Practical Collision Cryptanalysis of Waterfall

MiniMax

November 2026

Abstract

Waterfall is a cryptographic hash function family submitted by Hattersley to Round 1 of the NIST SHA-3 competition in 2008. Its compression function updates three coupled feedback shift registers through a byte-cascade permutation, feeds two write-only accumulator pools, and injects periodic block counters. Prior cryptanalysis was limited to Fluhrer's 2008 theoretical note proposing a 2^{70} two-excursion strategy, leaving open whether practical collision generators or short single-excursion colliding pairs exist.

We resolve this question with a practical algebraic differential collision attack on the full Waterfall family across all four NIST digest sizes (224, 256, 384, and 512 bits). First, we characterize the global maximum differential fixed-point sets of the byte-cascade permutation and isolate two differences whose 2^{13} fixed points form a single affine coset with an unconstrained top byte, converting nonlinear step conditions into linear constraints. Second, we cancel accumulator pool differences within a single excursion via even-parity activation per pool index while steering mandatory block-counter steps through message modification and inverse permutation lookups. Third, we project step recurrences onto the low three bytes to accelerate message solving by two orders of magnitude. Our solver generates verified 8-block (512-byte) full-state collisions in 3.5 ms and minimal steerable 7-block (448-byte) single-excursion collisions in 52.1 ms on a single core, exposing structural weaknesses in trunk-plus-accumulator hash designs.

1 Introduction

During the NIST SHA-3 competition (2008–2012) [11], 51 candidate hash families entered first-round evaluation [13], and several designs—such as Waterfall [8] and StreamHash [15]—sought high software hashing speed by pairing lightweight state updates during message absorption with a heavier mixing phase executed once during finalization. Earlier stream-oriented hash designs such as PANAMA [5] and RadioGatún [1] continually fed auxiliary buffer or belt words back into the active state during absorption. By contrast, Waterfall [8] and fellow first-round candidate Blender [4] isolate their auxiliary word accumulators as strictly write-only pools throughout message absorption (mixed only during finalization in Waterfall, and output directly without finalization mixing after trailing checksum rounds in Blender).

Designed by Hattersley [8], Waterfall maintains a 93-word (2,976-bit) internal state partitioned into two subsystems:

1. **Three Coupled Feedback Shift Registers (The Trunk):** Three cyclic 32-bit word registers of mutually coprime lengths 16, 7, and 6 (29 words or 928 bits in total), whose feedback trinomials are pairwise coprime over $\mathbb{F}_2[x]$. At each compression step, all three streams absorb the current 32-bit input word (offset by fixed stream constants $0, 1, 2$), add a two-tap feedback from the register, and pass the result through a four-round AES-derived byte-cascade permutation.

2. **Two Write-Only Accumulator Pools:** Two 32-word (1,024-bit) accumulator arrays (64 words or 2,048 bits in total) that XOR-accumulate the outputs of the second and third trunk

Table 1: Comparison of collision attacks on the full Waterfall hash family (224/256/384/512 bits).

Attack	Variants	Message Length	Cost / Time
Fluhrer [7]	All 4	Unspecified	$\approx 2^{70}$ ops (theoretical)
Ours (9-block)	**All 4**	576 **B (9 blocks)**	3.8 **ms**
Ours (8-block)	**All 4**	512 **B (8 blocks)**	3.5 **ms**
Ours (7-block)	**All 4**	448 **B (7 blocks)**	52.1 **ms**

streams at the step index modulo 32. Both pools remain strictly write-only throughout message absorption and are mixed only during finalization.

Whereas low-weight state differences in PANAMA enabled a 2^{82} collision attack by Rijmen et al. [14] and subsequent practical collisions by Daemen and Van Assche [6], Waterfall incorporates three explicit structural defenses [8]:

- **Coupled Dual-Tap Recurrences:** Any non-zero input difference enters all three streams simultaneously and propagates along feedback recurrences of mutually coprime register lengths 16, 7, and 6 with pairwise coprime feedback polynomials. The designer argued that no low-weight message difference can simultaneously extinguish differences across all three streams.
- **Mandatory Block-Counter Injection Every 16 Words:** After absorbing the 16 message words of each 64-byte message block, Waterfall executes a 17-th update step whose input is fixed to the 32-bit block counter. Because an attacker cannot inject a message difference at every 17-th step and cannot freely choose the counter value, periodic differential trails are disrupted every 17 steps.
- **Dual 32-Word Accumulator Pools and 8-Round Finalization:** Even if a differential trail returns the 29 trunk words to zero difference, its intermediate active steps deposit non-zero differences into the two 32-word accumulator pools. Those deposits are mixed into the first stream across 8 rounds of finalization before the 224/256/384/512-bit digest is output.

1.1 Prior Cryptanalysis

Prior to our work, the only public cryptanalysis of Waterfall was a theoretical collision note by Fluhrer [7], posted during Round 1 of the SHA-3 competition shortly after the submission deadline. Fluhrer proposed a two-phase differential strategy that repeats a 34-step truncated register-collision pattern across two excursions separated by a multiple of 32 steps to cancel pool differences, estimating an overall cost of 2^{70} operations across all digest sizes without giving word-level differentials or colliding message pairs. Prompted by this note, Hattersley withdrew Waterfall on December 20, 2008 [9]. No practical collision generator, verified colliding message pair, or short (< 1 kB) single-excursion collision for Waterfall was previously known.

1.2 Our Contributions and Technical Overview

We present the first practical full-state collision attacks on the full Waterfall hash family across all four NIST digest sizes ($224, 256, 384$, and 512 bits). Table 1 summarizes our results alongside Fluhrer's theoretical bound [7]. Our closed-form collision framework simultaneously minimizes the steerable colliding message length and accelerates message solving by two orders of magnitude through three technical contributions:

1. **Maximal Affine Coset Linearization of the Byte-Cascade Permutation (Section 3.1):** Because each micro-round of Waterfall's four-round permutation applies the AES S-box to a single byte and rotates the word by one byte, choosing an input difference with a zero top byte can leave the fourth-round S-box inactive. We show across all $2^{32} - 1$ non-zero word differences that the maximum number of differential fixed points is $2^{13} = 8{,}192$, attained by 8 differences (2 per inactive micro-round), and that uniquely for the 2 differences whose fourth micro-round is inactive, the 2^{13} valid inputs form a single 13-dimensional $\mathbb{F}_2$-affine coset whose top byte is free and whose low three bytes span a 32-element subspace (Theorem 1). Restricting active transitions to this coset converts every nonlinear permutation condition into an exact 19-bit linear syndrome constraint and yields 13 neutral coset bits per single-stream step in the spirit of Biham and Chen [2].

2. **Minimal Steerable Single-Excursion Collisions via Active Counter Modification (Section 3.2):** Restricting all word differences to a single fixed-point difference makes both the three trunk feedback recurrences and the 32-word accumulator updates linear over $\mathbb{F}_2$. Each accumulator word then cancels within a *single* differential excursion whenever the second and third streams are active an even number of times per step index modulo 32. Whereas forbidding active streams at block-counter steps restricts 3-step look-back steerable single-excursion trails to at least 9 blocks (576 bytes), we allow a single active stream at counter steps and apply message modification [16] through the inverse permutation. Modifying the immediate predecessor step yields 8-block (512-byte) collisions, and pre-steering one register cycle earlier yields 7-block (448-byte) collisions, matching our SAT lower bound that 6 or fewer blocks are unsatisfiable within the 3-step look-back steerable trail class (Theorem 2).

3. 256**-Fold Accelerated Closed-Form Step Solving via Byte-Cascade Projection (Section 3.3):** We derive exact step-to-step recurrences for the inter-stream state differences and use four unconstrained predecessor steps before the initial difference injection to decouple the three-stream injection constraints in closed form. At subsequent constrained steps, we exploit the fact that the free top byte of the fixed-point coset enters only the fourth micro-round of the permutation. Projecting the recurrence onto the low three bytes reduces every 8,192-element coset search to 32 three-round evaluations paired with a constant-time lookup table recovery of the top byte (Theorem 3). This 256-fold reduction yields verified 512-byte full-state collisions in 3.5 ms across all four NIST digest sizes ($224, 256, 384$, and 512 bits).

1.3 Paper Organization

Section 2 establishes our algebraic notation and differential preliminaries, followed by the complete mathematical specification of Waterfall. Section 3 develops our unified full-state collision cryptanalysis, proving the maximal 2^{13}-coset structure (Theorem 1), the minimal single-excursion block lengths under counter steering (Theorem 2), and the 32-element byte-cascade coset reduction (Theorem 3). Section 4 analyzes the structural flaws in Waterfall's design rationale and proposes lightweight mitigations. Section 5 reports experimental collision benchmarks and verified colliding message fingerprints.

2 Preliminaries and Specification

2.1 Notation and Algebraic Preliminaries

Let $\mathbb{F}_2 = \{0, 1\}$ denote the finite field of two elements, where field addition coincides with bitwise exclusive-OR ($\oplus$) and multiplication coincides with bitwise logical AND ($\wedge$). For $n \geq 1$, $\mathbb{F}_2^n$ denotes the n-dimensional vector space of n-bit strings over $\mathbb{F}_2$. Following the official NIST SHA-3

Round 1 specification of Waterfall [8], all words are 32-bit unsigned integers in $\mathbb{F}_2^{32}$ stored in little-endian byte order. Each word $w \in \mathbb{F}_2^{32}$ decomposes into four bytes as $w = (b_0, b_1, b_2, b_3) \in (\mathbb{F}_2^8)^4$ corresponding to the integer value $w = \sum_{j=0}^{3} b_j 2^{8j}$, where $b_j(w) \in \mathbb{F}_2^8$ extracts byte j (bits $8j$ through $8j+7$), so $b_0(w)$ is the least significant byte and $b_3(w)$ is the most significant byte.

