

Higher-order differential attacks on the full DuX

Guoqiang Liu^{1*} and Bing Sun¹

^{1*} College of Science, National University of Defense Technology,
Changsha, 410073, Hunan, China.

*Corresponding author(s). E-mail(s): liuguoqiang87@hotmail.com;

Abstract

DuX is a family of substitution-permutation block ciphers over $\mathbb{F}_q^{16}$ with $q \in \{2^8, 2^{16}, 65537\}$ and twelve rounds. For the two large-word instances its designers estimate that integral and higher-order differential distinguishers in the encryption direction reach at most six rounds, and fix the number of rounds accordingly. We bound the word-wise algebraic degree of DuX by exponent sets in the decryption direction, where the decryption S-box has coordinate degrees $(2, 3, 4, 2)$ against the $(5, 3, 2, 8)$ of the encryption S-box. Each row of the inverse diffusion matrix is supported on two residue classes of word indices modulo four. An active set equal to one class therefore keeps the vector of degree bounds constant within each class at every layer, and we prove that the four class-wise bounds then obey an exact recursion with base $2 + \sqrt{3}$ instead of four. This yields higher-order differential distinguishers for eleven rounds of DuX(2^{16}) and DuX(65537) with q^4 chosen ciphertexts, and for seven rounds of DuX(2^8) with 2^{88} . Extending one round towards the plaintext gives three successive systems of equations. Every unknown there has a coefficient computed from the returned plaintext words, and the key words recovered at one stage are substituted into the next; for DuX(65537) the last two systems are replaced by bivariate interpolation. Once each system attains the maximal rank that its structure permits, a condition that can be checked during the attack, solving the systems recovers all sixteen master-key words of the full twelve-round DuX(2^{16}) and DuX(65537). The data and time complexity is $2^{67.32}$ and $2^{67.59}$ with constant memory, and eight rounds of DuX(2^8) are covered with $2^{91.32}$. The two diffusion layers differ by a rotation of four words, so the analysis is the same for each of the 2^{11} key-dependent choices of diffusion layers.

Keywords: DuX, Algebraic degree, Higher-order differential attack, Key-recovery attack

MSC Classification: 94A60 , 11T71 , 68P25

1 Introduction

Homomorphic encryption [1–3] lets a server compute on encrypted data, but a homomorphic ciphertext is much larger than the plaintext it encrypts. The hybrid framework of Naehrig, Lauter and Vaikuntanathan [4] avoids that expansion on the client side: the client sends symmetric ciphertexts together with a homomorphic encryption of the symmetric key, and the server evaluates the symmetric primitive homomorphically. The cost of that evaluation is governed by the multiplicative depth and the number of multiplications of the primitive. These constraints have produced a line of dedicated designs, including LowMC [5], Kreyvium [6], FLIP [7], HERA [8], Rubato [9] and Pasta [10], and, over large fields, the MiMC and GMiMC families [11, 12].

Related work. Designs of this kind obtain their efficiency from components of low algebraic degree, and are therefore exposed to integral and higher-order differential cryptanalysis [13–16] and to the division property [17–19]. YuX [20] is a family of such block ciphers over $\mathbb{F}_q^{16}$ whose round function is built from a four-branch generalized Feistel network; it has three members, Yu₂X-8 and Yu₂X-16 over $\mathbb{F}_{2^8}$ and $\mathbb{F}_{2^{16}}$ and Yu_pX over $\mathbb{F}_{65537}$. Ni, Wang and Li [21] track the algebraic degree of its two binary-field members, which they denote collectively by Yu₂X, by exponent sets. For each coordinate of each round they record the set of exponent vectors that occur in its reduced polynomial, and propagate those sets layer by layer instead of a single degree. Their bounds are taken in the decryption direction, where the coordinate degrees of the S-box are smaller than in the encryption direction. The resulting upper bounds give higher-order differential distinguishers for ten of the fourteen rounds of Yu₂X-16 and recover the key of twelve of its fourteen rounds with $2^{120.4}$ data and time. Twelve rounds is the number that [20] recommends for Yu₂X-16 in homomorphic applications.

Motivation. DuX [22] is a more recent member of the same line. It is a substitution–permutation network over $\mathbb{F}_q^{16}$ with $q \in \{2^8, 2^{16}, 65537\}$ and twelve rounds, and its S-box layer is made of four parallel copies of a four-word generalized Feistel permutation. One bit extracted from the round key selects the diffusion layer of that round from two 16×16 matrices, which are maximum-distance-separable at the level of four-word blocks rather than at the level of the sixteen words. All three instances claim 128-bit security. The designers evaluate integral and higher-order differential attacks in the encryption direction, estimate that distinguishers of that kind reach at most six rounds, and record a security bound of seven rounds in their round-number table, leaving a margin of five rounds. The generalized Feistel permutation and its inverse do not have the same coordinate degrees, so the two directions of DuX admit different degree bounds, and the search reported in [22] covers the direction in which they are larger.

Our contribution. We bound the word-wise algebraic degree of DuX in the decryption direction, where the S-box S^{-1} has coordinate degrees (2, 3, 4, 2) against the (5, 3, 2, 8) of S . Since a word sums to zero as soon as its degree stays below a threshold proportional to q , while these degrees do not depend on q , the cheaper direction accommodates substantially more rounds on the two large-word instances. Three results follow.

First, a class-wise degree recursion. Each row of the inverse diffusion matrix has exactly five nonzero entries, and their indices occupy only two classes modulo four, so an output class draws from two input classes. Taking the active set to be one full class therefore keeps the vector of degree bounds constant on classes at every layer, and the four class bounds satisfy an exact recursion and grow exponentially with base $2 + \sqrt{3}$ instead of four (Theorem 2). The smallest of the four class-wise bounds is four times the previous one minus the one before that, so it equals 262087 after eleven layers, against a threshold of 262140 (Corollary 1). This yields eleven-round higher-order differential distinguishers for DuX(2^{16}) and DuX(65537) with q^4 chosen ciphertexts and a seven-round distinguisher for DuX(2^8) with 2^{88} .

Second, key recovery on the full cipher, in the chosen-ciphertext model: the subspaces are chosen on the ciphertext side and every coefficient is a sum of returned plaintext words, so the attack assumes an oracle that decrypts and outputs plaintexts. The eleven-round zero-sum state is the output of the second key addition, so one plaintext-side round separates it from the plaintext. That round is removed by linear masks rather than by guessing: the masks supported on the four zero-sum words give equations whose only unknown is the first whitening key, which the design sets equal to the master key. Sorted by the classes on which their images under the diffusion layer are supported, these masks yield three systems, and the key words recovered at one stage are substituted into the next (Theorem 3). The coefficient of every unknown is a sum of observed plaintext words (Proposition 5), so the rank of each system can be checked during the attack. All sixteen master-key words of the full twelve-round DuX(2^{16}) and DuX(65537) are recovered with $2^{67.32}$ and $2^{67.59}$ chosen ciphertexts, equal time and constant memory, and eight rounds of DuX(2^8) with $2^{91.32}$ (Corollary 4).

Third, results in the model used in [22] and in the known-key model. In the chosen-plaintext direction the criterion guarantees a seven-round distinguisher with q^3 texts (Corollary 3), one round beyond the estimate recorded in [22]. The round-number table of [22] counts covered rounds. There this moves integral and higher-order differential attacks from seven rounds to eight, and the difference from the twelve rounds of the cipher from five to four. In the known-key model, zero-sum partitions into sets of q texts cover thirteen rounds of DuX(2^{16}) and DuX(65537), and partitions into sets of q^4 texts cover sixteen rounds (Corollary 5). Along the way, the two diffusion layers are shown to differ by a rotation of four words. The branch pattern therefore only rotates the degree bounds by a multiple of four words, and the attack is the same for each of the 2^{11} patterns (Proposition 2).

Table 1 collects the main results. Time there is measured in computations of the full twelve-round DuX decryption, and the memory of every distinguisher and key recovery listed is a constant number of field elements. All figures with two decimals are rounded, and the key-recovery rows assume the rank condition of Corollary 4. For DuX(65537) the entries 2^{16} , 2^{48} and 2^{64} are rounded values of q , q^3 and q^4 , the exact quantities being powers of q . The zero-sum partitions exceed the twelve rounds of the cipher, so their round entries carry no denominator. The six-round estimate of [22] is stated there for DuX(2^{16}) and DuX(65537). Corrections to four formulas and examples in [22] are given in Appendix B.

Table 1 Summary of higher-order differential distinguishers and key-recovery attacks on DuX

Instance	Attack type	Round	Time	Data	Reference
DuX(2^8)	HD distinguisher	7/12	2^{88}	2^{88} CC	Section 4.1
	Key recovery	8/12	$2^{91.32}$	$2^{91.32}$ CC	Section 5.4
	Zero-sum partition	10	2^{96}	2^{96} KK	Section 6
DuX(2^{16})	Integral distinguisher	6/12	2^{16}	2^{16} CP	[22]
	HD distinguisher	7/12	2^{48}	2^{48} CP	Section 4.2
	HD distinguisher	11/12	2^{64}	2^{64} CC	Section 4.1
	Key recovery	12/12	$2^{67.32}$	$2^{67.32}$ CC	Section 5.4
	Zero-sum partition	13	2^{16}	2^{16} KK	Section 6
	Zero-sum partition	16	2^{64}	2^{64} KK	Section 6
	Zero-sum partition	16	2^{64}	2^{64} KK	Section 6
DuX(65537)	HD distinguisher	7/12	2^{48}	2^{48} CP	Section 4.2
	HD distinguisher	11/12	2^{64}	2^{64} CC	Section 4.1
	Key recovery	12/12	$2^{67.59}$	$2^{67.59}$ CC	Section 5.4
	Zero-sum partition	13	2^{16}	2^{16} KK	Section 6
	Zero-sum partition	16	2^{64}	2^{64} KK	Section 6

a/b : a rounds attacked out of b . HD: higher-order differential. CP: chosen plaintext. CC: chosen ciphertext. KK: known key.

Concurrent work. Ren, Xu, Xiong, Li and Wang [23] analyse DuX in the same chosen-ciphertext model, concurrently with and independently of the present work. Their degree bounds coincide with ours. The class-wise recursion with base $2 + \sqrt{3}$, the eleven-layer distinguisher on q^4 chosen ciphertexts, the 2^{88} seven-layer distinguisher on DuX(2^8), the identity between the two diffusion layers and the correction of the six-round examples of [22] were obtained on both sides. The key recoveries differ. The attack of Section 5 removes a single round in front of the eleven-layer distinguisher by linear masks sorted into three stages, so the memory stays constant and the rank of each stage is observable during the attack. They remove two rounds in front of a ten-layer distinguisher by structured linearisation, and recover the master key of the full twelve-round DuX(65537) from one structure of 2^{32} chosen ciphertexts, executed on random keys. For DuX(2^{16}) they report the full twelve rounds at 2^{47} , marked there as not executed at full scale, and for DuX(2^8) eight rounds at $2^{34.32}$. They also attack the YuX family and extend the known-key zero-sum partitions of Section 6 to it. The present text is the second version of Cryptology ePrint Archive, Paper 2026/1907, whose first version was posted on 10 September 2026; [23] was posted on 17 September 2026.

Outline. Section 2 fixes notation, recalls the framework of exponent sets and higher-order differentials, and describes DuX. Section 3 tracks the word-wise algebraic degree of DuX: the zero-sum criterion and the propagation rules, the structure of the two diffusion layers and the class-wise recursion. Section 4 derives the distinguishers, Section 5 the key recovery and Section 6 the zero-sum partitions. Section 7 reports the experimental verification and Section 8 concludes.

2 Preliminaries

2.1 Notation

Throughout, $q \in \{2^8, 2^{16}, 65537\}$ and $\mathbb{F}_q$ is the field with q elements. A state is a vector $X = (X_0, \dots, X_{15}) \in \mathbb{F}_q^{16}$ and X_p is its p -th word. A word index p is written $p = 4k + j$ with $0 \leq k \leq 3$ and $0 \leq j \leq 3$; k is the index of the four-word *S-box block* that contains p and $j = p \bmod 4$ is the *coordinate* of p in that block, the sixteen words together forming the cipher block. For $0 \leq j \leq 3$ we write

$$\mathcal{C}_j = \{p \mid 0 \leq p \leq 15, p \equiv j \pmod{4}\} = \{j, j+4, j+8, j+12\}$$

for the *class* of the four words of coordinate j , one in each block, and $\mathcal{C}_C = \bigcup_{j \in C} \mathcal{C}_j$ for $C \subseteq \{0, 1, 2, 3\}$. For $0 \leq \ell \leq 15$, the word-wise left rotation Rot_ℓ acts by $(\text{Rot}_\ell X)_p = X_{(p+\ell) \bmod 16}$; thus Rot_ℓ maps each class onto itself whenever $4 \mid \ell$. For a vector v over $\mathbb{F}_q$, $\text{supp}(v)$ is the set of indices where v is nonzero. All sums over sets of texts are sums in $\mathbb{F}_q$.

