

Practical Key Recovery Attacks on Full DuX and Reduced-Round YuX

Xingwei Ren^{1,2}, Bo Xu^{1,2}, Zhenyu Xiong^{1,2}, Yongqiang Li^{1,2}✉, Xichao Hu⁴,
Lin Jiao³ and Mingsheng Wang^{1,2}

¹ State Key Laboratory of Cyberspace Security Defense, Institute of Information Engineering,
Chinese Academy of Sciences, Beijing, China

² School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

³ State Key Laboratory of Cryptology, Beijing, China

⁴ Zhongguancun Laboratory, Beijing, China

Abstract

DuX and YuX are two recent families of FHE-friendly block ciphers, designed for efficient evaluation under fully homomorphic encryption. Both keep a state of sixteen words of a large finite field, arranged in four blocks, and each round applies a block-wise S-box of low degree followed by a circulant linear layer. We show that, in the chosen-ciphertext model, the algebraic degree of their decryption functions grows far more slowly than the designers' evaluation of the encryption direction suggests, and we turn this observation into practical key-recovery attacks.

The starting point is a sufficient criterion for zero sums. It treats affine subspaces in characteristic 2, full prime fields, multiplicative cosets and arbitrary point sets in one statement, and over prime fields it is tight on every cell we could compute exactly. Three techniques turn the criterion into attacks. Full block structures make the first S-box layer free. Cheap coordinate elimination extracts equations from states that are only partially balanced, helped in characteristic 2 by a second cheap coordinate that exists only there. Weighted moments yield thousands of equations from a single structure, and weights with several indices take the place of further structures.

We recover the master key of the full twelve-round DuX over $\mathbb{F}_{65537}$ from 2^{29} chosen ciphertexts and over $\mathbb{F}_{2^{16}}$ from 2^{32} , in both cases in complete runs on random keys, the former in 20 hours on one processor core. For YuX, we recover the key of eleven of the fourteen rounds of Yu_pX-65537 and of Yu₂X-16 from 2^{32} chosen ciphertexts, again in complete runs, where the best previous attack on Yu₂X-16 needed 2^{96} . Our distinguishers on DuX coincide with those of concurrent and independent work by Liu and Sun (eprint 2026/1907). Furthermore, our key recovery attacks lower the data of their full-round attacks from $2^{67.59}$ and $2^{67.32}$ chosen ciphertexts to 2^{29} and 2^{32} .

Keywords: DuX, YuX, higher-order differential attack, zero-sum distinguisher, key recovery

1 Introduction

A block cipher evaluated under fully homomorphic encryption pays for every multiplication it performs, so ciphers built for that setting keep both the number and the depth of their multiplications small. The deployment that motivates them is hybrid homomorphic encryption [LNV11], in which the client encrypts with a symmetric cipher and the server evaluates

Emails: renxingwei@iie.ac.cn (Xingwei Ren), xubo2024@iie.ac.cn (Bo Xu), xiongzhenyu@iie.ac.cn (Zhenyu Xiong), liyongqiang@iie.ac.cn (Yongqiang Li), xchao_h@163.com (Xichao Hu), jiaolin_jl@126.com (Lin Jiao), wangmingsheng@iie.ac.cn (Mingsheng Wang)

the decryption of that cipher homomorphically. LowMC [ARS⁺15], Kreyvium [CCF⁺18], FLIP [MJSC16], MiMC [AGR⁺16], HERA [CHK⁺21], Rubato [HKL⁺22], Pasta [DGH⁺23] and Chaghri [AMT22] all follow this pattern, and all of them draw their efficiency from round functions of low multiplicative complexity, over $\mathbb{F}_2$ or over a large field.

An algebraic degree that stays low and grows slowly with the number of rounds is also the property that integral and higher-order differential attacks exploit, and the original design of Chaghri fell to a higher-order differential attack for that reason [LAW⁺23].

DuX [WDL⁺26] and YuX [LLC⁺24] are two recent families of FHE-friendly block ciphers of this kind. Both are substitution-permutation networks whose state consists of sixteen words of a finite field $\mathbb{F}_q$, arranged in four blocks of four, and each round applies a 4×4 S-box to every block and a circulant linear layer to the sixteen words. DuX comes in three instances, over $\mathbb{F}_{2^8}$, $\mathbb{F}_{2^{16}}$ and $\mathbb{F}_{65537}$, each with twelve rounds and a claimed security level of 128 bits. YuX has two members over binary fields, Yu₂X-8 over $\mathbb{F}_{2^8}$ and Yu₂X-16 over $\mathbb{F}_{2^{16}}$, and one over the prime field $\mathbb{F}_{65537}$, Yu_pX-65537. Yu₂X-8 has twelve rounds, the other two have fourteen, and all three claim 128 bits of security.

The two directions of these ciphers are far from symmetric. The S-box of DuX is the fourth power of a four-branch generalised Feistel round, and the coordinates of its inverse have degrees (2, 3, 4, 2) where those of the S-box itself have (5, 3, 2, 8). For YuX the inverse has degrees (2, 2, 3, 4) against (8, 5, 3, 2).

The designers of both families evaluate integral and higher-order differential attacks in the encryption direction only. We work in the chosen-ciphertext model instead, in which the structures are chosen on the ciphertext side and every coefficient of our equations is a sum of returned plaintext words.⁵ Throughout, complexities are compared with the claimed level of 2^{128} rather than with the size of the key space.

Our Approach. We bound the formal degree of every state word in the decryption direction by a max-plus recurrence and prove a single sufficient criterion for zero sums, Theorem 1. A word is balanced over a structure as soon as its degree bound stays below a threshold that depends only on the active words, whether these range over affine subspaces in characteristic 2, over a full prime field or over multiplicative cosets. The same threshold holds when cosets and fields are mixed in one structure and, under a divided-difference mask, for arbitrary point sets, whose size then becomes a free parameter of the data.

For DuX the bound takes the same value on every position of a class and obeys an exact linear recursion, Corollary 1, so it grows by a factor $2 + \sqrt{3}$ per layer rather than 4. After eleven layers its value at position 3 is 262 087, just below the four-word threshold 262 140, which gives an eleven-layer distinguisher on q^4 chosen ciphertexts. Over prime fields the bound agrees with the exact degree on every cell we computed. In characteristic 2 it does not, and Observation 2 identifies the mechanism.

A distinguisher becomes an attack only if the rounds in front of it can be removed at a cost that the available data supports, and this is achieved by three techniques. A *full block structure* lets one whole block range over $\mathbb{F}_q^4$, and by Lemma 2 the first inverse S-box layer then costs nothing.

Cheap coordinate elimination replaces the four equations per block of a two-round extension by the left kernel of the linear layer restricted to the expensive coordinates, so that by Lemma 3 a state in which only some positions are balanced still yields equations. In characteristic 2 a one-dimensional kernel provably collapses, because Theorem 3 shows its combined row to be a relative trace over a group of order four, in which $4 = 0$. The same

⁵The designers of DuX exclude the chosen-ciphertext model from their evaluation on the grounds that the intended deployment is hybrid homomorphic encryption, in which the server only ever decrypts. We do not present our results as a violation of their chosen-plaintext claim. Both families are published as block ciphers with a security claim in bits, and the decryption direction is where their algebra is weakest. The previous analysis of YuX [NWL26] and the concurrent analysis of DuX [LS26] use the same chosen-ciphertext model.

characteristic also offers the remedy. Squaring is additive there, so the cubic coordinate of the DuX S-box becomes a second cheap coordinate, and by Lemma 5 it turns the collapsed kernel into a usable one of dimension four.

Weighted moments, which are inner product masked sums in the sense of Wang, Tang and Wang [WTW25] with monomial values as the mask, turn one structure into one batch of equations per admissible weight. Theorem 4 gives the admissible range, and a slice decomposition computes every batch from a single pass over the structure followed by one matrix product. With a weight that carries several indices, one structure yields the equations that several structures would.

With these tools we recover the master key of the full twelve-round DuX(65537) from a single structure of 2^{29} chosen ciphertexts per key, in 20 hours on one processor core, and the master key of the full twelve-round DuX(2^{16}) from a single structure of 2^{32} . Both attacks were run to completion on three random keys. Eleven rounds of DuX fall from 2^{15} chosen ciphertexts over $\mathbb{F}_{65537}$ and from $2^{15.35}$ over $\mathbb{F}_{2^{16}}$. For YuX, we recover the key of eleven of the fourteen rounds of Yu_pX-65537 and of Yu₂X-16 from 2^{32} , and eight of the twelve rounds of Yu₂X-8 and of DuX(2^8) fall from one structure of 2^{32} and of 2^{24} respectively.

Twelve of the fourteen rounds of Yu₂X-16 and of Yu_pX-65537 would cost 2^{64} chosen ciphertexts. That route and a thirteen-round route on DuX(65537) were not executed at full scale, and they are marked as theoretical wherever they appear. Table 1 collects the results, together with the number of field multiplications of every attack and the time and memory of one run of every executed one.

Our Contributions. The contributions of this paper are summarised as follows.

- We prove a sufficient criterion for zero sums, Theorem 1 in Section 4, which covers affine subspaces, full prime fields, multiplicative cosets, mixed structures and arbitrary point sets in one statement. We show that it is tight over prime fields on every cell we could compute exactly, explain the mechanism behind its slack in characteristic 2, and identify the key-independent constants that appear at its boundary over prime fields.
- We develop three techniques that turn a distinguisher into a key-recovery attack, presented in Section 5, namely full block structures, cheap coordinate elimination and weighted moments. They include the characteristic-2 collapse theorem for cheap coordinate elimination and its remedy through the cubic coordinate, weighted moments on cosets, mixed structures and point sets, and a template bound that predicts the rank of the linearised systems without a single decryption.
- We present the attacks of Sections 6 and 7, namely the executed full-round attacks on DuX(65537) at 2^{29} and on DuX(2^{16}) at 2^{32} chosen ciphertexts, and reduced-round attacks on the other four instances, among them eleven-round Yu_pX-65537 and Yu₂X-16 at 2^{32} and eight-round DuX(2^8) at 2^{24} .
- We record four corrections to the literature in Sections A and C. The Boolean degree of the two-round extension used by [NWL26] is 7 rather than 6, four distinguishers are missing from their published tables, a six-round integral example in the DuX design paper does not hold, and the key-dependent linear layer of DuX is transparent to this kind of analysis.

Related and Concurrent Work. Ni, Wang and Li [NWL26] analyse the two members of YuX over binary fields in the same chosen-ciphertext model, tracking the algebraic degree through sets of achievable exponents. They attack seven rounds of Yu₂X-8 with 2^{48} data

Table 1: Master key recovery attacks of this paper, in bold, next to the best previous or concurrent result on the same instance and round number.

Instance	Rounds	Model	Data	Mult.	Time	Proc.	Memory	Reference
DuX(65537)	11/12	CC	$2^{35.32}$	—	$2^{35.32}$ decr.	—	—	[LS26]
DuX(65537)	11/12	CC	2^{15}	$2^{44.15}$	1.6 h	1	20.8 GB	Section 6.2
DuX(65537)	12/12	CC	$2^{67.59}$	—	$2^{67.59}$ decr.	—	—	[LS26]
DuX(65537)	12/12	CC	2^{29}	$2^{46.50}$	19.7 h	1	13.8 GB	Section 6.1
DuX(2^{16})	11/12	CC	$2^{35.32}$	—	$2^{35.32}$ decr.	—	—	[LS26]
DuX(2^{16})	11/12	CC	$2^{15.35}$	$2^{40.59}$	2 039 s	1	6.21 GB	Section 6.2
DuX(2^{16})	12/12	CC	$2^{67.32}$	—	$2^{67.32}$ decr.	—	—	[LS26]
DuX(2^{16})	12/12	CC	2^{32}	$2^{48.12}$	7.5 h	19	26.1 GB	Section 6.3
DuX(2^8)	7/12	CC	$2^{19.32}$	—	$2^{19.32}$ decr.	—	—	[LS26]
DuX(2^8)	7/12	CC	$2^{12.70}$	$2^{38.60}$	824 s	1	2.06 GB	Section 6.5
DuX(2^8)	8/12	CC	$2^{91.32}$	—	$2^{91.32}$ decr.	—	—	[LS26]
DuX(2^8)	8/12	CC	2^{24}	$2^{44.79}$	9.1 h	1	16.3 GB	Section 6.5
DuX(2^{16})	8/12	CP	2^{16}	$2^{36.58}$	1 190 s	1	1.94 GB	Appendix D
DuX(65537)	8/12	CP	2^{16}	$2^{36.59}$	125 s	1	3.73 GB	Appendix D
Yu_pX-65537	11/14	CC	2^{32}	$2^{48.38}$	4.6 h	24	19.4 GB	Section 7.2
Yu_pX-65537	12/14[†]	CC	2^{64}	$2^{80.31}$	—	—	1.87 GB	Section 7.3
Yu ₂ X-16	11/14	CC	2^{96}	—	2^{96} decr.	—	—	[NWL26]
Yu₂X-16	11/14	CC	2^{32}	$2^{47.32}$	4.1 h	36	5.48 GB	Section 7.2
Yu ₂ X-16	12/14	CC	$2^{120.4}$	—	$2^{120.4}$ decr.	—	—	[NWL26]
Yu₂X-16	12/14[†]	CC	2^{64}	$2^{79.29}$	—	—	0.53 GB	Section 7.3
Yu ₂ X-8	7/12	CC	2^{48}	—	2^{49} decr.	—	—	[NWL26]
Yu₂X-8	7/12	CC	$2^{17.58}$	$2^{42.45}$	2.8 h	1	6.68 GB	Section 7.4
Yu₂X-8	8/12	CC	2^{32}	$2^{47.32}$	3.6 h	19	4.56 GB	Section 7.4
Yu ₂ X-8	10/12	CC	$2^{95.5}$	—	$2^{115.3}$ decr.	—	—	[NWL26]

CC: chosen ciphertext, CP: chosen plaintext. Data is per key in $\log_2$ chosen texts. Mult. is the number of field multiplications of the attack in $\log_2$, counted by the formula of Section 8, the same for executed and theoretical rows. Time and Proc. describe one executed run on the machine of Section 8: the elapsed time, in seconds below one hour and in hours above, and the number of processes, where 1 is a single thread on one reserved core. Memory is the peak resident size of that run in gigabytes. Rows marked [†] are theoretical: every component was executed, but the assembly of the equations over the full structure was not. [LS26] count time in full decryptions, equal to their data, and [NWL26] in decryptions of the attacked rounds, and one twelve-round DuX decryption is about $2^{7.6}$ field multiplications. The ten-round attack of [NWL26] on Yu₂X-8 remains the round record on that instance.

and ten rounds with $2^{95.5}$ data and $2^{115.3}$ time, and on Yu₂X-16 they reach eleven rounds with 2^{96} and twelve with $2^{120.4}$, leaving the member over the prime field as future work. Their ten-round attack on Yu₂X-8 remains the round record on that instance.

Liu and Sun [LS26] analyse DuX concurrently and independently of this work, in the same model, and their degree bounds coincide with ours. The class-wise recursion with base $2 + \sqrt{3}$, the eleven-layer distinguisher on q^4 chosen ciphertexts with margins 53 and 57, the 2^{88} seven-layer distinguisher on DuX(2^8), the identity between the two linear layers and the correction of the six-round examples of the design paper were all obtained on both sides.

The key-recovery attacks, however, differ. They remove one round in front of the eleven-layer distinguisher with linear masks sorted into three stages, which needs ten to twelve subspaces of q^4 ciphertexts, $2^{67.32}$ to $2^{67.59}$ data and time, and was verified on reduced-round configurations. We remove two rounds in front of the ten-layer distinguisher by structured linearisation, survive its partially balanced pattern with the cheap coordinate kernel, and extract with weighted moments every equation a single structure can yield. The full cipher then falls to one structure of 2^{29} chosen ciphertexts over $\mathbb{F}_{65537}$ and of 2^{32} over $\mathbb{F}_{2^{16}}$. Weighted moments alone would already reduce their data to a single subspace of 2^{64} , and on their own reduced-round configuration one subspace with weights attains

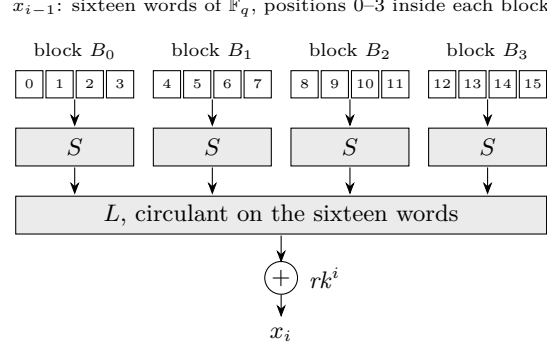


Figure 1: One round of DuX and YuX: a block-wise S-box on the four blocks of the sixteen-word state, the circulant linear layer and the round key addition.

the ranks of their three stages, as Section C reports.

We also reproduce their known-key zero-sum partitions and extend them to YuX in Table 10. Section C gives the correspondence in detail, together with the relation to the wider literature on integral attacks over large fields.

Organisation. Section 2 fixes the notation, the two specifications and the model. Section 3 proves the structural properties on which the analysis rests. Section 4 states and proves the criterion and studies its tightness. Section 5 turns distinguishers into key-recovery equations. Sections 6 and 7 give the attacks on DuX and YuX, Section 8 reports what was executed and how it was measured, and Section 9 concludes. Section A collects the specification notes, Section B the deferred proofs and Section C the comparison with prior and concurrent work. Section D gives the chosen-plaintext results with Table 9, Section E the known-key partitions, Section F the optimality of the structures inside our framework with Table 11, and Section G the numerical tables.

2 Preliminaries

2.1 Notation

Both ciphers keep a state of sixteen words $X_0, \dots, X_{15} \in \mathbb{F}_q$, grouped into four *blocks* $B_b = (X_{4b}, \dots, X_{4b+3})$. The index $p = i \bmod 4$ of a word is its *position* and the set $C_p = \{i : i \equiv p \pmod{4}\}$ is its *class*. We write Rot_j for the rotation $(\text{Rot}_j X)_i = X_{(i+j) \bmod 16}$. A circulant matrix $\text{Circ}(m_0, \dots, m_{15})$ then acts as $(MX)_i = \sum_j m_j X_{(i+j) \bmod 16}$. This convention reproduces the printed prime field matrices of both design papers.

Round keys are $rk^0, \dots, rk^r$, with $rk^0 = K$ the master key in both families. One round applies a block-wise S-box layer SL, a linear layer L on the sixteen words and a round key addition, and the last round has no linear layer, as Figure 1 shows:

$$x_0 = P + rk^0, \quad x_i = L(\text{SL}(x_{i-1})) + rk^i \quad (1 \leq i \leq r-1), \quad C = \text{SL}(x_{r-1}) + rk^r. \quad (1)$$

Decryption applies the inverse operations in the reverse order.

The state after ℓ inverse S-box layers of decryption is $x_{r-\ell}$, and we call ℓ the number of *layers*. The first layer is S^{-1} applied to $C - rk^r$ directly, so an ℓ -layer computation contains $\ell - 1$ inverse linear layers.

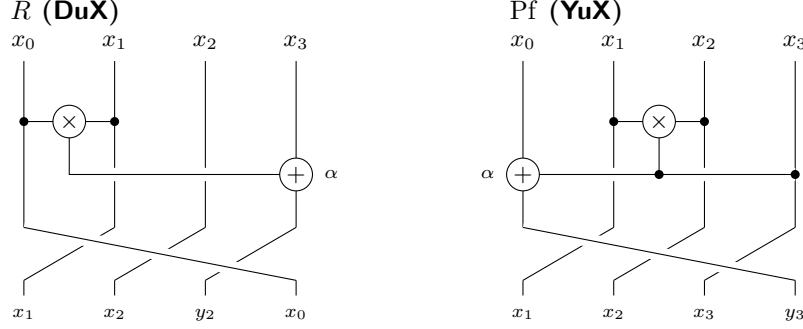


Figure 2: The four-branch Feistel rounds $R(x) = (x_1, x_2, y_2, x_0)$ of DuX, with $y_2 = x_0x_1 + x_3 + \alpha$, and $\text{Pf}(x) = (x_1, x_2, x_3, y_3)$ of YuX, with $y_3 = x_0 + x_1x_2 + x_3 + \alpha$, whose fourth powers are the two decryption S-boxes. A dot is a tap or a junction, a gap is a crossing without contact, the adder on the target branch also adds the constant α , and the diagonals at the bottom are the branch rotation.

2.2 DuX

DuX [WDL⁺26] has three instances, over $\mathbb{F}_{2^8}$ with the polynomial $x^8 + x^4 + x^3 + x + 1$, over $\mathbb{F}_{2^{16}}$ with $x^{16} + x^{12} + x^3 + x + 1$ and over $\mathbb{F}_{65537}$, all with $r = 12$ rounds, the constant $\alpha = 179$ and a claimed 128-bit security level. The master key is sixteen words, so DuX(2^{16}) and DuX(65537) have a key space larger than their claim. The S-box is built from the four-branch generalised Feistel round of Figure 2

$$\begin{aligned} R(x_0, x_1, x_2, x_3) &= (x_1, x_2, x_0x_1 + x_3 + \alpha, x_0), \\ R^{-1}(y_0, y_1, y_2, y_3) &= (y_3, y_0, y_1, y_2 - y_0y_3 - \alpha), \end{aligned}$$

as $S^{-1} = R^4$ for decryption and $S = (R^{-1})^4$ for encryption.

