whenever $b \neq 0$. Finally $\mathrm{DDT}_S[0][0] = 16$ and $\mathrm{DDT}_S[a][0] = \mathrm{DDT}_S[0][b] = 0$ for $a, b \neq 0$, which is the direct-sum structure.

Lemma 3 (Round kernel is doubly stochastic) *For any bijective linear layer L, the round matrix M is doubly stochastic. The zero state is isolated and absorbing, and M restricted to the $2^n - 1$ nonzero states is doubly stochastic.*

Proof Row sums equal one by probability conservation. For column sums, for fixed B the S-box output y is fixed and $z \mapsto A$ is a bijection, so

$$\sum_{A} M[A, B] = \sum_{z} \prod_{m} \frac{\mathrm{DDT}_S[z_m, y_m]}{16}. \tag{7}$$

The product factorizes over nibbles, and each factor sums to 16 by Lemma 2; hence the column sum is one.

Theorem 7 (Long-round spectral floor) *Assume M is primitive on the nonzero states. Then $DP_r(\Delta) \to 1/(2^n-1)$. If, in addition, M is diagonalizable on the nonzero states, then*

$$DP_r(\Delta) - \frac{1}{2^n-1} = \sum_{i\geq 2} \lambda_i^r \, r_i(\Delta)\, \ell_i(\Delta), \tag{8}$$

with λ_i the eigenvalues of the nonzero-block transition matrix and r_i, ℓ_i the right/left eigenvectors with $\ell_i^{\mathsf{T}} r_j = \delta_{ij}$. Moreover, for any bijective linear layer and $N \geq 2$, $\sigma_2(M) = 1$, so the standard operator-norm bound $\|M^r - \Pi\|_2 \leq \mu^r$ is vacuous.

Proof For a doubly stochastic irreducible aperiodic chain the uniform distribution $1/(2^n-1)$ on nonzero states is the unique stationary distribution, so $DP_r(\Delta) \to 1/(2^n-1)$. Under diagonalizability, the spectral decomposition of the primitive transition matrix gives the displayed identity; a Jordan block would only add polynomial prefactors and would not affect convergence. For the operator-norm clause, any bijective linear layer L induces a bijection $A \mapsto L(A)$ on the state set, and the matrix of this bijection on the delta-state basis is a permutation matrix, hence orthogonal, even when L is not a bit permutation; so under the convention $M[A,B] = \Pr(X_{t+1} = B \mid X_t = A)$, write $M = G^{\otimes N} P$ when the linear layer follows the S-box layer and $M = P G^{\otimes N}$ when it precedes it (as in CRAFT); P is the permutation matrix of L^{-1} in the first case and of L in the second, and $G^{\otimes N} = G \otimes \cdots \otimes G$ (N factors of the 16×16 kernel, hence a $2^n \times 2^n$ matrix). Because P is orthogonal, M and $G^{\otimes N}$ have identical singular values. By Lemma 2, $G = [1] \oplus G'$ with G' a 15×15 doubly stochastic matrix, so $G'\mathbf{1} = \mathbf{1}$ and $G'^{\mathsf{T}}\mathbf{1} = \mathbf{1}$, whence $(G'^{\mathsf{T}}G')\mathbf{1} = \mathbf{1}$; together with the isolated [1] block, $G^{\mathsf{T}}G = [1] \oplus (G'^{\mathsf{T}}G')$ has the eigenvalue 1 with multiplicity at least two, hence G has at least two unit singular values. The singular values of a Kronecker product are the pairwise products of the singular values of its factors, so $G^{\otimes N}$—and therefore M—has at least 2^N unit singular values. The zero state is an isolated 1×1 block [1] in both $G^{\otimes N}$ and M, contributing exactly one unit singular value; removing it leaves at least $2^N - 1$ unit singular values on the nonzero block; for $N \geq 2$ this is at least two, so the second singular value of the nonzero block is 1, i.e. $\sigma_2(M) = 1$. Since M is doubly stochastic, the stationary projection Π satisfies $M\Pi = \Pi M = \Pi$ and $\Pi^2 = \Pi$, so $M^r - \Pi = (M - \Pi)^r$, and the operator-norm route collapses to

$$\|M^r - \Pi\|_2 = \|(M-\Pi)^r\|_2 \leq \|M - \Pi\|_2^r = \sigma_2(M)^r = 1^r = 1,$$

whence the standard bound $\|M^r - \Pi\|_2 \leq \mu^r$ is vacuous.

The structural vacuity of operator-norm bounds.

Remark 6 (Algebraic necessity, not an accident) The equality $\sigma_2(M) = 1$ for $N \geq 2$ is an algebraic necessity of the Kronecker structure $M = G^{\otimes N} P$, not an accident of a particular S-box or linear layer. This observation clarifies a

subtlety of the spectral approach: the per-nibble contraction μ^r is an appealing but non-transferable estimate for full ciphers, and Theorem 7 makes this boundary precise. Every bijective 4-bit S-box makes the normalized DDT block G' doubly stochastic, so $G = [1] \oplus G'$ carries at least two unit singular values, and the singular spectrum of a Kronecker product multiplies these units outward to at least 2^N unit singular values of M. Consequently no closed-form analytical spectral bound—whether operator-norm or singular-value based, and however tightly the per-nibble contraction is estimated—can produce a nontrivial quantitative bound on DP_r for a full cipher with $N \geq 2$ at any finite number of rounds r. Note that our own structural interval of Theorem 4 and combinatorial tools such as MILP/SAT operate on entirely different principles and are not affected by this algebraic obstruction. This eliminates the wrong spectral route and identifies $\lambda_2(M)$ as the only meaningful quantity for future spectral analysis, a conclusion the experiments of Section 6.1 and Section A.1 corroborate. The discrepancy between $\sigma_2(M) = 1$ and $|\lambda_2(M)| < 1$ arises because M is non-normal: the singular values are forced to account for the worst-case direction across all inputs, while the eigenvalues follow the intrinsic eigendirections of M, with weights that depend on the specific Δ. The $\sigma_2(M) = 1$ clause is established through the factorizations of Theorem 7: $M = G^{\otimes N} P$ when the linear layer follows the S-box layer and $M = PG^{\otimes N}$ when it precedes it (as in CRAFT), with P an orthogonal permutation matrix in either case; the lower-bound and interval results of Theorems 3–4 hold for every bijective linear layer, including CRAFT's $L = \mathrm{PN} \circ \mathrm{MC}$, whereas the convergence to the uniform floor holds under the primitivity assumption of Theorem 7.