2.2 Algebraic degree and higher-order differential attacks

Definition 1 Let $F: \mathbb{F}_q^s \rightarrow \mathbb{F}_q$ be written as its reduced polynomial $\sum_e \varphi_e \prod_{t=1}^s x_t^{e_t}$ with $0 \leq e_t \leq q-1$. The *exponent set* of F is $E(F) = \{e \mid \varphi_e \neq 0\}$ and the *word-wise algebraic degree* of F is

$$D(F) = \max_{e \in E(F)} \sum_{t=1}^s e_t.$$

For a vectorial function the exponent sets and the degrees are taken coordinatewise, and E_j and D_j denote those of coordinate j .

Definition 2 Let $A \subseteq \{0, \dots, 15\}$ with $|A| = s \geq 1$; the words indexed by A are *active*. The associated *affine subspace* $V + c$ of dimension s is the set of q^s texts whose active words take every value of $\mathbb{F}_q^s$ exactly once and whose remaining words are equal to the constants c . For a function defined on the subspace, an output word has the *zero-sum property* if its values sum to zero over the whole subspace, and n_z denotes the number of output words with that property, $0 \leq n_z \leq 16$.

A higher-order differential attack [15, 16] takes an affine subspace whose dimension exceeds the algebraic degree of the primitive and uses the resulting zero sum as a distinguisher. Both quantities are there those of the binary vector space, the dimension being a number of bits and the degree the Boolean one. Integral attacks [13, 14] and the division property [17] give the same zero sums from a counting argument on the exponents. The two ingredients are therefore an upper bound on the degree and a rule that propagates it through the layers. We use the exponent sets of Definition 1, as in [21], where the degree is the total degree over $\mathbb{F}_q$ of the reduced polynomial. The criterion we apply is the word-wise one $D_j < s(q-1)$ of Theorem 1. The word dimension s is not the bit dimension: a subspace of word dimension s has bit dimension ns when $q = 2^n$.

The two notions do meet in one direction. For $q = 2^n$ the criterion of Theorem 1 implies that the Boolean algebraic degree of the corresponding coordinate function of ns variables is smaller than ns : $\sum_t \text{wt}(e_t) = ns$ forces $e_t = q - 1$ for every t and hence total degree $s(q - 1)$. The distinguishers below are therefore higher-order differentials in the sense of [15, 16, 21] as well; over $\mathbb{F}_{65537}$ the word-wise notion is the one available. The input structures of Definition 2 are affine subspaces over $\mathbb{F}_q$, so the data complexities below are powers of q .

2.3 Description of DuX

DuX [22] is a substitution–permutation network on $\mathbb{F}_q^{16}$ with $r = 12$ rounds. Its three instances, one for each of $q = 2^8$, $q = 2^{16}$ and $q = 65537$, are listed in Table 2 with the entries of the specification table of [22]; the prime-field instance needs no defining polynomial. The state and the master key are both elements of $\mathbb{F}_q^{16}$, so the key space has size q^{16} , which is 2^{128} for DuX(2^8), 2^{256} for DuX(2^{16}) and larger than 2^{256} for DuX(65537). The S-box layer SL applies to each S-box block k the same permutation S of $\mathbb{F}_q^4$, so that $\text{SL}(X)_{4k+j} = S(X_{4k}, X_{4k+1}, X_{4k+2}, X_{4k+3})_j$. The permutation S is the fourth power of the inverse of the four-branch generalized Feistel round

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

that is, $S = (R^{-1})^4$ and $S^{-1} = R^4$. The constant α is taken from $\mathbb{F}_q$; the specification table of [22] sets $\alpha = 179$ and the accompanying implementation sets $\alpha = 205$. The derivations below retain α symbolically and therefore hold for both values; the experiments of Section 7 use $\alpha = 179$.

Table 2 The three instances of DuX [22]

Instance	Block	Defining polynomial	α	r	Claimed security
DuX(2^8)	$\mathbb{F}_{2^8}^{16}$	$x^8 + x^4 + x^3 + x + 1$	179	12	128
DuX(2^{16})	$\mathbb{F}_{2^{16}}^{16}$	$x^{16} + x^{12} + x^3 + x + 1$	179	12	128
DuX(65537)	$\mathbb{F}_{65537}^{16}$	—	179	12	128

α is the constant of (1), r the number of rounds and the security levels are in bits.

The diffusion layer is one of two $\mathbb{F}_q$ -linear maps L_0, L_1 on $\mathbb{F}_q^{16}$, given by circulant matrices M_0, M_1 that are maximum-distance-separable at the level of four-word blocks. Writing $c^{(0)}$ and $c^{(1)}$ for their first rows, both layers are sums of word-wise rotations,

$$L_t(X) = \sum_{\ell=0}^{15} c_\ell^{(t)} \text{Rot}_\ell(X), \quad (2)$$

in which the coefficients of the two binary-field instances are all 0 or 1. With $\text{RKA}^{(i)}(X) = X + k^{(i)}$ for the round keys $k^{(0)}, \dots, k^{(12)}$, the round function of round

i and the cipher are

$$R^{(i)} = \text{RKA}^{(i)} \circ L_{t^i} \circ \text{SL}, \quad \text{DuX} = \text{RKA}^{(12)} \circ \text{SL} \circ R^{(11)} \circ \dots \circ R^{(1)} \circ \text{RKA}^{(0)}, \quad (3)$$

so the cipher consists of twelve S-box layers, eleven diffusion layers and thirteen key additions. We write $W_0 = P + k^{(0)}$ for the state after the first key addition and $W_i = R^{(i)}(W_{i-1})$ for $1 \leq i \leq 11$, so that $C = \text{SL}(W_{11}) + k^{(12)}$ is the ciphertext of the plaintext P . The branch bit of round i is

$$t^i = (k_{15}^{(i)} \wedge 1) \oplus ((k_{15}^{(i)} \gg 1) \wedge 1), \quad (4)$$

computed from the two least significant bits of the integer representative of word 15 of $k^{(i)}$, so the eleven branch bits give 2^{11} diffusion patterns. The key schedule of [22] sets $k^{(0)} = K$, hence the first whitening key is the master key itself.

Decryption inverts (3) layer by layer:

$$\begin{aligned} (R^{(i)})^{-1} &= \text{SL}^{-1} \circ L_{t^i}^{-1} \circ (\text{RKA}^{(i)})^{-1}, \\ \text{DuX}^{-1} &= (\text{RKA}^{(0)})^{-1} \circ (R^{(1)})^{-1} \circ \dots \circ (R^{(11)})^{-1} \circ \text{SL}^{-1} \circ (\text{RKA}^{(12)})^{-1}, \end{aligned} \quad (5)$$

where $(\text{RKA}^{(i)})^{-1}(X) = X - k^{(i)}$. Each decryption round thus applies its inverse diffusion layer before its inverse S-box layer, whereas the step from C to W_{11} has no diffusion layer, and u inverse S-box layers applied to C reach W_{12-u} .

2.4 The two directions

Both S and S^{-1} are compositions of four generalized Feistel rounds, but the multiplication in (1) sits in the third output of R and in the fourth output of R^{-1} , so the two directions have different coordinate degrees. Figure 1 shows one iteration of each of the two maps, written as a multiplication and an addition followed by a rotation of the four branches by one position. The two rotations go opposite ways: in R^{-1} the branch that receives the product becomes an operand of the multiplication of the next iteration, whereas in R two iterations pass before the product is used again. This is why the largest coordinate degree grows along 1, 2, 3, 5, 8 under iteration of R^{-1} and along 1, 2, 2, 3, 4 under iteration of R . Writing $y = S^{-1}(x) = R^4(x)$ and expanding (1) gives, in any characteristic,

$$\begin{aligned} y_0 &= x_0 + x_1x_2 + \alpha, \\ y_1 &= x_1 + x_0x_1x_2 + x_2x_3 + \alpha x_2 + \alpha, \\ y_2 &= (x_0x_1 + x_3 + \alpha)(x_1x_2 + x_0 + \alpha) + x_2 + \alpha, \\ y_3 &= x_0x_1 + x_3 + \alpha, \end{aligned} \quad (6)$$

while $y = S(x) = (R^{-1})^4(x)$ satisfies the chain

$$y_2 = x_2 - x_0x_3 - \alpha, \quad y_1 = x_1 - x_3y_2 - \alpha, \quad y_0 = x_0 - y_2y_1 - \alpha, \quad y_3 = x_3 - y_1y_0 - \alpha. \quad (7)$$

Lemma 1 *The coordinates of S have word-wise algebraic degrees $(5, 3, 2, 8)$ and the coordinates of S^{-1} have word-wise algebraic degrees $(2, 3, 4, 2)$. The exponent sets of S^{-1} , from which the constant exponent vector $(0, 0, 0, 0)$ is omitted here and throughout, are*

$$\begin{aligned} E_0 &= \{(1, 0, 0, 0), (0, 1, 1, 0)\}, \\ E_1 &= \{(1, 1, 1, 0), (0, 1, 0, 0), (0, 0, 1, 1), (0, 0, 1, 0)\}, \\ E_2 &= \{(2, 1, 0, 0), (1, 2, 1, 0), (1, 1, 0, 0), (1, 0, 0, 1), (1, 0, 0, 0), \\ &\quad (0, 1, 1, 1), (0, 1, 1, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}, \\ E_3 &= \{(1, 1, 0, 0), (0, 0, 0, 1)\}. \end{aligned}$$

These sets are independent of q and have 2, 4, 9 and 2 elements. The corresponding sets for S have 9, 4, 2 and 22 elements in characteristic two and 12, 4, 2 and 34 elements for DuX(65537).

Proof Expand (6) and (7). The listed exponent vectors are those of (6) after collecting terms; the degrees are the coordinatewise maxima of $\sum_t e_t$. The constant term of a coordinate carries the exponent vector $(0, 0, 0, 0)$ and contributes nothing to a degree. Counting it as well raises the four sizes of S^{-1} to 3, 5, 10, 3, and those of S to 10, 5, 3, 23 in characteristic two and to 13, 5, 3, 35 for DuX(65537). The dependence on the characteristic occurs only for S , and only in its coordinates 0 and 3. Three exponent vectors in coordinate 0 and twelve in coordinate 3 carry an even integer coefficient, among the latter $(0, 1, 1, 0)$ and $(2, 1, 0, 3)$; these vectors survive for DuX(65537) and cancel in characteristic two. The degrees $(5, 3, 2, 8)$ and $(2, 3, 4, 2)$ are the same in both characteristics. $\square$

The degrees in Lemma 1 agree with the value $(5, 3, 2, 8)$ stated for S in [22]. The degree of a permutation and the degree of its inverse jointly control the degree growth of an iterated construction [24]; here the two maxima differ by a factor of two, and that factor is the lever used in Sections 4 and 5. The security bounds attributed to [22] below are the values stated in the corresponding subsections of that work.

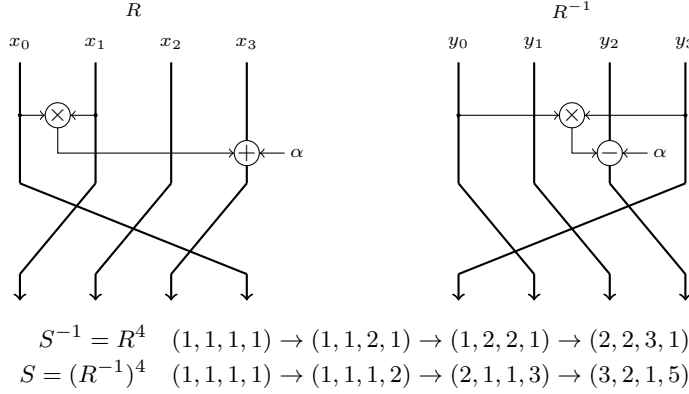


Fig. 1 One iteration of R and of R^{-1} , and the coordinate degrees after each iteration

3 Tracking the algebraic degree upper bounds of DuX

3.1 The zero-sum criterion and the propagation rules

Theorem 1 *Let F be a composition of layers of the forms (3) and (5), viewed as a map on $\mathbb{F}_q^{16}$, let A be an active set with $|A| = s$, and write coordinate j of F as a polynomial in the s active words reduced by $x^q = x$. If the word-wise algebraic degree D_j of that polynomial satisfies*

$$D_j < s(q-1),$$

then word j of F has the zero-sum property on every affine subspace of Definition 2, for every choice of the constant words and of the round keys. The conclusion also holds when D_j is replaced by any upper bound on it.