Expanding,

$$f = x_0x_1 + x_3 + \alpha, \quad g = x_1x_2 + x_0 + \alpha, \quad S^{-1}(x) = (g, x_2f + x_1 + \alpha, fg + x_2 + \alpha, f), \quad (2)$$

of coordinate degrees $(2, 3, 4, 2)$, and

$$a = x_2 - x_0x_3 - \alpha, \quad b = x_1 - x_3a - \alpha, \quad c = x_0 - ab - \alpha, \quad S(x) = (c, b, a, x_3 - bc - \alpha), \quad (3)$$

of degrees $(5, 3, 2, 8)$. The unique quadratic coordinate of S is the one at position 2, and in the following we refer to it as the *cheap* coordinate.

The linear layer is one of two circulants selected by a round key bit. In the decryption direction

$$L_0^{-1} = \sum_{j \in \{1, 4, 8, 9, 13\}} \text{Rot}_j, \quad L_1^{-1} = \sum_{j \in \{1, 5, 8, 12, 13\}} \text{Rot}_j$$

over both kinds of field. In the encryption direction L_0 is the sum of the rotations $\{0, 1, 3, 4, 7, 8, 9, 10, 12, 14, 15\}$ over $\mathbb{F}_{2^n}$ and the dense circulant inverse of L_0^{-1} over $\mathbb{F}_p$. Round i uses L_t with t the XOR of the two low bits of rk_{15}^i .

The key schedule iterates the four-branch map $\text{Rf}(x_0, x_1, x_2, x_3; rc) = (S(x_0) + x_2 + rc, S(x_1) + x_0 + x_3 + rc, S(x_1) + x_3 + 2x_0 + rc, x_1)$ on blocks of four words, four times per round key with four sub-constants, the constants being $S(t+1, t+2, t+3, t+4)$ for $t = 16i + 4j$. Because it is invertible, recovering any round key recovers the master key. Section A lists the typographical points of the design paper and the readings we adopt.

2.3 YuX

YuX [LLC⁺24] uses the same state and round structure of equation (1) with a single linear layer and no key-dependent selection. Its S-box comes from $\text{Pf}(x_0, x_1, x_2, x_3) = (x_1, x_2, x_3, x_0 + x_1x_2 + x_3 + \alpha)$ as $S^{-1} = \text{Pf}^4$ and $S = \text{Pf}^{-4}$, with $\alpha = 205$:

$$\begin{aligned} y_0 &= x_0 + x_1x_2 + x_3 + \alpha, & y_1 &= x_1 + x_2x_3 + y_0 + \alpha, \\ y_2 &= x_2 + x_3y_0 + y_1 + \alpha, & y_3 &= x_3 + y_0y_1 + y_2 + \alpha, \end{aligned} \quad (4)$$

of degrees $(2, 2, 3, 4)$, and

$$\begin{aligned} z_0 &= x_3 - x_0x_1 - x_2 - \alpha, & z_1 &= x_2 - z_0x_0 - x_1 - \alpha, \\ z_2 &= x_1 - z_1z_0 - x_0 - \alpha, & z_3 &= x_0 - z_2z_1 - z_0 - \alpha, \end{aligned} \quad (5)$$

with $S(x) = (z_3, z_2, z_1, z_0)$ of degrees $(8, 5, 3, 2)$. The cheap coordinate of S is at position 3.

The decryption linear layer is $L^{-1} = \sum_{j \in \{0, 3, 4, 8, 9, 12, 14\}} \text{Rot}_j$ over both kinds of field. Its inverse is the rotation sum $\{1, 2, 3, 5, 6, 7, 8, 12, 13, 14, 15\}$ over $\mathbb{F}_{2^n}$ and a dense circulant over $\mathbb{F}_p$. The key schedule is a four-branch non-linear feedback register on blocks, $X_4 = X_0 + S((X_1 + X_2 + X_3) \lll 3) + rc$, also invertible.

The instances are Yu₂X-8 over $\mathbb{F}_{2^8}$ with twelve rounds, Yu₂X-16 over $\mathbb{F}_{2^{16}}$ with fourteen rounds and Yu_pX-65537 over $\mathbb{F}_{65537}$ with fourteen rounds, all with a claimed 128-bit security level. In characteristic 2 the S-boxes of the two families are affine equivalent word by word, as Appendix B of [WDL⁺26] shows, whereas over prime fields they are not. The two families therefore differ mainly in their linear layers.

2.4 Structures, Zero Sums and the Model

A *structure* is a set $\mathcal{P}$ of chosen ciphertexts in which s words range over prescribed subsets $U_1, \dots, U_s \subseteq \mathbb{F}_q$ and the other $16 - s$ words are constant. A state word Z after ℓ layers is *balanced* on $\mathcal{P}$ if $\sum_{x \in \mathcal{P}} Z(x) = 0$. We record a layer's behaviour as a four-bit *pattern*, bit p set when position p is balanced in all four blocks. The pattern 1111 is a full zero sum, and 1101 means that positions 0, 1 and 3 are balanced. All attacks are in the chosen-ciphertext model. The structure is submitted to a decryption oracle, and every coefficient of the equations below is a sum of returned plaintext words. The attacks address the security claim of the ciphers themselves, and what they imply for the homomorphic setting, in which the designers exclude chosen ciphertexts, is left to the designers.

2.5 The Two Extension Regimes

Given an ℓ -layer zero sum on an r -round cipher we recover the key by peeling $r_{\text{KR}} = r - \ell$ rounds off the plaintext side. With $r_{\text{KR}} = 1$ the balanced state is $x_1 = L(\text{SL}(P + rk^0)) + rk^1$. With $r_{\text{KR}} = 2$ it is $x_2 = L(W') + rk^2$ where $W' = \text{SL}(x_1)$ is the *outer* S-box layer and $Y = \text{SL}(P + rk^0)$ the *inner* one.

Summing over the structure removes the round keys, because $|\mathcal{P}| \equiv 0$ in $\mathbb{F}_q$. The equations then involve only rk^0 and, through x_1 , rk^1 . The unknowns are linearised over the monomials that occur.

For the two-round extension these are the monomials of the products of two inner S-box outputs of *different* blocks, and their number is

$$M = 38\,051 \text{ (DuX, } \mathbb{F}_p), \quad 18\,025 \text{ (DuX, } \mathbb{F}_{2^n}), \quad 53\,857 \text{ (YuX, } \mathbb{F}_p), \quad 27\,895 \text{ (YuX, } \mathbb{F}_{2^n}),$$

against $\binom{256}{\leq 7} \approx 2^{43.6}$ for the Boolean linearisation used in [NWL26]. Section 5 gives the equations.

3 Structural Properties of DuX and YuX

The key-dependent choice of linear layer in DuX is invisible to the analysis, the inverse layer of DuX mixes only two positions into each output position, and the inverse layer of YuX contains a summand of rank four. Everything that follows rests on these properties of the linear layers.

Proposition 1 (transparency of the key-dependent linear layer). $L_1 = \text{Rot}_{12} \circ L_0$, equivalently $L_1^{-1} = \text{Rot}_4 \circ L_0^{-1}$. Since Rot_4 is a whole block rotation it commutes with SL and with every circulant, and $\text{Rot}_4(x) + k = \text{Rot}_4(x + \text{Rot}_{12}k)$. Pushing every rotation to the output,

$$\begin{aligned} \text{DuX}(P; rk^0, \dots, rk^r) &= \text{Rot}_{12T} \left(\text{DuX}_{L_0}(P; rk^0, \text{Rot}_{4T_1} rk^1, \dots, \text{Rot}_{4T} rk^r) \right), \\ T_i &= t_1 + \dots + t_i, \quad T = T_{r-1}, \end{aligned}$$

where DuX_{L_0} uses L_0 in every round.

Proof. The rotation sets satisfy $\{1, 5, 8, 12, 13\} = \{1, 4, 8, 9, 13\} + 4 \pmod{16}$, which is the first identity. The rest is the commutation relations and induction on the round. Over $\mathbb{F}_p$ the design paper itself notes that M_1 is M_0 shifted by one block. $\square$

As a consequence, the choice of linear layer is irrelevant to any zero-sum distinguisher, and for key recovery it amounts to a rotation of the output by two bits, so all analyses below fix L_0 . In particular, the claim of the design document that multiplying the state by key material adds non-linearity does not hold in view of this identity.

Proposition 2 (positional lag in DuX). Every rotation offset of L_0^{-1} and L_1^{-1} is $\equiv 0$ or $1 \pmod{4}$: output position p receives only positions p and $p+1$ of the S-box layer, as Figure 3 illustrates. YuX's offsets $\{0, 3, 4, 8, 9, 12, 14\}$ cover all four residues. Consequently the formal degree of DuX grows at $2 + \sqrt{3} \approx 3.73$ per layer by Corollary 1 while that of YuX grows at 4, and in DuX position 3 lags about one layer behind the others.

Proof. The offsets 1, 4, 8, 9, 13 are $1, 0, 0, 1, 1 \pmod{4}$ and the offsets 1, 5, 8, 12, 13 are $1, 1, 0, 0, 1 \pmod{4}$, so $(L_t^{-1}Y)_i$ is a sum of words of positions i and $i+1 \pmod{4}$ only. For YuX the offsets 0, 3, 4, 8, 9, 12, 14 reduce to $0, 3, 0, 0, 1, 0, 2$. The growth rates are Corollary 1 and the recurrence of Section 4. $\square$

Proposition 3 (low-rank part of the YuX layers). With $A_0 = \sum_{b=0}^3 \text{Rot}_{4b}$ the operator of rank 4 that replaces every word by the sum of the four words in its class, $L^{-1} = A_0 + \text{Rot}_3 + \text{Rot}_9 + \text{Rot}_{14}$, and over $\mathbb{F}_{2^n}$, $L = J + I + \text{Rot}_4 + \text{Rot}_9 + \text{Rot}_{10} + \text{Rot}_{11}$ with J the all-ones matrix of rank 1.

Proof. The offsets 0, 4, 8, 12 of L^{-1} are the four whole block rotations, whose sum is A_0 , and the remaining offsets are 3, 9 and 14. Over $\mathbb{F}_{2^n}$ the rotation set $\{1, 2, 3, 5, 6, 7, 8, 12, 13, 14, 15\}$ of L is the complement of $\{0, 4, 9, 10, 11\}$ in $\mathbb{Z}_{16}$, and the sum of all sixteen rotations is J . $\square$

It is these low-rank summands that make the cheap coordinate kernel of Section 5.4 non-trivial for YuX, and the sparse $\{p, p+1\}$ mixing plays the same role for DuX.

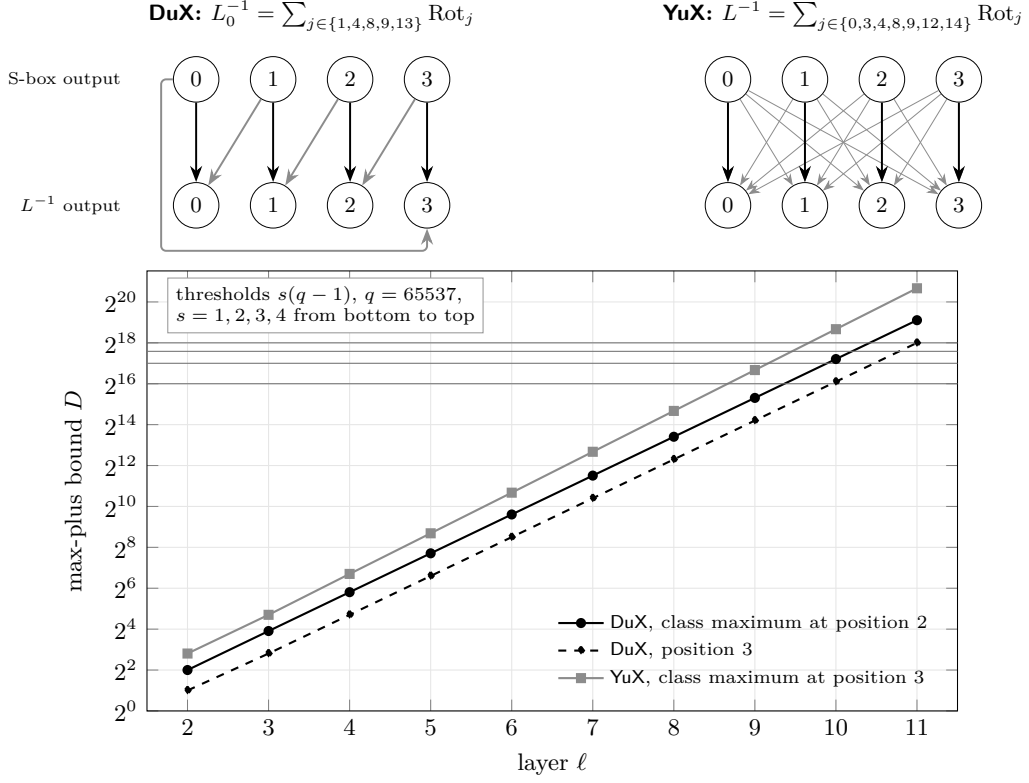


Figure 3: Positional lag: in DuX every output position of the inverse linear layer receives only positions p and $p+1$ of the S-box layer, the black and grey arrows, in YuX all four, and the resulting degree bounds for one active word from Table 2 cross the thresholds $s(q-1)$, $q = 65537$, one layer later in DuX than in YuX, with position 3 of DuX trailing its class maximum.

4 A Degree Criterion for Zero Sums

4.1 The Bound and the Threshold

Let the s active ciphertext words range over $U_1 \times \dots \times U_s$. For every state word we track an upper bound D on its *formal total degree*, the total degree of the unreduced polynomial in the s active words, by the max-plus recurrences of the two inverse S-boxes,

$$\begin{aligned} \text{DuX: } d_0 &= \max(e_1 + e_2, e_0), & d_3 &= \max(e_0 + e_1, e_3), \\ d_1 &= \max(e_2 + d_3, e_1), & d_2 &= \max(d_3 + d_0, e_2), \end{aligned} \quad (6)$$

$$\begin{aligned} \text{YuX: } d_0 &= \max(e_0, e_1 + e_2, e_3), & d_1 &= \max(e_1, e_2 + e_3, d_0), \\ d_2 &= \max(e_2, e_3 + d_0, d_1), & d_3 &= \max(e_3, d_0 + d_1, d_2), \end{aligned} \quad (7)$$

where $(e_0, \dots, e_3)$ are the bounds on the four inputs of a block, followed by $e'_i = \max_j d_{(i+j) \bmod 16}$ over the rotation offsets of the inverse linear layer.

Each recurrence is read off the closed forms of equations (2) and (4). A coordinate is a sum of products of earlier coordinates, and the degree of a product is at most the sum of the degrees. The encryption direction uses the same rule on equations (3) and (5) with the forward linear layer.

Theorem 1 (unified zero-sum criterion). *Let Z be a state word with formal total degree*

at most D in the active words. Then $\sum_{x \in U_1 \times \dots \times U_s} Z(x) = 0$ whenever $D < T$, where

$$T = \begin{cases} \sum_{i=1}^s (2^{m_i} - 1) & \text{over } \mathbb{F}_{2^n}, U_i \text{ an affine subspace of dimension } m_i, \\ s(p-1) & \text{over } \mathbb{F}_p, U_i = \mathbb{F}_p, \\ \sum_{i=1}^s 2^{k_i} & \text{over } \mathbb{F}_p, U_i \text{ a coset of the subgroup of order } 2^{k_i}, \end{cases}$$

where $m_i = n$ is the full field, and in the coset case the sum is the signed 2^s -term inclusion-exclusion over two cosets per variable.

Proof. Z is an $\mathbb{F}_q$ -linear combination of monomials, and for a monomial the sum factorises over the product set: $\sum_x \prod_i x_i^{a_i} = \prod_i (\sum_{x_i \in U_i} x_i^{a_i})$. A factor with $a_i = 0$ equals $|U_i|$, which is 0 in $\mathbb{F}_q$ for an affine subspace of positive dimension in characteristic 2 and for the full prime field, and is cancelled by the signed difference in the coset case, where the signed inclusion-exclusion over the 2^s coset tuples subtracts, for each variable, two cosets of the same subgroup.

A factor with $a_i > 0$ is non-zero only in three cases. Over $\mathbb{F}_p$ it requires $(p-1) \mid a_i$, hence $a_i \geq p-1$. Over a coset aH of the subgroup H of order 2^{k_i} , $\sum_{x \in aH} x^{a_i} = a^{a_i} \sum_{h \in H} h^{a_i}$ is non-zero only if $2^{k_i} \mid a_i$, hence $a_i \geq 2^{k_i}$. Over an affine subspace U_i of dimension m_i in characteristic 2, $\sum_{x \in U_i} x^{a_i}$ is the m_i -th order derivative of x^{a_i} , which vanishes when the Boolean degree $\text{wt}(a_i)$ is below m_i , and $0 < a_i < 2^{m_i} - 1$ implies $\text{wt}(a_i) \leq m_i - 1$, hence $a_i \geq 2^{m_i} - 1$.

All s lower bounds must hold simultaneously for the product to survive. A surviving monomial therefore has total degree at least T , and a word of formal degree below T sums to zero. $\square$

The full field prime case is Proposition 1 and Corollary 1 of *Out of Oddity* [BCD⁺20] and Theorem 1 of [LS26], and sums over multiplicative subgroups appear in Proposition 2 of [BCD⁺20]. Theorem 1 adds the joint bookkeeping.

A single threshold thus covers the three types of structure, both characteristics and both directions, and it depends only on one number per word. Note that the threshold grows linearly in the number of active words, whereas the bound D of a word that already ranges over the whole field does not depend on the field at all. Once a word is saturated, adding a further active word is therefore the only way to extend a distinguisher, and each such word gains whole layers.

The theorem provides a sufficient condition only, since it ignores the reduction $x^q = x$, and Sections 4.3 and 4.5 quantify the cases in which it is not tight.

4.2 The Class Recursion of DuX

Corollary 1 (class-wise recursion). *For DuX in the decryption direction let the active set be a whole class C_p , or one word from the second layer on with the block-symmetric bound. Then the bounds are constant on classes, and writing (d_0, d_1, d_2, d_3) for the class values after a layer, the next layer satisfies*

$$d'_0 = 2d_2, \quad d'_3 = d_1 + d_2, \quad d'_1 = d_1 + 2d_2, \quad d'_2 = d_1 + 3d_2$$

as soon as $d_2 \geq d_1 \geq d_0 \geq d_3$ and the sums dominate in equation (6), which holds from the second layer on. Hence each class value obeys $d^{(\ell+1)} = 4d^{(\ell)} - d^{(\ell-1)}$ and grows like $(2 + \sqrt{3})^\ell$.

Proof. By Proposition 2, $e'_p = \max(d_p, d_{p+1})$, so under the stated order $(e_0, e_1, e_2, e_3) = (d_1, d_2, d_2, d_0)$. Substituting into equation (6) with the sums dominating gives the four formulas. The pair (d_1, d_2) is then multiplied by the matrix $\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ of characteristic polynomial $\lambda^2 - 4\lambda + 1$, whose roots are $2 \pm \sqrt{3}$. The values d'_0 and d'_3 are linear in (d_1, d_2) and

Table 2: Per-position max-plus bounds, the maximum over the four blocks, in the decryption direction for one active ciphertext word, from the recurrences of Section 4: DuX started at position 3 and at position 0, and YuX started at position 0. The thresholds are $q - 1$ for one full word, $2(q - 1)$ for two, $3(q - 1)$ for three and $4(q - 1)$ for four words or a full block.

ℓ	DuX, word 3				DuX, word 0				YuX, word 0			
	pos. 0	1	2	3	pos. 0	1	2	3	pos. 0	1	2	3
1	0	1	1	1	1	1	2	1	1	1	1	2
2	2	3	4	2	4	5	7	3	3	4	5	7
3	8	11	15	7	14	19	26	12	13	14	20	26
4	30	41	56	26	52	71	97	45	52	52	78	104
5	112	153	209	97	194	265	362	168	204	208	308	412
6	418	571	780	362	724	989	1351	627	816	824	1228	1633
7	1560	2131	2911	1351	2702	3691	5042	2340	3265	3266	4898	6530
8	5822	7953	10864	5042	10084	13775	18817	8733	13023	13060	19553	26083
9	21728	29681	40545	18817	37634	51409	70226	32592	52041	52166	78124	104178
10	81090	110771	151316	70226	140452	191861	262087	121635	208260	208356	312438	416520
11	302632	413403	564719	262087	524174	716035	978122	453948	832800	833040	1249320	1665840

satisfy the same recursion one step later. The order and the domination are checked at layer 2, where the class vector is $(2, 3, 4, 2)$, and are preserved by the formulas. $\square$

Table 2 lists the resulting numbers for one active word of each cipher. The position-3 value of DuX after eleven layers is 262 087, against the four-word threshold $4(q - 1) = 262\,140$ over $\mathbb{F}_{2^{16}}$ and 262 144 over $\mathbb{F}_{65537}$. This yields the eleven-layer distinguisher on q^4 chosen ciphertexts.

Tables 3 and 4 collect the zero-sum distinguishers the attacks use. Each row gives the structure, its size, the first layer that is not fully balanced with its pattern, and the margin $T - D$ at the position that matters. Each attack below refers to the cell from which it starts.

4.3 Tightness

Observation 1 (equality of the bound over prime fields). *For $Yu_pX-65537$ with one active ciphertext word, the exact formal degree of every state word, computed by dense symbolic expansion without truncation, equals the max-plus value of equation (7) for ten layers, all sixteen words and three random master keys, 480 cells in all. For $DuX(65537)$ the exact univariate polynomials of the first nine layers attain the bound on three keys, and the character sum spectrum confirms equality at layers 4 to 9 for $DuX(2^{16})$ as well. At layer 10 of $DuX(2^{16})$ the exact degrees are $(81\,090, 110\,771, 151\,316, 70\,226)$, equal to the bound.*

This observation rests on computations for three keys and is not a theorem valid for all keys. On these cells the criterion is exact, and the three-word structure of $DuX(2^{16})$ in Section 6.3 relies on the exact value 151 316. In characteristic 2 the situation is different for YuX.