On single bytes $y \in \mathbb{F}_2^8$, we write $y \ll k$ and $y \gg k$ for 8-bit logical left and right shifts by k bit positions, and $\mathsf{ROTL}_k(y) = (y \ll k) \oplus (y \gg (8-k))$ for 8-bit circular left rotation. On 32-bit words $w = (b_0, b_1, b_2, b_3) \in (\mathbb{F}_2^8)^4$, we define the 8-bit right and left byte-rotation operators $\mathsf{ROT}_8, \mathsf{ROTL}_8 : (\mathbb{F}_2^8)^4 \to (\mathbb{F}_2^8)^4$ by

$$\mathsf{ROT}_8(b_0, b_1, b_2, b_3) = (b_1, b_2, b_3, b_0), \qquad \mathsf{ROTL}_8(b_0, b_1, b_2, b_3) = (b_3, b_0, b_1, b_2).$$

For a 16-word array $W = (W_0, \ldots, W_{15}) \in (\mathbb{F}_2^{32})^{16}$ and $m \in \{1, \ldots, 16\}$, $\mathrm{Trunc}_m(W) = (W_0, \ldots, W_{m-1}) \in (\mathbb{F}_2^{32})^m$ denotes truncation to the first m words.

For any pair (X, X') in an $\mathbb{F}_2$-vector space, we write $\Delta X = X \oplus X'$ for its XOR difference. Given a byte S-box $S : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ and input/output byte differences $\delta, \Delta \in \mathbb{F}_2^8$, the Difference Distribution Table (DDT) entry is defined by [3, 12]

$$\mathsf{DDT}[\delta][\Delta] := \big|\{u \in \mathbb{F}_2^8 : S(u) \oplus S(u \oplus \delta) = \Delta\}\big|.$$

For a 32-bit word permutation $F : \mathbb{F}_2^{32} \to \mathbb{F}_2^{32}$, the directional derivative at $Y \in \mathbb{F}_2^{32}$ along difference $\alpha \in \mathbb{F}_2^{32}$ is $\Phi_Y(\alpha) := F(Y) \oplus F(Y \oplus \alpha)$. The *differential fixed-point set* of F for a non-zero difference $\beta \in \mathbb{F}_2^{32} \setminus \{0\}$ is the set of inputs on which F acts as the identity map on difference β:

$$T_\beta := \{Y \in \mathbb{F}_2^{32} : F(Y) \oplus F(Y \oplus \beta) = \beta\}.$$

An $\mathbb{F}_2$-affine coset in $\mathbb{F}_2^{32}$ has the form $c \oplus V = \{c \oplus v : v \in V\}$ for a shift vector $c \in \mathbb{F}_2^{32}$ and an $\mathbb{F}_2$-linear subspace $V \subseteq \mathbb{F}_2^{32}$ of dimension $d = \dim_{\mathbb{F}_2}(V)$. We write $V = U \oplus W$ for the internal direct sum of subspaces $U, W \subseteq \mathbb{F}_2^{32}$ satisfying $U \cap W = \{0\}$. Given any surjective $\mathbb{F}_2$-linear quotient (syndrome) map $\sigma : \mathbb{F}_2^{32} \to \mathbb{F}_2^{32-d}$ with kernel $\ker(\sigma) = V$ and syndrome vector $\tau = \sigma(c) \in \mathbb{F}_2^{32-d}$, membership in the coset $c \oplus V$ is equivalent to the linear syndrome equation $\sigma(Y) = \tau$. Finally, for a binary sequence $(a_t)_{t \geq 0}$ with $a_t \in \mathbb{F}_2$, we represent its formal generating polynomial in $\mathbb{F}_2[x]$ as $A(x) = \sum_{t \geq 0} a_t x^t$, and write $\mathrm{lcm}(\cdot)$ for the least common multiple of polynomials in $\mathbb{F}_2[x]$ or of positive integers. Figure 1 illustrates Waterfall's single-step compression and 8-round finalization pipeline.

2.2 Internal State and Initialization

For a target digest bit length $h \in \{224, 256, 384, 512\}$, the internal state $\mathcal{S}$ of Waterfall consists of 93 words (2,976 bits):

$$\mathcal{S} = \Big(S_1[0 \ldots 15],\ S_2[0 \ldots 6],\ S_3[0 \ldots 5],\ \mathsf{Pool}_2[0 \ldots 31],\ \mathsf{Pool}_3[0 \ldots 31]\Big) \in (\mathbb{F}_2^{32})^{93}. \tag{1}$$

Here $S_1 \in (\mathbb{F}_2^{32})^{16}$, $S_2 \in (\mathbb{F}_2^{32})^7$, and $S_3 \in (\mathbb{F}_2^{32})^6$ denote the three cyclic shift registers of mutually coprime lengths $(K_1, K_2, K_3) = (16, 7, 6)$, whose feedback trinomials $1 + x + x^{K_i}$ are pairwise coprime over $\mathbb{F}_2[x]$. The arrays $\mathsf{Pool}_2, \mathsf{Pool}_3 \in (\mathbb{F}_2^{32})^{32}$ denote the two 32-word write-only accumulator pools. Hattersley [8] refers to them as Streams 1–3 and Pools 2–3, since Stream 1 has no associated pool. At initialization (Init), all 93 words are zeroed except for the initial word of S_2, which is set to the digest bit length:

$$S_2[0] \leftarrow h \in \{224, 256, 384, 512\}, \qquad \text{all other words} \leftarrow 0.$$

2.3 Four-Round Byte-Cascade Permutation Xbox

The sole nonlinear component of Waterfall is the 32-bit bijective word permutation $\mathsf{Xbox} : \mathbb{F}_2^{32} \to \mathbb{F}_2^{32}$, constructed from the FIPS-197 AES S-box $\mathsf{Sbox} : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ [10]. First, a 256-entry lookup table $\mathsf{XB}[0 \ldots 255] \subset \mathbb{F}_2^{32}$ is defined for each input byte $u \in \mathbb{F}_2^8$ by

$$\mathsf{XB}[u] = \big(y \oplus u,\ L_1(y),\ L_2(y),\ L_3(y)\big) \in (\mathbb{F}_2^8)^4, \qquad y = \mathsf{Sbox}(u), \tag{2}$$

where the three affine byte maps $L_1, L_2, L_3 : \mathbb{F}_2^8 \to \mathbb{F}_2^8$ are defined via their $\mathbb{F}_2$-linear parts $\tilde{L}_1, \tilde{L}_2, \tilde{L}_3$. These parts are 8-bit logical shifts $\ll, \gg$ on $\mathbb{F}_2^8$, combined with affine constants $(\texttt{0x3b}, \texttt{0x95}, \texttt{0x6a})$ as follows.

$$\begin{aligned}
L_1(y) &= \mathsf{ROTL}_7\big(y \oplus (y \ll 2) \oplus (y \gg 6) \oplus (y \ll 3)\big) \oplus \texttt{0x3b},\\
L_2(y) &= y \oplus (y \ll 5) \oplus (y \gg 3) \oplus (y \gg 7) \oplus \texttt{0x95},\\
L_3(y) &= y \oplus (y \ll 7) \oplus (y \gg 1) \oplus (y \gg 5) \oplus \texttt{0x6a}.
\end{aligned}$$

Given a 32-bit word $w = (b_0, b_1, b_2, b_3) \in (\mathbb{F}_2^8)^4$, the low byte b_0 occupies bits 0–7 and the top byte b_3 occupies bits 24–31. Then $\mathsf{Xbox}(w)$ applies four identical micro-rounds of table lookup and 8-bit right byte rotation $\mathsf{ROT}_8(b_0, b_1, b_2, b_3) = (b_1, b_2, b_3, b_0)$.

$$\text{For } r = 1, 2, 3, 4: \qquad w \leftarrow \mathsf{ROT}_8\big(w \oplus \mathsf{XB}[b_0(w)]\big). \tag{3}$$

The low byte of $w \oplus \mathsf{XB}[b_0]$ is $b_0 \oplus (\mathsf{Sbox}(b_0) \oplus b_0) = \mathsf{Sbox}(b_0)$, which ROT_8 immediately shifts into the most significant byte (bits 24–31). Each micro-round is therefore directly invertible via 8-bit left byte rotation $\mathsf{ROTL}_8(b_0, b_1, b_2, b_3) = (b_3, b_0, b_1, b_2)$ and the inverse AES S-box Sbox^{-1}:

$$\mathsf{Xbox}^{-1}(w): \quad \text{For } r = 1, 2, 3, 4: \qquad w \leftarrow \mathsf{ROTL}_8(w), \quad w \leftarrow w \oplus \mathsf{XB}\big[\mathsf{Sbox}^{-1}(b_0(w))\big].$$

2.4 Compression Step and Block Schedule

Let $t \geq 0$ index the sequential calls to $\mathsf{UpdateStep}(\mathcal{S}, x_t)$ starting from $t = 0$. Write the cyclic register lengths as $(K_1, K_2, K_3) = (16, 7, 6)$ and the stream constants as $(c_1, c_2, c_3) = (0, 1, 2)$. At step t, define the two-tap feedback sums from the three trunk registers by

$$\begin{aligned}
P_a(t) &= S_1[(t-1) \bmod 16] \oplus S_1[t \bmod 16], \qquad P_b(t) = S_2[(t-1) \bmod 7] \oplus S_2[t \bmod 7],\\
P_c(t) &= S_3[(t-1) \bmod 6] \oplus S_3[t \bmod 6],
\end{aligned} \tag{4}$$

and write $(P_1(t), P_2(t), P_3(t)) := (P_a(t), P_b(t), P_c(t))$. Step t combines these feedback sums with the input word x_t, updates the three trunk registers at cyclic index t mod K_i, and accumulates the new outputs of S_2 and S_3 into Pool_2 and Pool_3 at index t mod 32:

$$\begin{aligned}
Y_a(t) &= P_a(t) \oplus x_t, \qquad Y_b(t) = P_b(t) \oplus x_t \oplus 1, \qquad Y_c(t) = P_c(t) \oplus x_t \oplus 2,\\
S_1[t \bmod 16] &\leftarrow \mathsf{Xbox}\big(Y_a(t)\big), \qquad S_2[t \bmod 7] \leftarrow \mathsf{Xbox}\big(Y_b(t)\big),\\
S_3[t \bmod 6] &\leftarrow \mathsf{Xbox}\big(Y_c(t)\big), \qquad \mathsf{Pool}_2[t \bmod 32] \leftarrow \mathsf{Pool}_2[t \bmod 32] \oplus S_2[t \bmod 7],\\
\mathsf{Pool}_3[t \bmod 32] &\leftarrow \mathsf{Pool}_3[t \bmod 32] \oplus S_3[t \bmod 6].
\end{aligned} \tag{5}$$

For uniform indexing across $i \in \{1, 2, 3\}$ in (5), we write $(Y_1(t), Y_2(t), Y_3(t)) := (Y_a(t), Y_b(t), Y_c(t))$.