Remark 7 (Computed gaps) The per-nibble contraction is $\mu = \sigma_2(G') = 0.638171$ for PRESENT and CRAFT and 0.636359 for GIFT. By Theorem 7, however, μ^r does *not* bound the full nonzero chain for any bijective linear layer: for $N \geq 2$ the nonzero-block matrix has at least $2^N - 1$ unit singular values. The correct entrywise rate is the second eigenvalue $\lambda_2(M)$. A universal bound $|\lambda_2(M)| \leq \mu$ fails already for permutation matrices: a single-cycle permutation on $N = 2$ nonzero states has all non-principal eigenvalues on the unit circle, so $|\lambda_2| = 1 > \mu$. On the 16-bit toy of Section 6.1, $\lambda_2(M) = 0.238$ and $c_2 = |r_2(\Delta)\ell_2(\Delta)| \approx 1.8 \times 10^{-3}$; on the 4-nibble CRAFT-like instance $\lambda_2(M) = 0.4045$. The Dobrushin ergodicity coefficient is also trivial ($\delta(G') = 1$ for all three S-boxes). A bound on $\lambda_2(M)$ for the 64-bit ciphers is the main open tightening step.

4.6 K-Optimality Theory

Definition 4 (K, Δ, r) is saturated if $P_{\mathrm{comp}} = DP_r(\Delta)$.

Theorem 8 (Saturation criterion) *(K, Δ, r) is saturated iff no r-round path from Δ to Δ has an intermediate state of activity greater than K.*

Proof $DP_r(\Delta)$ sums all Δ-to-Δ paths; P_{comp} sums exactly those whose intermediate states lie in W_K. If every path lies in W_K, the sums coincide; otherwise the missing paths have strictly positive weight.

Lemma 4 (State-count complexity) $|W_K| = \sum_{j=0}^{K} \binom{N}{j}(2^4-1)^j$. *Thus the state count grows exponentially in K; the K-active propagator is fixed-parameter tractable in K and independent of 2^n.*

Proof For each j-subset of active nibbles there are 15^j nonzero values.

Remark 8 (CRAFT $\Delta^\star$) For the CRAFT differential $\Delta^\star$ of Section 6, exact computation gives $K_{\min}(\Delta^\star, r) = 4$ for $r \leq 3$: the strict lower bounds for $K = 4, 5, 6$ coincide at $r = 1, 2, 3$. Exact K=5 propagation completes through $r = 12$, its state count saturating at 902,400 states against 152,250 for K=4, and its lower bounds coincide with K=4 at every round $1 \leq r \leq 12$; this is a numerical agreement, not a convergence proof. Exact K=6 propagation still exceeds available memory at round 3, so K=4 remains the practical Pareto-optimal budget. Saturation in K is thus an experimentally supported hypothesis, not a theorem.

4.7 Cross-Cipher Generalization

Theorem 9 *The Markov identity of Theorem 1, the doubly-stochastic floor and the spectral identity of Theorem 7, and the lower-bound/interval results of Theorems 3–4 hold unchanged for both (i) a bit permutation $L = P$ and (ii) a nibble-wise MixColumns followed by a nibble permutation $L = \mathrm{PN} \circ \mathrm{MC}$ (CRAFT). The $\sigma_2(M) = 1$ clause of Theorem 7 holds for any bijective linear layer.*

Proof In both cases the transition kernel has the DDT-product form $M[A][B] = \prod_m \mathrm{DDT}_S[z_m, y_m]/16$ with bijective L and (z, y) the S-box input/output difference pair ($z = A$, $y = L^{-1}(B)$ for a bit permutation; $z = L(A)$, $y = B$ for CRAFT). The proofs only use the product form of M and the support operation supp, never the internal representation of L. For a bit permutation, $\mathrm{supp}(B)$ is read after gathering bits across nibbles; for CRAFT, each active S-box input nibble produces a nonzero output nibble. Therefore the K-active truncation and the interval bound apply verbatim.

5 Algorithms

1. **Candidate enumeration.** Enumerate all one-round iterative differentials with $\mathrm{act}(\Delta) \leq k_0$ using the ψ-closure FPT enumerator of Theorem 2. A maximum-remaining-value (MRV) order is only a solver acceleration and does not change the enumeration result.
2. **Activity-parameterized propagation.** Run Algorithm 1 and output P_{comp}, $\bar{P}_{\text{drop}}$, and the state counts $\{|q_t|\}_{t \leq r}$.

3. **Symmetry acceleration.** Run Algorithm 2 to fold the propagation over $\mathrm{Stab}_G(\Delta)$; this preserves P_{comp} and $\bar{P}_{\mathrm{drop}}$ exactly while shrinking the state set by the mean orbit size.

Algorithm 1 Activity-parameterized propagation

Require: Δ, K, r, $\varepsilon \geq 0$
Ensure: P_{comp}, $\bar{P}_{\mathrm{drop}}$, state counts $\{|q_t|\}_{t \leq r}$
$q_0 \leftarrow \{\Delta : 1\}$, $D \leftarrow 0$
for $t = 1$ **to** r **do**
  for each $(A, p) \in q_{t-1}$ **do**
    $z \leftarrow L(A)$ for CRAFT, $z \leftarrow A$ for a bit permutation
    for each y with $\mathrm{DDT}_S[z_m][y_m] > 0$ for all m **do**
      $B \leftarrow y$ for CRAFT, $B \leftarrow L(y)$ for a bit permutation
      if $\mathrm{act}(B) \leq K$ **then**
        $q_t[B] \leftarrow q_t[B] + p \prod_m \mathrm{DDT}_S[z_m][y_m]/16$
      else
        $D \leftarrow D + p \prod_m \mathrm{DDT}_S[z_m][y_m]/16$
      end if
    end for
  end for
  if $\varepsilon > 0$ **then**
    remove from q_t all states with mass $< \varepsilon$ and add the removed mass to D
  end if
end for
return $P_{\mathrm{comp}} = q_r(\Delta)$, $\bar{P}_{\mathrm{drop}} = D$, and the state counts $\{|q_t|\}_{t \leq r}$

Proof Correctness follows from Theorems 3 and 4. The optional ε-pruning only removes nonnegative path weights, so P_{comp} remains a strict lower bound.