Observation 2 (key-independent exact degrees in characteristic 2). *For Yu_2X-16 with one active word, the exact formal degrees computed by dense symbolic expansion up to layer 10 are identical for three random master keys at every layer and every word, and fall short of the max-plus bound by a gap that grows from 4.7% at layer 2 to 7.4% at layer 9 and 7.5% at layer 10. At layer 9, block 0 has exact degrees $(48\,128, 48\,224, 72\,240, 96\,352)$ against the bound $(52\,041, 52\,166, 78\,124, 104\,178)$.*

The gap accumulates only for YuX in the decryption direction in characteristic 2. Over $\mathbb{F}_p$ it is zero by Observation 1, and the encryption direction of YuX and the decryption direction of DuX each lose once, at layer 3, by 1.4% and 6.7% respectively, and nothing

Table 3: Zero-sum distinguishers on DuX in the decryption direction: the structure, its size, the first layer that is not fully balanced with its pattern, and the margin $T - D$ at the position that matters.

Instance	Structure	Data	Layers / next pattern	Margin
65537, 2^{16}	1 word, position 2 or 3	2^{16}	9 / 0000	—
65537, 2^{16}	1 word, position 0 or 1	2^{16}	8 / 1101	—
65537	difference of two order- 2^{10} cosets	2^{11}	6 / —	—
65537	difference of two order- 2^{15} cosets	2^{16}	8 / 1101	178
65537, 2^{16}	2 words (3, 7)	2^{32}	9 / 1101	20 301 at position 1
65537	3 words (3, 7, 11)	2^{48}	10 / 0000	45 292
2^{16}	3 words, dims (16, 16, 15)	2^{47}	10 / 0000	12 521 ^a
65537	4 words or a full block	2^{64}	10 / 0001	57 at position 3
2^{16}	4 words or a full block	2^{64}	10 / 0001	53 at position 3
2^8	1 word	2^8	5 / —	—
2^8	2 words	2^{16}	5 / 1001	—
2^8	3 words	2^{24}	5 / 1101	—
2^8	4 words	2^{32}	6 / —	240
2^8	4 words, dims (8, 8, 8, 5)	2^{29}	6 / 0000	16
2^8	full block + words 6, 10, dims (8, 7)	2^{47}	6 / 0001	179
2^8	11 words	2^{88}	6 / 0001	103

“Layers” counts inverse S-box layers from the ciphertext, and a pattern has bit p set when position p is balanced in all four blocks. ^a $T - D = 163\,837 - 151\,316$, where D is the exact degree of Observation 1. Which cells were measured, and on how many keys, is stated in Section 8.

afterwards. Along the structure axis the gap is 7.4% for one active word, at most about 2% for two, where the scan ran to completion, and zero for a full block.

The mechanism behind this gap is explained by Lemma 6 in Section B. With one active word the outputs of the first layer on the active block satisfy $Y_1 = Y_0 + g$ for a constant g , and the row sum $u_1 + u_3$ of L^{-1} that feeds block 1 equals g in characteristic 2. The quadratic term of one coordinate degenerates, and the exact degree drops below the bound. Over $\mathbb{F}_p$ the same row sum is $Y_0 + Y_1 + 2Y_3$. This sum is not constant, so no cancellation occurs. After the substitution of Lemma 2 the four variables of a full block are independent, the relation disappears, and the gap vanishes. The observation does not change any attack, and the gaps that would be required for a change are listed after Lemma 6.

4.4 Boundary Cells over Prime Fields

Over a prime field the criterion agrees with the exact degree on every cell we computed, and the cell at which D reaches T therefore deserves a closer look. For a full block, the sum at such a cell turns out to be a constant.

Proposition 4 (key-independent constants at the boundary). *Let $q = p$ be prime, let one full block of the ciphertext range over $\mathbb{F}_p^4$ with the other twelve words fixed, and write $y_0, \dots, y_3$ for the outputs of the first layer on that block after the substitution of Lemma 2. For a state word Z after ℓ layers let P be the homogeneous part of degree D of Z as a polynomial in y . (i) P depends neither on the round keys nor on α nor on the twelve inactive words. (ii) If $D = T = 4(p-1)$ then $\sum_{\mathcal{P}} Z = \sum_{y \in \mathbb{F}_p^4} P(y)$, a constant independent of the key. (iii) For YuX in the decryption direction the leading position is 3 with $D = 4^{\ell-1}$, so the boundary $p-1 = 4^{\ell-2}$ is met at layer 3 for $p = 5$, layer 4 for $p = 17$, layer 6 for*

Table 4: Zero-sum distinguishers on YuX in the decryption direction, with layers, patterns and margins as in Table 3.

Instance	Structure	Data	Layers / next pattern	Margin
Yu ₂ X-16	1 word, position 0	2 ¹⁶	8 / 1100	13 494, 13 369
Yu _p X-65537	1 word, position 0	2 ¹⁶	8 / 1100	13 495, 13 370
Yu ₂ X-16	1 word, position 3	2 ¹⁶	8 / 1100	1 327, 1 283
Yu _p X-65537	1 word, position 3	2 ¹⁶	8 / 1100	1 328, 1 284
Yu ₂ X-16	2 words (0, 4) or (0, 1)	2 ³²	9 / 0000	10 223
Yu _p X-65537	2 words (0, 4) or (0, 1)	2 ³²	9 / 0000	10 225
Yu ₂ X-16	2 words, dims (16, 14)	2 ³⁰	9 / 1110	—
Yu ₂ X-16	full block	2 ⁶⁴	10 / 1110	(131 068, 131 068, 65 532, -4)
Yu _p X-65537	full block	2 ⁶⁴	10 / 1110	(131 072, 131 072, 65 536, 0 ^d)
Yu ₂ X-8	full block	2 ³²	6 / 1110	(508, 508, 252, -4)
Yu ₂ X-8	4 words (0, 4, 8, 12)	2 ³²	6 / 1100 ^b	-4, -4
Yu ₂ X-8	1 word	2 ⁸	4 / 1100	1 ^c
Yu _p X-65537	one order-2 ¹⁰ coset, weights $a \not\equiv 0$	2 ¹⁰	5 / —	—
Yu _p X-65537	one order-2 ¹² coset, weights $a \not\equiv 0$	2 ¹²	6 / —	—
Yu _p X-65537	one order-2 ¹⁴ coset, weights $a \not\equiv 0$	2 ¹⁴	7 / —	—
Yu _p X-65537	one order-2 ¹⁵ coset, weights $a \not\equiv 0$	2 ¹⁵	8 / —	—

Two margins are the positions 0 and 1 of the pattern 1100, four margins are the four positions. The coset cells hold for the weighted sums of Proposition 8. ^b Measured pattern, where the criterion gives 0000, a Frobenius boundary. ^c Margin at a position-3 start, still balanced. ^d A boundary cell, whose sum is a key-independent constant by Proposition 4.

$p = 257$ and layer 10 for $p = 65\,537$, and the constants of the four blocks are $(3, 0, 0, 2)$, $(1, 0, 4, 13)$ and $(86, 231, 222, 244)$ for the first three primes.

Proof. (i) The maximiser of each max-plus recurrence of equations (6) and (7) is unique except for y_1 of YuX, whose two maximal terms are both free of constants, and the top parts of the layer-2 inputs are the 0/1 linear forms of L^{-1} in y ; the claim follows by induction on ℓ . (ii) By the proof of Theorem 1 every monomial of degree below T sums to zero, and $\sum_{y \in \mathbb{F}_p^4} y^e \neq 0$ requires every e_i to be a positive multiple of $p - 1$, so at total degree T only $e = (p - 1, \dots, p - 1)$ survives, with sum 1. That coefficient lies in P , and it equals $\sum_{\mathbb{F}_p^4} P$ by the same count. (iii) is the recurrence and an evaluation of P , which is homogeneous of degree T , on the $p^3 + p^2 + p + 1$ points of $\mathbb{P}^3(\mathbb{F}_p)$. $\square$

Since these constants are non-zero, the boundary cell of Yu_pX-65537 at layer 10 gives a sum with a known right-hand side rather than a balanced word, and its value for $p = 65\,537$ has not been computed. In the encryption direction the boundary cells behave differently for the two families.

Proposition 5 (boundary cells in the encryption direction). *In the encryption direction the leading position is 0 for YuX and 3 for DuX, with $D = 8^{\ell-1}$, so the boundary $8^{\ell-1} = 4(p-1)$ is met only for $p = 3$ at layer 2, $p = 17$ at layer 3 and $p = 65\,537$ at layer 7. For YuX the sum over a full block at the boundary cell is 0 for every such p , and the layer-7 full block pattern of Yu_pX-65537 is 1111.*

Proof. The top form of $S = \text{Pf}^{-4}$ involves only the inputs at offsets 0 and 1 of its block, so every top form from the second layer on lies in the subring generated by the rows 0, 1, 4, 5, 8, 9, 12, 13 of the forward circulant restricted to $(y_0, \dots, y_3)$. Each of these rows has equal y_2 - and y_3 -coefficients, hence is of the form $ay_0 + by_1 + c\sigma$ with $\sigma = y_0 + y_1 + y_2 + y_3$,

and the eight rows have rank 3. The map $y \mapsto (y_0, y_1, \sigma)$ is onto $\mathbb{F}_p^3$ with fibres of size p , so the sum over $\mathbb{F}_p^4$ of any polynomial in these three forms is 0. $\square$

For DuX the corresponding eight rows have rank 4, so the argument yields nothing beyond the second layer, where each block involves only two linear forms. The proposition is used for YuX-65537 in Section D.

4.5 Frobenius Cells and the Next-to-Leading Coefficient

In the encryption direction the degree grows by a factor 8 per layer and crosses q quickly, so a criterion based on the unreduced degree soon becomes uninformative in the direction the designers consider. For one active plaintext word the reduction is governed by the classical identity $\sum_{X \in \mathbb{F}_q} f(X) = \varepsilon \sum_{k \geq 1} c_{k(q-1)}$, where ε is 1 or -1 according to the characteristic and c_e is the coefficient of X^e . When $D = q$ the sum has the single term $c_{q-1} = c_{D-1}$.

Theorem 2 (vanishing next-to-leading coefficient). *For DuX in the encryption direction with one active plaintext word at position 1, from the fourth S-box layer on the coefficient of X^{D-1} in every state word, D its max-plus degree, is identically zero, for every master key. Consequently a word with $D = q$ is balanced over the full field.*

The proof is given in Section B.2. It tracks the leading pair of coefficients through one S-box layer, where vanishing next-to-leading coefficients of the inputs force those of all four outputs to vanish and only the degree- $8m$ coordinate can reach the top after the linear layer, and it establishes the base case at layer 4 from the rotation set of L_0 , including the degenerate keys for which the leading coefficient itself vanishes. The consequence is one more chosen-plaintext layer for DuX(2^{16}) than for DuX(65537). At layer 6 position 3 has $D = 65\,536$ on both instances, and this value is q for $\mathbb{F}_{2^{16}}$, where the sum picks up the next-to-leading coefficient, which is zero, and $q - 1$ for $\mathbb{F}_{65537}$, where it picks up the leading coefficient, which is not.

5 Key Recovery from Zero Sums

A zero sum after ℓ layers gives equations in the round keys of the $r_{\text{KR}} = r - \ell$ rounds between the balanced state and the plaintext. In this section we first derive these equations for one and for two extension rounds, and then present the three techniques that determine how much data the equations require, that is, full block structures, cheap coordinate elimination and weighted moments.

5.1 One Extension Round

Lemma 1 (vanishing of the pure-key terms). *Let $|\mathcal{P}| \equiv 0$ in $\mathbb{F}_q$, which holds for a full product set in characteristic 2 and for a full prime field. In $\sum_{P \in \mathcal{P}} f(P+k)$ every monomial whose plaintext exponent vanishes carries the factor $|\mathcal{P}| = 0$. Hence in the equation $[L(\sum_{\mathcal{P}} \text{SL}(P+rk^0))]_i = 0$ of a balanced word i of x_1 the degree in rk^0 drops by one in every monomial. The coordinate-2 equation of a DuX block, $-k_0 \sum P_3 - k_3 \sum P_0 + \sum (P_2 - P_0 P_3) = 0$, is linear in (k_0, k_3) , and one-round key recovery is linear algebra.*

Which key words a given coordinate determines can be read off a small table, Table 5, and this table matters because a fully balanced layer makes some of the equations degenerate.

Proposition 6 (determinacy, DuX). *Write $W = S(P+k)$ for one block and k its four whitening words, and sum the four coordinates of W over a structure. Table 5 lists, for each coordinate, its degree, the way k_1 occurs in it and the whitening words it determines.*

Table 5: The four coordinates of one DuX block $W = S(P + k)$ after summation over a structure: their degree, how the whitening word k_1 occurs in them and which whitening words they determine, from Proposition 6.

Coordinate	Degree	How k_1 occurs	Words determined
2 (a)	2	not at all	k_0, k_3 , linearly
1 (b)	3	only as $ \mathcal{P} k_1$, killed by the sum	k_0, k_2, k_3
0 (c)	5	only through $-k_1 \sum_{\mathcal{P}} a(P + k)$	k_1 iff position 2 of W is <i>not</i> balanced
3 (y_3)	8	always	all four

In particular the degree-2 coordinate a is linear in (k_0, k_3) and free of k_1 , the degree-3 coordinate b determines k_0, k_2 and k_3 , the degree-5 coordinate c determines k_1 if and only if position 2 of W is not balanced, and the degree-8 coordinate determines all four words.

The rank of the assembled system has a closed form in terms of the template of the coordinates used and the span of the reachable moment vectors, proved in Section B. When W is fully balanced, coordinates 0, 1 and 2 together cannot determine k_1 , because the only place k_1 appears in coordinate 0 is multiplied by $\sum_{\mathcal{P}} a$. That sum is the coordinate-2 equation itself.

Either the degree-8 coordinate is linearised as well, with 39 unknowns, or the words are substituted sequentially, $2 \rightarrow (k_0, k_3) \rightarrow 1 \rightarrow k_2 \rightarrow 3 \rightarrow k_1$. The sequential route needs two or three structures because the coefficient of k_1^2 in coordinate 3 is again $\sum_{\mathcal{P}} a = 0$. For YuX the order is $3 \rightarrow (k_0, k_1) \rightarrow 2 \rightarrow k_3 - k_2 \rightarrow 1 \rightarrow k_2 \rightarrow 0$.

When the next layer has the pattern 1100, both coefficients of the position-3 equation vanish at once. This can be remedied by choosing a weight of Theorem 4 beyond the margin of the next layer, which makes the two sums non-zero again. A one-round extension from the layer-8 zero sum of YuX uses this remedy, and in characteristic 2 the usable weights are then measured rather than predicted, for the reason given in Observation 2.

The dense L^{-1} of YuX means that no word of $W = \text{SL}(P + rk^0)$ is balanced when only some positions of x_1 are. A partially balanced layer such as the layer-9 pattern 1100 is not usable backwards.

It is usable forwards, through *forward combined equations*: for each balanced word i ,

$$\left[L \left(\sum_{\mathcal{P}} W \right) \right]_i = 0,$$

one equation per balanced word, linear in the union of the four coordinates' monomials over the four blocks, that is $4 \cdot 80 = 320$ unknowns over $\mathbb{F}_p$ and 200 in characteristic 2, as counted in Section B. With the weights of Theorem 4 one 2^{16} structure gives $60 \times 8 = 480$ equations of rank 136 over $\mathbb{F}_p$ and 104 in characteristic 2, enough to pin the sixteen degree-one columns. This route recovers the master key of Yu_pX-65537 and Yu₂X-16 at ten rounds in seconds.

5.2 Two Extension Rounds

With $r_{\text{KR}} = 2$ the balanced state is $x_2 = L(W') + rk^2$, $W' = \text{SL}(x_1)$, $x_1 = L(Y) + rk^1$, $Y = \text{SL}(P + rk^0)$. Reading the cheap coordinate of outer block j of DuX, $W'_{4j+2} = z_2 - z_0 z_3 - \alpha$ with $z = (x_1)_{4j..4j+3}$, and summing over the structure,

$$\sum_s w_s T_s - \sum_{s,t} A_s B_t Q_{s,t} - k'_{4j+3} \sum_s A_s T_s - k'_{4j} \sum_s B_s T_s = 0, \quad (8)$$

where A, B and w are the rows $4j, 4j+3$ and $4j+2$ of L , $k' = rk^1$, $T_s = \sum_{\mathcal{P}} Y_s$ and $Q_{s,t} = \sum_{\mathcal{P}} Y_s Y_t$ are the *moments* of the structure. For YuX the rows are $4j, 4j+1$ and $4j+3$ minus $4j+2$.

The unknowns are the monomials of rk^0 inside T_s and $Q_{s,t}$, together with the products $k' \cdot \text{mon}(rk^0)$ from the last two terms. A moment $Q_{s,t}$ with s and t in different blocks is a product of two block polynomials, and this keeps M at 38 051 rather than a Boolean $\binom{256}{\leq 7}$.

Each structure gives four rows, one per outer block, or $\dim K$ rows with Lemma 3, and one such batch per weight with Theorem 4. The assembly is dominated by the block pair moments Q , computed as a handful of matrix products per slice as in Section 5.5, and the solve is one elimination over $\mathbb{F}_q$.

The rank of the system is determined almost entirely by the template.

Proposition 7 (template bound). *Every equation row is a fixed linear image $\Phi_j m_{\mathcal{P}}$ of the moment vector $m_{\mathcal{P}}$ of its structure, so $\text{rank}_{\max} \leq \text{rank}(\Phi|V_0)$, where V_0 is the span of the realisable signed measures of total weight zero. The right-hand side can be measured with pseudo-structures, random points with random signed weights summing to zero, that is, without a single encryption or decryption, and it gives $N_{\min} \leq \lceil \text{rank}(\Phi|V_0)/n_{\text{eq}} \rceil$ for the number of equation batches needed.*

The measured bounds are 4 766 in characteristic 2 and 8 618 over $\mathbb{F}_{65537}$ for DuX, and 4 950 and 8 618 for YuX, the same prime field value for the two ciphers despite a 42% larger monomial set. The attacks solve after moving the 11 normalisation columns, the constant column and the monomials with zero inner exponent, to the right-hand side, of which 10 carry rank, and the ceilings measured the same way are then 4 756 and 8 608. No real structure and no combined row fell short of its ceiling, as Table 13 in Section G shows.

The template bound predicted the number of weights an attack needs three times, namely $N_w = \lceil 4950/4 \rceil = 1\,238$ for Yu₂X-16 at ten and eleven rounds and $\lceil 8618/4 \rceil = 2\,155$ for Yu_pX-65537 at ten rounds, and the point set attacks of Sections 6.2 and 7.1 reach their ceilings with $N_w = \text{rank}_{\max}/4$ and not one weight more. We treat this as an empirical rule whose upper bound the proposition proves. The rule holds when the rows of successive weights stay independent until the system saturates. The extended system of Lemma 5 shows that this is not automatic. Its four combined rows per weight are not independent up to saturation, the sixteen key columns are determined only at 14 400 rows, 1.19 times its ceiling of 12 056, and before that point the system rests on a plateau of 1 122 determined monomials and no key word, which is the saturation value of the collapsed row of Corollary 2. The same plateau and the same crossing point appear on pseudo-structures over three fields and on the two real instances that use the system. Whether a system determines the key must therefore be judged from a rank curve rather than from a single point.

The cost of an attack is counted in field multiplications by one formula, the same for executed and for theoretical rows. Let the structures hold N points in all, streamed in n_{sl} slices along the weight axis, and let W be the width of the moment layout, that is, the number of point sums a structure keeps. The assembly costs

$$\text{passes} \cdot N \cdot W + N_w \cdot W \cdot n_{\text{sl}} + n_{\text{rows}} \cdot n_{\text{cols}} \quad (9)$$

multiplications, one term each for the per-point moments, the Vandermonde products of Section 5.5 and the row assembly, where the number of passes over the structure is one for a one-dimensional weight and one per inner index otherwise, and the elimination costs $n_{\text{rows}} \cdot n_{\text{cols}} \cdot \text{rank}$. The data is N and the memory is one slice table plus N_w moment vectors. The widths are $W = 56\,550$ for DuX over $\mathbb{F}_{65537}$ and $25\,200$ over $\mathbb{F}_{2^n}$, $65\,400$ for the extended system of Lemma 5, and $81\,360$ and $39\,942$ for YuX in the two characteristics. Section 8 gives the measured time per multiplication, which depends on the field and on the phase, and the tables of Sections 6 and 7 give the count of every attack.

5.3 Full Block Structures

Lemma 2 (full block). *Let a whole four-word block of the ciphertext range over $\mathbb{F}_q^4$, the other twelve words being constant. Since S^{-1} is a bijection of $\mathbb{F}_q^4$, the substitution $y = S^{-1}(x+k)$ turns the four outputs of the first layer on that block into four free variables. The first S-box layer is free on that block, the max-plus recurrence starts at degrees $(1, 1, 1, 1)$ there, and the threshold of Theorem 1 counts $s = 4$ full field variables. The same holds in the encryption direction with S in place of S^{-1} .*

Proof. Reparametrising the point set by y is a bijection of the structure onto $\mathbb{F}_q^4$, and every state word after the first layer is a polynomial in y whose formal degree is the value that the recurrence computes from $(1, 1, 1, 1)$. The sum over the structure is the sum over $y \in \mathbb{F}_q^4$, to which Theorem 1 applies with $s = 4$. $\square$

The restriction to a whole block is essential, since S^{-1} maps an affine subspace of dimension below 4 to a variety that is not affine. For DuX a full block reaches the same layer as four words in different blocks, namely layer 11 with pattern 0001. For YuX it is the only 2^{64} structure that reaches layer 10 with a usable pattern, 1110, by the search of Section F, and it makes the eight-round Yu₂X-8 attack possible.