Proof Let $v = (v_1, \dots, v_s)$ range over $\mathbb{F}_q^s$. For a monomial $\prod_t v_t^{e_t}$ with all $e_t \leq q-1$ one has $\sum_{v \in \mathbb{F}_q^s} \prod_t v_t^{e_t} = \prod_t \sum_{v_t \in \mathbb{F}_q} v_t^{e_t}$, and $\sum_{z \in \mathbb{F}_q} z^e$ vanishes unless $e = q-1$, in which case it equals -1 . Hence the only monomial with a nonzero sum over $\mathbb{F}_q^s$ is $\prod_t v_t^{q-1}$, of degree $s(q-1)$, and a polynomial of word-wise algebraic degree smaller than $s(q-1)$ sums to zero. Reduction by $x^q = x$ only lowers exponents, so an upper bound on the unreduced degree bounds the reduced degree as well. Additions of round keys and of the constant α contribute constants, whose sum over the subspace is q^s times the constant and therefore zero, because the characteristic of $\mathbb{F}_q$ divides q^s . $\square$

We call $s(q-1)$ the threshold of the criterion. Theorem 1 is used together with the following propagation rule, which produces upper bounds on the D_j . Let $d \in \mathbb{Z}_{\geq 0}^{16}$ hold the current degree bounds, initialized to $d_p = 1$ for $p \in A$ and $d_p = 0$ otherwise. By (2) a diffusion layer whose first row has support J acts by

$$d'_p = \max_{\ell \in J} d_{(p+\ell) \bmod 16}, \quad (8)$$

and an S-box layer with exponent sets $E_0, \dots, E_3$ acts blockwise by

$$d'_{4k+j} = \max_{e \in E_j} \sum_{t=0}^3 e_t d_{4k+t}. \quad (9)$$

Rule (8) is the word-level division property of Todo [17] specialized to an $\mathbb{F}_q$ -linear map with the stated support, and (9) is the exponent-set propagation of Ni, Wang and Li [21]. Both rules take the word and not the bit as the unit, and both produce upper bounds, so Theorem 1 applies to their output. The word is the unit of the field-based division property of Cui, Hu, Wang and Wei [25] as well. They carry monomial prediction over to $\mathbb{F}_{2^n}$ and obtain tighter degree bounds for MiMC, Feistel-MiMC and GMiMC. On the early rounds, before the active words are spread over all sixteen positions, [21] checks its own bounds against that tool.

What (8) does not do is take the coefficients of the diffusion layer into account: it depends on J alone, and two layers with the same support give the same bound. That the linear layer influences the growth of the algebraic degree at all, and not only

the S-box layer, is the subject of Cid, Grassi, Gunsing, L uftenegger, Rechberger and Sch ofnegger [26]. Their bound is for an S-box layer of t parallel large S-boxes over $\mathbb{F}_{2^n}$ followed by an $\mathbb{F}_2$ -linear layer on nt bits. It makes the growth exponential at first and then linear, at a rate proportional to the algebraic degree of that layer over $\mathbb{F}_{2^n}^t$. The diffusion layers of DuX are $\mathbb{F}_q$ -linear, hence of degree one over the word space, so that case does not arise here. The effect that Section 3.3 isolates is a different one: the growth stays exponential, and the classes occupied by the support of M_t^{-1} lower its base from four to $2 + \sqrt{3}$.

3.2 Structure of the two diffusion layers

Proposition 1 *For all three instances of DuX,*

$$L_1 = \text{Rot}_{12} \circ L_0, \quad \text{equivalently} \quad L_1(X)_p = L_0(X)_{p-4} \quad \text{for all } p,$$

the first row of M_0^{-1} has support $\{1, 4, 8, 9, 13\}$ and the first row of M_1^{-1} has support $\{1, 5, 8, 12, 13\}$. In both cases the support indices occupy the classes $\{0, 1\}$ modulo four, so for $t \in \{0, 1\}$ the inverse diffusion layer L_t^{-1} maps

$$\text{output class } \mathcal{C}_c \longleftarrow \text{input classes } \mathcal{C}_c \cup \mathcal{C}_{c+1 \bmod 4}.$$

The first row of M_0 has eleven nonzero entries, with support $\{0, 1, 3, 4, 7, 8, 9, 10, 12, 14, 15\}$, for the two binary-field instances and sixteen nonzero entries for DuX(65537); in both cases the support meets all four classes.

Proof The first rows of [22] satisfy $c_\ell^{(1)} = c_{(\ell+4) \bmod 16}^{(0)}$ for $0 \leq \ell \leq 15$, so by (2) $L_1(X)_p = \sum_\ell c_{\ell+4}^{(0)} X_{p+\ell} = L_0(X)_{p-4}$, which is $\text{Rot}_{12} \circ L_0$ because $12 \equiv -4 \pmod{16}$. The supports of M_0^{-1} and M_1^{-1} are obtained by inverting the matrices over $\mathbb{F}_q$; the second follows from the first because $M_1^{-1} = (\text{Rot}_{12} M_0)^{-1} = M_0^{-1} \text{Rot}_4$. For the class map, note that by (8) the output word p of L_t^{-1} depends only on input words $p + \ell$ with ℓ in the support, and $\{(c + \ell) \bmod 4 \mid \ell \in \{1, 4, 8, 9, 13\}\} = \{c, c + 1\}$. $\square$

The class map of Proposition 1 is the structural fact behind the recursion of Section 3.3. The relation $L_1 = \text{Rot}_{12} \circ L_0$ has a second consequence: the branch pattern only rotates the degree bounds, and does nothing else to them.

Proposition 2 *Fix an active set A and a branch pattern $t = (t_1, \dots, t_{11})$, and let $d_t^{(r)}$ be the vector of degree bounds produced by (8) and (9) after the r -th S-box layer. Then*

$$d_t^{(r)} = \text{Rot}_{-4\sigma_r} d_0^{(r)}, \quad \sigma_r = t_1 + \dots + t_{r-1},$$

for every active set and every number of layers. Consequently the number of words guaranteed by Theorem 1 is the same for all 2^{11} patterns, and the guaranteed positions move by a rotation that maps every class onto itself. If $d_0^{(r)}$ is constant on classes, the guaranteed set itself is the same for all patterns. Moreover $(v M_1)_p = (v M_0)_{p+4}$ for every $v \in \mathbb{F}_q^{16}$, so a mask whose image under the diffusion layer is supported on a union of classes has that property for both layers.

Proof By Proposition 1 the first row of M_1 has support $J - 4$, where J is that of M_0 , so (8) with the branch bit 1 returns $\max_{\ell \in J} d_{p+\ell-4}$, which is Rot_{-4} of the vector returned with the branch bit 0. The S-box layer acts identically on the four blocks, so (9) commutes with $\text{Rot}_{4\ell}$ for every ℓ . Composing the layers moves each rotation to the front and adds the exponents. Round keys do not interfere: $\text{Rot}_{4\ell}(X+k) = \text{Rot}_{4\ell}(X) + \text{Rot}_{4\ell}(k)$, so the conjugation replaces the round-key sequence by its rotation, and (8) and (9) are valid for every value of the round keys, an added constant contributing nothing to a degree. Finally $\text{Rot}_{4\ell}$ maps each class onto itself, and a vector constant on classes is fixed by it. For the last statement, $c_\ell^{(1)} = c_{\ell+4}^{(0)}$ gives $(v M_1)_p = \sum_i v_i c_{p-i+4}^{(0)} = (v M_0)_{p+4}$, and $\mathcal{C}_C - 4 = \mathcal{C}_C$. $\square$

Section 5.1 of [22] argues that the key-dependent choice between L_0 and L_1 prevents an attacker from ignoring the nonlinearity introduced by the key. Proposition 2 shows that for the distinguishers of Section 4 and the key recovery of Section 5 a single degree-propagation computation covers all 2^{11} diffusion patterns, and the choice does not affect the complexity. Two of the active sets used below, namely $\{1, 5, 9\}$ and the eleven-word set of Table 3, are not unions of classes. The degree bounds they produce are nevertheless constant on classes at the layer at which the criterion is applied, so for these sets too the guaranteed positions are the same for all 2^{11} patterns.

3.3 The class-wise degree recursion

Theorem 2 *Consider the decryption direction (5) with active set $A = \mathcal{C}_2 = \{2, 6, 10, 14\}$, so $s = 4$. Then the vector of degree bounds produced by (8) and (9) is constant on each class at every layer. Writing $d^{(r)} = (d_0, d_1, d_2, d_3)$ for the class-wise degree bounds after the r -th inverse S-box layer, one has $d^{(2)} = (2, 3, 4, 2)$, and whenever $d_3 \leq d_0 \leq d_1 \leq d_2$,*

$$d_0^{(r+1)} = 2d_2^{(r)}, \quad d_1^{(r+1)} = d_1^{(r)} + 2d_2^{(r)}, \quad d_2^{(r+1)} = d_1^{(r)} + 3d_2^{(r)}, \quad d_3^{(r+1)} = d_1^{(r)} + d_2^{(r)}, \quad (10)$$

and $d^{(r+1)}$ satisfies the same ordering. Consequently (10) holds for all $r \geq 2$, the smallest of the four bounds is d_3 , and all four grow exponentially with base $2 + \sqrt{3}$ instead of four, by Corollary 1.

Proof If d is constant on classes, say $d_p = \delta_{p \bmod 4}$, then (8) together with Proposition 1 gives $d'_p = \max(\delta_{p \bmod 4}, \delta_{(p+1) \bmod 4})$, which again depends only on $p \bmod 4$; and (9) gives the same value for all four blocks because SL applies the same S^{-1} to each. The initial vector for $A = \mathcal{C}_2$ is $\delta = (0, 0, 1, 0)$, so class-uniformity holds at every layer.

Write $e_c = \max(d_c, d_{c+1 \bmod 4})$ for the class-wise bounds after the inverse diffusion layer. Under $d_3 \leq d_0 \leq d_1 \leq d_2$ this gives $e = (d_1, d_2, d_2, d_0)$. Substituting into (9) with the sets E_j of Lemma 1,

$$\begin{aligned} d'_0 &= \max(e_0, e_1 + e_2) = \max(d_1, 2d_2) = 2d_2, \\ d'_1 &= \max(e_0 + e_1 + e_2, e_1, e_2 + e_3, e_2) = d_1 + 2d_2, \\ d'_2 &= \max(e_0 + 2e_1 + e_2, 2e_0 + e_1, e_1 + e_2 + e_3, \dots) = d_1 + 3d_2, \\ d'_3 &= \max(e_0 + e_1, e_3) = \max(d_1 + d_2, d_0) = d_1 + d_2, \end{aligned}$$

where the maxima are attained as stated because $d_1 \leq d_2$ gives $d_1 + 3d_2 \geq 2d_1 + d_2$, and $d_0 \leq d_1$ gives $d_1 + 3d_2 \geq 2d_2 + d_0$ and $d_1 + d_2 \geq d_0$. The ordering is preserved: $d'_3 \leq d'_0$ is $d_1 + d_2 \leq 2d_2$, and $d'_0 \leq d'_1 \leq d'_2$ is immediate.

For the initial values, the step from C to W_{11} carries no inverse diffusion layer, so (9) applied to $\delta = (0, 0, 1, 0)$ gives $d^{(1)} = (1, 1, 1, 0)$, and one further round with $e = (1, 1, 1, 1)$ gives $d^{(2)} = (2, 3, 4, 2)$, which satisfies the ordering. $\square$

Corollary 1 *The smallest of the four class degree bounds obeys the recursion*

$$d_3^{(2)} = 2, \quad d_3^{(3)} = 7, \quad d_3^{(r+1)} = 4d_3^{(r)} - d_3^{(r-1)} \quad (r \geq 3),$$

so that $d_3^{(r)}$ takes, for $r = 2, \dots, 12$, the values

$$2, 7, 26, 97, 362, 1351, 5042, 18817, 70226, 262087, 978122.$$

Equivalently, $d_3^{(r)} = \frac{1}{2}[(2 + \sqrt{3})^{r-1} + (2 - \sqrt{3})^{r-1}]$ and $d_3^{(r+1)} + \sqrt{3}d_2^{(r)} = (2 + \sqrt{3})^r$ for all $r \geq 2$.

Proof The recursion is obtained from (10) by eliminating the two largest of the four bounds. Subtracting the second identity of (10) from the third gives $d_2^{(r)} - d_1^{(r)} = d_2^{(r-1)}$, hence $d_1^{(r-1)} = d_1^{(r)} - 2d_2^{(r-1)} = 3d_1^{(r)} - 2d_2^{(r)}$ and

$$d_3^{(r)} = d_1^{(r-1)} + d_2^{(r-1)} = 2d_1^{(r)} - d_2^{(r)} \quad (r \geq 2), \quad (11)$$

these steps being available for $r \geq 3$ and the remaining case $r = 2$ being $2 \cdot 3 - 4 = 2$ from the initial vector $d^{(2)} = (2, 3, 4, 2)$. Applying (11) at $r - 1$ and then (10) twice,

$$\begin{aligned} 4d_3^{(r)} - d_3^{(r-1)} &= 4(d_1^{(r-1)} + d_2^{(r-1)}) - (2d_1^{(r-1)} - d_2^{(r-1)}) = 2d_1^{(r-1)} + 5d_2^{(r-1)} \\ &= d_1^{(r)} + d_2^{(r)} = d_3^{(r+1)} \end{aligned}$$

for every $r \geq 3$, which is the recursion; the values follow from $d_3^{(2)} = 2$ and $d_3^{(3)} = 7$ by one multiplication and one subtraction per layer, so that $d_3^{(11)} = 4 \cdot 70226 - 18817 = 262087$ is reached without leaving the integers. The second-order recursion applies from $r = 3$ onward and not below: the vector $d^{(1)} = (1, 1, 1, 0)$ is the initial condition of the propagation and not an output of (10), and $4d_3^{(2)} - d_3^{(1)} = 8$ whereas $d_3^{(3)} = 7$.