Proof Each $g \in H$ fixes Δ and satisfies $M[gA][gB] = M[A][B]$, so g is an automorphism of the transition matrix. The H-orbits therefore form a quotient chain whose folded kernel $Q([A],[B]) = \sum_{B' \in [B]} M[A][B']$ is independent of the representative A. Algorithm 2 evolves the orbit-summed mass exactly; because Δ is fixed by H, the H-orbit of Δ is the singleton $\{\Delta\}$, so the returned value $q_r(\mathrm{can}(\Delta)) = q_r(\Delta) = P_{\mathrm{comp}}$ and the accumulated D equal the output of Algorithm 1. The speed-up is the mean H-orbit size.

Remark 9 (CRAFT instantiation) The full symmetry group of the CRAFT one-round transition is the Klein four-group $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$: the identity, the left–right half-swap of the 4×4 state, the S-box value automorphism swapping $\{1 \leftrightarrow 4,\ 3 \leftrightarrow 6,\ 9 \leftrightarrow 12,\ 11 \leftrightarrow 14\}$, and their composition. For the differential $\Delta^\star$ the stabilizer is $\mathrm{Stab}_G(\Delta^\star) = \{\mathrm{id}, \rho_1\}$ of order 2, where ρ_1 is the value automorphism fixing the active nibbles of $\Delta^\star$. Algorithm 2 therefore halves the reachable state set; at $r = 3$ the K=4 propagation collapses 98,337 states to 51,053 orbits with identical P_{comp} and $\bar{P}_{\mathrm{drop}}$. This factor-two reduction does not remove the exponential growth in K, so it does not by itself make K=6 tractable.

Algorithm 2 Symmetry-folding acceleration of activity-parameterized propagation

Require: Δ, K, r, and a group G of transition symmetries with $M[gA][gB] = M[A][B]$ for all $g \in G$
Ensure: P_{comp}, $\bar{P}_{\mathrm{drop}}$; identical to the output of Algorithm 1
$H \leftarrow \mathrm{Stab}_G(\Delta) = \{g \in G : g\Delta = \Delta\}$
define $\mathrm{can}(A) = \min_{g \in H} \mathrm{enc}(gA)$, the lexicographic minimum of the H-orbit of A
$q_0 \leftarrow \{\mathrm{can}(\Delta) : 1\}$, $D \leftarrow 0$
for $t = 1$ **to** r **do**
  for each $(\bar{A}, p) \in q_{t-1}$ **do**
    $z \leftarrow L(\bar{A})$ for CRAFT, $z \leftarrow \bar{A}$ for a bit permutation
    for each y with $\mathrm{DDT}_S[z_m][y_m] > 0$ for all m **do**
      $B \leftarrow y$ for CRAFT, $B \leftarrow L(y)$ for a bit permutation
      if $\mathrm{act}(B) \le K$ **then**
        $q_t[\mathrm{can}(B)] \leftarrow q_t[\mathrm{can}(B)] + p \prod_m \mathrm{DDT}_S[z_m][y_m]/16$
      else
        $D \leftarrow D + p \prod_m \mathrm{DDT}_S[z_m][y_m]/16$
      end if
    end for
  end for
end for
return $P_{\mathrm{comp}} = q_r(\mathrm{can}(\Delta))$, $\bar{P}_{\mathrm{drop}} = D$

6 Numerical Results

Every K-active entry in Table 1 is a strict lower bound on the closed-loop return probability $DP_r(\Delta^\star) = (M^r)[\Delta^\star, \Delta^\star]$: the inequality $M_K \le M$ is entrywise, so the truncated power never overestimates the true probability; the rightmost column reports the exact residual mass $\bar{P}_{\mathrm{drop}}$ that guarantees the lower bound is never an overestimate. The state counts in the table are the realized support sizes $|q_t|$ of the propagator—the states reachable from $\Delta^\star$ within activity K—which are far below the worst-case bound $|W_K|$ of Lemma 4: the propagator materializes only reachable states, so for CRAFT the K=4 count saturates at 152,250 rather than $|W_4| = 94{,}054{,}741$. Against the best single trail 2^{-8r}, the K=4 bound is larger by a factor of about $2^{36.1}$ at round 11 and $2^{40.1}$ at round 12, so the multi-path clustering effect is *proved*, not merely estimated. We emphasize that $\Delta^\star = [0, 0, \mathrm{A}, 0,\ 0, \mathrm{A}, 0, 0,\ 0, \mathrm{D}, 0, 0,\ 0, \mathrm{D}, 0, 0]$ is a one-round iterative differential of CRAFT (edp $= 2^{-8}$), found by the low-activity enumerator of Section 4.2; it is an exact closed-loop object independent of the Hadipour et al. analysis. Hadipour et al. [2] report an 11-round open-loop differential of CRAFT with heuristic probability $2^{-49.79}$, clustered from $2^{29.66}$ optimal trails. We cite this value only as a yardstick of how strongly clustering concentrates probability in CRAFT—the regime in which a single trail most understates the differential—not as a value to be compared with our closed-loop bound, since it is a heuristic estimate of a *different* (open-loop) differential. What matters is that our strict lower bound $2^{-51.8503}$ is non-trivial in exactly this hard-clustering regime, where the best single trail of $\Delta^\star$ is only 2^{-88}.