5.4 Cheap Coordinate Elimination

The two-round extension reads the outer S-box through its cheap coordinate only. For DuX this is the coordinate $a = x_2 - x_0x_3 - \alpha$ at position 2, and for YuX it is $z_0 = x_3 - x_0x_1 - x_2 - \alpha$ at position 3. Both are bilinear, so their linearisation set stays at M columns. The other coordinates have degrees 3, 5 and 8. Over a prime field they would blow the system up, and in characteristic 2 the cubic coordinate of DuX is the one exception, the subject of Lemma 5 below.

The equation of outer block j needs the four words of W' at position 2 for DuX and position 3 for YuX, and the backward form $W' = L^{-1}(x_2 - rk^2)$ reads x_2 at the positions the rotation offsets reach, positions 2 and 3 for DuX by Proposition 2 and all four for YuX. If the layer- ℓ pattern has a zero at one of those positions the plain equations do not hold. The following lemma resolves this problem.

Lemma 3 (cheap coordinates). *Let $\mathcal{S}$ be the set of balanced positions of x_ℓ and $R = \{i : i \bmod 4 \in \mathcal{S}\}$ the corresponding rows of the forward linear layer L , and call a column of L expensive if its position is not the cheap one. Every vector of the left kernel*

$$K = \{y \in \mathbb{F}_q^{|R|} : y \cdot L[R, \text{expensive columns}] = 0\}$$

yields one linearised equation containing only cheap coordinates, $\sum_b (yL)_{4b+c} \sum_{\mathcal{P}} W'_{4b+c} = 0$ with c the cheap position, whose unknowns are the ones the system already has. Each structure yields $\dim K$ equations, and so does each weight under Theorem 4.

Proof. For $i \in R$, $\sum_{\mathcal{P}} (x_\ell)_i = \sum_{\mathcal{P}} [L(W')]_i + |\mathcal{P}| rk_i^\ell = \sum_s L_{i,s} \sum_{\mathcal{P}} W'_s = 0$. Combining these rows with the coefficients y kills every expensive column by definition of K , and what remains is the stated equation. Its unknowns are the linearisation of the cheap coordinate over the inner monomials, the same for every row. $\square$

The kernel dimensions of all fifteen non-trivial patterns, both ciphers, both directions and six fields are tabulated in Table 12 in Section G. The attacks use the following cells.

For DuX in the decryption direction the patterns 0001, 1001 and 1101 have $\dim K = 1$, spanned by $y = (1, -1, 1, -1)$ over $\mathbb{F}_p$ and by $(1, 1, 1, 1)$ over $\mathbb{F}_{2^n}$. The combined row is the signed sum of the four position-3 rows, $-W'_2 + W'_6 - W'_{10} + W'_{14} = 0$ over $\mathbb{F}_p$, drawn in Figure 4. For YuX the pattern 1110 has $\dim K = 3$, with kernel vectors $(1, 1, 0, 0)$, $(0, 1, 0, 1)$ and $(0, 1, 1, 0)$ on $(W'_3, W'_7, W'_{11}, W'_{15})$, whose rows are $W'_3 + W'_7$, $W'_7 + W'_{15}$ and

	inner block 0				inner block 1				inner block 2				inner block 3			
column s	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$y_0 = +1$, row 3	m_{13}	m_{14}	m_{15}	m_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}
$y_1 = -1$, row 7	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_0	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
$y_2 = +1$, row 11	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_0	m_1	m_2	m_3	m_4
$y_3 = -1$, row 15	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_0
$\sum_j y_j L[4j+3, \cdot]$	$-\Delta_1$ $=_0$	$-\Delta_2$ $=_0$	$-\Delta_3$ $=_0$	$+\Delta_0$ $=_0$	$+\Delta_1$ $=_0$	$+\Delta_2$ $=_0$	$+\Delta_3$ $=_0$	$-\Delta_0$ $=_0$	$-\Delta_1$ $=_0$	$-\Delta_2$ $=_0$	$-\Delta_3$ $=_0$	$+\Delta_0$ $=_0$	$+\Delta_1$ $=_0$	$+\Delta_2$ $=_0$	$+\Delta_3$ $=_0$	$-\Delta_0$ $=_0$

$\Delta_c = m_c - m_{c+4} + m_{c+8} - m_{c+12}$, and for the DuX layer $\Delta_0 = \Delta_1 = \Delta_2 = 0$ while $\Delta_3 \neq 0$.

Figure 4: Cheap coordinate elimination on the DuX pattern 0001 over $\mathbb{F}_p$: the four rows $i = 3, 7, 11, 15$ of the circulant L , with entry $m_{(s-i) \bmod 16}$ in column s , are block shifts of one another, so their combination with $y = (1, -1, 1, -1)$ has the alternating block sums Δ_c as entries. Every expensive column cancels and only the cheap columns, shaded at position 2 of each inner block, keep a coefficient $\pm\Delta_3$, which gives the row $-W'_2 + W'_6 - W'_{10} + W'_{14} = 0$ of Section 5.4 up to a common factor. Over $\mathbb{F}_{2^n}$ the only kernel vector is $(1, 1, 1, 1)$, the four surviving coefficients are equal, and only the orbit sums $k_c + k_{4+c} + k_{8+c} + k_{12+c}$ of the key columns remain, as Theorem 3 states.

$W'_7 + W'_{11}$ up to sign. In the encryption direction the patterns 0001 and 1110 of DuX and 0111 and 1110 of YuX have $\dim K = 4$.

The two-round extension of DuX therefore needs only position 3 balanced. The two-word 2^{32} structure with layer-10 pattern 1101 supports twelve rounds, and the 2^{64} structures with layer-11 pattern 0001 support thirteen. For YuX, whose dense L^{-1} makes “one position balanced” equivalent to “all sixteen words balanced” in the backward form, the requirement drops to the pattern 1110 that the full block delivers.

5.4.1 Collapse of the Combined Row in Characteristic 2

Lemma 3 counts equations rather than rank. In characteristic 2 the combined row of a one-dimensional kernel turns out to be useless, for a reason that comes from representation theory. Let ρ be the rotation of the inner blocks by one, acting on the moment vector of a structure, the vector of all sums $\sum_{\mathcal{P}} Y_s Y_t$ and $\sum_{\mathcal{P}} Y_s$ the equations are built from, and π the same rotation acting on the linearisation columns, inside each monomial as $(b, u) \mapsto (b+1, u)$ and on the outer key slot as $(j, i) \mapsto (j+1, i)$.

Lemma 4 (block equivariance). *Let Φ_j be the template of outer block j : the fixed linear map sending a structure’s moment vector to that block’s equation row. Then $\Phi_j = \pi^j \circ \Phi_0 \circ \rho^j$.*

Proof. The coefficient attached to inner word s in the equation of outer block j is the entry $L[(s - 4j - \text{off}) \bmod 16]$ of a fixed row of the circulant, for a fixed offset off determined by the cheap coordinate. Replacing (j, s) by $(j+1, s+4)$ leaves it unchanged. So the equation of block $j+1$ is that of block j with every inner block relabelled $b \mapsto b+1$ and the outer key slot relabelled $j \mapsto j+1$, which is the claimed conjugation. $\square$

Consequently the combined map for a kernel vector y is the twisted relative trace

$$\Psi_y = \sum_{j \in \mathbb{Z}_4} y_j \pi^j \circ \Phi_0 \circ \rho^j, \quad (10)$$

and relative traces behave differently according to whether the group order 4 is invertible.

Theorem 3 (characteristic-2 collapse). (i) Over $\mathbb{F}_p$ with p odd, the DuX kernel vector is the non-trivial character $y_j = \chi(j) = (-1)^j$ of $\mathbb{Z}_4$. Decompose the moment space into the isotypic components V^ψ on which ρ acts by the character ψ . Then

$$\Psi|_{V^\psi} = \left(\sum_j (\chi\psi)(j) \pi^j \right) \circ \Phi_0 = 4 P_{(\chi\psi)^{-1}} \circ \Phi_0,$$

4 times a projector onto one π -eigenspace, and as ψ ranges over the four characters all four eigenspaces are reached.

(ii) Over $\mathbb{F}_{2^n}$ the only kernel vector is $y = (1, 1, 1, 1)$ and $4 = 0$. Put $\tau = 1 + \pi$ and $\sigma = 1 + \rho$, both nilpotent of order 4. Then

$$\Psi = \sum_{\substack{a, b \leq 3 \\ a|b=3}} \tau^a \circ \Phi_0 \circ \sigma^b,$$

the sum over the pairs whose bitwise or is 3. Every term sits at depth $a + b \geq 3$ of the τ/σ filtration, and the only term with $b = 0$ is $\tau^3 \circ \Phi_0$. On the ρ -invariant part of the moment space the combined row is therefore $\tau^3 \Phi_0 = (\sum_j \pi^j) \Phi_0$, whose image consists of π -orbit sums: the sixteen degree-one key columns appear only through the four orbit sums $k_c + k_{4+c} + k_{8+c} + k_{12+c}$.

Proof. (i) is the expansion of equation (10) on V^ψ , where ρ^j acts as $\psi(j)$, together with the fact that $\frac{1}{4} \sum_j \phi(j) \pi^j$ is the projector onto the ϕ^{-1} -eigenspace of π for any character ϕ . For (ii) write $\pi^j = (1 + \tau)^j = \sum_a \binom{j}{a} \tau^a$ and $\rho^j = \sum_b \binom{j}{b} \sigma^b$. By Lucas' theorem $\binom{j}{a} \equiv [a \subseteq j] \pmod{2}$ for $j \leq 3$, so $\sum_{j=0}^3 \binom{j}{a} \binom{j}{b} \equiv \#\{j : a|b \subseteq j\} = 2^{2-\text{wt}(a|b)}$, which is odd exactly when $a|b = 3$. On a ρ -invariant vector $\sigma = 0$, leaving $\tau^3 \Phi_0$, and $\tau^3 = 1 + \pi + \pi^2 + \pi^3$ in characteristic 2, whose image on a length-four orbit is the orbit sum and on shorter orbits is zero. $\square$

Corollary 2 (usability criterion). $\mathbb{F}_2[\mathbb{Z}_4] = \mathbb{F}_2[\tau]/(\tau^4)$ is a local ring, so every kernel vector is τ^v times a unit. The combined row restricted to the ρ -invariant moments then has rank $\text{rank}(\tau^v) = \sum_{\text{orbits}} \max(0, \Lambda - v)$ over the π -orbits of length Λ . A combined row is usable in characteristic 2 only if some kernel vector has τ -valuation $v \leq 2$, that is, only if the kernel is not spanned by $(1, 1, 1, 1)$. DuX's 0001 kernel is spanned by $(1, 1, 1, 1) = \tau^3$, of valuation 3, leaving one coefficient per orbit. YuX's 1110 kernel vectors $(1, 1, 0, 0)$, $(0, 1, 0, 1)$ and $(0, 1, 1, 0)$ have valuations 1, 2 and 1. When the cheap set has two coordinates a kernel vector has one such component per coordinate, its valuation is the smaller of the two, and a kernel that contains a unit component is usable in full.

The proof is in Section B. Measured against the template bound of Proposition 7, that is, for every possible set of real structures, the combined system of DuX in characteristic 2 saturates at rank 2168 of 18025 columns after normalisation. It determines 1122 monomials and none of the sixteen key words, where the four plain rows reach 4756. Over $\mathbb{F}_p$ the same construction keeps 8188 of 8608 and determines all sixteen. This is the full-round DuX(65537) attack of Section 6 with these numbers. For YuX the three combined rows keep 3822 of 4940 in characteristic 2 and 98.4% over $\mathbb{F}_p$, in line with Corollary 2.

5.4.2 The Cubic Coordinate in Characteristic 2

As a result of the collapse, the two-word structures of DuX(2^{16}) and DuX(2^8) have no usable row in front of the pattern 1101, and a second cheap coordinate is needed to restore one. Since the squaring map is additive in characteristic 2, the cubic coordinate of the DuX S-box becomes a sum of products of two sums, as the following lemma states.

Lemma 5 (cubic coordinate). *Over $\mathbb{F}_{2^n}$ the coordinate $b = x_1 + x_3a + \alpha = x_1 + x_2x_3 + x_0x_3^2 + \alpha x_3 + \alpha$ of the DuX S-box, read on the outer block j with $x = (x_1)_{4j..4j+3}$ and summed over a structure, is linear in the moments T_s , $Q_{s,t}$ of equation (8) and in the two further families $Q_{s,t}^{(2)} = \sum_{\mathcal{P}} Y_s Y_t^2$ and $T_t^{(2)} = \sum_{\mathcal{P}} Y_t^2$, with unknowns the monomials of rk^0 inside these moments and the products of k'_{4j+3} , k'_{4j+2} , $(k'_{4j+3})^2$ and k'_{4j} with them. Together with the rows of the coordinate a the linearisation has $M_b = 47749$ columns. For the cheap set $\{1, 2\}$ the kernel of Lemma 3 for the pattern 1101 has dimension four and is generated under block rotation by the vector that combines rows 3, 1 and 9 of the balanced positions. Its combined equation reads $\sum_{\mathcal{P}} (W'_1 + W'_2 + W'_6 + W'_{10} + W'_{13}) = 0$, the four kernel vectors have valuations 2, 0, 0 and 2 in the sense of Corollary 2, and all four combined rows are usable.*

The expansion of the summed coordinate into its six terms, of which three are new, is written out in Section B.4. The term $x_0x_3^2$ is a product of two sums because $x_3^2 = \sum_t W_3[t]^2 Y_t^2 + (k'_3)^2$ is additive in characteristic 2, and every pure-key term still carries the factor $|\mathcal{P}| = 0$. Over a prime field the same coordinate is a genuine cubic, and there the one-dimensional kernel is a projector by Theorem 3 and needs no second coordinate. The moments $Q^{(2)}$ are sixteen further block pair products per slice, computed by the same kernel as Q with one factor squared, so the layout grows from $W = 25\,200$ to $65\,400$ and the assembly costs about 2.6 times more per point.

Measured against the template bound of Proposition 7, the extended system saturates at rank 12 056 with 5 812 monomials determined, the sixteen degree-one key columns among them, on pseudo-structures over $\mathbb{F}_{2^4}$, $\mathbb{F}_{2^8}$ and $\mathbb{F}_{2^{16}}$ alike. The bound is independent of n as it is for the original system, and the two real instances that use the rows, twelve-round DuX(2^{16}) and eight-round DuX(2^8) in Section 6, reach it with the same two numbers. The same construction for YuX, with the cheap set $\{2, 3\}$, raises the kernel of 1110 from three to four but opens no pattern that 1111 does not already open, since every new usable pattern needs position 3 balanced.

5.5 Weighted Moments

Theorem 4 (weighted moments). *In the setting of Theorem 1, for every $a \in \mathbb{N}^s$ with $|a| = \sum_i a_i < T - D$,*

$$\sum_{x \in U_1 \times \dots \times U_s} x^a Z(x) = 0.$$

The key-recovery equations of Section 5 remain valid with every sum $\sum_{\mathcal{P}} h(P)$ replaced by $\sum_{\mathcal{P}} x^a h(P)$, with the same unknowns. Consequently one structure yields one batch of equations per admissible weight.

Proof. Each monomial of $x^a Z(x)$ has exponent vector $a + e$ with $|e| \leq D$, so $|a + e| < T$ and the proof of Theorem 1 applies verbatim. The key-recovery equations are obtained from the identity $\sum_{\mathcal{P}} x_\ell = L(\sum_{\mathcal{P}} W') + |\mathcal{P}| rk^\ell$ and its expansion in the round keys. Every pure-key coefficient there is a sum $\sum_{\mathcal{P}} x^a = 0$ for $a \geq 1$ as it is for $a = 0$, whether over the full field, over a coset of order $2^k \nmid a$ or over an affine subspace, where it vanishes because $\text{wt}(a) < m$. The linearisation is therefore unchanged, and only the moments become weighted. $\square$

The sums of Theorem 4 are inner product masked integral sums in the sense of Wang, Tang and Wang [WTW25], whose Definition 2, in the full version, lets every point of the structure carry its own coefficient. Here the coefficient of the point x is the monomial value x^a , a mask fixed by the structure and independent of the key, and for a coset difference or the inclusion-exclusion of Theorem 1 the sign of the term is part of the mask. What the theorem adds is the admissible range $|a| < T - D$, read off the degree bound of Theorem 1,

and the fact that the key-recovery equations of Section 5 survive the masking with the same unknowns. The batches of different weights are not independent in general. Their joint rank is the subject of Propositions 6 and 7 and was measured in every run of Section 8.

Computing the weighted moments one weight at a time would require one pass over the structure per weight, which a *slice decomposition* avoids. Fix the weight axis x_j and write G_v for the ordinary moment vector of the slice $\{x_j = v\}$ of the structure. Then $M_a = \sum_v v^a G_v$, a Vandermonde matrix times the per-slice ordinary moments. All weights together cost one ordinary pass plus one matrix product, with no number-theoretic transform. This reduces the eleven-round attacks from thousands of structures to one.

The weight may be a multi-index on several active words. The slices are then the fibres of the weighted coordinates, the Vandermonde matrix is the tensor product of one matrix per axis, and a joint decomposition along the first m weighted words computes every $(a_0, \dots, a_{m-1})$ from one pass, while each further index costs one more pass with its monomial folded into one factor of the moment product. In this way, multi-index weights take the place of additional structures. A margin that admits n_1 one-dimensional weights admits about $n_1^2/2$ two-dimensional ones, and the rows of one structure under two-dimensional weights span the same space as the rows of several structures under one-dimensional weights, with the same ceiling and the same determined monomials, as every attack of Sections 6 and 7 that uses them shows. The data complexity decreases by the number of structures saved, while the time complexity remains the same.

Proposition 8 (single coset zero sums). *Over $\mathbb{F}_p$ let one active word range over a single coset aH of the subgroup H of order 2^k , and let Z have formal degree D . Then $\sum_{x \in aH} x^{a'} Z(x) = 0$ for every weight $0 < a' < 2^k - D$ with $2^k \nmid a'$. The plain sum $a' = 0$ is not zero in general.*

Proof. $\sum_{x \in aH} x^e = a^e \sum_{h \in H} h^e$ is non-zero exactly when $2^k \mid e$. For a monomial $x^{a'+e}$ with $0 \leq e \leq D$ and $0 < a' < 2^k - D$ we have $0 < a' + e < 2^k$, so it never vanishes as a multiple of 2^k unless $a' + e = 0$, which is excluded. For $a' = 0$ the constant term $e = 0$ survives with $\sum_{h \in H} 1 = 2^k \neq 0$. $\square$

This halves the data of the coset variants. The ten-round $\text{Yu}_p\text{X-65537}$ attack runs on one 2^{15} coset instead of the 2^{16} field, with the same number of weights, as Section 7 reports. Proposition 2 of [BCD⁺20] obtains zero sums over a subgroup by subtracting the constant term $|H|F(0)$. The non-zero weight removes it without any additional cost.

A coset may also be combined with full field words in one structure.

Proposition 9 (mixed structures). *Over $\mathbb{F}_p$ let the active words range over $U_1 \times \dots \times U_s$ with $U_i = \mathbb{F}_p$ for $i \in I$ and $U_i = g_i H_i$ a coset of the subgroup of order 2^{k_i} for $i \in J$, and let Z have formal degree D . Then $\sum_x x^a Z(x) = 0$ for every weight a with $|a| < T - D$, where*

$$T = |I|(p-1) + \sum_{i \in J} 2^{k_i},$$

provided that $a_i \geq 1$ and $2^{k_i} \nmid a_i$ for every $i \in J$. Without a weight the sum vanishes only below the threshold $|I|(p-1)$ of the full field words.

Proof. For a monomial x^{a+e} with $|e| \leq D$ the sum factorises over the axes. A full field factor $\sum_{x \in \mathbb{F}_p} x^{a_i+e_i}$ is non-zero only if $(p-1) \mid (a_i + e_i)$ and $a_i + e_i \geq 1$, hence only if $e_i \geq (p-1) - a_i$. A coset factor $\sum_{x \in g_i H_i} x^{a_i+e_i} = g_i^{a_i+e_i} \sum_{h \in H_i} h^{a_i+e_i}$ is non-zero only if $2^{k_i} \mid (a_i + e_i)$, which under $a_i \geq 1$ and $2^{k_i} \nmid a_i$ forces $a_i + e_i \geq 2^{k_i}$, hence $e_i \geq 2^{k_i} - a_i$. Summing, a surviving monomial has $|e| \geq T - |a|$, which contradicts $|e| \leq D < T - |a|$. Without a weight a coset factor with $e_i = 0$ equals $|H_i| = 2^{k_i} \neq 0$, so only the full field factors can vanish and the threshold drops to $|I|(p-1)$. $\square$

Proposition 8 is the case $I = \emptyset$, $s = 1$. The twelve-round attack on DuX(65537) in Section 6.1 uses one coset of order 2^{13} or 2^{14} times one full field, at 2^{29} or 2^{30} points instead of 2^{32} , and Section 8 reports the threshold as tight on the toy instance over $\mathbb{F}_{193}$, where the weighted sums vanish for every weight below $T - D$ and fail from $T - D$ on. Since $x^{2^{k_i}}$ is constant on the coset, two weights whose coset exponents agree modulo 2^{k_i} give proportional rows. The same holds on every axis: the rows of the exponents a and $a + m$ on an axis of m points are dependent, because $1, x, \dots, x^{m-1}$ already span the functions on it, so an axis contributes at most as many independent weights as it has points.

In fact, the active words need not fill a field, a subspace or a coset at all.