Both identities of the statement solve that recursion. For the first, put $u_r = \frac{1}{2}[(2 + \sqrt{3})^{r-1} + (2 - \sqrt{3})^{r-1}]$. Each of $2 + \sqrt{3}$ and $2 - \sqrt{3}$ satisfies $\lambda^2 = 4\lambda - 1$, so each of the two powers, and hence u_r , is four times the previous one minus the one before that; and $u_2 = 2$, $u_3 = 7$ are the two initial values, so $u_r = d_3^{(r)}$ for every $r \geq 2$. For the second identity put $z_r = d_3^{(r+1)} + \sqrt{3}d_2^{(r)}$; by (10), $d_3^{(r+2)} = 2d_1^{(r)} + 5d_2^{(r)}$ and $d_2^{(r+1)} = d_1^{(r)} + 3d_2^{(r)}$, so

$$z_{r+1} = (2 + \sqrt{3})d_1^{(r)} + (5 + 3\sqrt{3})d_2^{(r)} = (2 + \sqrt{3})(d_1^{(r)} + d_2^{(r)} + \sqrt{3}d_2^{(r)}) = (2 + \sqrt{3})z_r,$$

and $z_2 = 7 + 4\sqrt{3} = (2 + \sqrt{3})^2$. $\square$

4 Higher-order differential distinguishers for DuX

4.1 The decryption direction

Corollary 2 *Let $B = C_3 = \{3, 7, 11, 15\}$. For $DuX(2^{16})$ and $DuX(65537)$, the words indexed by B of the state reached by eleven inverse S -box layers have the zero-sum property on every affine subspace of dimension $s = 4$ with active set C_2 , that is on q^4 chosen ciphertexts, for every key and every branch pattern. Within the propagation model of (8) and (9), the criterion of Theorem 1 guarantees no larger number of rounds for any active set.*

Proof By Theorem 2 the class degree bounds after eleven inverse S-box layers are $d^{(11)} = (302632, 413403, 564719, 262087)$. The threshold is $s(q-1) = 4 \cdot 65535 = 262140$ for DuX(2^{16}) and $4 \cdot 65536 = 262144$ for DuX(65537), so class 3 satisfies Theorem 1 with margins 53 and 57 respectively while the other three classes do not. At twelve layers the bound is $d_3^{(12)} = 978122$ by Corollary 1, which exceeds 262140 and therefore guarantees nothing at that number of layers. Exhaustive enumeration of all $2^{16} - 1$ nonempty active sets, with degree bounds propagated by (8) and (9), returns eleven layers as the maximum for which some word is guaranteed. The enumeration is complete because if a proper subset of an active set already satisfies the criterion it requires less data. $\square$

Table 3 lists the optimal active sets per number of rounds, as returned by the same enumeration; the class degree bounds underlying its rows are those of Table A1. The data of a row is one affine subspace of q^s chosen ciphertexts and equals its time, and for DuX(65537) the entries 2^{16} , 2^{32} and 2^{64} denote q , q^2 and q^4 . For DuX(2^8) the model guarantees up to seven rounds: at a fixed number of active words its threshold $s(q-1)$ is smaller by a factor of about 256, while the degrees of S^{-1} do not depend on q . This is the sense in which the larger word size is the weaker choice against the criterion.

Table 3 Higher-order differential distinguishers of DuX in the decryption direction

Instance	Active words s	Data	Rounds	Guaranteed zero-sum words
DuX(2^8)	1	2^8	5	all 16
DuX(2^8)	2	2^{16}	6	$\mathcal{C}_{\{0,3\}}$
DuX(2^8)	11	2^{88}	7	$\mathcal{C}_3$
DuX(2^{16})	1	2^{16}	9	all 16
DuX(2^{16})	2	2^{32}	10	$\mathcal{C}_{\{0,1,3\}}$
DuX(2^{16})	4	2^{64}	11	$\mathcal{C}_3$
DuX(65537)	1	2^{16}	9	all 16
DuX(65537)	2	2^{32}	10	$\mathcal{C}_{\{0,1,3\}}$
DuX(65537)	4	2^{64}	11	$\mathcal{C}_3$

4.2 The encryption direction

In the encryption direction the first row of M_0 meets all four classes by Proposition 1. Once the vector of degree bounds is constant within each class, the maximum in (8) can therefore be computed from one representative bound per class, and one diffusion layer makes the vector constant across all sixteen words. Write D_r for that common value after the r -th diffusion layer. Then (9) and Lemma 1 give

$$d^{(r+1)} = (5, 3, 2, 8) \cdot D_r, \quad D_{r+1} = 8D_r, \quad (12)$$

where $d^{(r+1)}$ is the vector after the $(r+1)$ -st S-box layer. In this case the common value is multiplied by eight at every round, and the smallest of the four bounds is $2D_r$, carried by class 2.

A single active word does not make the vector constant on classes, so (12) governs the trajectory only from the layer at which it becomes so. For the active sets of Corollary 3 that layer is the first diffusion layer for $A = \{1, 5, 9\}$, with $D_1 = 2$, and the second for $A = \{0\}$ and $A = \{1, 5\}$, with $D_2 = 24$ and $D_2 = 16$; the earlier layers are propagated by (8) and (9) directly.

Corollary 3 *For $\text{DuX}(2^{16})$ and $\text{DuX}(65537)$, the words indexed by C_2 after seven rounds of encryption have the zero-sum property on every affine subspace of dimension $s = 3$ with active set $\{1, 5, 9\}$, that is on q^3 chosen plaintexts. After six rounds, twelve words, namely $C_{\{0,1,2\}}$, have the zero-sum property on every subspace of dimension $s = 1$ with active set $\{0\}$, that is on q chosen plaintexts. For $\text{DuX}(2^8)$ the criterion reaches four rounds, with eight zero-sum words on a subspace of dimension $s = 2$ with active set $\{1, 5\}$.*

Proof Exhaustive enumeration of all nonempty active sets gives $D_6 = 2^{16}$ for $A = \{1, 5, 9\}$, hence $d^{(7)} = (327680, 196608, 131072, 524288)$ by (12), against the threshold $3(q - 1)$, which is 196605 for $\text{DuX}(2^{16})$ and 196608 for $\text{DuX}(65537)$. Class 2 is guaranteed with 131072, while class 1 attains 196608 and is guaranteed for neither instance. For $A = \{0\}$ one has $D_5 = 12288$ and $d^{(6)} = (61440, 36864, 24576, 98304)$ against the threshold $q - 1$, so classes 0, 1 and 2 are guaranteed. For $\text{DuX}(2^8)$ and $A = \{1, 5\}$, $D_3 = 128$ and $d^{(4)} = (640, 384, 256, 1024)$ against the threshold 510, so classes 1 and 2 are guaranteed. $\square$

Section 5.3 of [22] estimates that division-property distinguishers for the encryption of $\text{DuX}(2^{16})$ and $\text{DuX}(65537)$ do not exceed six rounds, and Section 5.5.3 of [22] states that seven or more rounds resist higher-order differential attacks, higher-order differentials being covered there by the integral framework. Those statements rest on two types of input structures, and the subspace of our Corollary 3 is neither of them: it is a coordinate subspace of word dimension three whose active set lies in a single class. The first structure is the whole space, in whose image the monomial $\prod_{i=0}^{15} x_i^{q-1}$ is reported after six rounds, so that the threshold $16(q - 1)$ is attained there. The second is a line $(a_0x + b_0, \dots, a_{15}x + b_{15})$ with constants a_i and b_i , along which all sixteen words vary and every output word is reported to have degree $q - 1$ in the parameter x after four rounds, attaining the threshold $q - 1$. Corollary 3 gives seven rounds with q^3 chosen plaintexts. Recorded in the round-number table of [22] in place of the estimate, that value moves integral and higher-order differential attacks from seven rounds to eight, and the difference from the twelve rounds of the cipher from five to four. Both figures are covered-round counts in the convention of that table; the criterion of Theorem 1 guarantees seven rounds in this direction.

Remark 1 The designers of DuX report four six-round integral distinguishers for $\text{DuX}(2^{16})$ and $\text{DuX}(65537)$, each with a single active word and a single balanced position, and estimate on that basis that distinguishers of this kind do not exceed six rounds. Corollary 3 guarantees twelve words at six rounds from the same q chosen plaintexts, which provides a theoretical basis for a distinguisher at that number of rounds. The four positions listed in [22] are not among the guaranteed ones, and the sums at those four positions over the whole subspace are nonzero (Appendix B).

4.3 The twelve-round boundary

Proposition 3 *For every permutation E of $\mathbb{F}_q^{16}$ and every word index j , $\sum_{X \in \mathbb{F}_q^{16}} E(X)_j = q^{15} \sum_{z \in \mathbb{F}_q} z = 0$ for all three values of q . Moreover, for the full codebook as active set, $s = 16$, the degree bounds after twelve inverse S-box layers are $(4215120, 5757961, 7865521, 3650401)$. The threshold $16(q - 1)$ is 4080 for $\text{DuX}(2^8)$, 1048560 for $\text{DuX}(2^{16})$ and 1048576 for $\text{DuX}(65537)$, so no word is guaranteed for any instance.*

Proof Each word of E takes every value of $\mathbb{F}_q$ exactly q^{15} times, and $\sum_{z \in \mathbb{F}_q} z = 0$ because q is even in two instances and the sum is $\frac{q(q-1)}{2} \equiv 0$ for $q = 65537$. The degree bounds follow from (8) and (9) started from $d_p = 1$ for all p . $\square$

The zero sums of Proposition 3 hold for every permutation of $\mathbb{F}_q^{16}$, so the eleven-round distinguisher of Corollary 2 is the one that separates DuX from a random permutation. Section 5 reaches the full cipher from it: a single round separates it from the plaintext, and that round can be removed without guessing.

5 Key recovery on the full DuX

5.1 Strategy of the attack

An r_{HD} -round higher-order differential distinguisher is placed at the ciphertext side and extended by r_{KR} rounds towards the plaintext, with $r = r_{\text{HD}} + r_{\text{KR}}$. Here $r_{\text{HD}} = 11$ and $r_{\text{KR}} = 1$. The attack has three steps.

1. *Choosing the subspaces.* Take n affine subspaces of dimension $s = 4$ with active set $\mathcal{C}_2$, each containing q^4 chosen ciphertexts, and request the corresponding plaintexts. By (5), eleven inverse S-box layers applied to a ciphertext reach W_1 , and by (3),

$$W_1 = L_{t^1}(\text{SL}(P + k^{(0)})) + k^{(1)}, \quad k^{(0)} = K. \quad (13)$$

Corollary 2 says that $\sum (W_1)_p = 0$ for $p \in B = \mathcal{C}_3$, the sum being over one subspace.

2. *Constructing the key-dependent equation system.* Let v be a row vector over $\mathbb{F}_q$ with $\text{supp}(v) \subseteq B$, that is the coefficient vector of an $\mathbb{F}_q$ -linear form $x \mapsto v \cdot x$ on the zero-sum words, and let $w = v M_{t^1}$ be the mask obtained by transporting v through the diffusion layer. Summing $v \cdot W_1$ over the subspace, the contribution of $k^{(1)}$ is $q^s (v \cdot k^{(1)}) = 0$ because the characteristic divides q^s , and (13) gives the key equation

$$G_w(a) = \sum_{\text{texts}} \sum_{k=0}^3 \sum_{j=0}^3 w_{4k+j} S(P^{(k)} + a^{(k)})_j = 0, \quad a = K. \quad (14)$$

Here $P^{(k)}$ and $a^{(k)}$ denote the four words of block k of the plaintext and of the unknown key, and the two inner sums run over the four S-box blocks and the four coordinates of each, as in $p = 4k + j$. Equation (14) involves the encryption S-box S , whose coordinate degrees are $(5, 3, 2, 8)$ by Lemma 1. Only the plaintexts of the

subspace and the vector w enter it; no key material other than a appears, and no round key of the r_{HD} covered rounds is guessed.

3. *Solving the system.* We choose masks v for which the transformed mask $w = v M_{t1}$ has its support in coordinates of low degree. For $C \subseteq \{0, 1, 2, 3\}$ set

$$V_C = \{ v \in \mathbb{F}_q^{16} \mid \text{supp}(v) \subseteq B, \text{supp}(v M_t) \subseteq C_C \}. \quad (15)$$

Solving these conditions over $\mathbb{F}_q$ gives $\dim V_{\{2\}} = 1$, $\dim V_{\{1,2\}} = 2$, $\dim V_{\{0,1,2\}} = 3$ and $\dim V_{\{0,1,2,3\}} = 4$ for the two binary-field instances, and 1, 1, 1, 4 for DuX(65537). The spaces themselves are the same for $t = 0$ and $t = 1$ by the last statement of Proposition 2. Sorting the masks by the support of w splits (14) into the three systems of Theorem 3, which are solved in turn.

5.2 Three-stage elimination

Write $\Sigma P_0^{(k)} = \sum P_{4k}$ and $\Sigma P_3^{(k)} = \sum P_{4k+3}$ for the sums, over the q^s plaintexts of the subspace, of the plaintext word of block k in coordinates 0 and 3, and $\Sigma y_j^{(k)} = \sum S(P^{(k)} + a^{(k)})_j$.