The 17-Step Block Schedule. Each 64-byte (512-bit) message block B_b ($b = 1, 2, \ldots, N_{\text{blk}}$) is parsed as 16 little-endian 32-bit words $m_{b,0}, \ldots, m_{b,15}$. The block is then compressed in 17 update steps:

$$x_{17(b-1)+q} = \begin{cases} m_{b,q} & \text{for } q \in \{0, 1, \ldots, 15\} \quad \text{(message-word step)},\\ b \bmod 2^{32} & \text{for } q = 16 \quad \text{(mandatory block-counter step } t \equiv 16 \pmod{17}). \end{cases} \tag{6}$$

Thus, B full 64-byte message blocks correspond to $T = 17B$ calls to $\mathsf{UpdateStep}$. Every step $t \equiv 16 \pmod{17}$ has its input fixed to $\mathrm{ctr}_t = (\lfloor t/17 \rfloor + 1) \bmod 2^{32}$.

2.5 Padding and Finalization

Let $\ell_{\text{rem}} \in \{0, \ldots, 511\}$ denote the residual bit count of any partial trailing block, with $\ell_{\text{rem}} = 0$ for block-aligned messages. If $\ell_{\text{rem}} > 0$, the partial block is zero-padded to 64 bytes and compressed via (6). Let $N_{\text{slab}} = \lfloor N_{\text{blk}}/2^{32} \rfloor$ denote the upper 32-bit block-counter carry word across all N_{blk} compressed blocks. In Final, Waterfall executes 18 post-absorption update steps: $\mathsf{UpdateStep}(\mathcal{S}, \ell_{\text{rem}})$, $\mathsf{UpdateStep}(\mathcal{S}, N_{\text{slab}})$, and 16 constant steps $\mathsf{UpdateStep}(\mathcal{S}, i)$ for $i = 0, 1, \ldots, 15$.

Let $A = (S_1[0], \ldots, S_1[15]) \in (\mathbb{F}_2^{32})^{16}$ denote the resulting 16-word state of S_1. Partition Pool_2 and Pool_3 into four 16-word half-pools:

$$H_1 = \mathsf{Pool}_2[0 \ldots 15], \qquad H_2 = \mathsf{Pool}_2[16 \ldots 31] \in (\mathbb{F}_2^{32})^{16},$$
$$H_3 = \mathsf{Pool}_3[0 \ldots 15], \qquad H_4 = \mathsf{Pool}_3[16 \ldots 31] \in (\mathbb{F}_2^{32})^{16}.$$

Define the 16-word bit-transposition bijection $L_{\text{fin}} : (\mathbb{F}_2^{32})^{16} \to (\mathbb{F}_2^{32})^{16}$ and finalization round permutation $\mathsf{FinalStep} = \Phi : (\mathbb{F}_2^{32})^{16} \to (\mathbb{F}_2^{32})^{16}$ for $i \in \{0, \ldots, 15\}$ (with indices reduced modulo 16) by

$$L_{\text{fin}}(W)_i = (W_i \wedge M_0) \oplus (W_{i+3} \wedge M_1) \oplus (W_{i+7} \wedge M_2) \oplus (W_{i+12} \wedge M_3),$$
$$\Phi(W)_i = \mathsf{Xbox}\big(L_{\text{fin}}(W)_i\big),$$

using the pairwise disjoint bitmasks $M_0 = \texttt{0x14141414}$, $M_1 = \texttt{0x28282828}$, $M_2 = \texttt{0x42424242}$, and $M_3 = \texttt{0x81818181}$, and set $\Psi := \Phi^4$. Starting from $W^{(0)} = A$, the 8-round finalization (Figure 1) computes

$$W^{(k)} = \Phi\big(W^{(k-1)}\big) \oplus H_k \quad (k = 1, 2, 3, 4), \qquad W^{(8)} = \Phi^4\big(W^{(4)}\big) = \Psi\big(W^{(4)}\big),$$
$$\mathsf{Digest}_h = \text{Trunc}_{h/32}\left(W^{(8)} \oplus A \oplus \bigoplus_{k=1}^{4} H_k \right).$$

The map Final is a deterministic function of the 93-word internal state (1) and the message bit length. Any two distinct B-block messages $M_1 \neq M_2$ whose 93 internal state words collide at step $T = 17B$ therefore produce identical digests. The collision holds across all subsequent message extensions and all output sizes $h \in \{224, 256, 384, 512\}$.

3 Full-State Collision Cryptanalysis

We now present our unified algebraic differential collision framework, combining single-excursion pool cancellation, active counter-step message modification, and byte-cascade coset projection into a closed-form solver that reaches the minimal steerable single-excursion block lengths $B = 7$ (448 bytes) and $B = 8$ (512 bytes).

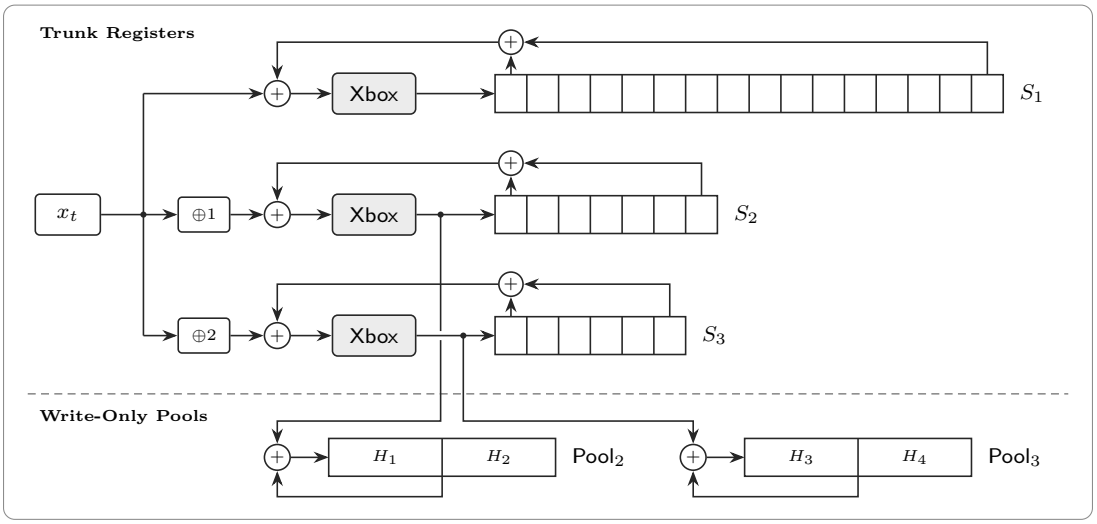
3.1 Differential Fixed-Point Cosets of Xbox

Our differential analysis follows the standard framework of Biham and Shamir [3], specialized to the inversion-based AES S-box [12]. For any word $Y \in \mathbb{F}_2^{32}$ and input difference $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3) \in (\mathbb{F}_2^8)^4$, denote the Xbox output difference by

$$\Phi_Y(\alpha) := \mathsf{Xbox}(Y) \oplus \mathsf{Xbox}(Y \oplus \alpha).$$

Figure 2 depicts the 4-round byte-cascade structure (2)–(3). Write $Y = (b_0, b_1, b_2, b_3) \in (\mathbb{F}_2^8)^4$ and let $i_r \in \mathbb{F}_2^8$ denote the S-box input byte in micro-round $r \in \{1, 2, 3, 4\}$. Each micro-round then maps $(u_0, u_1, u_2, u_3) \mapsto (u_1 \oplus L_1(\mathsf{Sbox}(i_r)),\ u_2 \oplus L_2(\mathsf{Sbox}(i_r)),\ u_3 \oplus L_3(\mathsf{Sbox}(i_r)),\ \mathsf{Sbox}(i_r))$ with $i_r = u_0$. Now take a pair $(Y, Y \oplus \alpha)$ with S-box input differences $(\delta_1, \delta_2, \delta_3, \delta_4) \in (\mathbb{F}_2^8)^4$ and S-box output differences $(\Delta_1, \Delta_2, \Delta_3, \Delta_4) \in (\mathbb{F}_2^8)^4$ across the four micro-rounds. Any target

(a) Single-Step Compression



(b) 8-Round Finalization

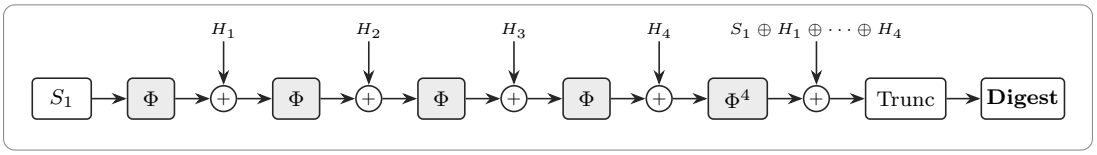


Figure 1: Overall trunk-plus-accumulator architecture of Waterfall [8]. **(a)** Each compression step $\mathsf{UpdateStep}(\mathcal{S}, x_t)$ absorbs the 32-bit input word x_t (16 message words followed by block counter $\mathrm{ctr}_b = b \bmod 2^{32}$ every 17-th step), updates the three cyclic trunk registers S_1, S_2, S_3 of coprime lengths $(16, 7, 6)$ via two-tap feedback and Xbox, and XOR-accumulates the new outputs of S_2 and S_3 across a one-way write-only boundary into the 32-word pools $\mathsf{Pool}_2 = (H_1, H_2)$ and $\mathsf{Pool}_3 = (H_3, H_4)$ at index $t \bmod 32$. **(b)** The 8-round finalization Final absorbs the four 16-word half-pools $H_1, \ldots, H_4$ into S_1 through $\Phi(W) = \mathsf{Xbox}^{\otimes 16}(L_{\mathrm{fin}}(W))$, applies Φ^4, XORs the feed-forward sum $S_1 \oplus \bigoplus_{k=1}^{4} H_k$, and truncates to $h/32$ words.