Proposition 10 (point sets). *Let $U_i \subseteq \mathbb{F}_q$ be any set of $n_i \geq 2$ points and $w_i(u) = \prod_{v \in U_i, v \neq u} (u - v)^{-1}$ its divided-difference weights. Then for a state word Z of formal degree D and every a with $|a| < T' - D$, where $T' = \sum_{i=1}^s (n_i - 1)$,*

$$\sum_{x \in U_1 \times \dots \times U_s} \left(\prod_{i=1}^s w_i(x_i) \right) x^a Z(x) = 0,$$

in both characteristics.

Proof. $\sum_{u \in U_i} w_i(u) u^e$ is the leading coefficient of the Lagrange interpolant of u^e on U_i , that is, the divided difference of order $n_i - 1$ of u^e , which is 0 for $e < n_i - 1$ and 1 for $e = n_i - 1$. A monomial x^{a+e} therefore survives only if $a_i + e_i \geq n_i - 1$ on every axis, hence only if $|e| \geq T' - |a| > D$. $\square$

The masked sum is an inner product masked integral sum of [WTW25] whose characterisation of the vanishing exponents is this T' . What we use here is the freedom to choose n_i as the smallest size whose margin $T' - D$ still admits the N_w weights a key recovery needs, so that the data becomes a quantity that can be chosen freely between two powers of q rather than a power of q . The eleven-round attacks on DuX and the ten-round attack on YuX-65537 in Sections 6.2 and 7.1 use point sets of 41 735 and 28 239 points in place of the field. The key-recovery equations survive the mask with the same unknowns, and the moments carry the weights w_i as a factor.

The admissible range of weights, $|a| < T - D$, must use the largest D among the state positions the rows read, which for a combined row of Lemma 3 are the balanced positions only, and five rules fix the range in the cases that occur:

- Full domain words: the range is $T - D$.
- Full block structures: the weights can only be put on the *ciphertext* words of the block, and after the substitution $y = S^{-1}(x + k)$ the word at position p has degree $\deg_p$ in y , namely (5, 3, 2, 8) for DuX and (8, 5, 3, 2) for YuX. A weight $\prod_p x_p^{a_p}$ costs $\sum_p a_p \deg_p$ degree units, and a one-dimensional weight on the cheapest word, position 2 for DuX and position 3 for YuX, ranges over the a with $2a < T - D$.
- Subspace words: the range is $(2^m - 1) - D$.
- Cosets: the range is $2^k - D$ with the multiples of 2^k excluded as in Proposition 8, and in a mixed structure T is the sum of Proposition 9.
- Point sets: the range is $T' - D$ of Proposition 10.
- Chosen-plaintext rows on a Frobenius cell: a word balanced by Theorem 2 has $D = q$, and nothing is known about the coefficient a weight would select. Such rows admit $a = 0$ only.

These rules give 126 usable one-dimensional weights for the eight-round Yu_2X -8 route, half its margin of 252, 29 for the thirteen-round $\text{DuX}(65537)$ route from its margin of 57, and 32 768 for the twelve-round Yu_pX -65537 route, half of 65 536. With multi-index weights the same two margins admit 5 376 vectors on the two cheapest words of Yu_2X -8 and 3 158 on the four words of $\text{DuX}(65537)$. The full block cost and the coset exclusion are the two rules that are most easily overlooked. The control experiment on the YuX toy instance in Section 8 shows what happens when a weight beyond the margin is used. The system can still be solved, but the solution is wrong and nothing in the computation indicates it.

6 Key Recovery on DuX

Table 6 lists the chosen-ciphertext key recoveries on DuX . We describe the full-round attacks on the two large instances first and the other rows after them.

Table 6: Master key recovery on DuX in the chosen-ciphertext model, by instance and number of attacked rounds.

Instance	Rounds	Data	Mult.	Time	Proc.	Memory
$\text{DuX}(65537)$	10/12	2^{16}	$2^{22.75}$	4.3 s	1	0.06 GB
$\text{DuX}(65537)$	11/12	2^{15}	$2^{44.15}$	1.6 h	1	20.8 GB
$\text{DuX}(65537)$	12/12	2^{29}	$2^{46.50}$	19.7 h	1	13.8 GB
$\text{DuX}(65537)$	12/12	2^{30}	$2^{46.03}$	1.1 h	19	19.6 GB
$\text{DuX}(65537)$	13 [†]	$2^{65.59}$	$2^{81.38}$	—	—	—
$\text{DuX}(2^{16})$	10/12	2^{16}	$2^{22.75}$	9.3 s	1	1.06 GB
$\text{DuX}(2^{16})$	11/12	$2^{15.35}$	$2^{40.59}$	2 039 s	1	6.21 GB
$\text{DuX}(2^{16})$	12/12	2^{32}	$2^{48.12}$	7.5 h	19	26.1 GB
$\text{DuX}(2^8)$	7/12	$2^{12.70}$	$2^{38.60}$	824 s	1	2.06 GB
$\text{DuX}(2^8)$	8/12	2^{24}	$2^{44.79}$	9.1 h	1	16.3 GB

Data is one structure per key except for $\text{DuX}(2^8)$ at seven rounds, which uses 26 structures. The two twelve-round rows of $\text{DuX}(65537)$ are the two mixed structures of Section 6.1, and the ten-round times are per key. Mult., Time, Proc. and Memory are as in Table 1, and [†] marks a theoretical row. The thirteen-round row holds over the prime field only, and nine rounds of $\text{DuX}(2^8)$ are out of reach, as Section 6.5 shows.

6.1 The Full-Round Attack on $\text{DuX}(65537)$

The attack uses one structure per key. In its simplest form, ciphertext words 3 and 7 range over $\mathbb{F}_{65537}$ and the other fourteen words are random constants, so the structure has $65\,537^2 = 2^{32}$ chosen ciphertexts, all decrypted through the full twelve rounds. The two mixed structures below reduce the data to 2^{30} and 2^{29} with the same equations.

After ten layers the pattern is 1101, as Table 3 shows. The margins at the four positions are 49 982, 20 301, $-20\,244$ and 60 846, so position 2 is not balanced while position 3 is balanced with a margin of 60 846. The four plain rows of equation (8) read position 2 of x_{10} and therefore do not hold.

The kernel of Lemma 3 for the pattern 0001 is one-dimensional, spanned by $y = (1, 65\,536, 1, 65\,536) \propto (1, -1, 1, -1)$, and it gives one row per weight that reads position 3 only. By Theorem 4 the admissible weights of that row are $a < 60\,846$ on the word x_3 , of which $N_w = 9\,000$ are used. That number is the ceiling 8 608 of Proposition 7 plus about 5%. Figure 5 shows the whole attack in this form. Its upper row is the cipher between the plaintext and the structure: the two extension rounds, read through the cheap coordinate $a = x_2 - x_0x_3 - \alpha$ of the outer S-box as in Section 5.2, stand in front of the ten inverse layers that lead from the ciphertexts to x_2 . Its lower row is the path from the

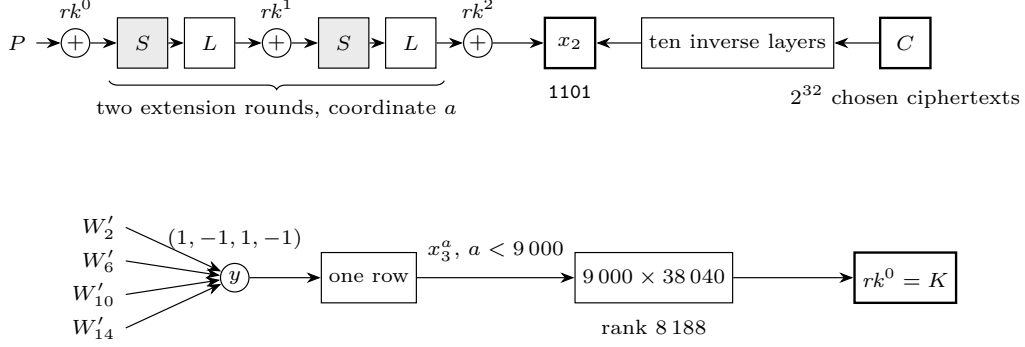


Figure 5: The twelve-round attack on DuX(65537) in its full field form. Above, the cipher between the plaintext and the structure of 2^{32} chosen ciphertexts, with the two extension rounds in front of the ten inverse layers. Below, the four position-3 rows of the blocks combined by the kernel vector y into one row, multiplied by the 9000 weights into the linear system, whose degree-one columns are the master key. Section 6.1 explains each step and the two mixed structures that reduce the data to 2^{30} and 2^{29} .

four position-3 rows of the blocks, through the kernel vector y and the 9000 weights, to the linear system whose degree-one columns are the master key.

The assembly streams the structure in slices $\{x_3 = v\}$ of 65 537 points. For each slice the sixteen plaintext words are decrypted, the $n_P = 75$ plaintext monomials per block are formed and the ten block pair moments Q are computed as one matrix product. The slice moments are then accumulated into the 9000 weighted moments by the Vandermonde product of Section 5.5, and the 9000 rows are formed from the weighted moments and combined with y .

The system to solve is 9000×38040 over $\mathbb{F}_{65537}$. Its columns are the 38051 monomials with the 11 normalisation columns moved to the right-hand side. Its rank is 8188, and 2592 monomials are determined uniquely. Among them are the sixteen degree-one columns rk_i^0 , and these equal the master key. The recovered key is checked by re-encryption.

The word x_3 need not range over the whole field. By Proposition 9 a coset of the subgroup of order 2^{14} on word 3 times the full field on word 7 has threshold $T = 2^{14} + 65536$, and the layer-10 pattern is then 1001 with a margin of 11 694 at position 3. The same 0001 row with the 9000 weights $a = 1, \dots, 9000$ on the coset word gives the same system, of the same rank 8188 with the same 2592 monomials determined, from $2^{14} \cdot 65537 = 2^{30}$ chosen ciphertexts. A coset of order 2^{13} leaves a margin of 3502 only, below the 8608 weights one axis would need, and a two-dimensional weight (a_0, a_1) with $a_0 \geq 1$ on the coset word and $a_1 \in \{0, 1, 2\}$ on the full field word restores the count. The three passes give $3501 + 3500 + 3499 = 10500$ rows from 2^{29} chosen ciphertexts, and the system has again rank 8188 with the same 2592 monomials determined. The three structures give the same linear system up to the choice of rows, and each recovers the sixteen key words.

The 2^{29} run on a single thread took 19.7 h at a peak of 13.8 GB, of which the three passes over the structure took 23 922, 23 150 and 21 212 s and the elimination 1 127 s. The 2^{30} run on 19 processes took 1.1 h and 11.5 core-hours at 19.6 GB. The structure with the smallest data therefore requires the most time, since the passes are repeated and the elimination does not become smaller, and Table 6 lists both runs.

6.2 Ten and Eleven Rounds

Eleven rounds start from the layer-9 zero sum of one active word with the four plain rows, and the ceilings are 8608 over $\mathbb{F}_{65537}$ and 4756 over $\mathbb{F}_{2^{16}}$. On $\text{DuX}(2^{16})$ a point set of $n = 41\,735$ values on word 3, with the divided-difference mask of Proposition 10, leaves the $N_w = 1\,189 = \text{rank}_{\max}/4$ weights the ceiling needs and not one more, and the 4756 rows have rank 4756 with 2188 monomials determined. The data is $2^{15.35}$ chosen ciphertexts, against 2^{16} for the full field, and a single-thread run takes 2039 s at 6.21 GB. On $\text{DuX}(65537)$ the layer-9 pattern of a single coset of order 2^{15} is 1101, so the plain rows do not hold there, and the 0001 row of Lemma 3 with the weights $a \in [1, 13\,951)$, $2^{15} \nmid a$, of Proposition 8 gives 9000 rows of rank 8188 with 2592 monomials determined. The data is 2^{15} chosen ciphertexts, and the run takes 1.6 h on one thread at 20.8 GB. Without weighted moments the full field attack on $\text{DuX}(65537)$ would need 2152 structures and $2^{27.07}$ chosen ciphertexts. In other words, one weight takes the place of one structure.

Ten rounds use the layer-9 zero sum with the one-round extension of Lemma 1. Two structures of 2^{16} suffice without weights, and one structure with eight weights, $a = 0, \dots, 7$, does the same from 2^{16} chosen ciphertexts in a few seconds per key. Over $\mathbb{F}_{2^{16}}$ the weighted sums go through logarithm tables and the single structure costs nine times the two, still under ten seconds per key.

6.3 Twelve Rounds of $\text{DuX}(2^{16})$

The two-word structure over $\mathbb{F}_{2^{16}}$ has the same layer-10 pattern 1101 as over the prime field, but its combined row for the cheap set $\{2\}$ is the all-ones kernel vector, of τ -valuation 3, and by Theorem 3 the system cannot separate the four key words of a rotation orbit. The measured ceiling is rank 2178, or 2168 after normalisation. That is 45.8% of the 4756 of the four plain rows and independent of n , and zero of the sixteen key words are determined.

The cheap set $\{1, 2\}$ of Lemma 5 replaces that row by four. The margins at positions 0, 1 and 3 of the layer-10 pattern are 49980, 20299 and 60844, so the weights on word 3 range below 20299, and $N_w = 3\,780$ of them, the crossing point 3600 of Section 5.2 plus 5%, give 15120 rows over the 47738 columns that remain after normalisation. The rank is 12056 and 5812 monomials are determined, the sixteen key words among them, the two numbers of the template bound of Section 5.4.2. The rows of the first 3600 weights already give them, and the rows of the first 3200 stop at rank 11768 with 1122 monomials and no key word, the plateau of Section 5.2. The data is one structure of 2^{32} chosen ciphertexts per key, as for $\text{DuX}(65537)$, and the moments $Q^{(2)}$ make the pass over the structure the dominant cost. The run took 7.5 h on 19 processes and 133.6 core-hours at a peak of 26.1 GB.

The route of the first version of this paper made all four positions balanced at layer 10 with three active words of subspace dimensions (16, 16, 15), for which $T = 163837 > D = 151316$ with D exact by Observation 1, at 2^{47} chosen ciphertexts and about 2^{62} multiplications. That structure was never assembled, and the cubic coordinate makes it unnecessary.

6.4 Thirteen Rounds over the Prime Field

A full block at 2^{64} has the layer-11 pattern 0001 with margin 57, as listed in Table 3. The combined row of Lemma 3 gives one equation per weight. The weights must sit on the ciphertext words of the block and cost 2, 3, 5 and 8 degree units per power on the four positions, which leaves 29 one-dimensional weights on the cheapest word or 3158 multi-index weights on all four. Recovering the key takes $\lceil 8608/3158 \rceil = 3$ structures, or $2^{65.59}$ chosen ciphertexts and $2^{81.38}$ multiplications, where one-dimensional weights would

need 308 structures and $2^{72.3}$, and it applies to a thirteen-round DuX over $\mathbb{F}_{65537}$, one round more than the cipher has.

Over $\mathbb{F}_{2^{16}}$ the same combined row is unusable by Corollary 2, the cheap set $\{1, 2\}$ leaves the pattern 0001 with a kernel of valuation 2 that determines none of the key words, and the row holds for the prime field only. None of this was executed. Nevertheless, in this model the twelve-round design has no round margin left.

6.5 DuX(2^8)

Seven rounds fall from $2^{12.70}$ chosen ciphertexts. The data consists of 26 single-word structures of 2^8 , used with the layer-5 zero sum and 46 weights each, and a single-thread run takes 824 s.

Eight rounds use three active words at 2^{24} , whose layer-6 pattern is 1101 with margins 347, 194, -15 and 403, the cheap set $\{1, 2\}$ of Lemma 5 and a two-dimensional weight. The first index a_0 sits on word 3 and the second, $a_1 \in \{0, \dots, 19\}$, on word 7, with $a_0 + a_1 < 194$, which gives 3 690 weights and 14 760 rows from one structure. The system has rank 12 056 with 5 812 monomials determined, the two numbers of the twelve-round DuX(2^{16}) attack, and recovers all sixteen key words from 2^{24} chosen ciphertexts. A single-thread run takes 9.1 h at 16.3 GB, of which the twenty passes take 29 182 s and the elimination 3 641 s.

With one-dimensional weights the same structure gives 776 rows and 17 structures are needed, or $2^{28.09}$ chosen ciphertexts. Twelve to sixteen structures stop at 1 122 determined monomials and no key word, and the seventeenth, at 13 192 rows, crosses to 12 056 and 5 812, on two keys with the same curve. Four active words at 2^{32} with the layer-6 zero sum, the four plain rows and 240 weights need five structures for the ceiling 4 756 of the four plain rows, or $2^{34.32}$, and the subspace structure of dimensions (8, 8, 8, 5) at 2^{29} reaches that ceiling from one structure with a weight of four indices and 1 445 vectors. The attack with one structure at 2^{24} is the cheapest of the four.

Nine rounds are out of reach. The only pattern any structure achieves at layer 7 is 0001, cheapest at 2^{47} with a full block plus two words of dimensions (8, 7) and margin 179, and that pattern is unusable in characteristic 2 by Corollary 2, with either cheap set.

7 Key Recovery on YuX

Table 7 lists the chosen-ciphertext key recoveries on YuX. We describe ten to twelve rounds of the two large instances first and then Yu₂X-8.

7.1 Ten Rounds

Ten rounds start from the layer-9 pattern 1100 at 2^{16} of Table 4, and three routes reach the key. The first uses the forward combined equations of Section 5.1 with 60 weights, which give 480 equations from one structure and recover the key in 115 s per key on Yu_pX-65537 and 55 s on Yu₂X-16, on one thread.

The second is the two-round extension from the layer-8 zero sum. On Yu_pX-65537 a point set of $n = 28\,239$ values on word 0 with the mask of Proposition 10 leaves the $N_w = 2\,155 = \lceil 8\,618/4 \rceil$ weights the ceiling needs, and the 8 620 rows have rank 8 608, the normalised ceiling, with 2 088 monomials determined. The data is $2^{14.79}$ chosen ciphertexts, and a single-thread run takes 1 706 s at 16.7 GB. The full field gives the same attack from 2^{16} with rank 8 576, and on Yu₂X-16 the full field with $N_w = 1\,238 = \lceil 4\,950/4 \rceil$ gives rank 4 916 in minutes.

The third is the same two-round extension on a single coset of the order- 2^{15} subgroup of $\mathbb{F}_{65537}^\times$, as Proposition 8 allows. It uses 2^{15} chosen ciphertexts per key, weights $a \in [1, 6\,685]$ with $2^{15} \nmid a$ and the same $N_w = 2\,155$, and a single-thread run takes 1.4 h at 22.4 GB. The

Table 7: Master key recovery on YuX in the chosen-ciphertext model, by instance and number of attacked rounds.

Instance	Rounds	Data	Mult.	Time	Proc.	Memory
Yu _p X-65537	10/14	2^{16}	$2^{30.25}$	115 s	1	0.10 GB
Yu _p X-65537	10/14	$2^{14.79}$	$2^{43.02}$	1 706 s	1	16.7 GB
Yu _p X-65537	10/14	2^{15}	$2^{43.30}$	1.4 h	1	22.4 GB
Yu _p X-65537	11/14	2^{32}	$2^{48.38}$	4.6 h	24	19.4 GB
Yu _p X-65537	12/14 [†]	2^{64}	$2^{80.31}$	—	—	1.87 GB
Yu ₂ X-16	10/14	2^{16}	$2^{29.57}$	55 s	1	0.09 GB
Yu ₂ X-16	11/14	2^{32}	$2^{47.32}$	4.1 h	36	5.48 GB
Yu ₂ X-16	12/14 [†]	2^{64}	$2^{79.29}$	—	—	0.53 GB
Yu ₂ X-8	7/12	$2^{17.58}$	$2^{42.45}$	2.8 h	1	6.68 GB
Yu ₂ X-8	8/12	2^{32}	$2^{47.32}$	3.6 h	19	4.56 GB

The three ten-round Yu_pX-65537 rows are the three routes of Section 7.1. Data is one structure per key except for Yu₂X-8 at seven rounds, which uses three structures. Mult., Time, Proc. and Memory are as in Table 1, and [†] marks a theoretical row. Ten rounds of Yu₂X-8 are beyond the routes of this paper, and the attack of [NWL26] at $2^{95.5}$ data and $2^{115.3}$ time stands.

coset structure halves the data, and it also reaches the normalised ceiling 8 608 where the full field single-word structure stops at 8 576, as the point set does. Removing $a = 0$ from the full field run does not change its rank, so the missing 32 dimensions belong to the full field single-word structure and not to the weight set.

7.2 Eleven Rounds

Eleven rounds use two active words (0, 4) at 2^{32} , whose layer-9 zero sum has thinnest margin 10 225 on Yu_pX-65537 and 10 223 on Yu₂X-16, the two-round extension and the four plain rows. On Yu_pX-65537, $N_w = 2\,400$ weights give 9 600 equations from one structure and rank 8 608. A run takes 4.6 h on 24 processes, 92.9 core-hours, at 19.4 GB peak.

On Yu₂X-16 the arithmetic is over $\mathbb{F}_{2^{16}}$, with $N_w = 1\,300$ weights against the minimum 1 238 of the template bound. The resulting $5\,200 \times 27\,884$ system has rank 4 940, the normalised ceiling, and 1 606 monomials are determined. One two-word structure therefore spans V_0 exactly. A run takes 4.1 h on 36 processes, 144.4 core-hours, at 5.48 GB.

7.3 Twelve Rounds

Twelve rounds need layer 10, and only the full block reaches it. Its pattern is 1110 at 2^{64} , with margins (131 072, 131 072, 65 536, 0) for Yu_pX-65537 and (131 068, 131 068, 65 532, -4) for Yu₂X-16.

The three-dimensional kernel of Lemma 3 gives three equations per weight. By Corollary 2 its vectors have valuations 1, 2 and 1, so the rows are usable in characteristic 2, with ceiling 3 832 or 77.4%. The halved full block weight budget leaves 32 768 weights against the $N_w = \lceil 8\,618/3 \rceil = 2\,873$ required on Yu_pX-65537 and $\lceil 4\,950/3 \rceil = 1\,650$ on Yu₂X-16. One structure therefore suffices.