Table 1 CRAFT K-active ablation for the iterative differential $\Delta^\star = [0,0,A,0,\ 0,A,0,0,\ 0,D,0,0,\ 0,D,0,0]$, edp $= 2^{-8}$. Entries are "LB $(-\log_2)$ / #States"; K=5 is exact through $r = 12$ and reproduces the K=4 bound at every round ("$--$" marks rounds not computed), while K=6 exhausts memory at round 3. The "single-trail" column is edp$^r = 8r$, the best single-trail bound. The rightmost column is the K=4 cumulative dropped mass $\bar{P}_{\mathrm{drop}}$, shown as $1 - \mu_r$, where μ_r is the total kept mass after r rounds; it is reported only as a structural reference, not as a tight upper bound. For reference, Hadipour et al. [2] report an 11-round *open-loop* differential—a different difference, not an iterative differential—with heuristic probability $2^{-49.79}$, clustered from $2^{29.66}$ optimal trails; it is cited only as a yardstick of how strongly clustering concentrates probability in CRAFT, not as a value comparable with the closed-loop bounds tabulated here.

r	K=3	K=4	K=5	K=6	single-trail $8r$	$\bar{P}_{\mathrm{drop}}$ (K=4)
1	0/0	8.0000/864	8.0000/864	8.0000/864	8	0
2	0/0	12.4150/21,112	12.4150/412,384	12.4150/412,384	16	0.8125
3	0/0	17.1545/98,337	17.1545/842,464	17.1545/13,513,864	24	$1-1.6632\times10^{-2}$
4	0/0	21.6919/152,238	21.6919/902,388	OOM	32	$1-1.7593\times10^{-3}$
5	0/0	26.2212/152,250	26.2212/902,400	OOM	40	$1-1.1618\times10^{-4}$
6	0/0	30.6491/152,250	30.6491/902,400	OOM	48	$1-7.5982\times10^{-6}$
7	0/0	35.0294/152,250	35.0294/902,400	OOM	56	$1-4.7943\times10^{-7}$
8	0/0	39.3214/152,250	39.3214/902,400	OOM	64	$1-3.0236\times10^{-8}$
9	0/0	43.5525/152,250	43.5525/902,400	OOM	72	$1-1.8773\times10^{-9}$
10	0/0	47.7232/152,250	47.7232/902,400	OOM	80	$1-1.1663\times10^{-10}$
11	0/0	51.8503/152,250	51.8503/902,400	OOM	88	$1-7.2082\times10^{-12}$
12	0/0	55.9438/152,250	55.9438/902,400	OOM	96	$1-4.4548\times10^{-13}$
13	0/0	60.0134/152,250	—	OOM	104	$1-2.7482\times10^{-14}$
14	0/0	64.0668/152,250	—	OOM	112	$1-1.6953\times10^{-15}$
15	0/0	68.1090/152,250	—	OOM	120	$1-1.0451\times10^{-16}$

The strength of the bound is best seen against the single-trail baseline. Because the bound is strict, the differential probability is at least $2^{-51.8503}$ at round 11 and $2^{-55.9438}$ at round 12; a single-trail analysis would only guarantee 2^{-88} and 2^{-96}, so the activity decomposition recovers factors of about 2^{36} and 2^{40} that any single-trail argument misses, while a heuristic cluster estimate carries no guarantee at all. The lower bound is therefore a certified statement about the probability itself rather than a plausibility argument.

K=3 cannot propagate $\Delta^\star$ beyond the first round. Exact K=5 propagation completes through round 12 and reproduces the K=4 bound at every round, so K=4 is the Pareto-optimal budget: it already captures the return mass of $\Delta^\star$, while K=6 exhausts memory at round 3. Enumerating all one-round iterative differentials of CRAFT shows that this saturation is not an artifact of $\Delta^\star$: there are exactly 544 such differentials, all of activity 4, and on a random sample of six of them the K=4 and K=5 bounds coincide, confirming that the short-round saturation is a family-wide structural phenomenon.

The same framework applies unchanged to the bit-permutation ciphers PRESENT-64 and GIFT-64; the resulting lower bounds, and the gain over the single-trail baseline, are reported in Appendix A (Table 2 and Table 3).

The two iterative differentials are located by the same ψ-closure enumerator of Section 4.2 that produced $\Delta^\star$; neither is claimed optimal. The larger GIFT

state count is a structural property of GIFT's denser bit permutation, not a shortcoming of the algorithm.

6.1 Exact Validation on a 16-bit Toy SPN

To validate the framework against ground truth we instantiate a 16-bit SPN (N=4 nibbles) with the PRESENT S-box and a random bit permutation. The full state space has 65536 states, so the exact DP_r is computed by full iteration and compared with the K-active lower bounds. The optimal iterative $\Delta = \texttt{0x0089}$ has edp $= 2^{-5}$. For every K and every $r \leq 8$ the lower bound $\mathrm{LB}_K(r) \leq DP_r(\Delta)$ is verified exactly and converges monotonically to the truth as K increases. On this small instance the reverse-reachability tightening of Theorem 5 removes the residual mass exactly at short rounds, so the containment $\mathrm{LB}_K(r) \leq DP_r(\Delta) \leq P_{\mathrm{comp}} + \bar{P}_{\mathrm{drop}}$ holds at every round; Table 4 reports the exact values. The toy validation therefore confirms the strictness and the monotone convergence of the lower bound, and the containment of Theorem 4. The upper bound is a structural guarantee rather than a tight estimate: here the tightening is exact through round 3 and sharp at round 4, then degrades—exactly the behavior that makes the lower bound the operative quantity on full-size ciphers. The exact DP_r converges to the floor $1/65535$ predicted by Theorem 7 ($DP_{14} = 2^{-15.999978}$). Block orthogonal iteration on the deflated chain $N = M - \Pi_X - e_0 e_0^{\mathsf{T}}$ yields $\lambda_2(M) = 0.238$ and $c_2 = |r_2(\Delta)\ell_2(\Delta)| \approx 1.8 \times 10^{-3}$; a two-eigenvalue truncation $|DP_r - \mathrm{floor}| \approx c_2|\lambda_2|^r + (1-c_2)|\lambda_3|^r$ matches the exact deviation within a factor 4 at $r = 14$. This is the correct eigenvalue bound; the singular-value bound is vacuous here ($\sigma_2(M) = 1$ by Theorem 7).

Reproducibility. The K-active propagator and the reverse-reachability bound are implemented in a single Python code base with the S-boxes and linear layers of CRAFT, PRESENT, and GIFT; no cipher-specific assumption is used. The code and the raw tables used to produce every table in this paper are available from the corresponding author upon reasonable request. On a single core the full CRAFT K=4 propagation to round 12 completes in about 35 seconds with 152,250 states; the K=6 run reaches 13.5 million states at round 3 before exhausting memory, which is the practical ceiling behind the K=4 budget.

