Cycle structure of generalized and closed loop invariants

Yongzhuang Wei[1], René Rodríguez[2], and Enes Pasalic[1,2]

[1] Guilin University of Electronic Technology, Guilin, Guangxi Province 541004, China
walker_wyz@guet.edu.cn
[2] University of Primorska, FAMNIT, Koper 6000, Slovenia
enes.pasalic6@gmail.com, rene7ca@gmail.com

Abstract. This article gives a rigorous mathematical treatment of generalized and closed loop invariants (CLI) which extend the standard notion of (nonlinear) invariants used in the cryptanalysis of block ciphers. Employing the cycle structure of bijective S-box components, we precisely characterize the cardinality of both generalized and CLIs. We demonstrate that for many S-boxes used in practice quadratic invariants (especially useful for mounting practical attacks in cases when the linear layer is an orthogonal matrix) might not exist, whereas there are many quadratic invariants of generalized type (alternatively quadratic CLIs). In particular, it is shown that the inverse mapping $S(x) = x^{-1}$ over $GF(2^4)$ admits quadratic CLIs that additionally possess linear structures. The use of cycle structure is further refined through a novel concept of *active cycle set*, which turns out to be useful for defining invariants of the whole substitution layer. We present an algorithm for finding such invariants provided the knowledge about the cycle structure of the constituent S-boxes used.

Keywords: Block ciphers · Generalized nonlinear invariants · Permutation cycles · Closed loop invariants · Linear structures · Distinguishing attacks · SP networks.

1 Introduction

The design of block ciphers, used as symmetric key encryption algorithms, is well understood and their security has been traditionally evaluated using some standard cryptanalytic techniques such as differential attack [BS90], linear attack [Mat93], and their diverse variations [LH94] [HTW15]. During the last few years some other possibilities concerning the cryptanalysis of certain families of block ciphers have emerged. Nevertheless, whereas most of the well-established designs are quite robust to these cryptanalytic methods it appears that almost exclusively lightweight block ciphers show certain vulnerability in this context. This feature is primarily due to a rather simplified design strategy of implementation-constrained block ciphers and in particular the main weakness seems to be their simple key schedule.

Nonlinear invariant attacks were introduced at ASIACRYPT 2016 by Todo *et al.* [TLS16] and they gained a lot of attention due to their efficient application in breaking full-round block ciphers such as SCREAM, iSCREAM [GLSV14] and Midori64 [BBI$^+$15]. The nonlinear invariant attack can be seen as a further extension of the invariant subspace attack introduced in [LAAZ11], which identifies the property of having inputs and outputs that belong to the same affine subspace through (many) encryption rounds under the so-called *weak key assumption*. The core idea of nonlinear invariant attacks is to look for a nonlinear Boolean function $g : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ for which the evaluation of $g(x) + g(E_k(x))$ is constant for any x, where $E_k(x)$ is the encryption function of a considered n-bit block cipher performed using the secret key k. The function g is then called a nonlinear invariant for $E_k(x)$ and those

keys $k \in \mathcal{K}$ for which g is a nonlinear invariant are called *weak keys*. In general, for a random permutation this property holds with a probability of about 2^{1-N} if N plaintext/ciphertext pairs are considered (assuming g is a balanced Boolean function). Consequently, any cipher admitting nonlinear invariants covering all encryption rounds can be easily distinguished from a random permutation, assuming that the used invariant has a certain bias.

In general, nonlinear invariants for a full-round block cipher are derived by finding nonlinear invariants for each separate round (if these exist) which are then merged together assuming that the round keys all belong to the family of weak keys. The crucial point here is that many lightweight block ciphers, motivated by efficiency of implementation, use a simplified round key schedule and therefore deriving the round keys from the master key using round constants. Even though the assumption on finding invariants for a whole cipher appears to be quite unrealistic, it was demonstrated in [TLS16] that many recently proposed lightweight block ciphers have serious weaknesses in this context. Another important point is the fact that, apart from the assumption on weak keys, the success of this attack heavily relies on the choice of round constants since their proper selection can protect ciphers against these attacks [BCLR17].

The concept of nonlinear invariants was extended in [YWP19] to consider *generalized invariants* for which $g(x + a_1) + g(E_k(x) + a_2) = c$, with $c \in \mathbb{F}_2$, thus employing two n-bit vectors a_1 and a_2 for the purpose of eliminating the effect from the round constants. This method was demonstrated to be efficient in mounting a distinguishing attack on iSCREAM, using a different family of weak keys than the one identified for iSCREAM in the context of standard nonlinear invariant attacks in [TLS16]. Extending this approach further [YWP19], two Boolean functions $g_1, g_2 : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ were used to define the notion of *closed loop invariants* (CLIs) which is another useful criterion related to the robust choice of round constants. More precisely, for a given block cipher $E_k(x)$ the adversary may also try to identify two different nonlinear Boolean functions $g_1, g_2 : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ such that $g_1(x) + g_2(E_k(x)) = c$ and $g_2(x) + g_1(E_k(x)) = c$ are simultaneously satisfied for any x, for some class of weak keys. It was then shown that the conclusion drawn in [BCLR17], that a choice of round constants is independent of the substitution layer, is quite inadequate when these generalized concepts are employed.