The bottleneck lies in the assembly. By equation (9) the attack costs $2^{80.31}$ multiplications on Yu_pX-65537 and $2^{79.29}$ on Yu₂X-16, with 1.87 GB of memory for Yu_pX-65537 and 0.53 GB for Yu₂X-16, and at the $43\,\mu\text{s}$ per point of the prime field pass measured in Section 8 the assembly on Yu_pX-65537 would take about $2.6 \cdot 10^7$ core-years. Thirteen rounds are out of reach, because no admissible pattern exists at layer 11 for any structure with at most one free block and any set of words, by the search of Section F.

7.4 Yu₂X-8

Seven rounds fall from $2^{17.58}$ chosen ciphertexts. The attack uses two active words at 2^{16} and the layer-5 zero sum with margin 34, the four plain rows and a two-dimensional weight on the two words, whose $\binom{35}{2} = 595$ vectors give 2380 rows per structure. Two structures give 4760 rows of rank 4760 and no key word, three reach the ceiling 4940 with 1606 monomials determined, and a fourth adds nothing, so three is the minimum. With one-dimensional weights the margin admits 34 per structure and 40 structures are needed, or $2^{21.32}$. A single-thread run of the attack with three structures takes 2.8 h at 6.68 GB.

Eight rounds are the one key recovery in this paper that uses a full block. At 2^{32} the layer-6 pattern is 1110 with margin 252, and the three combined rows give three equations per weight. A two-dimensional weight on the two cheapest ciphertext words of the block, positions 3 and 2 at costs 2 and 3, admits the 5376 vectors with $2a_0 + 3a_1 < 252$, of which the 1400 cheapest are used, all computed in one pass by the joint decomposition of Section 5.5. The 4200 rows have rank 3822, the normalised ceiling, with 1296 monomials determined, and one structure recovers the key. With the 126 one-dimensional weights each structure gives 378 equations and eleven structures are needed, or $2^{35.46}$, where nine and ten stop at ranks 3402 and 3780, equal to their row counts. The run with one structure took 3.6 h on 19 processes, 67.2 core-hours, at 4.56 GB.

The ten-round attack of [NWL26], a Boolean linearisation over four extension rounds at $2^{95.5}$ data and $2^{115.3}$ time, remains the round record on this instance. Our route stops at eight rounds because layer 7 is unreachable, as Section F shows.

8 Experimental Verification

Every key recovery in Tables 1, 6 and 7 without a dagger was executed end to end on the real instance with its real key schedule and recovered all sixteen words of the master key. The three rows marked [†] are thirteen-round DuX(65537), twelve-round Yu_pX-65537 and twelve-round Yu₂X-16, for which every component was executed but the assembly of the equations over the full structure of 2^{64} points was not. The twelve-round attacks on DuX(65537) and DuX(2^{16}) are executed attacks on full parameter sets. All other executed rows concern reduced-round versions.

Table 8 lists, for every executed attack, the one run whose figures the tables report and the number of independent random master keys the attack was repeated on. The ten-round attacks on DuX, which take seconds per key, were run on fifty keys each. The attacks of the bold rows of Table 1 were repeated on three keys and the remaining ones on two, and on every key of an attack the rank, the number of determined monomials and the saturation curve over structures or weights were the same. Data is counted per key, and a row with one structure per key has data equal to the structure size.

Distinguishers. On the real instances we measured by direct summation, in the decryption direction, the single-word and two-word structures of the two large DuX instances and the coset differences of DuX(65537), the one- to four-word structures of DuX(2^8), the single-word and two-word structures of Yu₂X-16 and Yu_pX-65537, the single-word, four-word and full block structures of Yu₂X-8, and the single coset sums of Yu_pX-65537. In the encryption direction we measured the single-word structures of both large DuX instances, the two-word structures of DuX(2^{16}) and the six-round examples of the design paper. Each cell was measured on three independent random master keys, except the single-word cells of DuX(65537) at positions 2 and 3 on five and the single coset cells and the six-round examples on two, and the exact degrees of Observations 1 and 2 were computed on three keys. The constants of Proposition 4 were checked on three real master keys with two sets of inactive constants each for $p = 5$ and $p = 17$, and by direct summation over $\mathbb{F}_{257}^4$ for

Table 8: The run reported for every executed attack: the number of processes, the elapsed time, the processor time, the peak resident memory, the rank of the linear system and the number of independent random master keys the attack was repeated on.

Attack	Model	Data	Proc.	Elapsed (s)	CPU (h)	Memory (GB)	Rank	Keys
DuX(65537) 10/12	CC	2^{16}	1	215	0.060	0.06	—	50
DuX(65537) 11/12	CC	2^{15}	1	5 597	1.555	20.76	8 188	3
DuX(65537) 12/12	CC	2^{29}	1	70 753	19.523	13.84	8 188	3
DuX(65537) 12/12	CC	2^{30}	19	4 069	11.510	19.63	8 188	3
DuX(2^{16}) 10/12	CC	2^{16}	1	464	0.129	1.06	—	50
DuX(2^{16}) 11/12	CC	$2^{15.35}$	1	2 039	0.566	6.21	4 756	2
DuX(2^{16}) 12/12	CC	2^{32}	19	26 840	133.631	26.13	12 056	3
DuX(2^8) 7/12	CC	$2^{12.70}$	1	824	0.229	2.06	4 756	2
DuX(2^8) 8/12	CC	2^{24}	1	32 834	8.484	16.33	12 056	3
Yu _p X-65537 10/14 (i)	CC	2^{16}	1	115	0.032	0.10	136	3
Yu _p X-65537 10/14 (ii)	CC	$2^{14.79}$	1	1 706	0.474	16.67	8 608	3
Yu _p X-65537 10/14 (iii)	CC	2^{15}	1	4 988	1.337	22.40	8 608	2
Yu _p X-65537 11/14	CC	2^{32}	24	16 427	92.9	19.4	8 608	3
Yu ₂ X-16 10/14	CC	2^{16}	1	55	0.015	0.09	104	3
Yu ₂ X-16 11/14	CC	2^{32}	36	14 636	144.4	5.48	4 940	3
Yu ₂ X-8 7/12	CC	$2^{17.58}$	1	9 947	2.763	6.68	4 940	3
Yu ₂ X-8 8/12	CC	2^{32}	19	13 108	67.172	4.56	3 822	3
DuX(65537) 7/12	CP	$2^{18.8}$	1	42	0.012	0.06	44	2
DuX(2^{16}) 7/12	CP	$2^{18.6}$	1	2.8	0.001	0.06	40	2
DuX(65537) 8/12	CP	2^{16}	1	125	0.035	3.73	1 234	2
DuX(2^{16}) 8/12	CP	2^{16}	1	1 190	0.331	1.94	1 086	2
DuX(2^{16}) 8/12	CP	$2^{23.38}$	1	774	0.215	0.15	1 336	3

CC: chosen ciphertext, CP: chosen plaintext. Proc. is the number of processes, where 1 is a single thread pinned to one reserved physical core. Elapsed is the wall-clock time and CPU the user plus system time of the whole process tree in core-hours, both from the operating system, and Memory the peak resident size. The ten-round DuX rows report one run over all fifty keys and their elapsed time is the total. The ten-round Yu_pX-65537 routes are those of Section 7.1.

$p = 257$.

The remaining eight cells of Tables 3 and 4 rest on Theorem 1 alone. They are the three-word, (16, 16, 15) and four-word cells of the large DuX instances, the (8, 8, 8, 5), full block and eleven-word cells of DuX(2^8), and the (16, 14) and full block cells of YuX.

The criterion of Theorem 1 reproduced every measured cell of our own grid with one exception. The grid has 39 cells, all but two on the real instances, and 21 more on the toy instances with thin margins, -12 , 5 and 22 at the thinnest, together with 17 YuX cells on three keys each. The exception is the four-word layer-6 cell of Yu₂X-8, which is balanced at positions 0 and 1 where the criterion gives 0000, a Frobenius boundary at margin -4 . The threshold of Proposition 9 is tight on the toy instance over $\mathbb{F}_{193}$ with one coset of order 2^6 times the full field: at layer 5, where $D = 209$ against $T = 256$, the weighted sums vanish for every weight below 47 and fail from 47 on, on three keys.

The multi-word Frobenius cells, namely the two-word layer-7 pattern 0010 of DuX(2^{16}), the four-word layer-4 zero sum of DuX(2^8) and that Yu₂X-8 cell, are supported by measurement on three keys only, while the single-word cells are covered by Theorem 2 for all keys. The pattern of each attack's structure was checked on the instance before the attack was run.

Machine and Measurement. All runs used one server with two Xeon Gold 6230R processors, 52 physical cores and 104 hardware threads at a base frequency of 2.10 GHz and a turbo frequency of up to 4.00 GHz, and 251 GB of memory, shared with other jobs. Every run reported in the tables was pinned to physical cores reserved for it, with the sibling hardware threads left idle. An attack that one thread completes within 36 hours is reported from a single-thread run: one process on one core, every BLAS and OpenMP

thread pool set to one thread and the elimination single-threaded, the number of threads sampled every ten seconds and never above one, and the ratio of elapsed time to processor time at most 1.1, a run that exceeded it being discarded and repeated. An attack beyond that limit that consists of one structure is reported from a run on P single-threaded worker processes plus a P -threaded elimination, completed in one piece without a restart, with P stated next to its elapsed time. Processor time is the user plus system time of the whole process tree as reported by the operating system, and memory is the peak resident size of the tree, which for a single-thread run counts the shared slice tables once and can exceed the figure of a multi-process run of the same attack.

A reserved core does not run at a fixed speed. The same single-thread run, same command and same key, took 1 222 s with 51 cores of the machine active and 824 s with 27, at a ratio of elapsed to processor time of 1.000 both times, because the all-core turbo frequency falls with the number of active cores. The single-thread times of the tables were therefore all measured in one window in which three to eight cores were in use for them, and comparisons with other machines rest on the multiplication counts of equation (9) rather than on times. Processes beyond the number of cores cost processor time without saving elapsed time. For instance, the 2^{30} attack of Section 6.1 took 1.1 h on 19 processes and 11.5 core-hours, and 1.2 h on 32 processes and 20.8 core-hours, and the eleven-round Yu₂X-16 attack cost 245.5 and 207.0 core-hours in two runs on 72 processes against 144.4 on 36.

The time per field multiplication of equation (9), measured on the single-thread runs, is 1.3 to 1.7 ns for the moment assembly in characteristic 2 and 0.14 to 0.73 ns over $\mathbb{F}_{65537}$, where the weighted accumulation is one matrix product against a gather through logarithm tables, and 0.40 to 0.48 ns and 0.23 to 0.35 ns for the elimination in the two characteristics. Per point of a structure, decryption and moments included, the passes cost 43 μ s over $\mathbb{F}_{65537}$ and 87 μ s over $\mathbb{F}_{2^8}$ with the wider layout of Lemma 5. The five multiplication counts of the theoretical rows were the same, to their stated precision, when recomputed with equation (9) from the single term of the pass over the structure.

Controls. On a truncated structure, a proper subset of the 2^{32} points, the assembly detects the truncation. Coefficients appear outside the predicted monomial set, because the monomials dropped by the linearisation are those whose coefficients vanish by the structure identities.

The toy instances over $\mathbb{F}_{193}$, $\mathbb{F}_{257}$ and $\mathbb{F}_{24}$ serve to measure tightness and to test the attacks at small scale, and they are not attacks on DuX or YuX. The twelve-round attack of Section 6.1 has an exact scaled-down analogue on the toy instance of DuX over $\mathbb{F}_{257}$, with the same S-box and linear layers, $\alpha = 179$ and twelve rounds. It attacks eight rounds with two active ciphertext words at $2^{16.01}$ per structure, the layer-6 pattern 1101, the 0001 combined row of Lemma 3 and one-dimensional weights $a < 150$ on word 3, and 60 structures with 150 weights each give 9 000 rows over 38 040 columns. The system has rank 8 188 with 2 592 monomials determined and all sixteen inner key words correct, the same rank and the same number of determined monomials as in the real attack and equal to the pseudo-structure ceiling of the combined row. On the toy instance over $\mathbb{F}_{193}$ one coset of order 2^6 times the full field gives, at 8 648 rows, the same rank 8 608 and the same 2 592 determined monomials as the full field times the full field. On the toy instance over $\mathbb{F}_{24}$ the rows of Lemma 5 recover the key from a single-word structure at the third layer, where the cheap set $\{2\}$ gives no usable row, with the rank 12 056 and the 5 812 determined monomials of the real attacks, and a two-dimensional weight on one structure reaches the ceiling 4 756 of the four plain rows at 4 756 rows where one-dimensional weights on 85 structures need 5 040.

On the YuX toy over $\mathbb{F}_{24}$, a weight beyond the admissible margin produces a system that solves to sixteen “determined” key words of which zero or one is correct, with no

visible sign of failure.

The identity of Lemma 4 and the eigenvalue statements of Theorem 3 were checked on random moment vectors over $\mathbb{F}_{2^4}$, the eigenvalue $+1$ case with orbit sums, and over $\mathbb{F}_{2^{57}}$, the eigenvalue -1 case. The identity of Proposition 1 was verified numerically on all six instances, and the rank formula of Section B matched the measured rank in all six cells tested.

Implementation. The implementation is written in Python with NumPy, with the inner loops of the moment assembly and the characteristic-2 elimination in C. All six instances and the toy instances share one code base, parametrised by the field, the S-box and the linear layer. The criterion, the kernel dimensions, the template bounds, the weight ranges and the multiplication counts of the tables are computed by the same code that runs the attacks. Every attack was run from a fresh random master key, and the recovered key was verified by re-encryption.

9 Conclusion

In the chosen-ciphertext model, the full twelve-round DuX over both large fields falls to a key-recovery attack that was executed on random keys, from 2^{29} chosen ciphertexts over $\mathbb{F}_{65537}$ and 2^{32} over $\mathbb{F}_{2^{16}}$, and eleven of fourteen rounds of Yu_pX-65537 and Yu₂X-16 as well as eight of twelve rounds of Yu₂X-8 and DuX(2^8) fall far below their claims. These results stem from three properties that the two FHE-friendly block cipher families share. The coordinate degrees of the decryption S-boxes, $(2, 3, 4, 2)$ and $(2, 2, 3, 4)$, are far below those of the encryption S-boxes, $(5, 3, 2, 8)$ and $(8, 5, 3, 2)$, which the designers evaluate. The linear layers have a low-rank part, described in Propositions 2 and 3, which slows the degree growth of DuX to $2 + \sqrt{3}$ per layer and makes the cheap coordinate kernel non-trivial in both families. No key material is added between the S-box layer and the linear layer within a round, so a sum over a structure removes every pure-key term, as in Lemma 1, and the two-round linearisation stays at about 2^{15} columns. We do not develop design recommendations here.

Several questions remain open. The full field single-word YuX structures fall short of the normalised rank ceiling by 24 dimensions in characteristic 2 and by 32 over the prime field, while every other structure reaches it, and we do not have an explanation for this gap. The prediction of N_w by the template bound is an empirical rule that holds only for systems whose rows stay independent until saturation, as the extended system of Lemma 5 shows, and the three theoretical rows await an assembly at full scale. Finally, the equality of the max-plus bound with the exact degree over prime fields rests on computations for three keys and is not a theorem for all keys.

References

- [AGR⁺16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 191–219. Springer, Berlin, Heidelberg, December 2016. doi:10.1007/978-3-662-53887-6_7.
- [AM09] Jean-Philippe Aumasson and Willi Meier. Zero-sum distinguishers for reduced Keccak- f and for the core functions of Luffa and Hamsi. Comment submitted to the NIST hash function competition, four-page note, 2009. URL: <https://www.aumasson.jp/data/papers/AM09.pdf>.

- [AMT22] Tomer Ashur, Mohammad Mahzoun, and Dilara Toprakhisar. Chaghri - A FHE-friendly block cipher. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *ACM CCS 2022*, pages 139–150. ACM Press, November 2022. doi:10.1145/3548606.3559364.
- [ARS⁺15] Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 430–454. Springer, Berlin, Heidelberg, April 2015. doi:10.1007/978-3-662-46800-5_17.
- [BC11] Christina Boura and Anne Canteaut. Zero-sum distinguishers for iterated permutations and application to Keccak-f and Hamsi-256. In Alex Biryukov, Guang Gong, and Douglas R. Stinson, editors, *SAC 2010*, volume 6544 of *LNCS*, pages 1–17. Springer, Berlin, Heidelberg, August 2011. doi:10.1007/978-3-642-19574-7_1.
- [BC13] Christina Boura and Anne Canteaut. On the influence of the algebraic degree of F^{-1} on the algebraic degree of $G \circ F$. *IEEE Transactions on Information Theory*, 59(1):691–702, 2013. doi:10.1109/TIT.2012.2214203.
- [BCD11] Christina Boura, Anne Canteaut, and Christophe De Cannière. Higher-order differential properties of Keccak and Luffa. In Antoine Joux, editor, *FSE 2011*, volume 6733 of *LNCS*, pages 252–269. Springer, Berlin, Heidelberg, February 2011. doi:10.1007/978-3-642-21702-9_15.
- [BCD⁺20] Tim Beyne, Anne Canteaut, Itai Dinur, Maria Eichlseder, Gregor Leander, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Yu Sasaki, Yosuke Todo, and Friedrich Wiemer. Out of oddity - new cryptanalytic techniques against symmetric primitives optimized for integrity proof systems. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part III*, volume 12172 of *LNCS*, pages 299–328. Springer, Cham, August 2020. doi:10.1007/978-3-030-56877-1_11.
- [BCP23] Clémence Bouvier, Anne Canteaut, and Léo Perrin. On the algebraic degree of iterated power functions. *DCC*, 91(3):997–1033, 2023. doi:10.1007/s10623-022-01136-x.
- [BV24] Tim Beyne and Michiel Verbauwhe. Ultrametric integral cryptanalysis. In Kai-Min Chung and Yu Sasaki, editors, *ASIACRYPT 2024, Part VII*, volume 15490 of *LNCS*, pages 392–423. Springer, Singapore, December 2024. doi:10.1007/978-981-96-0941-3_13.
- [BV25] Tim Beyne and Michiel Verbauwhe. Integral cryptanalysis in characteristic p . In Goichiro Hanaoka and Bo-Yin Yang, editors, *ASIACRYPT 2025, Part I*, volume 16245 of *LNCS*, pages 66–96. Springer, Singapore, December 2025. doi:10.1007/978-981-95-5018-0_3.
- [CCF⁺18] Anne Canteaut, Sergiu Carpov, Caroline Fontaine, Tancrede Lepoint, María Naya-Plasencia, Pascal Paillier, and Renaud Sirdey. Stream ciphers: A practical solution for efficient homomorphic-ciphertext compression. *Journal of Cryptology*, 31(3):885–916, July 2018. doi:10.1007/s00145-017-9273-9.
- [CGG⁺22] Carlos Cid, Lorenzo Grassi, Aldo Gunesing, Reinhard Lüttenegger, Christian Rechberger, and Markus Schofnegger. Influence of the linear layer on the algebraic degree in SP-networks. *IACR Trans. Symm. Cryptol.*, 2022(1):110–137, 2022. doi:10.46586/tosc.v2022.i1.110-137.

- [CHK⁺21] Jihoon Cho, Jincheol Ha, Seongkwang Kim, ByeongHak Lee, Joohee Lee, Jooyoung Lee, Dukjae Moon, and Hyojin Yoon. Transciphering framework for approximate homomorphic encryption. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT 2021, Part III*, volume 13092 of *LNCS*, pages 640–669. Springer, Cham, December 2021. doi:[10.1007/978-3-030-92078-4_22](https://doi.org/10.1007/978-3-030-92078-4_22).
- [CHWW22] Jiamin Cui, Kai Hu, Meiqin Wang, and Puwen Wei. On the field-based division property: Applications to MiMC, Feistel MiMC and GMiMC. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part III*, volume 13793 of *LNCS*, pages 241–270. Springer, Cham, December 2022. doi:[10.1007/978-3-031-22969-5_9](https://doi.org/10.1007/978-3-031-22969-5_9).
- [DGH⁺23] Christoph Dobraunig, Lorenzo Grassi, Lukas Helming, Christian Rechberger, Markus Schafnegg, and Roman Walch. Pasta: A case for hybrid homomorphic encryption. *IACR TCHES*, 2023(3):30–73, 2023. doi:[10.46586/tches.v2023.i3.30-73](https://doi.org/10.46586/tches.v2023.i3.30-73).
- [DKR97] Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The block cipher Square. In Eli Biham, editor, *FSE’97*, volume 1267 of *LNCS*, pages 149–165. Springer, Berlin, Heidelberg, January 1997. doi:[10.1007/BFb0052343](https://doi.org/10.1007/BFb0052343).
- [EGL⁺20] Maria Eichlseder, Lorenzo Grassi, Reinhard Lüftenegger, Morten Øygarden, Christian Rechberger, Markus Schafnegg, and Qingju Wang. An algebraic attack on ciphers with low-degree round functions: Application to full MiMC. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 477–506. Springer, Cham, December 2020. doi:[10.1007/978-3-030-64837-4_16](https://doi.org/10.1007/978-3-030-64837-4_16).
- [HKL⁺22] Jincheol Ha, Seongkwang Kim, ByeongHak Lee, Jooyoung Lee, and Mincheol Son. Rubato: Noisy ciphers for approximate homomorphic encryption. In Orr Dunkelman and Stefan Dziembowski, editors, *EUROCRYPT 2022, Part I*, volume 13275 of *LNCS*, pages 581–610. Springer, Cham, May / June 2022. doi:[10.1007/978-3-031-06944-4_20](https://doi.org/10.1007/978-3-031-06944-4_20).
- [HSWW20] Kai Hu, Siwei Sun, Meiqin Wang, and Qingju Wang. An algebraic formulation of the division property: Revisiting degree evaluations, cube attacks, and key-independent sums. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 446–476. Springer, Cham, December 2020. doi:[10.1007/978-3-030-64837-4_15](https://doi.org/10.1007/978-3-030-64837-4_15).
- [JK97] Thomas Jakobsen and Lars R. Knudsen. The interpolation attack on block ciphers. In Eli Biham, editor, *FSE’97*, volume 1267 of *LNCS*, pages 28–40. Springer, Berlin, Heidelberg, January 1997. doi:[10.1007/BFb0052332](https://doi.org/10.1007/BFb0052332).
- [Knu95] Lars R. Knudsen. Truncated and higher order differentials. In Bart Preneel, editor, *FSE’94*, volume 1008 of *LNCS*, pages 196–211. Springer, Berlin, Heidelberg, December 1995. doi:[10.1007/3-540-60590-8_16](https://doi.org/10.1007/3-540-60590-8_16).
- [KW02] Lars R. Knudsen and David Wagner. Integral cryptanalysis. In Joan Daemen and Vincent Rijmen, editors, *FSE 2002*, volume 2365 of *LNCS*, pages 112–127. Springer, Berlin, Heidelberg, February 2002. doi:[10.1007/3-540-45661-9_9](https://doi.org/10.1007/3-540-45661-9_9).
- [Lai94] Xuejia Lai. Higher order derivatives and differential cryptanalysis. In *Communications and Cryptography: Two Sides of One Tapestry*, pages 227–233. Springer, 1994. doi:[10.1007/978-1-4615-2694-0_23](https://doi.org/10.1007/978-1-4615-2694-0_23).