output difference $\Delta w = \Phi_Y(\alpha) = (t_0, t_1, t_2, t_3) \in (\mathbb{F}_2^8)^4$ uniquely determines $(\Delta_1, \Delta_2, \Delta_3, \Delta_4)$ backward and $(\delta_1, \delta_2, \delta_3, \delta_4)$ forward via the linear triangular system:

$$\begin{aligned}
&\Delta_4 = t_3, \qquad \Delta_3 = t_2 \oplus \tilde{L}_3(\Delta_4), \qquad \Delta_2 = t_1 \oplus \tilde{L}_2(\Delta_4) \oplus \tilde{L}_3(\Delta_3),\\
&\Delta_1 = t_0 \oplus \tilde{L}_1(\Delta_4) \oplus \tilde{L}_2(\Delta_3) \oplus \tilde{L}_3(\Delta_2), \qquad (7)\\
&\delta_1 = \alpha_0, \qquad \delta_2 = \alpha_1 \oplus \tilde{L}_1(\Delta_1), \qquad \delta_3 = \alpha_2 \oplus \tilde{L}_2(\Delta_1) \oplus \tilde{L}_1(\Delta_2),\\
&\delta_4 = \alpha_3 \oplus \tilde{L}_3(\Delta_1) \oplus \tilde{L}_2(\Delta_2) \oplus \tilde{L}_1(\Delta_3). \qquad (8)
\end{aligned}$$

In particular, when $\alpha = \beta = (\beta_0, \beta_1, \beta_2, \mathtt{0x00})$ has top byte $\beta_3 = \mathtt{0x00}$, the top byte $b_3 \in \mathbb{F}_2^8$ of $Y = (b_0, b_1, b_2, b_3)$ does not enter the S-box during micro-rounds $r = 1, 2, 3$; instead, three applications of ROT_8 shift b_3 into the fourth-round S-box input byte $i_4 = c_{3,0} \oplus b_3$, where $c_{3,0} \in \mathbb{F}_2^8$ denotes the low byte after three micro-rounds evaluated at $b_3 = 0$.

Theorem 1 (Maximal $\mathbb{F}_2$-Affine Differential Fixed-Point Cosets of Xbox)**.** *For any non-zero difference $\beta = (\beta_0, \beta_1, \beta_2, \beta_3) \in (\mathbb{F}_2^8)^4 \setminus \{0\}$, define the differential fixed-point set*

$$T_\beta := \big\{ Y \in \mathbb{F}_2^{32} : \Phi_Y(\beta) = \beta \big\}.$$

1. ***Closed-Form Size and Global Optimality:*** *For every non-zero β, the cardinality is $|T_\beta| = \prod_{r=1}^{4} \mathsf{DDT}[\delta_r(\beta)][\Delta_r(\beta)]$. Since the AES S-box satisfies $\mathsf{DDT}[\delta][\Delta] \in \{0, 2, 4\}$ for*

$\delta \neq 0$ *and* $\mathsf{DDT}[0][0] = 256$*,* $|T_\beta|$ *is maximized when one micro-round* $r_0 \in \{1,2,3,4\}$ *is inactive (*$\delta_{r_0}(\beta) = \Delta_{r_0}(\beta) =$ `0x00`*) and the remaining three active micro-rounds realize* $\{\mathsf{DDT}[\delta_r][\Delta_r]\}_{r \neq r_0} = \{2,4,4\}$*. An exhaustive check over all* $2^{32}-1$ *non-zero differences confirms that the global maximum is*

$$\max_{\beta \neq 0} |T_\beta| = 2 \times 4 \times 4 \times 256 = 8{,}192 = 2^{13},$$

attained by 8 *differences in* $\mathbb{F}_2^{32}$ *(*2 *for each* $r_0 \in \{1,2,3,4\}$*), of which uniquely* $\beta =$ `0x00f9e70e` *and* $\beta' =$ `0x003f7638` *satisfy* $r_0 = 4$ *(*$\beta_3 =$ `0x00`*).*

2. ***Direct-Sum*** $\mathbb{F}_2$***-Affine Coset Structure:*** *Whereas the six maximizers with* $r_0 \in \{1,2,3\}$ *pass the unconstrained byte* i_{r_0} *through* Sbox *before later rounds and form nonlinear* 2^{13}*-element manifolds, for each of the two* $r_0 = 4$ *maximizers* $\beta \in \{$`0x00f9e70e`, `0x003f7638`$\}$ *the set* T_β *is a* single $\mathbb{F}_2$*-affine coset*

$$T_\beta = c \oplus V, \qquad V = H_{24} \oplus \mathcal{B}_3, \qquad \dim_{\mathbb{F}_2}(V) = 13, \qquad \beta \in H_{24},$$

where $H_{24} \subset \mathbb{F}_2^{24} \times \{0\}$ *is a* 5*-dimensional low-*24*-bit subspace (*$|H_{24}| = 32$*) and* $\mathcal{B}_3 = \{(0,0,0,b_3) : b_3 \in \mathbb{F}_2^8\} \subset (\mathbb{F}_2^8)^4$ *is the unconstrained* 8*-dimensional top-byte subspace.*

3. 19***-Bit Linear Syndrome Characterization:*** *For* $\beta \in \{$`0x00f9e70e`, `0x003f7638`$\}$*, let* $\sigma : \mathbb{F}_2^{32} \to \mathbb{F}_2^{19}$ *denote the* $\mathbb{F}_2$*-linear quotient map modulo* V *(supported entirely on the low* 24 *bits of its input) and let* $\tau = \sigma(c) \in \mathbb{F}_2^{19}$*. Then for any* $Y \in \mathbb{F}_2^{32}$*,*

$$\Phi_Y(\beta) = \beta \quad \Longleftrightarrow \quad Y \in T_\beta \quad \Longleftrightarrow \quad \sigma(Y) = \tau. \tag{9}$$

Proof. Setting $(t_0,t_1,t_2,t_3) = (\beta_0,\beta_1,\beta_2,\beta_3)$ in (7)–(8) uniquely fixes $(\delta_r(\beta),\Delta_r(\beta))_{r=1}^4$. Because the map $Y \mapsto (i_1,i_2,i_3,i_4)$ is triangular and bijective over $(\mathbb{F}_2^8)^4$, $|T_\beta| = \prod_{r=1}^4 \mathsf{DDT}[\delta_r(\beta)][\Delta_r(\beta)]$. Since $\beta \neq 0$, not all four rounds can be inactive; if two or more rounds had $\Delta_r = 0$, (7)–(8) can be checked directly across the $\binom{4}{2} \cdot (255)^2$ candidates to fall short of 2^{13}, and an exhaustive 2^{32}-word sweep confirms $\max_{\beta \neq 0} |T_\beta| = 2^{13}$ at 8 differences (2 per r_0). When $r_0 = 4$, $\Delta_4 = \beta_3 =$ `0x00` leaves i_4 unconstrained ($\mathsf{DDT}[0][0] = 256$). For the inversion-based AES S-box [12], each active solution set $A_r = \{u \in \mathbb{F}_2^8 : \mathsf{Sbox}(u) \oplus \mathsf{Sbox}(u \oplus \delta_r) = \Delta_r\}$ $(r = 1,2,3)$ is an $\mathbb{F}_2$-affine 1- or 2-flat, and on $A_1 \times A_2 \times A_3$ the S-box outputs $\mathsf{Sbox}(i_1), \mathsf{Sbox}(i_2)$ are $\mathbb{F}_2$-affine functions of (i_1,i_2) (having dimension ≤ 2). Pulling (A_1,A_2,A_3) back to (b_0,b_1,b_2) therefore yields a 5-dimensional $\mathbb{F}_2$-affine flat $c_0 \oplus H_{24}$ in $\mathbb{F}_2^{24} \times \{0\}$ with $\beta \in H_{24}$, while $b_3 \in \mathbb{F}_2^8$ never enters i_1,i_2,i_3. Part (3) follows immediately from $\ker(\sigma) = V$. □

3.2 Pool Parity Cancellation and Counter Steering

Fix $\beta =$ `0x00f9e70e` from Theorem 1 and restrict the difference of every message word and every trunk register word across a T-step prefix ($T = 17B$) to the two-element alphabet $\{0,\beta\}$. Let $d_t \in \mathbb{F}_2$ indicate whether $\Delta x_t = \beta$, and let $(e_1[t],e_2[t],e_3[t]) := (u_t,v_t,w_t) \in \mathbb{F}_2^3$ indicate whether the words written into $S_1[t \bmod 16]$, $S_2[t \bmod 7]$, and $S_3[t \bmod 6]$ at step t carry difference β.