Theorem 3 *Fix an affine subspace with active set C_2 . Then, in (14):*

- (i) *for $v \in V_{\{2\}}$, the equation is $\mathbb{F}_q$ -affine in the eight unknowns $a_{4k}, a_{4k+3}, 0 \leq k \leq 3$, the coefficient of a_{4k} being $-w_{4k+2}\Sigma P_3^{(k)}$ and that of a_{4k+3} being $-w_{4k+2}\Sigma P_0^{(k)}$;*
- (ii) *for $v \in V_{\{1,2\}}$, after substituting values for the unknowns of (i), the equation is $\mathbb{F}_q$ -affine in the four unknowns a_{4k+2} with coefficient $-w_{4k+1}\Sigma P_3^{(k)}$, and is independent of a_{4k+1} ;*
- (iii) *for $v \in V_{\{0,1,2,3\}}$, after substituting values for the unknowns of (i) and (ii), the equation is $\mathbb{F}_q$ -affine in the eight quantities a_{4k+1} and a_{4k+1}^2 , with*

$$\begin{aligned} \text{coeff}(a_{4k+1}) &= w_{4k+3}(2\Sigma(y_2\bar{y}_1)^{(k)} - \Sigma P_0^{(k)}) - w_{4k}\Sigma y_2^{(k)}, \\ \text{coeff}(a_{4k+1}^2) &= w_{4k+3}\Sigma y_2^{(k)}, \end{aligned} \quad (16)$$

where $\bar{y}_1$ is the value of y_1 in (7) at $a_{4k+1} = 0$. In even characteristic the first coefficient is $w_{4k+3}\Sigma P_0^{(k)} - w_{4k}\Sigma y_2^{(k)}$.

A mask in $V_{\{0,1,2,3\}}$ whose image meets both C_3 and C_0 receives both contributions to the linear coefficient of (16), and they add; the coefficient reduces to a single term exactly when $w_{4k} = 0$ or $w_{4k+3} = 0$. Every coefficient listed is computable from the plaintexts of the subspace together with the key values recovered at the earlier stages.

Proof All three parts expand (7) with $x_j = P_{4k+j} + a_{4k+j}$. A term that depends only on the key is constant across the subspace, so its sum is q^s times that constant and vanishes; the terms that retain a plaintext factor are summed explicitly below.

(i) By (7), $y_2 = x_2 - x_0x_3 - \alpha$. Summing, $\Sigma y_2 = \Sigma P_{4k+2} - \Sigma P_{4k}P_{4k+3} - a_{4k+3}\Sigma P_0^{(k)} - a_{4k}\Sigma P_3^{(k)}$, the terms $q^s a_{4k+2}$, $q^s a_{4k}a_{4k+3}$ and $q^s \alpha$ having vanished. Weighting by w_{4k+2} and summing over k gives the stated affine form; the unknown a_{4k+2} has dropped out because it occurs only linearly in y_2 .

(ii) Substituting y_2 into (7) gives $y_1 = x_1 - x_2x_3 + x_0x_3^2 + \alpha x_3 - \alpha$. The unknown a_{4k+1} occurs only in the linear term x_1 and drops out as above. Once a_{4k} and a_{4k+3} have been recovered, the only remaining unknown is a_{4k+2} , which occurs in $-x_2x_3$ with coefficient $-\Sigma P_{4k+3} = -\Sigma P_3^{(k)}$. Masks in $V_{\{1,2\}}$ also carry a component on $\mathcal{C}_2$, which is already determined by (i).

(iii) Write $y_1 = x_1 + \tilde{w}$ with $\tilde{w} = -x_3y_2 - \alpha$, so that $y_0 = x_0 - y_2y_1 - \alpha = u - y_2x_1$ with $u = x_0 - y_2\tilde{w} - \alpha$. Since y_2 and u involve only the unknowns of (i) and (ii), summing y_0 over the subspace leaves the coefficient $-\Sigma y_2^{(k)}$ for a_{4k+1} , whence the stated $\mathcal{C}_0$ coefficient. For $\mathcal{C}_3$, expanding $y_3 = x_3 - y_1y_0 - \alpha$ with the same substitutions gives $y_3 = x_3 + y_2x_1^2 + (2y_2\tilde{w} - x_0 + \alpha)x_1 - \tilde{w}u - \alpha$. Summing and using $x_1^2 = P_{4k+1}^2 + 2P_{4k+1}a_{4k+1} + a_{4k+1}^2$, the coefficient of a_{4k+1}^2 is $\Sigma y_2^{(k)}$ and that of a_{4k+1} is $2\Sigma(y_2(P_{4k+1} + \tilde{w}))^{(k)} - \Sigma P_0^{(k)}$, since $\Sigma(-x_0 + \alpha) = -\Sigma P_0^{(k)}$. The factor of y_2 there is $\bar{y}_1$. In even characteristic the doubled term vanishes and the coefficient is $\Sigma P_0^{(k)}$. Coordinates 1 and 2 contribute nothing to either coefficient: y_2 does not involve x_1 , and in y_1 the variable x_1 occurs only linearly and sums out. Hence in (14) the coefficient of a_{4k+1} collects the $\mathcal{C}_0$ term with weight w_{4k} and the $\mathcal{C}_3$ term with weight w_{4k+3} , which gives (16). $\square$

Stage (i) determines the eight words of K in classes $\mathcal{C}_0$ and $\mathcal{C}_3$, stage (ii) the four words in $\mathcal{C}_2$ and stage (iii) the four words in $\mathcal{C}_1$, so the full master key is obtained. Stage (iii) is solved by linearizing, that is, by treating a_{4k+1} and a_{4k+1}^2 as separate unknowns. The linearization carries one relation with it. Proposition 4 shows that the four square columns are proportional to a vector whose entries sum to zero on every subspace, so one of the eight columns is a combination of the others. The number of *effective columns*, by which we mean the columns remaining once such a dependency is removed, is therefore seven. If the accumulated matrix attains rank seven, its kernel is spanned by the vector η of Proposition 4(iii), which is supported on the square unknowns. The resulting ambiguity leaves the four quantities a_{4k+1} unchanged, and those are what stage (iii) has to determine.

For DuX(65537) with $B = \mathcal{C}_3$, $\dim V_{\{1,2\}} = 1$ and the single direction does not meet $\mathcal{C}_1$, so stage (ii) is unavailable. There, stages (ii) and (iii) are replaced by exact bivariate interpolation of (14) on a three-by-four grid of key candidates, which yields four masks per subspace. The grid carries $4 \cdot 11 = 44$ non-constant monomials, which are not 44 independent unknowns. Once a_{4k} and a_{4k+3} are determined by stage (i), the two remaining unknowns $u_1 = a_{4k+1}$ and $u_2 = a_{4k+2}$ enter (7) with total degree at most three: y_2 and y_1 are affine in (u_1, u_2) and $y_3 = x_3 - (x_0 - \alpha)y_1 + y_2y_1^2 - \alpha$. Hence $u_1u_2^3$, $u_1^2u_2^2$ and $u_1^2u_2^3$ do not occur, and the coefficient of $u_1^2u_2$ is a constant independent of the plaintexts, so it is multiplied by $q^s = 0$ on summing over a subspace. Four columns per block vanish identically and the effective number of columns is at most $44 - 16 = 28$.

The four masks of a subspace do not contribute four independent rows to that system. The mask whose direction lies in $V_{\{2\}}$ has its w supported on $\mathcal{C}_2$, so its equation involves only $y_2 = x_2 - x_0x_3 - \alpha$ of (7). There u_1 does not occur, and u_2 occurs

linearly with a constant coefficient, which is therefore multiplied by $q^s = 0$ on summing over the subspace. That row vanishes on all 28 columns, so a subspace contributes $\rho = \dim V_{\{0,1,2,3\}} - \dim V_{\{2\}} = 3$ rows and no more.

5.3 The rank condition

The linearization of stage (iii) carries one relation among its columns, and the relation is structural rather than accidental.

Proposition 4 *Let q be even, let $B = C_3$, let $t \in \{0, 1\}$, let $v \in V_{\{0,1,2,3\}}$, put $w = vM_t$ and $\tau = \sum_{i \in B} v_i$. Then*

- (i) $w_{4k+3} = \tau$ for every k , so the four square coefficients of (16) are $\tau \Sigma y_2^{(k)}$;
- (ii) $\sum_{k=0}^3 \Sigma y_2^{(k)} = 0$ on every subspace used by the attack;
- (iii) $\eta = (0, 0, 0, 0, 1, 1, 1, 1)$ lies in the kernel of the stage-(iii) coefficient matrix, whose rank is therefore at most 7;
- (iv) w_{4k} takes only two values, $w_0 = w_8$ and $w_4 = w_{12}$, and $\tau = w_0 + w_4$; a single subspace consequently contributes at most two independent rows.

Proof Both M_0 and M_1 are circulant with entries in $\{0, 1\}$, so $(M_t)_{i,j}$ depends only on $j - i \bmod 16$. On rows B and columns C_3 every entry equals 1, which is (i). On rows B and columns C_0 the rows are alternately $(0, 1, 0, 1)$ and $(1, 0, 1, 0)$, the two patterns exchanging with t ; the submatrix has rank two, so $w_0 = w_8$ and $w_4 = w_{12}$, and summing the two columns gives $w_0 + w_4 = \sum_{i \in B} v_i = \tau$. By (i) and (16) the pair (w_0, w_4) determines the whole row, whence (iv).

For (ii) take v to be the all-ones vector on B . By the previous paragraph vM_t is the indicator vector of C_2 , so this v spans $V_{\{2\}}$ and its stage-(i) equation reads $\sum_{k=0}^3 \Sigma y_2^{(k)} = 0$, the coordinate-2 terms of the four S-box blocks being the words indexed by C_2 . Every subspace used by the attack satisfies the stage-(i) equations. Pairing a row of (16) with η then gives $\tau \sum_k \Sigma y_2^{(k)} = 0$, which is (iii). $\square$

The kernel direction η vanishes on the first four coordinates, those of the unknowns a_{4k+1} . If the coefficient matrix attains rank 7 — the largest value permitted by Proposition 4(iii) — its kernel is spanned by η alone, so the solution set is a coset of $\langle \eta \rangle$ and is constant in those four coordinates, which determines a_1, a_5, a_9, a_{13} uniquely. What the attack has to reach is this maximal rank.

Proposition 5 *All coefficients of the unknowns listed in Theorem 3 are, up to the known factors w_p and up to quantities determined at earlier stages, the sums $\Sigma P_0^{(k)} = \sum P_{4k}$ and $\Sigma P_3^{(k)} = \sum P_{4k+3}$ of the plaintext words returned for the subspace. In particular they can be computed from the observed plaintext words, so the rank of each stage's coefficient matrix can be checked during the attack, and further subspaces can be requested while it falls short of the maximum.*

Proof This is Theorem 3, whose coefficients are exactly $w_{4k+2}\Sigma P_3^{(k)}$, $w_{4k+2}\Sigma P_0^{(k)}$, $w_{4k+1}\Sigma P_3^{(k)}$ and the two entries of (16); the vector $w = vM_t$ is known and $\Sigma y_2^{(k)}$ is determined once stages (i) and (ii) are solved. The plaintexts of a subspace are the decryptions of q^s chosen ciphertexts, so they are observed. $\square$

Nonzero coefficients alone do not guarantee maximal rank, as the all-ones matrix of size 10×7 shows. Maximal rank is therefore used as a condition checked during the attack, in the sense of Proposition 5: a subspace whose row does not increase the rank of the accumulated matrix is replaced by a further subspace. The complexity estimate below allows for two extra subspaces per stage.

5.4 Complexity

Time is measured in computations of the full twelve-round DuX decryption and splits as $T = T_c + T_s$, the cost of constructing the system and the cost of solving it. For each subspace the attack makes a single pass over its q^s texts, accumulating sums of products of plaintext words of total degree at most eight: at most 49 such sums per S-box block in characteristic two and at most 74 for DuX(65537). Among them are the constant terms of the three stages, the coefficients of the unknowns being the sums of Proposition 5. Hence $T_c = n q^s$ for n subspaces.

The solve is $\mathbb{F}_q$ -linear algebra on a system of at most eight columns in characteristic two, and of at most 44 columns for DuX(65537); in either case it is independent of q^s , so $T_s = O(1)$ and $T = T_c$ to the precision quoted. The data is likewise $n q^s$, and the pass being streaming, the memory of the algorithm is a constant number of field elements.

Write c_j for the number of effective columns at stage j and ρ_j for the number of independent rows that a single subspace contributes there. The number of subspaces needed to attain the maximal rank, with two spare subspaces, is

$$n = \max_j \lceil c_j / \rho_j \rceil + 2. \quad (17)$$

For $B = \mathcal{C}_3$, Theorem 3 and Proposition 4(iii) give $c = (8, 4, 7)$, while (15) and Proposition 4(iv) give $\rho = (\dim V_{\{2\}}, \dim V_{\{1,2\}} - \dim V_{\{2\}}, 2) = (1, 1, 2)$. Hence $n = \max\{8, 4, 4\} + 2 = 10$; the maximum is attained at the first stage, which therefore determines the number of subspaces.