- [LAW⁺23] Fukang Liu, Ravi Anand, Libo Wang, Willi Meier, and Takanori Isobe. Coefficient grouping: Breaking Chaghri and more. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023, Part IV*, volume 14007 of *LNCS*, pages 287–317. Springer, Cham, April 2023. doi:10.1007/978-3-031-30634-1_10.
- [LCC⁺25] Muzhou Li, Jiamin Cui, Longzheng Cui, Kai Hu, Chao Niu, and Meiqin Wang. Towards better integral distinguishers over $\mathbb{F}_p$ based on exact coefficients of monomials. Cryptology ePrint Archive, Report 2025/955, 2025. URL: <https://eprint.iacr.org/2025/955>.
- [LLC⁺24] Fen Liu, Yongqiang Li, Huiqin Chen, Lin Jiao, Ming Luo, and Mingsheng Wang. YuX: Finite field multiplication based block ciphers for efficient FHE evaluation. *IEEE Transactions on Information Theory*, 70(5):3729–3749, 2024. doi:10.1109/TIT.2024.3349414.
- [LNV11] Kristin Lauter, Michael Naehrig, and Vinod Vaikuntanathan. Can homomorphic encryption be practical? Cryptology ePrint Archive, Report 2011/405, 2011. URL: <https://eprint.iacr.org/2011/405>.
- [LS26] Guoqiang Liu and Bing Sun. Higher-order differential attacks on the full DuX. Cryptology ePrint Archive, Paper 2026/1907, 2026. Version of 17 September 2026, the second version, whose numbering of results is used here. URL: <https://eprint.iacr.org/2026/1907>.
- [MJSC16] Pierrick Méaux, Anthony Journault, François-Xavier Standaert, and Claude Carlet. Towards stream ciphers for efficient FHE with low-noise ciphertexts. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part I*, volume 9665 of *LNCS*, pages 311–343. Springer, Berlin, Heidelberg, May 2016. doi:10.1007/978-3-662-49890-3_13.
- [NWL26] Jianqiang Ni, Gaoli Wang, and Yingxin Li. Tracking algebraic degree with exponent sets: Higher-order differential attacks on FHE-friendly cipher Yu₂X. *Designs, Codes and Cryptography*, 94(6):art. 123, 2026. doi:10.1007/s10623-026-01852-8.
- [Tod15] Yosuke Todo. Structural evaluation by generalized integral property. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 287–314. Springer, Berlin, Heidelberg, April 2015. doi:10.1007/978-3-662-46800-5_12.
- [WDL⁺26] You Wu, Xinfeng Dong, Yongqiang Li, Fen Liu, Yuan Zhang, Rong Cheng, Hao Tan, and Wenzheng Zhang. DuX: a block cipher for efficient FHE evaluation. *Designs, Codes and Cryptography*, 94(7):art. 140, 2026. doi:10.1007/s10623-026-01855-5.
- [WTW25] Weizhe Wang, Deng Tang, and Haoyang Wang. Inner product masked integral distinguishers and integral sets over large finite fields: Applications to MiMC, CIMINION and Chaghri. In Willy Susilo and Josef Pieprzyk, editors, *ACISP 25, Part I*, volume 15658 of *LNCS*, pages 58–77. Springer, Singapore, July 2025. Full version: Cryptology ePrint Archive, Paper 2023/1872, revised 8 April 2025. The section and definition numbers cited in this paper are those of the full version. URL: <https://eprint.iacr.org/2023/1872>, doi:10.1007/978-981-96-9095-4_4.

- [XZBL16] Zejun Xiang, Wentao Zhang, Zhenzhen Bao, and Dongdai Lin. Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 648–678. Springer, Berlin, Heidelberg, December 2016. doi:10.1007/978-3-662-53887-6_24.

A Specification Notes and Errata

DuX. The specification is the definitional sections, the algorithms and the figures of the design paper [WDL⁺26], and its security analysis section is not normative. We read the constant as $\alpha = 179$, the round number as $r = 12$ and the security target as 128 bits for all three instances, the master key as sixteen field words, and the third branch of the key schedule map Rf as $S(x_1) + x_3 + 2x_0 + rc$, the formula of their Section 4.3 rather than the expansion of their Section 5.5.1, which drops every $2(\cdot)$ term and is right only in characteristic 2. These readings were confirmed with the design team, as was the position that the choice of security model is the analyst’s.

The printed expansions of R^{-2} and S in the design paper each lack a $-\alpha$ term, which we corrected against the definitions and confirmed with the designers, and the printed quadratic system, one line of their Algorithm 3 and one six-round example carry further slips that do not affect the cipher. The public implementation of the designers differs from the paper in several points, α among them, and we did not use it.

The Six-Round Integral Examples of the DuX Paper. Section 5.3 of [WDL⁺26] lists four six-round integral distinguishers with one active plaintext word and one balanced output word. They are active word 3 with balanced word 0, and then $7 \rightarrow 7$, $11 \rightarrow 4$ and $15 \rightarrow 12$. These four claims contradict each other, because the cipher is equivariant under the whole block rotation Rot_4 , so the balanced sets for active words 3, 7, 11, 15 must be translates of each other by four.

Our measurement by direct summation, on two keys and all sixteen active positions, gives the pattern 1111 at layer 5 on both large instances. At layer 6, DuX(2^{16}) has 1111 for active positions $\equiv 1$, 1110 for $\equiv 0, 2$ and 0110 for $\equiv 3$, and DuX(65537) has 1110 for $\equiv 0, 1, 2$ and 0110 for $\equiv 3$. At layer 7 no word is balanced. The four claimed positions are not balanced at layer 6 under either round-counting convention. Appendix B of [LS26] reports the same conclusion, with the balanced set $C_{1,2}$ for those active words.

YuX. The specification source is Sections III–IV, Algorithms 1–3, Table II and footnote 2 of [LLC⁺24]. One ambiguity is relevant here. Line 2 of Algorithm 1 indexes the master key by four *sliding* windows $(key_i, \dots, key_{i+3})$, $i = 0..3$, which use only seven key words and leave the round keys $rk^1, \dots, rk^r$ with $7n$ bits of entropy, whereas Section VI-D uses the non-overlapping blocks $k_{4i..4i+3}$. We adopt the latter reading. Neither reading affects $rk^0 = K$, which every attack recovers, and running the two-round extension under both readings gives byte-identical systems.

Five further ambiguities of the specification, namely the embedding of the round constants into the field, the sharing of α between the instances, the rotation in the key schedule, the round numbers and the range of n , do not affect rk^0 either and are resolved as in Section VI-D of the design paper. The four quadratic S-box relations printed in Section VI-D are the characteristic-2 form and have wrong signs and constants over $\mathbb{F}_p$, the first differing from z_0 by a sign.

B Deferred Proofs

B.1 Lemma 6

Lemma 6 (layer-2 collapse of YuX in characteristic 2). *Consider YuX with one active ciphertext word at position 0 of block 0, and write $Y_0, \dots, Y_3$ for the output of the first layer on that block. Then $Y_1 = Y_0 + g$ with $g = c_1 + c_2c_3 + \alpha$ a constant, and among the sixteen rows of L^{-1} exactly the two feeding positions 1 and 3 of block 1 satisfy $u_1 + u_3 = (Y_1 + Y_3) + (Y_0 + Y_3) = Y_0 + Y_1 = g$. In characteristic 2 the second coordinate of S^{-1} expands as $y_1 = u_0 + u_1 + u_3 + u_2(u_1 + u_3)$. On block 1 its quadratic term therefore degenerates to gu_2 and its degree drops from 3 to 1, and the degree of y_3 drops from 6 to 5. The exact layer-2 profile of block 1 is $(3, 1, 5, 5)$ against the max-plus $(3, 3, 5, 6)$. Over $\mathbb{F}_p$ the same row sum is $Y_0 + Y_1 + 2Y_3$, which is not constant, and no cancellation occurs.*

Proof. Substituting $(X + c_0, c_1, c_2, c_3)$ into equation (4) gives $Y_0 = X + (c_0 + c_1c_2 + c_3 + \alpha)$, $Y_1 = Y_0 + g$, $Y_2 = c_3Y_0 + Y_1 + c_2 + \alpha$ and $Y_3 = Y_0Y_1 + Y_2 + c_3 + \alpha$, of degrees $(1, 1, 1, 2)$, whence $Y_0 + Y_1 = g$ in characteristic 2. Keeping only the non-constant indices $0, \dots, 3$ in $(L^{-1}Y)_i = \sum_{j \in \{0,3,4,8,9,12,14\}} Y_{i+j}$, the four inputs of block 1 are $(Y_0 + Y_2, Y_1 + Y_3, Y_2, Y_0 + Y_3)$. Hence $u_1 + u_3 = Y_0 + Y_1 = g$, while for the other three blocks $u_1 + u_3$ has degree 2. On block 1, $u_0 + u_1 + u_3 = Y_1 + Y_2$ has degree 1 and $u_2(u_1 + u_3) = gY_2$ has degree 1. Hence $\deg y_1 = 1$. Then $y_0 = u_0 + u_1u_2 + u_3$ has degree 3 and $y_2 = u_2 + u_3y_0 + y_1$ has degree 5. For $y_3 = u_3 + y_0y_1 + y_2$ the product has $\deg(y_0y_1) = 4 < 5$, so $\deg y_3 = 5$. Over $\mathbb{F}_p$, $y_1 = u_1 + u_2u_3 + y_0 + \alpha$ with $y_0 \ni u_1u_2$ gives $u_2(u_1 + u_3)$ and $u_1 + u_3 = Y_0 + Y_1 + 2Y_3$. $\square$

The lemma also shows how far the gap would have to grow before any attack changed. Turning the layer-9 pattern 1100 of Yu₂X-16 into 1110 would need a 16.1% gap at position 2, where the exact degree 72 240 exceeds $T = 65\,535$, the two-word layer-10 cells need 45.7%, and the best three-word structure, $(0, 1, 8)$ by exhaustive search with $D = 218\,454$ against $T = 196\,605$, needs 11.1%.

What the gap does buy is an explanation of a practical effect. At layer 9 the exact margin at position 0 is $65\,535 - 48\,128 = 17\,407$, and the weight range that the ten-round attack measures on the instance is about 17 400. When comparing against SMT-based tables the gap is the reason the two can differ in the conservative direction.

B.2 Theorem 2

Write every state word, as a polynomial in the active plaintext word X , by its max-plus degree m and its top two coefficients (p, p') of X^m and X^{m-1} .

One S-box layer in the encryption direction maps four block inputs of degree at most m , with top pairs (p_0, p'_0) and (p_3, p'_3) at positions 0 and 3, to

output	degree	leading coefficient	next-to-leading coefficient
a	$2m$	p_0p_3	$p_0p'_3 + p'_0p_3$
b	$3m$	$p_0p_3^2$	$p_3^2p'_0$
c	$5m$	$p_0^2p_3^3$	$p_0^2p_3^2p'_3$
y_3	$8m$	$p_0^3p_3^5$	$p_0^2p_3^4(p_0p'_3 + p'_0p_3)$

valid for $m \geq 2$ and verified by symbolic computation over $\mathbb{F}_2$ for $m \in \{2, 3, 8, 128\}$. For $m = 1$ the leading coefficient of x_2 would land in the next-to-leading slot of a , and from the second layer on $m \in \{16, 128, 1024, \dots\}$. The cross terms cancel pairwise in characteristic 2. Hence if $p'_0 = p'_3 = 0$ then all four next-to-leading coefficients vanish.

After the linear layer the output degree $8m$ is attained only by the y_3 -type coordinates, since the next candidate, $5m$, is below $8m - 1$. The next-to-leading coefficient of every

output of the linear layer is therefore a sum of those of y_3 -type words, and these vanish. The property is self-sustaining once it holds for all sixteen inputs of a layer.

Base case, active word 1. After the first layer only word 3 has degree 2, with leading coefficient $\lambda = rk_2^0 + rk_0^0 rk_3^0 + \alpha$. Word 3 feeds L_0 -output i exactly when $(3-i) \bmod 16$ lies in the rotation set $\{0, 1, 3, 4, 7, 8, 9, 10, 12, 14, 15\}$, which fails for $i \in \{1, 6, 8, 13, 14\}$. Hence at the second layer the position-0 input of block 2 has degree ≤ 1 and that block's y_3 has degree 13, while the other three blocks have degree 16 with the *same* leading coefficient λ^8 .

At the third layer the leading coefficient of input u_i is $|T_i \cap \{0, 1, 3\}| \cdot \lambda^8$ where T_i is the set of full-degree y_3 blocks feeding i . For $i = 4, 8, 12$ the intersections $\{0, 1\}$, $\{1, 3\}$, $\{0, 3\}$ have even size, so those inputs lose their leading term and blocks 1, 2, 3 drop below full degree. Only block 0 keeps it, with $T_0 \cap \{0, 1, 3\} = T_3 \cap \{0, 1, 3\} = \{0, 1, 3\}$. Hence $p_0 = p_3$ and $p'_0 = p'_3$, and the next-to-leading coefficient of a , $p_0 p'_3 + p'_0 p_3$, vanishes.

Consequently all sixteen inputs of the fourth layer have vanishing next-to-leading coefficient, as checked on both instances for three random keys each, and the induction takes over.

In the degenerate case $\lambda = 0$ the true degrees collapse to $(0, 1, 0, 1)$ at the first layer, and the tracked top window is identically zero from the third layer on for both instances and three random round key sets. Hence $c_{D-1} = 0$ holds as well. $\square$

B.3 Corollary 2

In $\mathbb{F}_2[\tau]/(\tau^4)$ every non-zero element is $\tau^v u$ with u a unit, so $\Psi_y = u(\pi) \tau^v \Phi_0 \sigma$ -terms. The rank of the combined row on the ρ -invariant moments is therefore that of τ^v on the column space, that is, on each π -orbit of length Λ , which is a free $\mathbb{F}_2[\tau]/(\tau^\Lambda)$ -module of rank one, the rank of multiplication by τ^v . That rank is $\max(0, \Lambda - v)$.

For $v = 3$ every orbit of length four keeps one coefficient, the orbit sum, and the degree-one key columns, which form four orbits of length four, are seen only through their sums. The τ -valuations of $(1, 1, 0, 0) = 1 + t$, $(0, 1, 0, 1) = t(1 + t)^2$ and $(0, 1, 1, 0) = t(1 + t)$ in $\mathbb{F}_2[t]/((1 + t)^4)$ are 1, 2, 1, and that of $(1, 1, 1, 1) = (1 + t)^3$ is 3. $\square$

B.4 The Cubic Coordinate of Lemma 5

With $x = (x_1)_{4j..4j+3}$, $x_i = u_i + k'_i$ and $u = L(Y)$ restricted to the block, write W_i for the row of L that produces u_i , so that $u_i = \sum_s W_i[s] Y_s$. In characteristic 2, $b = x_1 + x_2 x_3 + x_0 x_3^2 + \alpha x_3 + \alpha$ and $x_3^2 = u_3^2 + (k'_3)^2$ with $u_3^2 = \sum_t W_3[t]^2 Y_t^2$. Summing over the structure and dropping every term with the factor $\sum_{\mathcal{P}} 1 = 0$,

$$\begin{aligned} \sum_{\mathcal{P}} b = & \sum_s (W_1 + \alpha W_3)[s] T_s + \sum_{s,t} W_2[s] W_3[t] Q_{s,t} + k'_3 \sum_s W_2[s] T_s + k'_2 \sum_t W_3[t] T_t \\ & + \sum_{s,t} W_0[s] W_3[t]^2 Q_{s,t}^{(2)} + (k'_3)^2 \sum_s W_0[s] T_s + k'_0 \sum_t W_3[t]^2 T_t^{(2)}, \end{aligned}$$

with $T_s = \sum_{\mathcal{P}} Y_s$, $Q_{s,t} = \sum_{\mathcal{P}} Y_s Y_t$, $Q_{s,t}^{(2)} = \sum_{\mathcal{P}} Y_s Y_t^2$ and $T_t^{(2)} = \sum_{\mathcal{P}} Y_t^2$. The first four terms are those of the coordinate a in equation (8) with other rows of L , and the last three are new. Under a weight x^a the same expansion holds with weighted moments, where $T^{(2)}$ is no longer the square of T and is accumulated separately. The unknowns are the monomials of rk^0 in T , Q , $Q^{(2)}$ and $T^{(2)}$, the latter two being the monomials of Q and T with one exponent doubled, together with the products of k'_3 , k'_2 , $(k'_3)^2$ and k'_0 with the monomials of T or $T^{(2)}$. Counting them as in Section 2 gives $M_b = 47\,749$ columns against $M = 18\,025$, with the same 50 monomials per block in Y .

The kernel statement is a computation. For the pattern 1101 the rows of L at positions 0, 1 and 3 are combined so that every column of the coordinates c and y_3 cancels, which leaves a four-dimensional space, and the vector that combines rows 3, 1 and 9 has the

stated support $\{1, 2, 6, 10, 13\}$ on the coordinates b and a . Its components on the four blocks, read on the coordinate a , are $(1, 1, 1, 0)$, a unit of $\mathbb{F}_2[\mathbb{Z}_4]$, and on the coordinate b they are $(1, 0, 0, 1)$, of valuation 1. The four block rotations of this vector span the kernel and have minimal valuations 2, 0, 0 and 2 over the two coordinates. For the pattern 0001 the same cheap set gives a two-dimensional kernel spanned by $(1, 0, 1, 0)$ and $(0, 1, 0, 1)$ on the position-3 rows, both of valuation 2, and the rows determine none of the key words. $\square$

B.5 Forward Combined Equations

With $W = \text{SL}(P + rk^0)$ and $x_1 = L(W) + rk^1$, summing word i of x_1 over a structure gives $\sum_{\mathcal{P}}(x_1)_i = [L(\sum_{\mathcal{P}} W)]_i + |\mathcal{P}| rk_i^1 = [L(\sum_{\mathcal{P}} W)]_i$, which is zero when word i is balanced. The unknowns are the monomials of rk^0 occurring in the four coordinates of W after summation, which number 80, 24, 7 and 3 by position for YuX over $\mathbb{F}_p$ and 80 in the union, replicated over the four blocks that row i of L touches. With a weight x^a the same derivation holds because $\sum_{\mathcal{P}} x^a = 0$ for $a \geq 1$. $\square$

B.6 The Rank Formula of Proposition 6

The rank of the assembled system is $\text{rank}(A) = \text{rank}(\Phi) - \dim(\text{im } \Phi \cap (V^\perp)^{|\text{coords}|})$, where Φ is the template map of the coordinates used, V the span of the reachable moment vectors, and each already-balanced coordinate of W contributes one vector to $V^\perp$. The prediction matches the measured rank in all six cells we tested.

C Relation to Prior and Concurrent Work

C.1 Related Work

Analyses of YuX and DuX. The exponent set technique of Ni, Wang and Li [NWL26] is the direct predecessor of this work. It tracks, for each state word of Yu₂X-8 and Yu₂X-16, the set of achievable exponents, reads the algebraic degree off as the maximum Hamming weight, refines the first layers with the general monomial prediction of [CHWW22], and builds key recoveries with Gröbner bases over one extension round and with a Boolean linearisation over two to four rounds. Our criterion replaces the exponent set by a single max-plus number per word and gains uniformity at the price of being looser on cross-block cells, as Section C.3 quantifies, and the structured linearisation of Section 5.2 replaces their Boolean one.

Liu and Sun [LS26] analyse DuX concurrently and independently in the same model. The relation is spelled out in Section 1 and below.

Integral and Zero-Sum Properties over Large Fields. The degree criterion $\sum_{x \in \mathbb{F}_q} f(x) = 0$ for $\deg f < q - 1$, its multivariate form on affine subspaces, and sums over multiplicative subgroups of $\mathbb{F}_q^\times$ are Proposition 1, Corollary 1 and Proposition 2 of *Out of Oddity* [BCD⁺20], where they give integral and zero-sum distinguishers on GMiMC and HadesMiMC. Theorem 1 is the joint bookkeeping of those facts with the characteristic-2 subspace threshold, in both directions.