Whenever every active stream at step t satisfies the linear syndrome condition (9), Xbox acts as the identity on the difference β ($\Phi_Y(\beta) = \beta$). As a result, the differential trajectory $(u_t,v_t,w_t)_{t=0}^{T-1}$ is governed by the exact $\mathbb{F}_2$-linear shift-register recurrences:

$$u_t = u_{t-1} \oplus u_{t-16} \oplus d_t, \qquad v_t = v_{t-1} \oplus v_{t-7} \oplus d_t, \qquad w_t = w_{t-1} \oplus w_{t-6} \oplus d_t, \tag{10}$$

with $u_s = v_s = w_s = 0$ for $s < 0$. Using standard polynomial delay notation over $\mathbb{F}_2[x]$ for shift-register trails [14, 6], (10) reads $(1+x+x^{16})U(x) = (1+x+x^7)V(x) = (1+x+x^6)W(x) = D(x)$. Because the three feedback trinomials are pairwise coprime over $\mathbb{F}_2[x]$, finite-support trunk

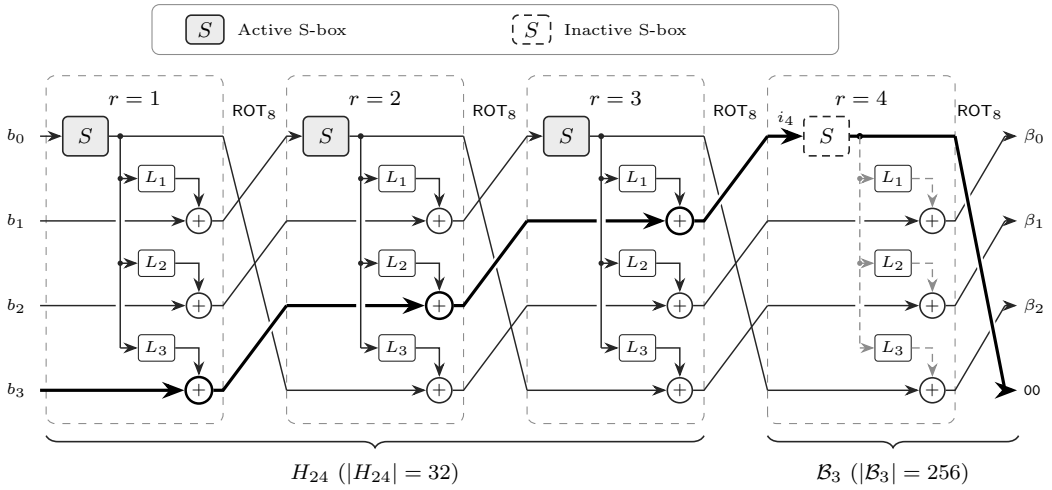


Figure 2: Byte-lane Generalized-Feistel structure of Xbox across its four micro-rounds ($r = 1, 2, 3, 4$) for a maximal fixed-point difference $\beta = (\beta_0, \beta_1, \beta_2, \texttt{0x00})$ (Theorems 1 and 3). Setting $\beta_3 = \delta_4 = \Delta_4 = \texttt{0x00}$ leaves the fourth-round AES S-box inactive ($\mathsf{DDT}[0][0] = 256$), while the unconstrained top byte $b_3 \in \mathbb{F}_2^8$ (heavy bold staircase) shifts undisturbed along lanes $3 \to 2 \to 1 \to 0$ into $i_4 = c_{3,0} \oplus b_3$, yielding $|T_\beta| = 32 \times 256 = 2^{13}$ and enabling the 256× coset-scan reduction.

excursions correspond precisely to multiples $D(x) = L(x)Q(x)$ of the degree-29 generator polynomial

$$L(x) = \operatorname{lcm}\left(1{+}x{+}x^{16},\, 1{+}x{+}x^{7},\, 1{+}x{+}x^{6}\right) = (1{+}x{+}x^{16})(1{+}x{+}x^{7})(1{+}x{+}x^{6}), \qquad \deg L = 29. \tag{11}$$

A complete 93-word internal state collision occurs at step T if and only if both the 29 trunk registers and the 64 accumulator pool words return to zero difference:

1. **Trunk Register Extinction:** The final K_i steps of each stream write zero differences, overwriting every register slot:

$$u_t = 0 \;\; \forall t \in [T-16, T), \qquad v_t = 0 \;\; \forall t \in [T-7, T), \qquad w_t = 0 \;\; \forall t \in [T-6, T). \tag{12}$$

2. **Single-Excursion In-Window Pool Parity Cancellation:** Every step t with $v_t = 1$ (resp. $w_t = 1$) XORs the *exact same difference* β into $\mathsf{Pool}_2[t \bmod 32]$ (resp. $\mathsf{Pool}_3[t \bmod 32]$). Repeating a register excursion twice at a step offset $\Delta s \equiv 0 \pmod{\operatorname{lcm}(17, 32) = 544}$ (requiring 44 blocks = 2,816 bytes) is therefore *unnecessary*. Instead, all 64 pool words cancel to zero difference within a *single excursion* if and only if each residue class $r \in \{0, \ldots, 31\}$ is visited an even number of times:

$$\bigoplus_{\substack{0 \le t < T \\ t \equiv r \pmod{32}}} v_t = 0, \qquad \bigoplus_{\substack{0 \le t < T \\ t \equiv r \pmod{32}}} w_t = 0 \qquad \forall r \in \{0, 1, \ldots, 31\}. \tag{13}$$

Condition Accounting and Counter Steering. Recall the pre-input trunk tap sums $P_a(t), P_b(t), P_c(t)$ defined in (4), where $Y_a(t) = P_a(t) \oplus x_t$, $Y_b(t) = P_b(t) \oplus x_t \oplus 1$, and $Y_c(t) = P_c(t) \oplus x_t \oplus 2$. The pairwise inter-stream input differences

$$A_t := Y_a(t) \oplus Y_b(t) = P_a(t) \oplus P_b(t) \oplus 1, \quad C_t := Y_a(t) \oplus Y_c(t) = P_a(t) \oplus P_c(t) \oplus 2, \quad G_t := A_t \oplus C_t \tag{14}$$

are *completely independent of the current message word* x_t. Let $\mathrm{act}_t = u_t + v_t + w_t \in \{0, 1, 2, 3\}$ denote the number of active streams at step t, and let T_0 denote the unique initial *injection*

step where the first non-zero message difference $d_{T_0} = 1$ activates all three streams ($u_{T_0} = v_{T_0} = w_{T_0} = 1$):

- **Message-Word Steps ($t \not\equiv 16 \pmod{17}$):**
  - If $\mathrm{act}_t = 0$: $x_t \in \mathbb{F}_2^{32}$ is completely free (32 bits of freedom, 0 conditions).
  - If $\mathrm{act}_t = 1$: choosing x_t in the 2^{13}-coset $P_i(t) \oplus c_i \oplus c \oplus V$ satisfies $\sigma(Y_i(t)) = \tau$ automatically (13 neutral bits of freedom [2] in V, 0 x_t-independent conditions).
  - If $\mathrm{act}_t = 2$: x_t absorbs one active stream, leaving one x_t-independent 19-bit linear condition on the state entering step t: $A_t \in V$ (if $u_t = v_t = 1$), $C_t \in V$ (if $u_t = w_t = 1$), or $G_t \in V$ (if $v_t = w_t = 1$).
  - If $\mathrm{act}_t = 3$ (at $t = T_0$): x_{T_0} absorbs one stream, leaving two x_{T_0}-independent 19-bit conditions ($A_{T_0} \in V$ and $C_{T_0} \in V$, 38 bits in total), which we decouple across the four predecessor steps $T_0 - 4, \ldots, T_0 - 1$ by pre-aligning $G_{T_0-2} = 0$ (Section 3.3).
- **Mandatory Block-Counter Steps ($t_{\mathrm{ctr}} \equiv 16 \pmod{17}$):** At any counter step $t = t_{\mathrm{ctr}}$, $d_t = 0$ and the input word is fixed to $\mathrm{ctr}_t = (\lfloor t/17 \rfloor + 1) \bmod 2^{32}$. Forcing $u_t = v_t = w_t = 0$ at every counter step restricts 3-step look-back steerable single-excursion patterns to $B \geq 9$ blocks (576 bytes). We break this 9-block barrier by combining single-excursion pool parity cancellation (13) with active block-counter message modification [16] ($\mathrm{act}_t \leq 1$ at $t \equiv 16 \pmod{17}$, illustrated in Figure 3). When a single stream $i \in \{1, 2, 3\}$ is active at counter step $t = t_{\mathrm{ctr}}$, it imposes a single 19-bit condition $\sigma(P_i(t) \oplus \mathrm{ctr}_t \oplus c_i) = \tau$. Here $P_i(t) = \mathsf{Xbox}(Y_i(t-1)) \oplus S_i[t \bmod K_i]$, where $S_i[(t-1) \bmod K_i] = \mathsf{Xbox}(Y_i(t-1))$ is written at step $t-1$ and $S_i[t \bmod K_i] = \mathsf{Xbox}(Y_i(t_{\mathrm{pre}}))$ was written K_i steps earlier at $t_{\mathrm{pre}} = t - K_i$. The counter condition is satisfied in closed form via two mechanisms:
  1. **Immediate Predecessor Modification ($e_i[t-1] = 0$):** When step $t-1$ has $\mathrm{act}_{t-1} = 0$, $Y_i(t-1)$ is unconstrained, so setting
  $$Y_i(t-1) \leftarrow \mathsf{Xbox}^{-1}\big(S_i[t \bmod K_i] \oplus \mathrm{ctr}_t \oplus c_i \oplus c\big) \tag{15}$$
  satisfies $Y_i(t) = c \in T_\beta$ deterministically in $O(1)$ using a single Xbox^{-1} evaluation. When $\mathrm{act}_{t-1} = 1$ on a distinct stream $p \neq i$, $Y_i(t-1)$ is steered across the 32-element low-24-bit coset of step $t-1$.
  2. **Delayed Ring-Buffer Pre-Steering ($e_i[t-1] = 1$):** Suppose the same stream i is active at both step $t-1$ and counter step t. The recurrence $e_i[t] = e_i[t-1] \oplus e_i[t-K_i] \oplus d_t$ with $d_t = 0$ guarantees that $e_i[t_{\mathrm{pre}}] = 0$, so stream i is inactive at step $t_{\mathrm{pre}} = t - K_i$. Fixing $Y_i(t-1) = c \in T_\beta$ at step $t-1$ fixes $S_i[(t-1) \bmod K_i] = \mathsf{Xbox}(c)$, transferring the counter condition backward to step $t_{\mathrm{pre}} = t - K_i$:
  $$\mathsf{Xbox}\big(Y_i(t_{\mathrm{pre}})\big) \in \big(\mathsf{Xbox}(c) \oplus \mathrm{ctr}_t \oplus c_i \oplus c\big) \oplus V, \tag{16}$$
  which is discharged at step $t_{\mathrm{pre}} = t - K_i$ via a precomputed counter-syndrome lookup table $\mathcal{T}_{\mathrm{ctr}}$ (Theorem 3) before the walk reaches step t.