At 15 rounds the provable lower bound is $2^{-68.109}$, below the 2^{-64} random threshold; this demonstrates the reach of the method beyond the 2^{-64} threshold.

7 Related Work

Classical differential cryptanalysis reports edp^r for a single trail [1], and truncated differentials [10] generalize the search to activity patterns but remain qualitative. Markov-chain formulations of differential probability date back to

Lai–Massey–Murphy [4]. Multi-trail clustering estimates the differential effect, as in the CRAFT analysis of Hadipour et al. [2]. Along the same direction, Kwon and Song [3] recently compute differential probabilities of word-based block ciphers by exactly enumerating the maximum differential characteristics over a multistage graph and by incorporating lower-probability characteristics, obtaining estimates closer to the true differential probability than any single characteristic; they observe the characteristic-based overestimation of security in particular for SKINNY, Midori, and CRAFT. The congruent differential cluster method of Cui et al. [14] computes the exact clustered differential probability of Midori-64, CRAFT-64, and SKINNY-64, while the state-test technique of M'Foukh et al. [15] yields a 26-round differential attack on CRAFT, the current record. CRAFT has also been analysed in the weak tweak-key [20] and impossible-differential [21] settings; these address different attack models and do not bear on the clustering of a fixed iterative differential, which is our object of study. These methods target a fixed input/output differential but do not return an entrywise-justified lower bound. The instantiated designs are the bit-permutation ciphers PRESENT [11] and GIFT-64 [12], and the MixColumns-based CRAFT [13]. Our **activity decomposition** turns the exact but high-dimensional Markov power into a principled finite approximation with an entrywise-justified lower bound.

Relative to these lines of work, the present paper advances the state of the art in three respects. First, it establishes a structural spectral fact: for a parallel SPN with $N \geq 2$ nibbles the second singular value of the round matrix is exactly $\sigma_2(M) = 1$, so the operator-norm route is vacuous and the correct convergence rate is the second eigenvalue $\lambda_2(M)$. Second, it replaces the qualitative output of classical truncated and single-trail analysis with an entrywise-justified lower bound: every reported number is a strict lower bound on DP_r, guaranteed by an exact structural residual-mass accounting, rather than a single-trail edp^r or a heuristic cluster estimate. Third, the activity decomposition is generic over any 4-bit bijective S-box and any bijective linear layer, and it transfers unchanged between bit-permutation and MixColumns-based SPNs.

8 Limitations and Discussion

Distinction between the decomposition and spectral estimation. The lower bound of Theorem 3 is a finite, constructive, per-instance result: it keeps every path that stays inside the low-activity subspace W_K, computed directly from M and W_K, irrespective of the spectral properties of M. The eigenvalue identity of Theorem 7 is instead an asymptotic, analytic description of the long-round convergence rate. The two are complementary and must not be conflated; the vacuity of the operator-norm bound for $N \geq 2$ concerns only the spectral estimation route and does not affect the validity of the decomposition lower bound. To avoid any misunderstanding: the failure of the singular-value route does not mean that no upper bound can be established.

Indeed, Theorem 4 itself provides a structural upper bound ($P_{\text{comp}} + \bar{P}_{\text{drop}}$), and combinatorial tools (e.g., MILP) can effectively bound single-trail probabilities. Our negative result specifically targets the analytical spectral bound $\|M^r - \Pi\|_2 \leq \mu^r$, which was historically expected to govern the convergence rate but is now proven to be algebraically vacuous.

Each item below states a boundary precisely, and each doubles as a research direction; none of them weakens the reported lower bounds, which are unconditional theorems.

- *Residual mass at long rounds.* $\bar{P}_{\text{drop}}$ counts every discarded prefix, most of which never returns to Δ; on the full-size ciphers it is loose at long rounds (for CRAFT at round 11 it is ≈ 1), so this paper treats P_{comp} as the operative quantity and keeps the two-sided containment of Theorem 4 only as the structural guarantee that the lower bound is strict, not as an experimental claim. On full-size ciphers the lower endpoint is therefore the mathematically informative quantity, while the upper endpoint serves only to certify that the lower bound is never an overestimate. The fact that the retained mass shrinks at long rounds is not a defect of the decomposition but a consequence of the missing analytic bound on $\lambda_2(M)$. Every route for strengthening the lower bound further is closed: the singular-value bound is vacuous for $N \geq 2$ (Theorem 7); exact reverse reachability (Theorem 5) removes all dropped mass at $r \leq 9$, 12.5% at $r = 10$ and 1% at $r = 11$, and then saturates because $\mathcal{R}_s$ reaches all 65535 nonzero patterns at $s = 11$; and raising K cannot help, since for any fixed K the mass retained in W_K tends to zero at long rounds. Even a probability-weighted refinement of Theorem 5—weighting the residual reachable mass by the maximal one-round return probability into $\Delta^\star$, $\rho_1 = \prod_{m \in \text{supp}(\Delta^\star)} \max_z \text{DDT}_S[z][\Delta^\star_m]/16 = 2^{-8}$—would only lower the upper endpoint to about 2^{-8}, because the DDT bound constrains the final step alone while the intermediate steps are bounded only by stochasticity, so the endpoint still saturates and remains loose. The operative provable results remain the lower bounds $2^{-51.8503}$ at round 11 and $2^{-55.9438}$ at round 12; strengthening them further requires a bound on $\lambda_2(M)$ for the full 64-bit cipher, which remains open.
- *Analytic bound on $\lambda_2(M)$.* The singular-value route is closed as a matter of algebra: for $N \geq 2$ we proved $\sigma_2(M) = 1$ (Theorem 7), so no operator-norm estimate can bound a full-size cipher. This isolates $\lambda_2(M)$ as the unique spectral quantity of interest. The eigenvalue is observed as 0.238 on the exact toy and 0.4045 on the 4-nibble CRAFT-like instance, so the correct rate is well below 1 in every measured case. A sharp analytic bound on $\lambda_2(M)$ and its projection coefficients for the 64-bit ciphers remains the central challenge for the next generation of spectral arguments in this area; it is a direction we have now made precise, not a shortcoming of the present results, whose lower bounds are unconditional.
- *Primitivity of the full-size chain.* The long-round floor of Theorem 7 is conditional on M being primitive (irreducible and aperiodic) on the nonzero

states. We verify this condition wherever the full support graph is tractable: for all three S-boxes the 15-node DDT support graph is strongly connected and contains a self-loop (a one-round iterative differential), hence aperiodic; and on the 16-bit toy and the 4-nibble CRAFT-like instance the complete $(2^{16} - 1)$-node support graph is strongly connected and contains a self-loop, so the chain is primitive. For the 64-bit ciphers primitivity is not proven and is adopted as the standing assumption of Theorem 7, the irreducibility/aperiodicity hypothesis that standard spectral treatments of these chains require.