Wang, Tang and Wang [WTW25] generalise the integral distinguisher over a large field to the inner product masked form $\sum_j c_j F(x_j) = 0$, in which every point of the structure carries its own coefficient, and they characterise a structure by the exponents whose masked power sums vanish, the IPM integral sets of their Section 4 in the full version. Additive subgroups, multiplicative subgroups and merged cosets are their examples, and the subgroup sum of *Out of Oddity* is the masked sum with the mask $(1, \dots, 1, -|G|)$ on $G \cup \{0\}$ in their Theorem 2. The weighted sums of Theorem 4 are masked sums in this

sense, with the monomial values x^a as the mask, and the identity $\sum_x x^a f(x) = -c_{q-1-a}$ they rest on is classical.

Here we use these masks as a source of key-recovery equations. Theorem 4 reads the admissible exponents off the degree bound of Theorem 1, carries the masked sums into the equations of Section 5 with unchanged unknowns, and combines them with the structured linearisation of Section 5.2 and the cheap coordinate rows of Lemma 3. The slice decomposition of Section 5.5 computes every admissible weight from one pass over the structure and one matrix product. How many independent equations the weights yield is the question of Propositions 6 and 7, and the ranks were measured in every attack of Sections 6 and 7. Proposition 8, the zero sum over a single coset with a non-zero weight, is the same mechanism on a coset: the weight $x^{a'}$ removes the constant term that the mask entry $-|G|$ removes in their Theorem 2 and in Proposition 2 of *Out of Oddity*. Its role in this paper is to halve the data of the ten-round YuX-65537 attack of Section 7.

Beysse and Verbaud generalise integral and ultrametric integral cryptanalysis [BV24] to rings of prime characteristic, with automated tools [BV25], and Li et al. [LCC⁺25] build integral distinguishers over $\mathbb{F}_p$ from exact coefficients of monomials. The leading coefficient argument of Theorem 2 works with exact coefficients as well, by hand and for one family of S-boxes.

Zero-sum distinguishers in the known-key setting go back to Aumasson and Meier [AM09] and Boura and Canteaut [BC11, BCD11], and the partitions of Section E apply the same idea, both directions from an intermediate state, over large fields. The classical references for the bit-level setting are the square and integral attacks [DKR97, KW02] and higher-order differentials [Lai94, Knu95].

Degree Growth and Its Exact Value. The division property [Tod15, XZBL16] and its algebraic reformulation as monomial prediction [HSWW20] are the standard bit-level tools. They are not competitive here because the state is sixteen large field words. The influence of the linear layer on the degree growth of an SPN is studied by Cid et al. [CGG⁺22]. Proposition 2 and Corollary 1 are a concrete instance of that influence, and Boura and Canteaut’s bound on the degree of $G \circ F$ in terms of the degree of F^{-1} [BC13] is the classical statement that the degrees of a permutation and of its inverse constrain each other, the asymmetry this paper exploits.

Exact degrees below the trivial bound have been established for iterated power functions by Bouvier, Canteaut and Perrin [BCP23], for MiMC by Eichlseder et al. [EGL⁺20], who also give the algebraic key-recovery model we follow in spirit, and by coefficient grouping in characteristic 2 by Liu et al. [LAW⁺23]. Observation 2 and Theorem 2 are the same question for an SPN with a Feistel-derived S-box, answered by dense symbolic expansion and by a leading pair recursion. Interpolation attacks [JK97] were the first to work with the exact polynomial of a cipher, the object the exact degree computations here compute.

FHE-Friendly Designs. The deployment setting is hybrid homomorphic encryption [LNV11], which motivated LowMC [ARS⁺15], Kreyvium [CCF⁺18], FLIP [MJSC16], MiMC [AGR⁺16], HERA [CHK⁺21], Rubato [HKL⁺22], Pasta [DGH⁺23] and Chaghri [AMT22], and is the designers’ reason for restricting attention to chosen-plaintext adversaries. DuX and YuX belong to the subfamily that evaluates the *decryption* circuit homomorphically. That is the direction in which their algebra is weak.

C.2 The Concurrent Analysis of DuX by Liu and Sun

Five findings were obtained independently on both sides. The zero-sum criterion is their Theorem 1 and our Theorem 1, which adds subspaces, cosets and the second direction. The class recursion with base $2 + \sqrt{3}$ is their Theorem 2 and Corollary 1 and our Corollary 1.

The identity $L_1 = \text{Rot}_{12} \circ L_0$ of the two linear layers is their Propositions 1 and 2 and our Proposition 1. The eleven-layer distinguisher at q^4 with margins 53 and 57, together with the ten-layer distinguisher at q^2 and the nine-layer one at q , is their Corollary 2 and Table 3 and our Table 3. The correction of the six-round examples of the design paper is their Appendix B and our Section A.

The key-recovery layers, however, differ. [LS26] take the deepest distinguisher, eleven layers at q^4 per subspace, and one extension round with masks. This costs ten to twelve subspaces, $2^{67.32}$ to $2^{67.59}$ data at twelve rounds and $2^{35.32}$ at eleven. We take the ten-layer distinguisher at q^2 , two extension rounds by structured linearisation, the cheap coordinate kernel to survive the pattern 1101, and weighted moments to get all the equations from one structure. This gives the 2^{29} and 2^{32} of Table 1 at twelve rounds and 2^{15} and $2^{15.35}$ at eleven. At ten rounds they use four subspaces of q and we use one structure of q with eight weights, 2^{18} against 2^{16} , and on DuX(2^8) the gap at eight rounds is 2^{67} , their $2^{91.32}$ against our 2^{24} .

Weighted moments alone, applied to their eleven-layer setting, would reduce their data to one subspace of 2^{64} , because the position-3 margin 53 leaves at least the eight rows of their first stage, and the further factor 2^{32} comes from the second extension round. On the reduced configuration of their Section 7, seven rounds of DuX(2^8) with three active words at 2^{24} per subspace and a margin of 41 at the balanced position, one subspace with the weights $a = 0, \dots, 40$ attains the ranks 8, 4 and 7 of the three stages of their Theorem 3 on three random master keys, where their runs use ten, four and four subspaces, and without weights one subspace gives the ranks 1, 1 and 2 of their Proposition 4. Their attack keeps a constant number of field elements in memory and solves systems of at most 44 columns, where our twelve-round DuX(65537) attack holds 13.8 GB for an elimination over 38 051 columns, both of which fit on a workstation.

The Mask Formulation. [LS26] phrase the one-round extension with masks, in their Theorem 3 and Proposition 4. A row vector v supported on the balanced words B of x_1 and its transport $w = vL$ through the linear layer give $\sum_P \sum_p w_p S(P^{(p)} + k^{(p)}) = 0$, and the spaces $V_C = \{v : \text{supp}(v) \subseteq B, \text{supp}(vL) \subseteq C_C\}$, sorted by the degrees of the coordinates they touch, give a three-stage elimination. In our language V_C is the left kernel of Lemma 3 for the one-round extension with “cheap” meaning “of position in C ”, the stages are the sequential substitution above, and the rank formula of Proposition 6 is a closed form for their rank condition, which they check during the attack. We adopt their formulation where it is cleaner and keep the closed form.

C.3 The Exponent Set Tables of Ni, Wang and Li

The tables of [NWL26] give, for each structure, layer and output word, an upper bound on the *Boolean* degree. For s active words in characteristic 2, the Boolean degree implied by a formal total degree D is the knapsack value $\text{hw}(D, s, n) = \max\{\sum_i \text{wt}(a_i) : \sum_i a_i \leq D, a_i \leq 2^n - 1\}$, and for $s = 1$ this is their own $\lfloor \log_2(D + 1) \rfloor$. Since $\text{hw}(D, s, n) = sn$ exactly when $D \geq s(2^n - 1) = T$, the Boolean criterion $\text{deg} < sn$ and $D < T$ are the same condition. We transcribed all 772 printed cells of their Tables 3, 4, 6 and 7 and compared them with the criterion. Of these, 353 agree exactly, 361 have a looser max-plus value and 58 a tighter one. The looser cells are all cross-block cells, in which the total degree cannot see that only same-block variables reach a word in the first layers, and for the 180 same-block cells the two agree in 156 and max-plus is tighter in 22. Only 21 cells change the balance verdict. In 17 of them the SMT table is stronger, every one at a Frobenius boundary, where the formal degree is exactly a power of two and the margin 0 or -4 , and in the other 4 the criterion is stronger and yields distinguishers the published tables do not contain, namely the layer-9 pattern 1100 of Yu₂X-16, the layer-5 pattern of Yu₂X-8 started at position 3, and two eight- and nine-word cells of Yu₂X-8 at layer 6. The single-word

Table 9: Master key recovery on DuX in the chosen-plaintext model, the model of the design paper.

Instance	Rounds	Data	Mult.	Time
DuX(65537)	7/12	$2^{18.8}$	$2^{24.00}$	42 s
DuX(2^{16})	7/12	$2^{18.6}$	$2^{23.77}$	2.8 s
DuX(2^{16})	8/12	2^{16}	$2^{36.58}$	1 190 s
DuX(65537)	8/12	2^{16}	$2^{36.59}$	125 s
DuX(2^{16})	8/12	$2^{35.46}$	$2^{42.46}$	—
DuX(2^{16})	8/12	$2^{23.38}$	$2^{36.33}$	774 s
DuX(65537)	$9^\dagger$	2^{64}	—	—

Data is in $\log_2$ chosen plaintexts, Mult. is as in Table 1, times are the elapsed seconds of one single-thread run, and † marks the theoretical row. The $2^{35.46}$ route was not run at full scale and carries its operation count only.

cell at position 0 of Table 4 agrees with their Table 3, and the single-word cell started at position 3 is not in their tables. Their four-word structure (0, 4, 8, 12) at 2^{63} gives the pattern 1100 for the full block, and the four-word Yu_2X -8 cell, balanced at 1100 where the criterion gives 0000, agrees with their Table 6. We also note that the two-variable rows of their Table 4 print output positions 0 and 1 on one line although their bounds differ, $D = 12$ against 14 at layer 3.

D Chosen-Plaintext Results

In the encryption direction the S-box degrees are (5, 3, 2, 8) and the linear layer is dense. The criterion reaches layer 6 with one word, for DuX(2^{16}) by Theorem 2 and for DuX(65537) with position 3 left unbalanced in the pattern 1110. It reaches layer 7 with two words at 2^{32} , in the pattern 0010 on DuX(2^{16}), or with three words at 2^{48} , and layer 7 with a full block at 2^{64} in the pattern 1110. Layer 8 is unreachable for any structure in the search space of Section F. The design paper estimates that integral distinguishers of this kind reach at most six rounds. Its four explicit six-round examples are not balanced, as Section A records, a finding shared with Appendix B of [LS26].

The same key-recovery machinery applies with the cheap coordinates f and g of S^{-1} at positions 3 and 0. Table 9 lists the routes.

Seven rounds fall to the one-round extension from the layer-5 zero sum, from $2^{18.8}$ chosen plaintexts on DuX(65537) and $2^{18.6}$ on DuX(2^{16}), in seconds. Eight rounds fall to the two-round extension from a single 2^{16} structure on both large instances. The layer-6 pattern 1110, the four-dimensional kernel and 400 weights give 1 600 equations of rank 1 086 on DuX(2^{16}) and of rank 1 234 on the $M = 3\,307$ columns of DuX(65537). A single-thread run takes 1 190 s on DuX(2^{16}) and 125 s on DuX(65537), the former slower because the weighted sums over $\mathbb{F}_{2^{16}}$ go through logarithm tables.

A route independent of the two-round machinery cross-validates that result on DuX(2^{16}). The two-word layer-7 pattern 0010 with the one-round extension on the single balanced position reaches eight rounds from $2^{35.46}$ chosen plaintexts in $2^{42.46}$ multiplications. So does the two-round extension without weights, which needs $N_{\min} = 167$ structures, or $2^{23.38}$ chosen plaintexts. Nine rounds would need the full block at 2^{64} and the four combined rows, and are theoretical.

YuX. In the encryption direction the criterion reaches layer 6 with one word, in the pattern 0111 with kernel dimension 4 for the cheap coordinates y_0 and $y_1 - y_0$ of S^{-1} , and layer 7 only with a full block at 2^{64} , in the pattern 0111 on Yu_2X -16 and, by Proposition 5, in the full zero sum 1111 on Yu_pX -65537, whose position-0 word sits at the boundary

$D = T$. The corresponding eight- and nine-round key recoveries are theoretical, because we did not implement the assembly for YuX in this direction, and no exact degree gap exists in this direction from layer 4 on. Yu₂X-8 caps at layer 4.

E Known-Key Zero-Sum Partitions

In the known-key model the intermediate state may be placed anywhere and the criterion applied in both directions [AM09, BC11, BCD11]. With the state immediately after a diffusion layer, as in [LS26], the cosets of the coordinate subspace spanned by an active set A of s words partition the state space into sets of q^s elements. Every backward group is (L^{-1}, S^{-1}) , with a leading inverse diffusion layer, and every forward group is (S, L) . The threshold is $s(q - 1)$ on both sides, and a zero-sum partition over $r_b + r_f$ rounds of the iterated structure, with arbitrary fixed round keys and branch bits, follows when all sixteen words clear it on both sides.

For DuX our splits reproduce Corollary 5 of [LS26] number by number, namely 13 and 16 rounds with q and q^4 texts per part on the large instances and 8 and 10 rounds with 2^{16} and 2^{96} on DuX(2^8), and an exhaustive search over active sets modulo block rotation confirms that their sets maximise $r_b + r_f$. Table 10 lists the partitions that are new in this paper.

Placing the state before the diffusion layer, which [LS26] do not consider, lets a whole block pass its first S^{-1} without cost by Lemma 2 and gives DuX(2^8) nine rounds with 2^{32} texts per part where [LS26] need 2^{96} for ten. YuX, which [LS26] do not treat, admits 13, 14 and 15 rounds with q , q^2 and q^3 texts.

All splits were checked by direct summation on the toy instances over $\mathbb{F}_{2^4}$, $\mathbb{F}_{2^{57}}$ and their YuX counterparts with two keys each. These are distinguishers of an idealised structure, not attacks on the twelve-round ciphers.

Table 10: Known-key zero-sum partitions that are new in this paper.

Instance	Active set A	Texts per part	$r_b + r_f$
DuX(2^{16}), (65537)	$\{0, 1\}$	q^2	$9 + 6 = 15$
DuX(2^8)	full block, before L	2^{32}	$6 + 3 = 9$
Yu _p X-65537, Yu ₂ X-16	$\{0\}$	q	$8 + 5 = 13$
Yu _p X-65537, Yu ₂ X-16	$\{1, 2\}$	q^2	$8 + 6 = 14$
Yu _p X-65537, Yu ₂ X-16	$\{0, 3, 8\}$	q^3	$9 + 6 = 15$
Yu _p X-65537, Yu ₂ X-16	$\{0, 2, 3, 6\}$	q^4	$9 + 6 = 15$
Yu _p X-65537, Yu ₂ X-16	full block, before L	q^4	$9 + 5 = 14$
Yu ₂ X-8	$\{1\}$	2^8	$4 + 3 = 7$
Yu ₂ X-8	$\{0, 1, 3, 5\}$	2^{32}	$5 + 3 = 8$

“Before L ” places the state between the S-box layer and the diffusion layer instead of after the diffusion layer, the placement of Lemma 2.

F Optimality of the Structures Inside the Framework

The data figures above are the smallest the framework allows. We searched exhaustively over structures with zero or one free block, any subset of the remaining words and, in characteristic 2, any allocation of subspace dimensions, for the cheapest structure making each pattern class hold at each layer, for both ciphers and both directions. Because the max-plus profile depends only on *which* words are active, the search factorises. We enumerate active sets modulo the four block rotations and then solve a one-dimensional convex allocation problem for the cheapest dimension budget clearing the degree. Since $2^m - 1$ is convex in m , the extreme allocation is optimal for a fixed bit budget.

The pattern classes are the ones the key recoveries consume. They are 1111, which gives the four plain rows, $\mathbf{xxx1}$ for DuX with $\dim K = 1$, $\mathbf{x1x1}$ for DuX in characteristic 2 with the cheap set $\{1, 2\}$ of Lemma 5 and $\dim K = 4$, 1110 for YuX with $\dim K = 3$ and, in the encryption direction, 0111 and 1110. Table 11 gives the result.

Table 11: The deepest layer with a usable pattern for each instance and direction, with the cheapest structure that reaches it.

Instance	Direction	Layer	Pattern	Cheapest structure	Data
DuX(65537)	decryption	11	0001	four words, or a full block	2^{64}
DuX(2^{16})	decryption	10	1101, cheap set $\{1, 2\}$	two words	2^{32}
DuX(2^8)	decryption	6	1101, cheap set $\{1, 2\}$	three words	2^{24}
Yu _p X-65537, Yu ₂ X-16	decryption	10	1110	full block	2^{64}
Yu ₂ X-8	decryption	6	1110	full block	2^{32}
DuX(65537), DuX(2^{16})	encryption	7	1110	full block	2^{64}
Yu _p X-65537, Yu ₂ X-16	encryption	7	0111	full block	2^{64}

Layer 11 of DuX(2^{16}) and layer 7 of DuX(2^8) admit only 0001, which is unusable in characteristic 2 by Corollary 2 with either cheap set, and full balance at layer 10 of YuX would need 2^{112} . With the cheap set $\{2\}$ alone the two DuX rows in characteristic 2 need 1111, at 2^{47} with dimensions (16, 16, 15) and at 2^{29} with dimensions (8, 8, 8, 5). Every other next layer is unreachable for any structure in the search space, and in the encryption direction DuX(2^8) and Yu₂X-8 cap at layer 4.

“Optimal” here means optimal *inside* this framework, that is, Theorem 1 plus the three tools, and not an absolute lower bound. It also refers to the shape of a structure, a full field, a subspace, a coset or a full block, since the point sets of Proposition 10 make the data of a given shape a quantity chosen freely between two powers of q , and multi-index weights change how many structures a shape needs and not which shapes reach a layer. Observation 2 shows where the framework is conservative, and Section C.3 where a finer bookkeeping of exponents is sharper. The cheapest *distinguisher* is also not always the cheapest *attack*, because the degree margin has to be budgeted together with the data. For DuX(2^8) at layer 6 the allocation (8, 8, 8, 5) costs 2^{29} instead of 2^{32} , but it shrinks the weight margin from 240 to 16. With one-dimensional weights the attack then needs 75 structures instead of five and ends at $2^{35.2}$ rather than $2^{34.32}$, and only a weight with four indices allows a single structure of it to suffice, as Section 6.5 reports. The three-word structure at 2^{24} with the cubic coordinate is cheaper than both.

G Numerical Tables

Table 12 gives $\dim K$ of Lemma 3 for every pattern, where bit p set means that position p is balanced, computed over $\mathbb{F}_{2^4}$, $\mathbb{F}_{2^8}$, $\mathbb{F}_{2^{16}}$, $\mathbb{F}_{193}$, $\mathbb{F}_{257}$ and $\mathbb{F}_{65537}$. The values agree across the six fields except in the two DuX/encryption cells marked, 0110 with 2 over $\mathbb{F}_{2^n}$ against 1 over $\mathbb{F}_p$ and 0111 with 6 against 5, because L^{-1} is a 0/1 matrix whose $\mathbb{F}_2$ -rank can be smaller. Neither of these cells is used in the attacks. The last row is the cheap set $\{1, 2\}$ of Lemma 5, which exists in characteristic 2 only, and its two cells 0001 and 1001 would be 1 over a prime field. Table 13 gives the pseudo-structure bounds $\text{rank}(\Phi|V_0)$ of Proposition 7 on the whole column space, and the ceilings after normalisation, which moves eleven columns to the right-hand side, ten of which carry rank. For the extended system of Lemma 5 the bound was measured after normalisation only.

Table 12: Kernel dimensions $\dim K$ of the cheap rows of Lemma 3 for the fifteen non-trivial patterns, both ciphers and both directions.

pattern	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
DuX, dec	1	0	4	0	1	0	4	0	1	0	4	0	1	0	4
DuX, enc	4	1	5	0	4	2/1	6/5	0	4	1	5	0	4	4	8
YuX, dec	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4
YuX, enc	0	0	0	0	0	0	4	0	0	0	4	4	4	4	8
DuX, dec, cheap set $\{1, 2\}$	2	1	5	0	4	4	8	0	2	1	5	0	4	4	8

2/1 and 6/5: characteristic 2 / prime field. The cheap set $\{1, 2\}$ row is the characteristic-2 count of Lemma 5.

Table 13: Template bounds of Proposition 7 and the ceilings after normalisation, with the attacks that attain them.

system	field	M	bound	normalised ceiling	attained by
DuX, four rows	$\mathbb{F}_{2^n}$	18 025	4 766	4 756	11-round DuX(2^{16}), 7-round DuX(2^8)
DuX, four rows	$\mathbb{F}_{65537}$	38 051	8 618	8 608	11-round DuX(65537)
DuX, 0001 row	$\mathbb{F}_{2^n}$	18 025	2 178	2 168	no attack, 0/16 key words
DuX, 0001 row	$\mathbb{F}_p$	38 051	8 198	8 188	12-round DuX(65537), 11-round coset route
DuX, 1101 rows, cheap set $\{1, 2\}$	$\mathbb{F}_{2^n}$	47 749	—	12 056	12-round DuX(2^{16}), 8-round DuX(2^8)
YuX, four rows	$\mathbb{F}_{2^n}$	27 895	4 950	4 940	11-round YuX-16, 7-round YuX-8
YuX, four rows	$\mathbb{F}_{65537}$	53 857	8 618	8 608	11-round YuX-65537, point set and coset 10-round YuX-65537
YuX, 1110 rows	$\mathbb{F}_{2^n}$	27 895	3 832	3 822	8-round YuX-8
YuX, 1110 rows	$\mathbb{F}_p$	53 857	8 478	—	12-round route, not run