To ensure that every step condition can be discharged sequentially without global multi-block backtracking, we call a trail $(d_t, u_t, v_t, w_t)_{t=0}^{T-1}$ 3-*step look-back steerable* if: (i) it has a unique 3-stream injection step $T_0 \geq 4$ preceded by four unconstrained ($\mathrm{act}_r = 0$) non-counter steps $T_0 - 4, \ldots, T_0 - 1$, with $d_t = 0$ and $\mathrm{act}_t \leq 1$ at every counter step $t \equiv 16 \pmod{17}$, $\mathrm{act}_{T_0+1} \leq 1$, and $\mathrm{act}_t \leq 2$ for all $t > T_0$; and (ii) every subsequent step $t > T_0$ bearing an x_t-independent condition ($\mathrm{act}_t = 2$ at a non-counter step, or $\mathrm{act}_t = 1$ at a counter step $t \equiv 16 \pmod{17}$) has a non-counter predecessor $t - 1 \not\equiv 16 \pmod{17}$ and admits a non-counter step $s \in \{t-1, t-2, t-3\} \cap (T_0, t)$ such that no step $r \in [s, t)$ bears an x_r-independent condition

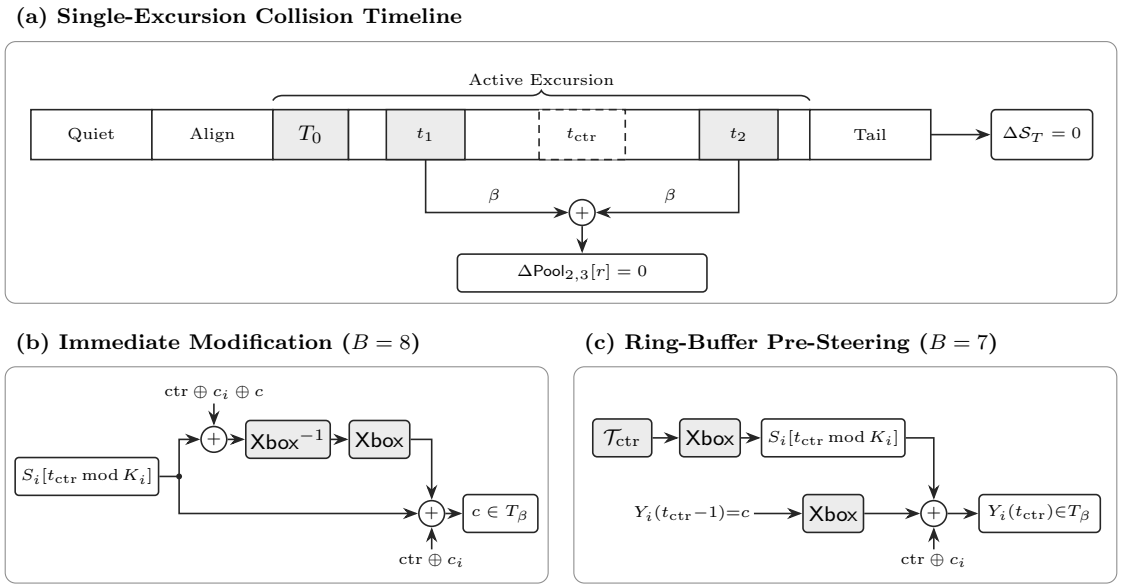


Figure 3: Single-excursion collision structure and active block-counter modification (Theorem 2). **(a)** Across $T = 17B$ steps, the quiet prefix $[0, T_0-4)$ and 4-step alignment $[T_0-4, T_0)$ prepare the 3-stream injection at T_0 ($\Delta x_{T_0} = \beta$); inside the active excursion ($\Delta \in \{0, \beta\}$, mandatory counter steps $t_{\mathrm{ctr}} \equiv 16 \pmod{17}$ with $\Delta x_{t_{\mathrm{ctr}}} = 0$), residue-paired active steps $t_1 \equiv t_2 \equiv r \pmod{32}$ cancel every pool slot ($\beta \oplus \beta = 0$), and $[T-16, T)$ extinguishes all 29 trunk words to yield $\Delta\mathcal{S}_T = 0$. **(b)** When $e_i[t_{\mathrm{ctr}} - 1] = 0$, step $t_{\mathrm{ctr}} - 1$ applies Xbox^{-1} to cancel $S_i[t_{\mathrm{ctr}} \bmod K_i] \oplus \mathrm{ctr} \oplus c_i$ at step t_{ctr}, placing $Y_i(t_{\mathrm{ctr}}) = c \in T_\beta$ ($B = 8$). **(c)** When $e_i[t_{\mathrm{ctr}} - 1] = 1$, $d_{t_{\mathrm{ctr}}} = 0$ forces $e_i[t_{\mathrm{pre}}] = 0$ at $t_{\mathrm{pre}} = t_{\mathrm{ctr}} - K_i$, pre-steering $S_i[t_{\mathrm{ctr}} \bmod K_i]$ via $\mathcal{T}_{\mathrm{ctr}}$ while pinning $Y_i(t_{\mathrm{ctr}} - 1) = c$ ($B = 7$).

(i.e., $\mathrm{act}_r \le 1$ if $r \not\equiv 16 \pmod{17}$ and $\mathrm{act}_r = 0$ if $r \equiv 16 \pmod{17}$). (Without condition (ii), the unconstrained linear system (10)–(13) already admits algebraic solutions at $B = 5$ blocks, but consecutive constrained steps create coupled multi-step nonlinear dependencies.)

Theorem 2 (Minimal Steerable Single-Excursion Block Length for Waterfall)**.** *Among all* 3*-step look-back steerable single-excursion trails satisfying* (10)*, tail extinction* (12)*, and modulo-*32 *pool parity cancellation* (13)*:*

1. *If active streams are forbidden on block-counter steps (*$\mathrm{act}_t = 0$ *for all* $t \equiv 16 \pmod{17}$*), all block lengths* $B \le 8$ *(*$T \le 136$*) are unsatisfiable, and the shortest steerable trail has* $B = 9$ *blocks (*576 *bytes,* $T = 153$ *steps).*

2. *With active block-counter modification (*$\mathrm{act}_t \le 1$ *for* $t \equiv 16 \pmod{17}$*), block lengths* $B \le 6$ *(*$T \le 102$*,* ≤ 384 *bytes) are unsatisfiable, while* $B = 8$ *blocks (*512 *bytes,* $T = 136$ *steps) is satisfiable using immediate* $O(1)$ *counter inversion* (15)*, and* $B = 7$ *blocks (*448 *bytes,* $T = 119$ *steps) is satisfiable and minimal among* 3*-step look-back steerable trails using ring-buffer pre-steering* (16)*.*

3.3 Closed-Form Step Recurrences and Coset Reduction

We now show how the x_t-independent conditions $A_t \in V$, $C_t \in V$, and $G_t \in V$ of (14) are solved in closed form without heuristic search. We also show how the direct-sum structure $V = H_{24} \oplus \mathcal{B}_3$ accelerates every constrained coset step by 256×.

Step-to-Step Difference Recurrences. At step t, define $z_t := P_a(t) \oplus x_t = Y_a(t) \in \mathbb{F}_2^{32}$, and define the three state-known constants entering step $t+1$:

$$K_A(t) := S_1[(t+1) \bmod 16] \oplus S_2[(t+1) \bmod 7] \oplus 1,$$
$$K_C(t) := S_1[(t+1) \bmod 16] \oplus S_3[(t+1) \bmod 6] \oplus 2, \qquad K_{AP}(t) := K_A(t) \oplus K_C(t).$$

Because $Y_b(t) = z_t \oplus A_t$ and $Y_c(t) = z_t \oplus C_t$, the new words written at step t satisfy $S_1[t \bmod 16] \oplus S_2[t \bmod 7] = \mathsf{Xbox}(z_t) \oplus \mathsf{Xbox}(z_t \oplus A_t) = \Phi_{z_t}(A_t)$, yielding the exact one-step recurrences:

$$\begin{aligned} A_{t+1} &= \Phi_{z_t}(A_t) \oplus K_A(t), \\ C_{t+1} &= \Phi_{z_t}(C_t) \oplus K_C(t), \\ G_{t+1} &= \Phi_{z_t \oplus A_t}(G_t) \oplus K_{AP}(t). \end{aligned} \tag{17}$$

Decoupling the Three-Stream Injection Step. At the single 3-stream injection step T_0, both $A_{T_0} \in V$ and $C_{T_0} \in V$ (equivalently, $G_{T_0} \in V$ and $C_{T_0} \in V$) must hold simultaneously (38 bits). Using four unconstrained predecessor steps $T_0-4, T_0-3, T_0-2, T_0-1$, the two 19-bit conditions decouple completely in closed form:

1. **Pre-Alignment ($G_{T_0-2} = 0$):** Using the closed-form differential constructor (7)–(8), choose z_{T_0-4} so that G_{T_0-3} lies in the solvable input-difference set of $K_{AP}(T_0-3)$, and then invert $\Phi_\eta(G_{T_0-3}) = K_{AP}(T_0-3)$ in $O(1)$ at step T_0-3 ($z_{T_0-3} = \eta \oplus A_{T_0-3}$), forcing $G_{T_0-2} = 0$, i.e., $A_{T_0-2} = C_{T_0-2} =: D$.