- *Pruning correctness.* Removing nonnegative path weights keeps a subset of DP_r, so the kept mass is still a strict lower bound.
- *Complexity of K.* The propagator is fixed-parameter tractable in the activity bound K: its state space has size $|W_K| = \sum_{j \le K} \binom{N}{j} 15^j$, independent of the full block size 2^n, so raising K is a parameter trade-off rather than a methodological ceiling. The ablation supports the chosen budget: K=3 collapses to the single-trail bound, while exact K=5 propagation reproduces the K=4 lower bounds at every round $1 \le r \le 12$; this saturation is strong evidence that K=4 already captures the return mass of $\Delta^\star$. K=6 exceeds available memory at round 3, so K=4 is the Pareto-optimal budget. Saturation in K is an experimentally supported hypothesis, not a theorem.
- *Scope: iterative differentials.* The decomposition of Theorem 4 holds for an arbitrary differential pair (Δ, Δ'). We restrict the present treatment to closed-loop iterative differentials because the ψ-graph closure, the first-return decomposition and the long-round floor exploit the closed-loop structure; extending the bound to open-loop differentials is left to future work.
- *No optimality claim.* We fix the one-round iterative differential $\Delta^\star$ of CRAFT and bound its closed-loop return probability; we do not search for the optimal iterative differential or trail, and the reported bounds are not claimed to be best possible over all differences.
- *Relation to state of the art.* The operative proved results are at 11–12 rounds. The contribution is the provable methodology rather than a longer distinguisher: Hadipour et al. report an 11-round open-loop differential with heuristic probability $2^{-49.79}$, a different object from the closed-loop return probability we bound, and we do not compare the two as if they coincided. Our closed-loop lower bounds $2^{-51.8503}$ and $2^{-55.9438}$ are strict and exceed the best single trail 2^{-8r} by factors of about 2^{36} and 2^{40}, proving that the clustering effect on this iterative differential is real and large.

9 Conclusion

We presented a provable Markov framework for iterative differentials of nibble-based SPNs. The activity decomposition gives strict, entrywise-justified lower bounds, is FPT in K, and transfers across bit-permutation and MixColumns-based ciphers. We proved that the normalized DDT kernel is doubly stochas-

tic for any bijective linear layer; under the primitivity assumption of Theorem 7 (irreducible and aperiodic on the nonzero states), the exact differential probability converges to the uniform floor $1/(2^n - 1)$. We proved that the operator-norm/singular-value bound is structurally vacuous for full ciphers with $N \geq 2$, and that the correct entrywise convergence rate is the second eigenvalue $\lambda_2(M)$. The framework is validated against the exact probability on a 16-bit toy instance by full state-space iteration, confirming the monotone convergence of the lower bound, and the K-ablation on CRAFT, PRESENT and GIFT-64 quantifies the tightness–complexity Pareto point. The operative experimental results are the provable CRAFT lower bounds $2^{-51.8503}$ and $2^{-55.9438}$ at rounds 11 and 12; these are strict, entrywise-justified bounds that exceed the best single trail 2^{-8r} by factors of about 2^{36} and 2^{40}, respectively, so they prove—rather than merely estimate—the multi-path clustering effect for this iterative differential. These lower bounds are strict statements about the differential probability itself, certified by the residual-mass accounting of Theorem 4, rather than plausibility arguments. These bounds are the lower endpoints of the two-sided containment of Theorem 4: they are exact retained-walk masses that increase with K, while the upper endpoint, computed from the discarded mass, decreases with K and meets the lower endpoint at $K = N$ (Remark 2). We emphasize that the aim of this work is not to extend the best known distinguisher, but to supply the missing mathematical guarantee for a fixed iterative differential. The remaining open step is an analytic bound on $\lambda_2(M)$ for the full 64-bit CRAFT.

Future work. The most immediate direction is an analytic bound on $\lambda_2(M)$ and its projection coefficients for the full 64-bit ciphers, which would sharpen the lower bound into a quantitative estimate of DP_r. A second direction is to tighten the dropped-mass bound further by exploiting cipher symmetries. Algorithm 2 already folds the stabilizer of Δ exactly, but for CRAFT this gives only a factor-two reduction (Remark 9) and does not remove the exponential growth in K; stronger structural or spectral tightening is needed to reduce $\bar{P}_{\text{drop}}$ at longer rounds. A third direction is to lift the bound from closed-loop iterative differentials to arbitrary open-loop pairs (Δ, Δ'); the decomposition bound and the reverse-reachability tightening apply verbatim once the propagator reads the off-diagonal entry $(M_K^r)[\Delta, \Delta']$, but the optimal open-loop differentials of the literature (e.g., the 11-round CRAFT differential of Hadipour et al., with input activity 7) lie beyond the feasible activity bound K, so reaching them is a genuine open problem rather than a routine extension. Finally, extending the framework to larger S-boxes, to other SPN and Feistel-type structures, and to the linear-hull setting would further test the generality of the activity decomposition.