Corollary 4 *Let each stage attain the maximal rank permitted by its structure, a condition that can be checked during the attack by Proposition 5. Then all sixteen words of the master key are recovered with the data and time of Table 4 and a constant number of field elements of memory. In particular the full twelve rounds of DuX(2^{16}) and DuX(65537) are covered with $2^{67.32}$ and $2^{67.59}$ chosen ciphertexts and equal time, against the claimed security level of 2^{128} , and DuX(2^8) is covered up to eight rounds with $2^{91.32}$.*

Proof Combine Corollary 2 and Table 3 with Theorem 3, and evaluate (17) on the dimensions of (15). A shallower distinguisher has a larger zero-sum set B , hence a larger $\dim V_{\{2\}}$ and

Table 4 Master-key recovery on DuX

Instance	Round	r_{HD}	Active words s	Subspaces	Data = time
DuX(2^8)	6/12	5	1	4	$2^{10.00}$
DuX(2^8)	7/12	6	2	10	$2^{19.32}$
DuX(2^8)	8/12	7	11	10	$2^{91.32}$
DuX(2^{16})	10/12	9	1	4	$2^{18.00}$
DuX(2^{16})	11/12	10	2	10	$2^{35.32}$
DuX(2^{16})	12/12	11	4	10	$2^{67.32}$
DuX(65537)	10/12	9	1	4	$2^{18.00}$
DuX(65537)	11/12	10	2	10	$2^{35.32}$
DuX(65537)	12/12	11	4	12	$2^{67.59}$

r_{HD} is the number of rounds of the distinguisher.

fewer subspaces, and a larger ρ_3 , the rank of $v \mapsto (w_{4k}, w_{4k+3})_{0 \leq k \leq 3}$ on $V_{\{0,1,2,3\}}$. For B equal to all sixteen words the dimensions are 4, 8, 16 with $\rho_3 = 8$ and $n = 4$; for $|B| = 8$ and $|B| = 12$ they are 1, 2, 8 with $\rho_3 = 6$ and 1, 4, 12 with $\rho_3 = 8$, and $n = 10$; for $B = C_3$ they are 1, 2, 4 with $\rho_3 = 2$ and $n = 10$. In odd characteristic $c_3 = 8$, which changes no line. For DuX(65537) with $B = C_3$ the interpolation route of Section 5.2 replaces stages (ii) and (iii) with a single system of at most $c = 28$ effective columns, to which each subspace contributes $\rho = \dim V_{\{0,1,2,3\}} - \dim V_{\{2\}} = 3$ independent rows, so that (17) gives $n = \max\{\lceil 8/1 \rceil, \lceil 28/3 \rceil\} + 2 = 12$. $\square$

Each subspace of Table 4 consists of q^s chosen ciphertexts. The subspace counts there are the values of (17) and include two spare subspaces per stage; they are the counts at which each stage attains the maximal rank permitted by its structure, a condition that can be checked during the attack by Proposition 5. In the last line the counts of the interpolation route of Section 5.2 enter (17) in place of those of stages (ii) and (iii). Time is in computations of the full twelve-round DuX decryption and the memory is a constant number of field elements.

The attack on DuX(2^8) covers eight of its twelve rounds: its deepest guaranteed distinguisher has seven rounds and the technique of Section 5.1 removes one more.

For comparison, Table 5 lists the attacks of Ni, Wang and Li [21] on Yu₂X, the family that precedes DuX. Appendix B of [22] states that in even characteristic the decryption S-box of DuX is affine equivalent to the corresponding S-box of YuX. On the small fields $\mathbb{F}_q$ with $q = 2^2$ and $q = 2^3$, the differential spectrum, the Walsh spectrum, the row distribution of the difference distribution table and the mask distribution of the Walsh table coincide for the two S-boxes. The diffusion layers and the key schedules of the two families differ, so Table 5 is a comparison of orders of magnitude rather than of the same primitive. The subspaces of [21] are $\mathbb{F}_2$ -subspaces, so its data complexities are not restricted to powers of the field size q . The denominator b there is the general recommendation of [20], which is twelve rounds for Yu₂X-8 and fourteen for Yu₂X-16. For homomorphic applications [20] recommends twelve rounds of Yu₂X-16, and the last line of Table 5 covers that number of rounds.

Table 5 Attacks on Yu₂X [21]

Instance	Attack type	Round	Time	Data
Yu ₂ X-8	HD distinguisher	6/12	2^{31}	2^{31} CC
	Key recovery	7/12	2^{49}	2^{48} CC
	Key recovery	10/12	$2^{115.3}$	$2^{95.5}$ CC
Yu ₂ X-16	HD distinguisher	10/14	2^{63}	2^{63} CC
	Key recovery	11/14	2^{96}	2^{96} CC
	Key recovery	12/14	$2^{120.4}$	$2^{120.4}$ CC

a/b , HD and CC are as in Table 1.

6 Zero-sum partitions

In the known-key model the propagation may start from an intermediate state and run in both directions, so the criterion applies to the two directions separately and the numbers of rounds add [27, 28]. The asymmetry of Lemma 1 is directly exploitable here: the backward direction, in which the degree bounds grow more slowly, is given more rounds than the forward one.

The alignment of the layers is what fixes the count. Taking the intermediate state immediately after a diffusion layer, the forward part consists of r_f groups SL, L , SL, $\dots$ and the backward part of r_b groups each carrying a leading inverse diffusion layer, as in (5). A zero-sum partition requires all sixteen words to have the zero-sum property simultaneously on the plaintext side and on the ciphertext side. The round function of (3) is iterated beyond twelve rounds below, while the key schedule of [22] produces twelve round keys. The statement is therefore about that iterated structure with arbitrary fixed round keys $k^{(i)}$ and arbitrary branch bits t^i , of which DuX itself is the twelve-round member.

Corollary 5 *Let $V_A \subseteq \mathbb{F}_q^{16}$ be the coordinate subspace spanned by the words of an active set A , and place the intermediate state immediately after a diffusion layer. The q^{16-s} cosets of V_A partition the intermediate-state space into sets of q^s elements, and their backward and forward images partition the plaintext space and the ciphertext space. For the following active sets, all sixteen words sum to zero on both images of every coset.*

- (i) *For DuX(2^{16}) and DuX(65537), $A = \{0\}$ and $s = 1$ give thirteen rounds, split as $r_b + r_f = 8 + 5$. The class degree bounds are (10084, 13775, 18817, 8733) after the eight backward groups and (7680, 4608, 3072, 12288) after the five forward groups, both against the threshold $q - 1$.*
- (ii) *For the same two instances, $A = C_0$ and $s = 4$ give sixteen rounds, split as $10 + 6$, with bounds (140452, 191861, 262087, 121635) backwards and (61440, 36864, 24576, 98304) forwards against the threshold $4(q - 1)$.*
- (iii) *For DuX(2^8), $A = \{0, 4\}$ and $s = 2$ give eight rounds, split as $5 + 3$, with bounds (194, 265, 362, 168) and (120, 72, 48, 192) against the threshold $2(q - 1) = 510$; and $A = C_{\{0,1,2\}}$ and $s = 12$ give ten rounds, split as $6 + 4$, with bounds*

(1560, 2131, 2911, 1351) and (960, 576, 384, 1536) against the threshold $12(q - 1) = 3060$.

Proof The cosets of V_A are exactly the affine subspaces of Definition 2 with active set A , and they partition $\mathbb{F}_q^{16}$. The backward map from the intermediate state to the plaintext and the forward map from it to the ciphertext are bijections, so the images of the cosets partition the plaintext space and the ciphertext space into sets of q^s texts each, one per coset. Applying Theorem 1 to a coset in the two directions separately, with (8) and (9) for S^{-1} backwards and for S forwards, guarantees all sixteen words of both images once the listed bounds fall below the threshold, which is the content of (i)–(iii). Those bounds hold for every coset, so every part of the partition has the property. The active sets are those maximizing $r_b + r_f$ among all sets of the stated size. The trivial bound $\deg(G \circ F) \leq \deg(G) \cdot \deg(F)$, iterated over the layers with the maximal coordinate degree of the S-box, reads $4^{r_b} < 16(q - 1)$ backwards and $8^{r_f} < 16(q - 1)$ forwards under the full-codebook threshold and gives $5 + 3 = 8$ for $\text{DuX}(2^8)$ and $9 + 6 = 15$ for the other two instances. The exact propagation is better in the backward direction, where the bounds grow exponentially with base $2 + \sqrt{3}$ instead of four by Theorem 2, and the listed splits are what it returns. $\square$

Both large-word instances therefore admit zero-sum partitions over thirteen rounds of the iterated structure, one more than the twelve rounds of DuX , using sets of only q texts. By Proposition 2 the branch bits do not affect this: the degree bounds for any branch pattern are a rotation of those for the all-zero pattern, and the condition that all sixteen words be guaranteed is invariant under every word rotation.

7 Experimental verification

The results of Sections 4–6 are analytic. The entries of Tables 3 and 4 are outputs of the propagation rules and of (17). The eleven-round distinguisher and the twelve-round key recovery hold at their stated data complexities, the key recovery under the rank condition that can be checked during the attack by Proposition 5. The attacks on Yu_2X in Table 5 are the analytic results of [21]. Four further checks are carried out on reduced-round DuX at sizes that can be enumerated: the splits of Corollary 5, the balanced positions after six rounds of encryption, the three stages of the elimination, and the extra round supplied by the class recursion.

At the splits of Corollary 5, the observed zero sums agree with the guaranteed ones in both directions on reduced-round DuX with round keys generated by the key schedule of [22]. At $5 + 3$ rounds all sixteen words sum to zero on both sides; at $6 + 3$ rounds the plaintext side no longer sums to zero while the ciphertext side still does, and at $5 + 4$ rounds the reverse holds. The side that fails is in each case the side for which Theorem 1 guarantees nothing, so the leading inverse diffusion layer of (5) is exercised by the check. The six outcomes are identical for t fixed to 0 and to 1, as Proposition 2 requires.

The guaranteed zero-sum sets were verified to be exactly the zero-sum sets by direct summation at six rounds of encryption of $\text{DuX}(2^{16})$ with $s = 1$, that is q plaintexts, for three random master keys. The results appear in Appendix B, where the criterion guarantees eight positions and direct summation finds those same eight.

The three-stage elimination of Theorem 3 was verified experimentally on a reduced-round configuration in which the zero-sum equations are evaluated at the analogue of the state W_1 of (13) and the set guaranteed there by Theorem 1 is exactly $\mathcal{C}_3$. The subspaces V_C of (15) then have the dimensions $\dim V_{\{2\}} = 1$, $\dim V_{\{1,2\}} = 2$ and $\dim V_{\{0,1,2\}} = 3$ that they have at twelve rounds, so the elimination passes through the same three systems in the same order, eliminating the same unknowns at each stage. The configuration is DuX(2^8) with the six-round distinguisher, $s = 3$ and seven rounds attacked, with 2^{24} ciphertexts per subspace and 10, 4 and 4 subspaces at the three stages. Six inverse S-box layers and five inverse diffusion layers lead from a ciphertext to W_1 here, against eleven and ten at twelve rounds.

The four states of (13) were summed directly over one subspace for three random master keys; Table C3 of Appendix C gives the class degree bounds, the sets guaranteed by Theorem 1 and the words observed to be balanced. The criterion guarantees $\mathcal{C}_3$ in W_1 and no word in any of the other three states, and the masks used are those supported on $\mathcal{C}_3$. The equations are therefore read at the layer at which the criterion guarantees a zero sum and with the masks that it provides. The two relations that the states satisfy hold exactly: the sums of P and of $P + K$ agree word by word, since the added key constant sums to $q^s K = 0$, and $\sum W_1 = L_{t^1}(\sum \text{SL}(P + K))$, since $q^s k^{(1)} = 0$.

For each of three independent random master keys, stage (i) returned the eight key words in $\mathcal{C}_{\{0,3\}}$, stage (ii) the four in $\mathcal{C}_2$ and stage (iii) the four in $\mathcal{C}_1$. All sixteen agree with the key in use, at a total cost of $10 \cdot 2^{24} = 2^{27.3}$ chosen ciphertexts. At every stage of all three runs the words that the stage is to determine were uniquely determined, the coefficient matrix having attained the maximal rank permitted by Proposition 4, checked as in Proposition 5.