2. **Decoupled Solve at T_0-2 and T_0-1:** Because $G_{T_0-2} = 0$ implies $\Phi_{z_{T_0-2} \oplus A_{T_0-2}}(0) = 0$, the state $G_{T_0-1} = K_{AP}(T_0-2)$ is *independent of* z_{T_0-2}. We therefore first solve $\Phi_Z(G_{T_0-1}) \in K_{AP}(T_0-1) \oplus V$ for $Z = z_{T_0-1} \oplus A_{T_0-1}$, then solve $\Phi_{Z \oplus G_{T_0-1}}(C_{T_0-1}) \in K_C(T_0-1) \oplus V$ for the required difference C_{T_0-1}, and finally invert $\Phi_{z_{T_0-2}}(D) = C_{T_0-1} \oplus K_C(T_0-2)$ for z_{T_0-2}.

Byte-Cascade Coset Reduction. At any 2-stream step t, (17) imposes a condition of the form $A_t = \Phi_{z_{t-1}}(D_0) \oplus K \in V$ with $D_0 = A_{t-1}$ and $K = K_A(t-1)$ (and similarly for C_t or G_t). When $\mathrm{act}_{t-1} = 0$, $z_{t-1} \in \mathbb{F}_2^{32}$ is unconstrained, so we construct z_{t-1} satisfying $\Phi_{z_{t-1}}(D_0) = K \oplus v$ (for the first admissible $v \in V$) in $O(1)$ via (7)–(8). When $\mathrm{act}_{t-1} = 1$, z_{t-1} is restricted to a 2^{13}-element coset $z_b \oplus V$. Rather than evaluating $\Phi_z(D_0)$ by brute force across all $2^{13} = 8{,}192$ elements $z \in z_b \oplus V$, we exploit the direct-sum decomposition $V = H_{24} \oplus \mathcal{B}_3$ of Theorem 1.

Theorem 3 (32-Element Byte-Cascade Coset Reduction)**.** *Let $z = z_0 \oplus (0,0,0,b_3) \in z_b \oplus V$, where $z_0 \in z_{b,0} \oplus H_{24}$ (with $z_{b,0} = (b_0(z_b), b_1(z_b), b_2(z_b), 0)$) ranges over only $|H_{24}| = 32$ values and $(0,0,0,b_3) \in \mathcal{B}_3$ carries the unconstrained top byte $b_3 \in \mathbb{F}_2^8$. Let $w_0(z_0)$ and $w_1(z_0, D_0)$ denote the states of z_0 and $z_0 \oplus D_0$ after the first three micro-rounds of* Xbox *(evaluated at $b_3 = 0$), and let $c_{3,0} = b_0(w_0(z_0))$ and $c_{3,1} = b_0(w_1(z_0, D_0))$. Then:*

1. *In micro-rounds $r = 1, 2, 3$, the S-box input bytes are independent of b_3, while after three right rotations b_3 shifts into the low byte (bits 0–7) of both states, canceling in $w_0 \oplus w_1$ and shifting the fourth-round S-box index to $i_4 = c_{3,0} \oplus b_3$ with fixed input difference $\delta_4 = c_{3,0} \oplus c_{3,1}$.*

2. *Therefore, the fourth-round S-box output difference $\Delta_4 = \mathsf{Sbox}(i_4) \oplus \mathsf{Sbox}(i_4 \oplus \delta_4)$ satisfies*

$$\begin{aligned} & \sigma\big(\Phi_{z_0 \oplus (0,0,0,b_3)}(D_0) \oplus K\big) = 0 \\ \iff \quad & \sigma\big(\mathsf{ROT}_8(\mathsf{XB}[i_4] \oplus \mathsf{XB}[i_4 \oplus \delta_4])\big) = \sigma\big(\mathsf{ROT}_8(w_0(z_0) \oplus w_1(z_0, D_0)) \oplus K\big). \end{aligned}$$

Since $\sigma(\mathsf{ROT}_8(\mathsf{XB}[i_4] \oplus \mathsf{XB}[i_4 \oplus \delta_4]))$ depends exclusively on $\Delta_4 \in \mathbb{F}_2^8$ (as its top byte $\Delta_4 \oplus \delta_4$ lies in $\mathcal{B}_3 \subset V = \ker(\sigma)$), a precomputed syndrome-to-byte-difference lookup table $\mathcal{T}_\Delta$ recovers the unique candidate Δ_4 in $O(1)$, and $b_3 = i_4 \oplus c_{3,0}$ exists if and only if $\mathsf{DDT}[\delta_4][\Delta_4] > 0$.

Algorithm 1: Unified Minimal-Length Full-State Collision Generator for Waterfall-224/256/384/512

Input: Target digest bit length $h \in \{224, 256, 384, 512\}$, block count $B \in \{7, 8, 9\}$ ($T = 17B$), seed $s \geq 0$.
Output: Colliding B-block message pair (M_1, M_2) ($M_1 \neq M_2$, $|M_1| = |M_2| = 64B$ bytes) with identical 93-word internal state at step T.
1. Construct $T_\beta = c \oplus V$ for $\beta = \texttt{0x00f9e70e}$ via (7)–(8), extract the 32-element low-24-bit subspace $H_{24} \subset V$ and quotient map $\sigma : \mathbb{F}_2^{32} \to \mathbb{F}_2^{19}$, and precompute the $O(1)$ syndrome lookup tables $\mathcal{T}_\Delta$ and $\mathcal{T}_{\text{ctr}}$.
2. Initialize state pair $(\mathcal{S}^{(1)}, \mathcal{S}^{(2)}) \leftarrow (\mathsf{Init}(h), \mathsf{Init}(h))$ and execute the 4-step closed-form injection alignment at steps $T_0 - 4, \ldots, T_0$ so that $A_{T_0}, C_{T_0} \in V$ and $\sigma(Y_a(T_0)) = \sigma(Y_b(T_0)) = \sigma(Y_c(T_0)) = \tau$.
3. for $t = T_0 + 1$ **to** $T - 1$ **do**
if t is a ring-buffer pre-steering step $t_{\text{pre}} = t_{\text{ctr}} - K_i$ for a counter step t_{ctr} with $e_i[t_{\text{ctr}}] = e_i[t_{\text{ctr}} - 1] = 1$ **then**
Steer $Y_i(t_{\text{pre}})$ via $\mathcal{T}_{\text{ctr}}$ so that $\mathsf{Xbox}(Y_i(t_{\text{pre}})) \oplus \mathsf{Xbox}(c) \oplus \text{ctr}_{t_{\text{ctr}}} \oplus c_i \in c \oplus V$ (16).
else if t owns an x_t-independent 19-bit condition ($\text{act}_t = 2$ or counter step with $e_i[t-1] = 0$) **then**
if $\text{act}_{t-1} = 0$ **then** solve step $t - 1$ in $O(1)$ closed form via (7)–(8) or Xbox^{-1} (15).
else redraw look-back word x_{t-k} ($k \in \{2, 3\}$) and run the 32-step byte-cascade solver via $\mathcal{T}_\Delta, \mathcal{T}_{\text{ctr}}$ (Theorem 3).
Place $x_t^{(1)}$ in its admissible coset (or $x_t^{(1)} \leftarrow \text{ctr}_t$ if $t \equiv 16 \pmod{17}$) and $x_t^{(2)} \leftarrow x_t^{(1)} \oplus d_t\beta$, advancing $(\mathcal{S}^{(1)}, \mathcal{S}^{(2)})$.
4. return the $64B$-byte messages M_1, M_2 formed by the non-counter words $(x_t^{(1)})_{t \not\equiv 16}$ and $(x_t^{(2)})_{t \not\equiv 16}$.

Thus, solving $\Phi_z(D_0) \oplus K \in V$ over $z \in z_b \oplus V$ requires only 32 three-round evaluations and $O(1)$ table lookups instead of 8,192 four-round Xbox evaluations, achieving a $256\times$ reduction. The exact same 32-element reduction applies to 1-stream counter and ring-buffer pre-steering steps via a precomputed syndrome lookup table $\mathcal{T}_{\text{ctr}}$ evaluated at $\tau \oplus \sigma(\mathsf{ROT}_8(w_0(y_0)) \oplus K_{\text{ctr}})$.

4 Security Analysis and Mitigations

4.1 Flaws in the Design Security Argument

The official Waterfall specification [8] rested its collision security argument on three structural pillars. Our algebraic differential framework demonstrates that all three pillars collapse:

1. **Failure of Coupled Stream Diffusion Over a Fixed-Point Coset:** Hattersley assumed that coupling three feedback streams of mutually coprime lengths $(16, 7, 6)$ through the nonlinear permutation Xbox would cause differences to avalanche uncontrollably across words and steps. However, Xbox is constructed as a 4-round byte cascade where only one byte is substituted per micro-round. Choosing a difference β with $\beta_3 = \texttt{0x00}$ and $\delta_4(\beta) = \Delta_4(\beta) = \texttt{0x00}$ therefore leaves the fourth-round S-box completely inactive ($\mathsf{DDT}[0][0] = 256$). This creates a 2^{13}-element affine fixed-point coset $T_\beta = c \oplus V$ on which Xbox behaves as the *identity map* on differences ($\Phi_Y(\beta) = \beta$). The three coupled nonlinear streams then become a linear code over $\mathbb{F}_2$, governed by the degree-29 generator polynomial $L(x) = \text{lcm}(1 + x + x^{16}, 1 + x + x^7, 1 + x + x^6)$ of (11).

2. **Failure of the 17-Step Block-Counter Defense:** Injecting the block counter ctr_b every 16 message words was intended to prevent differential control at step $t \equiv 16 \pmod{17}$. The permutation Xbox is a bijection, and the register slot $S_i[(t-1) \bmod K_i]$ (resp. $S_i[t \bmod K_i]$) read at counter step t is written at step $t - 1$ (resp. $t_{\text{pre}} = t - K_i$). Any single-stream condition at a counter step is therefore satisfied in advance via message modification [16] using Xbox^{-1} and $\mathcal{T}_{\text{ctr}}$ (15)–(16).