Table 2 PRESENT-64 K=3 versus K=4 for $\Delta = [3, 1, 0, 0, 3, 0, \ldots, 0]$, edp = 2^{-8}. The rightmost column is the K=4 cumulative dropped mass $\bar{P}_{\text{drop}}$ shown as $1 - \mu_r$, reported only as a structural reference, not as a tight upper bound.

r	K=3 LB	K=3 #States	K=4 LB	K=4 #States	Gain	$\bar{P}_{\text{drop}}$ (K=4)
1	8.0	6	8.0	47	0.00	0.7890625
2	16.0	6	16.0	113	0.00	$1 - 1.6403 \times 10^{-3}$
3	24.0	6	24.0	992	0.00	$1 - 1.0803 \times 10^{-5}$
4	32.0	6	31.9776	12,866	0.022	$1 - 1.6552 \times 10^{-7}$
5	40.0	6	39.6657	124,470	0.334	$1 - 4.4004 \times 10^{-9}$
6	48.0	6	47.3648	489,300	0.635	$1 - 1.9603 \times 10^{-10}$
7	56.0	6	54.7763	814,416	1.224	$1 - 1.5479 \times 10^{-11}$
8	64.0	6	62.1891	903,504	1.811	$1 - 7.5641 \times 10^{-13}$
9	72.0	6	68.9747	908,620	3.025	$1 - 2.5509 \times 10^{-14}$
10	80.0	6	72.6849	908,620	7.315	$1 - 1.6091 \times 10^{-15}$
11	88.0	6	78.1441	908,620	9.856	$1 - 1.0619 \times 10^{-16}$
12	96.0	6	83.5532	908,620	12.447	$1 - 8.0578 \times 10^{-18}$

Gain denotes the tightening of K=4 over K=3 in bits: $(-\log_2 LB_{K=3}) - (-\log_2 LB_{K=4})$.

Table 3 GIFT-64 K=4 provable lower bounds for $\Delta = [1, 0, 0, 9, 0, \ldots, 0, 1, 0, \ldots, 0]$, edp = 2^{-9}. These bounds are reported only for cross-cipher generality; they lie below the 2^{-64} random-probability threshold. The rightmost column is the K=4 cumulative dropped mass $\bar{P}_{\text{drop}}$ shown as $1 - \mu_r$, reported only as a structural reference, not as a tight upper bound.

r	LB ($-\log_2$)	#States	Status	$\bar{P}_{\text{drop}}$ (K=4)
1	9.0000	128	provable LB	0.75
2	18.0000	2,849	provable LB	$1 - 1.1039 \times 10^{-2}$
3	27.0000	43,892	provable LB	$1 - 4.9561 \times 10^{-4}$
4	34.8915	187,288	provable LB	$1 - 2.4386 \times 10^{-5}$
5	42.4075	570,459	provable LB	$1 - 1.1757 \times 10^{-6}$
6	50.0719	1,061,830	provable LB	$1 - 8.3388 \times 10^{-8}$
7	57.7989	1,210,748	provable LB	$1 - 4.7309 \times 10^{-9}$
8	64.8378	1,227,246	provable LB	$1 - 2.8230 \times 10^{-10}$
9	69.2002	1,232,394	provable LB	$1 - 1.6992 \times 10^{-11}$
10	71.1158	1,232,896	provable LB	$1 - 1.0177 \times 10^{-12}$
11	75.3790	1,232,906	provable LB	$1 - 6.4391 \times 10^{-14}$
12	80.4163	1,232,906	provable LB	$1 - 4.3853 \times 10^{-15}$

A Cross-cipher numerical results

A second toy with a nibble-permutation layer shows that when the layer preserves nibble activity, the active count is invariant, so K=1 already reproduces the exact DP; this isolates the structural condition under which the activity decomposition is exact.

A.1 Eigenvalue Bound and the CRAFT-like Reduced Instance

For all three ciphers the per-nibble contraction is $\mu = \sigma_2(G')$ (0.638171 for PRESENT and CRAFT, 0.636359 for GIFT). Theorem 7 shows, however, that μ^r is *not* a valid bound on the full nonzero chain for any bijective linear layer: for $N \geq 2$ the restricted transition matrix

Table 4 Sandwich verification of Theorem 4 on the 16-bit toy SPN (K=3, PRESENT S-box, random bit permutation; Δ = `0x0089`, edp = 2^{-5}). All entries are $-\log_2$. Every row satisfies $P_{\text{comp}} \le DP_r \le P_{\text{comp}} + \bar{P}_{\text{drop}}$; the rightmost column is the structural upper bound of Theorem 4, which is loose at long rounds, while the reverse-reachability tightening of Theorem 5 recovers the exact DP_r through $r = 3$.

r	P_{comp} (K=3)	exact DP_r	$P_{\text{comp}} + \bar{P}_{\text{drop}}$
1	5.0000	5.0000	5.0000
2	10.0000	10.0000	1.9944
3	15.0000	15.0000	0.7059
4	17.3039	16.4657	0.2740
5	17.6004	16.0352	0.1162
6	18.6633	16.0263	0.0480
7	19.7856	16.0080	0.0203
8	20.9758	16.0016	0.0086

Table 5 Cross-cipher K=4 provable lower bounds ($-\log_2$). CRAFT and GIFT columns are strict lower bounds; PRESENT K=4 is reported through round 12. No spectral column is reported because the operator-norm bound is vacuous (Theorem 7). The rightmost three columns are the corresponding K=4 dropped masses $\bar{P}_{\text{drop}}$ shown as $1 - \mu_r$, reported only as a structural reference, not as a tight upper bound.

r	CRAFT LB	PRESENT LB	GIFT LB	CRAFT $\bar{P}_{\text{drop}}$	PRESENT $\bar{P}_{\text{drop}}$	GIFT $\bar{P}_{\text{drop}}$
9	43.5525	68.9747	69.2002	$1-1.8773\times10^{-9}$	$1-2.5509\times10^{-14}$	$1-1.6992\times10^{-11}$
10	47.7232	72.6849	71.1158	$1-1.1663\times10^{-10}$	$1-1.6091\times10^{-15}$	$1-1.0177\times10^{-12}$
11	51.8503	78.1441	75.3790	$1-7.2082\times10^{-12}$	$1-1.0619\times10^{-16}$	$1-6.4391\times10^{-14}$
12	55.9438	83.5532	80.4163	$1-4.4548\times10^{-13}$	$1-8.0578\times10^{-18}$	$1-4.3853\times10^{-15}$
13	60.0134	–	–	$1-2.7482\times10^{-14}$	–	–
14	64.0668	–	–	$1-1.6953\times10^{-15}$	–	–

PRESENT values are the K=4 bounds of Table 2; K=3 collapses to the single-trail bound 2^{-8r}.

has $2^N - 1$ unit singular values, so the operator-norm bound is vacuous. The correct bound is the eigenvalue decomposition of Theorem 7. On a 4-nibble CRAFT-like instance (MC4 + PN4 + CRAFT Sb0, the same column matrix as CRAFT's MC) block orthogonal iteration gives $\lambda_2(M) = 0.4045$ (plus -0.2816 and a complex pair 0.1014 ± 0.3899i, $|\cdot| = 0.4028$), confirming geometric convergence of rate $0.40 < 1$. This isolates the eigenvalue, not the singular value, as the correct bound, and leaves the full 64-bit bound on $\lambda_2(M)$ as the main open step.