The interpolation route that replaces stages (ii) and (iii) over a prime field was verified experimentally on six configurations. In each of them the set guaranteed at the analogue of W_1 is again exactly $\mathcal{C}_3$, the criterion guarantees no word one layer earlier or at the plaintext, and the subspaces of (15) have the dimensions $\dim V_{\{2\}} = 1$ and $\dim V_{\{0,1,2,3\}} = 4$ that they have at twelve rounds. Three of them have a small field, namely DuX(29) with the four-round distinguisher, five rounds attacked and $s = 1$; DuX(101) with the five-round distinguisher, six rounds attacked and $s = 1$; and DuX(53) with the five-round distinguisher, six rounds attacked and $s = 2$. The other three have a field of the order of magnitude of 65537 itself. With the nine-round distinguisher, ten rounds attacked and $s = 1$, the three conditions hold exactly for the primes q with $18818 \leq q - 1 < 21728$: the class bounds after nine layers are (21728, 29681, 40545, 18817) by Table A1 and the threshold at $s = 1$ is $q - 1$. The smallest, a middle one and the largest of the 291 such primes are 18839, 20287 and 21727. For three independent random master keys in each configuration, stage (i) returned the eight key words in $\mathcal{C}_{\{0,3\}}$ and the interpolation stage the remaining eight, all sixteen agreeing with the key in use. In each of the eighteen runs the accumulated matrix had 44 columns and rank 28, and the unknowns it determined were the twenty-eight monomials $a_{4k+1}^i a_{4k+2}^l$ with $(i, l) \notin \{(1, 3), (2, 1), (2, 2), (2, 3)\}$. The four left undetermined in each block are exactly the four that Section 5.2 removes in passing from 44 columns to 28. Repeating each run with the number of subspaces increased

from eight, the sixteen words were determined at ten subspaces and at no smaller number, which is the $\lceil c/\rho \rceil = \lceil 28/3 \rceil$ of (17) for $c = 28$ effective columns and $\rho = 3$ rows per subspace. The count in Table 4 adds the two spare subspaces.

The condition that the guaranteed set be exactly $\mathcal{C}_3$ is what fixes c at 28. When the field is large enough for the criterion to guarantee the class 2 as well at the distinguisher layer, the column of a_{4k+1}^2 vanishes with the other four and c drops to 24. This happens from $q - 1 \geq 40546$ on, and in particular at $q = 65537$ with the nine-round distinguisher, which is therefore not a configuration in which the twelve-round system can be read. At twelve rounds the eleven-round distinguisher guarantees $\mathcal{C}_3$ and no more, which is the case for the six configurations above.

Table 6 collects the twenty-one key recoveries; in each of them all sixteen key words were determined and agreed with the key in use. Table C6 of Appendix C gives the growth of the accumulated rank with the number of subspaces. The per-stage counts that enter (17) in the two routes are listed in Table C4 of the same appendix, and the coefficients of Proposition 5 at stage (iii) in Table C5.

Table 6 Key recoveries measured on reduced-round configurations

Instance	Route	r_d	r	s	Subspaces	Data	Keys tested
DuX(2 ⁸)	three-stage	6	7	3	10	2 ^{27.3}	3
DuX(29)	interpolation	4	5	1	12	2 ^{8.4}	3
DuX(101)	interpolation	5	6	1	12	2 ^{10.2}	3
DuX(53)	interpolation	5	6	2	12	2 ^{15.0}	3
DuX(18839)	interpolation	9	10	1	12	2 ^{17.8}	3
DuX(20287)	interpolation	9	10	1	12	2 ^{17.9}	3
DuX(21727)	interpolation	9	10	1	12	2 ^{18.0}	3

r_d and r are the numbers of rounds of the distinguisher and attacked, s the dimension of a subspace of chosen ciphertexts and the last column the number of independent random master keys.

The rounds guaranteed by Theorem 2 were compared with direct summation on reduced word sizes, where the guaranteed number of rounds is small enough for a subspace to be enumerated. The reduced instance keeps the sixteen-word state, the S-box and the two diffusion matrices, whose entries are 0 and 1, so that the propagation rules depend only on supports that do not change with q . Reducing the word size to n bits gives a $16n$ -bit block with threshold $4(q - 1)$ for $q = 2^n$. The comparison is with the trivial bound, which is $4^{r-2}(2, 3, 4, 2)$ at $r \geq 2$ inverse S-box layers, the exact vector after two layers multiplied by the maximal coordinate degree 4 of S^{-1} once per further layer. Table 7 lists the two bounds and the words observed to sum to zero in every run of a line, at the word sizes $n = 3, 5, 7, 8, 9$, that is at block lengths from 48 to 144 bits. Each sum is over q^4 ciphertexts, namely 2^{12} , 2^{20} , 2^{28} , 2^{32} and 2^{36} .

At every word size the layer at which the two bounds part carries zero sums, and the words that sum to zero there are the words guaranteed by Theorem 2. These are, at $n = 3$, the four words of $\mathcal{C}_3$ at four layers, and at $n = 5, 7, 9$ the eight words of $\mathcal{C}_{\{0,3\}}$ at five, six and seven layers, with a ninth word summing to zero in one of the two runs

at $n = 7$. The trivial bound is above the threshold in all four classes at every one of those layers and guarantees no word there, so each of these zero sums is guaranteed by Theorem 2 alone. One layer further neither bound guarantees a word and none of the sixteen sums to zero. The active set $\mathcal{C}_3$ was measured with the same outcomes as $\mathcal{C}_2$: at every layer for $n = 3, 5, 7$, and at the parting layer for $n = 9$. At $n = 3$ the active sets $\mathcal{C}_0$ and $\mathcal{C}_1$ were measured as well, where neither bound guarantees a word at four layers and none of the sixteen sums to zero. At $n = 8$ the two bounds give the same number of layers, six, and there the class-wise bound guarantees four words more than the trivial bound. The same parting occurs at the full word size. At ten inverse S-box layers the trivial bound is $4^8(2, 3, 4, 2) = (131072, 196608, 262144, 131072)$, which guarantees the twelve words of $\mathcal{C}_{\{0,1,3\}}$ against the threshold 262140 of DuX(2^{16}); at eleven layers it is $4^9(2, 3, 4, 2)$, which guarantees none. The eleven-round distinguisher of Table 3 is therefore guaranteed by Theorem 2 alone.

Table 7 Guaranteed and observed zero sums on reduced word sizes, $A = \mathcal{C}_2$ and $s = 4$

n	$4(q-1)$	r	Keys tested	Class-wise bound	Trivial bound	Observed
3	28	3	16	all 16	all 16	all 16
		4	16	$\mathcal{C}_3$	none	$\mathcal{C}_3$
		5	16	none	none	none
5	124	4	3	all 16	all 16	all 16
		5	3	$\mathcal{C}_{\{0,3\}}$	none	$\mathcal{C}_{\{0,3\}}$
		6	3	none	none	none
7	508	5	2	all 16	all 16	all 16
		6	2	$\mathcal{C}_{\{0,3\}}$	none	$\mathcal{C}_{\{0,3\}}$
		7	2	none	none	none
8	1020	6	1	all 16	$\mathcal{C}_{\{0,1,3\}}$	all 16
		7	1	none	none	none
9	2044	6	1	all 16	all 16	all 16
		7	2	$\mathcal{C}_{\{0,3\}}$	none	$\mathcal{C}_{\{0,3\}}$
		8	1	none	none	none

r is the number of inverse S-box layers and the fourth column the number of independent random master keys. The class-wise bound is that of Theorem 2; the two columns give the words each bound guarantees.

8 Conclusion

Bounding the word-wise algebraic degree of DuX in the decryption direction guarantees eleven-round higher-order differential distinguishers for DuX(2^{16}) and DuX(65537) with q^4 chosen ciphertexts. Removing one round in front of such a distinguisher recovers all sixteen words of the master key of the full twelve-round versions with $2^{67.32}$ and $2^{67.59}$ chosen ciphertexts, equal time and constant memory. The key recovery requires each stage of the elimination to attain the maximal rank permitted by its structure, a condition that can be checked during the attack by Proposition 5.

Both statements are in the chosen-ciphertext model, and under them the round margin of these two instances against this criterion is zero.

Two structural features carry the result. The S-box has coordinate degrees $(2, 3, 4, 2)$ in the decryption direction against $(5, 3, 2, 8)$ in the encryption direction, and each row of the inverse diffusion matrix has five nonzero entries confined to two classes modulo four. The latter makes the class degree bounds satisfy an exact recursion under which they grow exponentially with base $2 + \sqrt{3}$ instead of four. Since the threshold $s(q - 1)$ grows linearly in q while these degrees do not depend on q , the larger word size is the weaker choice: the distinguisher covers seven rounds of DuX(2^8) and the key recovery eight, against eleven and twelve rounds of the two large-word instances.

In the chosen-plaintext model of [22] the criterion reaches seven rounds with q^3 texts. In the round-number table of [22], that value moves integral and higher-order differential attacks from seven rounds to eight, and the difference from the twelve rounds of the cipher from five to four. In the known-key model, zero-sum partitions reach thirteen rounds with sets of q texts and sixteen rounds with sets of q^4 , the two counts beyond twelve referring to the iterated structure of (3) with arbitrary fixed round keys. Finally, the two diffusion layers differ by a rotation of four words, so the branch pattern only rotates the degree bounds by a multiple of four words and the 2^{11} patterns affect neither the analysis nor the complexity.

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

Data availability The scripts that generate the propagation tables, the null-space dimensions and the verification results reported in Section 7 are available from the authors on request.

Appendix A Degree trajectories

Table A1 lists the class degree bounds of Theorem 2 for the active set $A = \mathcal{C}_2$, layer by layer, from which the entries of Table 3 are derived. Its entries are exact values of the recursion and are upper bounds on the reduced word-wise degree, and its last column is read against the threshold at $s = 4$, namely 262140 for DuX(2^{16}) and 262144 for DuX(65537). A reduced polynomial in s variables has total degree at most $s(q - 1)$, so an entry of Table A1 exceeding the threshold bounds the degree without being attained by it. Both propagation rules are maxima of sums with nonnegative coefficients and are therefore monotone in the vector of degree bounds, and the initial vector of an active set contained in $\mathcal{C}_2$ is componentwise at most that of $\mathcal{C}_2$. The rows of Table A1 are thus valid upper bounds for the smaller active sets as well. The lines of Table 3 with $s = 1$ and $s = 2$ use them in that way, with $A = \{2\}$ and $A = \{2, 6\}$ and the thresholds $q - 1$ and $2(q - 1)$, and the lines with $s = 4$ use the rows as they stand. For the eleven-word active set of DuX(2^8) at seven rounds, namely $A = \{0, 2, 3, 4, 6, 7, 8, 10, 11, 12, 14\}$, the vector after two layers is exactly twice the vector of $\mathcal{C}_2$ after two layers. The two propagation rules are positively homogeneous, so the factor two persists at every later layer, and the bounds at seven layers are (3120, 4262, 5822, 2702) against the threshold

$11 \cdot 255 = 2805$, class 3 being guaranteed. Sixteen eleven-word active sets give these same bounds; the one displayed is the complement of $\mathcal{C}_1 \cup \{15\}$.

Table A1 Class degree bounds after r inverse S-box layers for $A = \mathcal{C}_2$

r	d_0	d_1	d_2	d_3	Guaranteed zero sums at $s = 4$
0	0	0	1	0	–
1	1	1	1	0	all 16
2	2	3	4	2	all 16
3	8	11	15	7	all 16
4	30	41	56	26	all 16
5	112	153	209	97	all 16
6	418	571	780	362	all 16
7	1560	2131	2911	1351	all 16
8	5822	7953	10864	5042	all 16
9	21728	29681	40545	18817	all 16
10	81090	110771	151316	70226	all 16
11	302632	413403	564719	262087	$\mathcal{C}_3$
12	1129438	1542841	2107560	978122	none

In the encryption direction the corresponding data are given by (12). The class-wise degree bounds after the r -th S-box layer are $(5, 3, 2, 8)D_{r-1}$ with $D_r = 8D_{r-1}$, and the values entering Corollary 3 are $D_3 = 128$ for the active set $\{1, 5\}$, $D_5 = 12288$ for $\{0\}$ and $D_6 = 2^{16}$ for $\{1, 5, 9\}$.

Appendix B Corrections to formulas and examples in [22]

This appendix records four checks of [22]. The comparison with the explicit form of S^{-1} was run for every α in the field on the small instances $q \in \{2^2, 2^3\}$ and $q \in \{3, 5, 7\}$. The differential and linear bounds were checked on the same instances with $\alpha = 179 \bmod q$; the maximal Walsh value was not computed for $q = 7$. The chain form and the six-round examples were checked with $\alpha = 179$ on the instances for which they are stated.

Theorem 1 and Theorem 2 of [22]. The bound q^2 on the differential uniformity of S^{-1} and the bound q^3 on its linearity both hold. Equality is attained on every instance for which the corresponding quantity was evaluated, that is on $q \in \{2^2, 2^3, 3, 5, 7\}$ for the differential uniformity and on $q \in \{2^2, 2^3, 3, 5\}$ for the linearity, so the two bounds are tight there.

The explicit form of S^{-1} in [22]. The four expressions given there agree pointwise with (6) on every small instance and for every α , and the coordinate degrees $(5, 3, 2, 8)$ stated there for S agree with Lemma 1.

The chain form of S in the Gröbner basis section of [22]. The relations are printed there as

$$y_2 = x_0x_3 + \alpha - x_2, \quad y_1 = y_2x_3 - x_1, \quad y_0 = y_1y_2 + \alpha - x_0, \quad y_3 = y_0y_1 - x_3,$$

whereas the relations satisfied by $S = (R^{-1})^4$ are (7). The two sets agree when q is even and $\alpha = 0$, and differ otherwise; the first discrepancy usually occurs at $x = 0$. Both sets consist of four relations of the same degrees, so the counts that [22] substitutes into its complexity formula for the Gröbner basis analysis are the ones that (7) gives as well.