Table 2: Single-core full-state collision benchmarks and component ablations of our framework, verified against the official NIST reference implementation [8] (Xeon 6982P-C).

Our Configuration	Blocks	Bytes	Steps	Active	Solve Time	Variants	Success
Dual-excursion (w/o Eq. (13))	44	2,816 B	748	36	45.0 ms	256	18/20 (90%)
2^{13}-coset scan (w/o Thm. 3)	9	576 B	153	50	171.0 ms	256	20/20 (100%)
Quiet counter (9-block)	9	576 B	153	50	3.8 ms	All 4	20/20 (100%)
Step $t-1$ mod. (8-block)	**8**	**512 B**	136	56	**3.5 ms**	**All 4**	**20/20 (100%)**
Step $t-K_i$ pre-steering (7-block)	**7**	**448 B**	119	44	**52.1 ms**	**All 4**	**20/20 (100%)**

3. **Failure of the Write-Only Accumulator Pools:** Because Pool_2 and Pool_3 are *write-only* during message compression and never feed back into S_1, S_2, or S_3, the trunk trajectory is completely blind to the pool state. When all active stream steps output the fixed difference β, all 64 pool words cancel to zero difference within a single excursion via modulo-32 even parity (13).

4.2 Structural Mitigations and Overhead

To repair Waterfall's structural collision vulnerabilities, two modifications suffice:

1. **Pool-to-Trunk Feedback During Absorption:** Add $\mathsf{Pool}_2[(t-\eta_1) \bmod 32]$ and $\mathsf{Pool}_3[(t-\eta_2) \bmod 32]$ (for fixed tap offsets η_1, η_2) into $Y_a(t), Y_b(t)$, and $Y_c(t)$ inside $\mathsf{UpdateStep}$, introducing pool-to-state feedback analogous to PANAMA [5]. This breaks the write-only isolation of the pools, preventing decoupled modulo-32 pool parity cancellation at a cost of two 32-bit XORs per step ($\approx 3\%$ throughput overhead).

2. **Full 4-Byte Parallel Substitution and MDS Diffusion in Xbox:** Replace the single-byte substitution $w \oplus \mathsf{XB}[b_0(w)]$ in each round of Xbox with a full 4-byte parallel AES S-box layer followed by a 4×4 MDS matrix over $\mathbb{F}_{2^8}$ (branch number 5, as in an AES column round [10]). Because every two consecutive rounds then activate at least 5 byte S-boxes (≥ 10 active S-boxes across four rounds), this eliminates inactive-S-box rounds and large affine fixed-point cosets.

5 Experimental Verification

All collision solvers were implemented in C with standard compiler optimizations on an Intel Xeon 6982P-C processor. We verified them offline against the unmodified official NIST SHA-3 Round 1 reference implementation [8].

5.1 Collision Benchmarks and Component Ablations

Table 2 reports the single-core performance of our collision solvers across block lengths $B \in \{7, 8, 9\}$ (448, 512, and 576 bytes) and all four official NIST digest sizes ($h \in \{224, 256, 384, 512\}$). To isolate the impact of each technical contribution in Section 3, we also evaluate two ablated configurations of our solver on Waterfall-256: disabling single-excursion pool parity cancellation (13) (repeating a quiet-counter excursion twice at step shift $\Delta s = 544$, 44 blocks = 2,816 bytes), and disabling byte-cascade coset projection (scanning all 2^{13} coset elements on the 9-block quiet-counter trail).

8-Block (512-Byte) Solver. Our 8-block solver succeeds on Attempt 1 across all 20 tested seeds (20/20 distinct colliding pairs and digests), with a minimum runtime of 3.45 ms and median of 4.26 ms (mean 3.5 ms, walk time 1.3 ms), verified across Waterfall-224, −256, −384, and

Table 3: Cryptographic fingerprints of verified 448-byte (7-block) and 512-byte (8-block) full-state colliding message pairs (M_1, M_2) for Waterfall-256.

Field	Verified Value (Hexadecimal)
7-Block (448-Byte, 119-Step) Full-State Collision Pair	
M_1 SHA-256	`11dd59480b379f12158d86466f4535fcc43d1172ed455a5cec62bcc5bc5f61c5`
M_2 SHA-256	`9ad0f2ed6eb3188820e5dc19d9898351ec9009c76efbb69755a5bbe3100b32fd`
Waterfall-256	`f1b0894f55696afb7e52b15ceb5e9f1c74d22d02a626e4b95a9f880cc45e5a29`
Active words	44 message words differing by $\beta =$ `0x00f9e70e`
8-Block (512-Byte, 136-Step) Full-State Collision Pair	
M_1 SHA-256	`a0c2fcb4a4cca44951ac764a5e7c39b11897e8eb1ed771f81dbd3b518d92174f`
M_2 SHA-256	`579636a5831d61db67d526c89280a58eb2bdf23377f0311cf99d6bcf2445f341`
Waterfall-256	`18a75611cd796ed22ff29a54db3f0e2a7cf6b9a25e5b065db9e8071fa4f482d2`
Active words	56 message words differing by $\beta =$ `0x00f9e70e`

-512. Enabling single-excursion pool cancellation (13) and immediate counter modification (15) reduces message length by 5.5× relative to our 44-block dual-excursion configuration (2,816 B to 512 B), while byte-cascade coset reduction (Theorem 3) accelerates 9-block solving by 45× (171.0 ms to 3.8 ms; 49× to 3.5 ms at 8 blocks).

7-Block (448-Byte) Minimal Steerable Solver. Using delayed ring-buffer pre-steering (16) at step $t_{\text{pre}} = 50 - 16 = 34$ for counter step $t_{\text{ctr}} = 50$, together with immediate counter modification at $t_{\text{ctr}} = 84$, our 7-block solver generates verified 448-byte full-state collisions in 52.1 ms for Waterfall-256, matching the SAT lower bound for 3-step look-back steerable trails ($B \leq 6$ unsatisfiable). Across the remaining NIST digest sizes, the single-core solve times are 83.2 ms (Waterfall-224), 290.8 ms (Waterfall-384), and 126.0 ms (Waterfall-512).

Table 3 summarizes the cryptographic fingerprints and official NIST Waterfall digests of our verified 448-byte (7-block) and 512-byte (8-block) collision pairs.

Use of AI. The cryptanalytic results and artifacts were initially developed through an autonomous agent system powered by a MiniMax internal model. We validated the key technical claims and refined the presentation of the manuscript.

References

[1] G. Bertoni, J. Daemen, M. Peeters, and G. Van Assche, "RadioGatún, a Belt-and-Mill Hash Function," Presented at the NIST Second Cryptographic Hash Workshop, Santa Barbara, CA, Aug. 2006; Cryptology ePrint Archive, Paper 2006/369, 2006.

[2] E. Biham and R. Chen, "Near-Collisions of SHA-0," in *Advances in Cryptology – CRYPTO 2004*, LNCS, vol. 3152, pp. 290–305, Springer, 2004.

[3] E. Biham and A. Shamir, "Differential Cryptanalysis of DES-like Cryptosystems," *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.

[4] C. Bradbury, "BLENDER: A Proposed New Family of Cryptographic Hash Algorithms," Algorithm Specification and Supporting Documentation, Submission to the NIST SHA-3 Competition (Round 1), Oct. 2008.

[5] J. Daemen and C. Clapp, "Fast Hashing and Stream Encryption with PANAMA," in *Fast Software Encryption – FSE'98*, LNCS, vol. 1372, pp. 60–74, Springer, 1998.

[6] J. Daemen and G. Van Assche, "Producing Collisions for Panama, Instantaneously," in *Fast Software Encryption – FSE 2007*, LNCS, vol. 4593, pp. 1–18, Springer, 2007.

[7] S. Fluhrer, "Collision Attack on the Waterfall Hash Function," Cryptology ePrint Archive, Paper 2008/531, Dec. 2008.

[8] B. Hattersley, "Waterfall Hash — Algorithm Specification and Analysis," Version 1.0, Submission to the NIST SHA-3 Competition (Round 1), Oct. 15, 2008.

[9] B. Hattersley, "OFFICIAL COMMENT: Waterfall is broken," Official Public Comment, NIST SHA-3 Competition (Round 1), Dec. 20, 2008.

[10] National Institute of Standards and Technology, "Advanced Encryption Standard (AES)," Federal Information Processing Standards Publication FIPS PUB 197, Nov. 26, 2001.

[11] National Institute of Standards and Technology, "Announcing Request for Candidate Algorithm Nominations for a New Cryptographic Hash Algorithm (SHA-3) Family," *Federal Register*, vol. 72, no. 212, pp. 62212–62220, Nov. 2, 2007.

[12] K. Nyberg, "Differentially Uniform Mappings for Cryptography," in *Advances in Cryptology – EUROCRYPT'93*, LNCS, vol. 765, pp. 55–64, Springer, 1994.

[13] A. Regenscheid, R. Perlner, S.-j. Chang, J. Kelsey, M. Nandi, and S. Paul, "Status Report on the First Round of the SHA-3 Cryptographic Hash Algorithm Competition," NIST Interagency Report 7620, Sep. 2009.

[14] V. Rijmen, B. Van Rompay, B. Preneel, and J. Vandewalle, "Producing Collisions for PANAMA," in *Fast Software Encryption – FSE 2001*, LNCS, vol. 2355, pp. 37–51, Springer, 2002.

[15] M. Trojnara, "StreamHash Algorithm Specifications and Supporting Documentation," Submission to the NIST SHA-3 Competition (Round 1), Oct. 2008.

[16] X. Wang and H. Yu, "How to Break MD5 and Other Hash Functions," in *Advances in Cryptology – EUROCRYPT 2005*, LNCS, vol. 3494, pp. 19–35, Springer, 2005.