This closes the loop between theory and experiment. The obstruction to the operator-norm route is structural ($\sigma_2(M) = 1$ by Theorem 7), yet the numerically computed spectrum exhibits genuine geometric contraction with $|\lambda_2(M)| \approx 0.40 < 1$, consistent with the exact eigenvalue identity of Theorem 7. The eigenvalue, not the singular value, is therefore both the mathematically correct and the empirically operative rate, and this directs the remaining analytic effort squarely at $\lambda_2(M)$ for the full 64-bit ciphers. The cross-cipher comparison of the K=4 lower bounds is tabulated in Table 5.

References

1. E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *J. Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

2. H. Hadipour, S. Sadeghi, M. M. Niknam, L. Song, and N. Bagheri, "Comprehensive security analysis of CRAFT," *IACR Trans. Symmetric Cryptol.*, vol. 2019, no. 4, pp. 290–317, 2020.
3. D. Kwon and J. Song, "Computing the differential probability of a word-based block cipher," *Cryptography*, vol. 9, no. 2, Art. no. 42, 2025.
4. X. Lai, J. L. Massey, and S. Murphy, "Markov ciphers and differential cryptanalysis," in *Advances in Cryptology—EUROCRYPT '91*, ser. LNCS, vol. 547. Springer, 1991, pp. 17–38.
5. N. Mouha, Q. Wang, D. Gu, and B. Preneel, "Differential and linear cryptanalysis using mixed-integer linear programming," in *Information Security and Cryptology—Inscrypt 2011*, ser. LNCS, vol. 7537. Springer, 2012, pp. 57–76.
6. S. Sun, L. Hu, P. Wang, K. Qiao, X. Ma, and L. Song, "Automatic security evaluation and (related-key) differential characteristic search: application to SIMON, PRESENT, LBlock, DES(L) and other bit-oriented block ciphers," in *Advances in Cryptology—ASIACRYPT 2014*, ser. LNCS, vol. 8873. Springer, 2014, pp. 158–178.
7. A. Abdelkhalek, Y. Sasaki, Y. Todo, M. Tolba, and A. M. Youssef, "MILP modeling for (large) S-boxes to optimize probability of differential characteristics," *IACR Trans. Symmetric Cryptol.*, vol. 2017, no. 4, pp. 99–129, 2017.
8. J. Daemen and V. Rijmen, "Probability distributions of correlation and differentials in block ciphers," *J. Math. Cryptol.*, vol. 1, no. 3, pp. 221–242, 2007.
9. C. Blondeau and B. Gerard, "Multiple differential cryptanalysis: theory and practice," in *Fast Software Encryption—FSE 2011*, ser. LNCS, vol. 6733. Springer, 2011, pp. 35–54.
10. L. R. Knudsen, "Truncated and higher order differentials," in *Fast Software Encryption—FSE 1994*, ser. LNCS, vol. 1008. Springer, 1995, pp. 196–211.
11. A. Bogdanov et al., "PRESENT: An ultra-lightweight block cipher," in *CHES 2007*, ser. LNCS, vol. 4727. Springer, 2007, pp. 450–466.
12. S. Banik, S. K. Pandey, T. Peyrin, Y. Sasaki, S. M. Sim, and Y. Todo, "GIFT: A small present," in *CHES 2017*, ser. LNCS, vol. 10529. Springer, 2017, pp. 321–345.
13. C. Beierle, G. Leander, A. Moradi, and S. Rasoolzadeh, "CRAFT: Lightweight tweakable block cipher with efficient protection against DFA attacks," *IACR Trans. Symmetric Cryptol.*, vol. 2019, no. 1, pp. 5–45.
14. T. Cui, Y. Mao, Y. Yang, Y. Zhang, J. Zhang, and C. Jin, "Congruent differential cluster for binary SPN ciphers," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2385–2397, 2024.
15. D. M'Foukh, M. Naya-Plasencia, and P. Neumann, "The state-test technique on differential attacks: a 26-round attack on CRAFT and other applications," in *Advances in Cryptography—ASIACRYPT 2025*, ser. LNCS, vol. 16245. Springer, 2025, pp. 253–284.
16. Z. Xu, Y. Li, L. Jiao, M. Wang, and W. Meier, "Do not misuse the Markov cipher assumption: automatic search for differential and impossible differential characteristics in ARX ciphers," IACR Cryptology ePrint Archive, Paper 2022/135, 2022.
17. E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd ed. Springer, New York, 2006.
18. R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, Cambridge, 2013.
19. T. Peyrin, Q. Q. Tan, H. Zhang, and C. Zhou, "Trail-Estimator: An automated verifier for differential trails in block ciphers," IACR Cryptology ePrint Archive, Paper 2025/396, 2025.
20. G. Leander and S. Rasoolzadeh, "Weak tweak-keys for the CRAFT block cipher," *IACR Trans. Symmetric Cryptol.*, vol. 2022, no. 1, pp. 38–63, 2022.
21. F. Liu, Y. Li, H. Chen, L. Ma, Z. Liu, and W. Yang, "Impossible differential cryptanalysis of lightweight tweakable block cipher CRAFT," *Cybersecurity*, vol. 8, 2025. DOI: 10.1186/s42400-025-00393-2.

Declarations

Funding. The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests. The author declares that he has no competing interests that are relevant to the content of this article.

Ethics approval. This article contains no studies involving human participants or animals; accordingly, no ethics approval is required.

Data availability. All data generated or analysed during this study, together with the source code used to produce the numerical tables, are available from the corresponding author upon reasonable request.

Declaration of generative AI in scientific writing

During the preparation of this work, the author used generative AI tools (large language models) to assist with language editing, drafting, and improving readability. All mathematical derivations, proofs, and experimental results were carried out independently by the author. After using these tools, the author reviewed and edited the content and takes full responsibility for the accuracy and integrity of the manuscript.