The six-round integral examples in Section 5.3 of [22]. Four examples are given there, each with a single active word and one balanced position. In the indexing of Section 2.1 they are: active word 3 with balanced position 0; active word 7 with balanced position 7; active word 11 with balanced position 4; and active word 15 with balanced position 12. The guaranteed set depends on the active word. For $A = \{0\}$ twelve words are guaranteed at six rounds by Corollary 3, whereas for each of these four active words the criterion guarantees exactly the eight positions of $\mathcal{C}_{\{1,2\}}$; the four claimed positions lie in the classes 0, 3, 0 and 0. Since Theorem 1 is a sufficient condition, the four claims were settled by direct summation over the whole subspace of dimension $s = 1$, that is q plaintexts. Table B2 gives the resulting sums for three random master keys on DuX(2^{16}); every value there would be zero if the claim held, and none is zero. The set of positions that do sum to zero is $\mathcal{C}_{\{1,2\}} = \{1, 2, 5, 6, 9, 10, 13, 14\}$ in every case: for the tested keys and active words the observed balanced positions coincide with those guaranteed by the criterion. The same coincidence of balanced positions is observed with the branch bits of the key schedule and with t fixed to 0 and to 1, and on DuX(65537).

Table B2 Sums at the four positions claimed balanced in [22]

Active word	Claimed balanced position	Sum, key 1	Sum, key 2	Sum, key 3
3	0	17009	55950	119
7	7	54147	701	37818
11	4	59892	41967	17835
15	12	31431	54451	50461

Appendix C Experimental data

Table C3 gives the four states of (13) in the reduced-round configuration of Section 7, with the active set $A = \{0, 2, 3\}$ and $s = 3$ and one subspace of 2^{24} ciphertexts for each of three random master keys. Listed are their class degree bounds, the sets that Theorem 1 guarantees from those bounds against the threshold $s(q - 1) = 765$, and the words observed to sum to zero. The bounds for P and for W_0 coincide, as do the sums, since the added key constant sums to $q^s K = 0$. A zero sum is guaranteed at W_1 and at no earlier state, which is where the equations of Theorem 3 are read.

Table C4 lists the quantities that enter (17) at twelve rounds in the two routes of Section 5, its last column being the number of key words the stage determines. The three-stage route gives $n = \max\{8, 4, 4\} + 2 = 10$ and is the one used for DuX(2^{16}), the interpolation route $n = \max\{8, 10\} + 2 = 12$ and is the one used for DuX(65537).

Table C3 The four states of (13) on DuX(2⁸) with seven rounds attacked

State	Class degree bounds	Guaranteed	Key 1	Key 2	Key 3
P	(3120, 4262, 5822, 2702)	none	none	none	none
$W_0 = P + K$	(3120, 4262, 5822, 2702)	none	none	none	none
$SL(P + K)$	(1142, 1560, 1560, 836)	none	$\mathcal{C}_3 \cup \{0\}$	$\mathcal{C}_3$	$\mathcal{C}_3$
W_1	(836, 1142, 1560, 724)	$\mathcal{C}_3$	$\mathcal{C}_{\{0,3\}}$	$\mathcal{C}_{\{0,3\}}$	$\mathcal{C}_{\{0,3\}}$

Table C4 The counts entering (17) at twelve rounds

Route	Stage	c_j	ρ_j	$\lceil c_j/\rho_j \rceil$	Key words
Three-stage	(i)	8	1	8	8
	(ii)	4	1	4	4
	(iii)	7	2	4	4
Interpolation	(i)	8	1	8	8
	(ii), (iii)	28	3	10	8

Table C5 gives the coefficients of Proposition 5 at stage (iii) of the three-stage route, on DuX(2⁸) with seven rounds attacked, in the three runs of the first line of Table 6. The entry in column k is the coefficient of a_{4k+1} in the equation of one subspace, in the field representation of Section 2.1; it depends on the subspace drawn, and Proposition 5 requires it to be nonzero, which it is in all twelve entries.

Table C5 Coefficients of a_1 at stage (iii) of the three-stage route

	$k = 0$	$k = 1$	$k = 2$	$k = 3$
Key 1	206	110	185	165
Key 2	162	202	216	215
Key 3	221	219	227	161

Table C6 gives the rank of the interpolation system of Section 5.2 against the number of subspaces on the three small-field configurations of Table 6, its last three columns counting the unknowns the system determines. Each subspace supplies four equations and three independent rows, as $\rho = \dim V_{\{0,1,2,3\}} - \dim V_{\{2\}} = 3$ gives, and the rank, which is the same in the three configurations, stops at the 28 effective columns among the 44 columns of the system, the other 16 vanishing identically. The eight unknowns a_{4k+1} and a_{4k+2} that carry the eight key words of $\mathcal{C}_{\{1,2\}}$ are determined from ten subspaces on and at no smaller number in all three configurations, as they are in the three large-field configurations.

Table C6 Rank of the interpolation system against the number of subspaces

Subspaces	Equations	Rank	DuX(29)	DuX(101)	DuX(53)
1	4	3	0	0	0
2	8	6	0	0	0
3	12	9	0	0	0
4	16	12	0	0	0
5	20	15	0	0	0
6	24	18	0	0	0
7	28	21	0	0	0
8	32	24	0	0	0
9	36	27	2	0	1
10	40	28	28	28	28
11	44	28	28	28	28
12	48	28	28	28	28

References

- [1] Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, pp. 169–178. ACM, New York, NY, USA (2009)
- [2] Brakerski, Z., Gentry, C., Vaikuntanathan, V.: (leveled) fully homomorphic encryption without bootstrapping. In: Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, ITCS 2012, pp. 309–325. ACM, New York, NY, USA (2012)
- [3] Halevi, S., Shoup, V.: Algorithms in HELib. In: Advances in Cryptology – CRYPTO 2014. Lecture Notes in Computer Science, vol. 8616, pp. 554–571. Springer, Berlin, Heidelberg (2014)
- [4] Naehrig, M., Lauter, K., Vaikuntanathan, V.: Can homomorphic encryption be practical? In: Proceedings of the 3rd ACM Cloud Computing Security Workshop, CCSW 2011, pp. 113–124. ACM, New York, NY, USA (2011)
- [5] Albrecht, M.R., Rechberger, C., Schneider, T., Tiessen, T., Zohner, M.: Ciphers for MPC and FHE. In: Advances in Cryptology – EUROCRYPT 2015. Lecture Notes in Computer Science, vol. 9056, pp. 430–454. Springer, Berlin, Heidelberg (2015)
- [6] Canteaut, A., Carpov, S., Fontaine, C., Lepoint, T., Naya-Plasencia, M., Paillier, P., Sirdey, R.: Stream ciphers: a practical solution for efficient homomorphic-ciphertext compression. *Journal of Cryptology* **31**(3), 885–916 (2018) <https://doi.org/10.1007/s00145-017-9273-9>
- [7] Méaux, P., Journault, A., Standaert, F.-X., Carlet, C.: Towards stream ciphers

- for efficient FHE with low-noise ciphertexts. In: *Advances in Cryptology – EUROCRYPT 2016. Lecture Notes in Computer Science*, vol. 9665, pp. 311–343. Springer, Berlin, Heidelberg (2016)
- [8] Cho, J., Ha, J., Kim, S., Lee, B., Lee, J., Lee, J., Moon, D., Yoon, H.: Transciphering framework for approximate homomorphic encryption. In: *Advances in Cryptology – ASIACRYPT 2021. Lecture Notes in Computer Science*, vol. 13092, pp. 640–669. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92078-4_22
 - [9] Ha, J., Kim, S., Lee, B., Lee, J., Son, M.: Rubato: noisy ciphers for approximate homomorphic encryption. In: *Advances in Cryptology – EUROCRYPT 2022. Lecture Notes in Computer Science*, vol. 13275, pp. 581–610. Springer, Cham (2022)
 - [10] Dobraunig, C., Grassi, L., Helming, L., Rechberger, C., Schafnegg, M., Walch, R.: Pasta: a case for hybrid homomorphic encryption. *IACR Transactions on Cryptographic Hardware and Embedded Systems* **2023**(3), 30–73 (2023)
 - [11] Albrecht, M.R., Grassi, L., Rechberger, C., Roy, A., Tiessen, T.: MiMC: efficient encryption and cryptographic hashing with minimal multiplicative complexity. In: *Advances in Cryptology – ASIACRYPT 2016. Lecture Notes in Computer Science*, vol. 10031, pp. 191–219. Springer, Berlin, Heidelberg (2016)
 - [12] Albrecht, M.R., Grassi, L., Perrin, L., Ramacher, S., Rechberger, C., Rotaru, D., Roy, A., Schafnegg, M.: Feistel structures for MPC, and more. In: *Computer Security – ESORICS 2019. Lecture Notes in Computer Science*, vol. 11736, pp. 151–171. Springer, Cham (2019)
 - [13] Daemen, J., Knudsen, L.R., Rijmen, V.: The block cipher Square. In: *Fast Software Encryption, FSE 1997. Lecture Notes in Computer Science*, vol. 1267, pp. 149–165. Springer, Berlin, Heidelberg (1997)
 - [14] Knudsen, L.R., Wagner, D.: Integral cryptanalysis. In: *Fast Software Encryption, FSE 2002. Lecture Notes in Computer Science*, vol. 2365, pp. 112–127. Springer, Berlin, Heidelberg (2002)
 - [15] Lai, X.: Higher order derivatives and differential cryptanalysis. In: *Communications and Cryptography: Two Sides of One Tapestry*, pp. 227–233. Springer, Boston, MA (1994)
 - [16] Knudsen, L.R.: Truncated and higher order differentials. In: *Fast Software Encryption, FSE 1994. Lecture Notes in Computer Science*, vol. 1008, pp. 196–211. Springer, Berlin, Heidelberg (1995)
 - [17] Todo, Y.: Structural evaluation by generalized integral property. In: *Advances in Cryptology – EUROCRYPT 2015. Lecture Notes in Computer Science*, vol. 9056, pp. 287–314. Springer, Berlin, Heidelberg (2015)

- [18] Todo, Y., Morii, M.: Bit-based division property and application to Simon family. In: Fast Software Encryption, FSE 2016. Lecture Notes in Computer Science, vol. 9783, pp. 357–377. Springer, Berlin, Heidelberg (2016)
- [19] Xiang, Z., Zhang, W., Bao, Z., Lin, D.: Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers. In: Advances in Cryptology – ASIACRYPT 2016. Lecture Notes in Computer Science, vol. 10031, pp. 648–678. Springer, Berlin, Heidelberg (2016)
- [20] Liu, F., Li, Y., Chen, H., Jiao, L., Luo, M., Wang, M.: YuX: finite field multiplication based block ciphers for efficient FHE evaluation. *IEEE Transactions on Information Theory* **70**(5), 3729–3749 (2024) <https://doi.org/10.1109/TIT.2024.3349414>
- [21] Ni, J., Wang, G., Li, Y.: Tracking algebraic degree with exponent sets: higher-order differential attacks on FHE-friendly cipher Yu2X. *Designs, Codes and Cryptography* **94**(6), 123 (2026) <https://doi.org/10.1007/s10623-026-01852-8>
- [22] Wu, Y., Dong, X., Li, Y., Liu, F., Zhang, Y., Cheng, R., Tan, H., Zhang, W.: DuX: a block cipher for efficient FHE evaluation. *Designs, Codes and Cryptography* **94**(7), 140 (2026) <https://doi.org/10.1007/s10623-026-01855-5>
- [23] Ren, X., Xu, B., Xiong, Z., Li, Y., Wang, M.: Practical key recovery attacks on full DuX and reduced-round YuX. *Cryptology ePrint Archive*, Paper 2026/2045 (2026). <https://eprint.iacr.org/2026/2045>
- [24] Boura, C., Canteaut, A.: 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)
- [25] Cui, J., Hu, K., Wang, M., Wei, P.: On the field-based division property: applications to MiMC, Feistel MiMC and GMiMC. In: Advances in Cryptology – ASIACRYPT 2022. Lecture Notes in Computer Science, vol. 13793, pp. 241–270. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-22969-5_9
- [26] Cid, C., Grassi, L., Gunesing, A., Lüttenegger, R., Rechberger, C., Schefnegger, M.: Influence of the linear layer on the algebraic degree in SP-networks. *IACR Transactions on Symmetric Cryptology* **2022**(1), 110–137 (2022) <https://doi.org/10.46586/tosc.v2022.i1.110-137>
- [27] Aumasson, J.-P., Meier, W.: Zero-sum distinguishers for reduced Keccak- f and for the core functions of Luffa and Hamsi. Presented at the rump session of CHES 2009 (2009)
- [28] Boura, C., Canteaut, A., De Cannière, C.: Higher-order differential properties of Keccak and Luffa. In: Fast Software Encryption, FSE 2011. Lecture Notes in Computer Science, vol. 6733, pp. 252–269. Springer, Berlin, Heidelberg (2011)