One objective of this article is to provide a more comprehensive and rigorous treatment of generalized and closed loop invariants; their precise description and exact cardinality (based on the cycle structure of underlying S-box) as well as their relationship to standard invariants. In this context, we provide exact estimates on the cardinality of generalized and closed loop invariants in terms of the cycle structure of bijective S-boxes and identify quadratic such invariants. Most notably, we show that many S-boxes used in practice do not admit standard quadratic invariants whereas there exist subspaces of quadratic generalized invariants and CLIs. In particular, we show that the downscaled inverse S-box of AES defined over $GF(2^4)$ admits quadratic CLIs which furthermore possess linear structures. For this reason, the use of such S-boxes in the design of lightweight block ciphers may cause undesired security issues. Nevertheless, these quadratic CLIs provably do not exist for the inverse mapping $S(x) = x^{-1}$ over $GF(2^n)$ whenever $n > 4$.

A novel concept of *active cycle set* is introduced, and then used to specify nonlinear invariants that not only address a single S-box but rather can be used to handle the whole substitution layer in the (classical) case that this layer is implemented as a parallel application of several (not necessarily identical) S-boxes of relatively small size. In this context, we propose algorithm for determining the cycle structure of concatenated S-boxes which can be used (in certain cases) to determine the cycle structure of the entire substitution layer,

cf. Algorithm 2. This specification is possible in those cases when the number of cycles is relatively small which appears to be the case for some well known S-boxes, for instance those used in LED cipher and AES. The complexity of Algorithm 2 relies on the number of cycles of the two concatenated boxes. More precisely, if the number of cycles of the two concatenated boxes are r_1 and r_2, respectively, then the complexity of Algorithm 2 is $r_1 \times r_2 \times \gcd(l_1, l_2)$, where l_1 and l_2 are the lengths of two cycles considered.

1.1 Organization

The rest of the paper is organized as follows. The basic idea and principles of the generalized nonlinear invariant attack are described in Section 2. In Section 3, the exact cardinality of closed loop invariants in terms of the cycle decomposition of a given S-box is derived. It is shown that the inverse function $S(x) = x^{-1}$ over $GF(2^4)$ admits quadratic CLIs which additionally posses linear structures of dimension two. In Section 4, we perform an extensive theoretical analysis related to generalized invariants. For many S-boxes used in practice, not admitting standard invariants, we demonstrate the existence of generalized ones among these also quadratic ones. The concept of an active cycle set is introduced in Section 5 as a useful tool for specifying probabilistic invariants based on which an efficient algorithm for specifying invariants of the whole S-box layer is given. Some concluding remarks are given in Section 6.

2 Standard and generalized nonlinear invariant attacks

The encryption process of an iterative block cipher consisting of r rounds, the ciphertext C is derived by encrypting a plaintext P using the round subkey K_i, where $i = 0, \ldots, r-1$. More precisely,

$$x_0 = P, \quad x_{i+1} = F_{K_i}(x_i) = F(x_i) + K_i, \quad C = x_r, \tag{1}$$

where $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is the round function of the considered (n-bit) block cipher, and $F_{K_i}(x_i) = F(x_i) + K_i$ means that the output of the round function is XORed with the round subkey K_i. For simplicity, the pre-whitening key is ignored.

In difference to the standard invariant attack [TLS16, TLS18], based on the relationship $g(x) + g(F_{K_i}(x)) = c$, its recent generalization considers a nonlinear Boolean function $g : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ and a pair of n-bit constants (a_1, a_2) such that $g(x + a_1) + g(F_{K_i}(x) + a_2) = c$ (where c is a binary constant) holds for any $x \in \mathbb{F}_2^n$. This approach, employing a pair of constants, may be more flexible for the purpose of compensating the effect of round constant addition, see [YWP19] for more details.

We consider the following round function $F : \mathbb{F}_2^{t \times m} \longrightarrow \mathbb{F}_2^{t \times m}$, with $t \times m = n$, of an substitution permutation network (SPN) cipher that consists of an S-box layer $\mathcal{S} : \mathbb{F}_2^{t \times m} \longrightarrow \mathbb{F}_2^{t \times m}$ and a linear layer $\mathcal{L} : \mathbb{F}_2^{t \times m} \longrightarrow \mathbb{F}_2^{t \times m}$:

$$F(x) = \mathcal{L} \circ \mathcal{S}(x),$$

where $\mathcal{S}(x_1, \ldots, x_t) = (S_1(x_1), \ldots, S_t(x_t))$ and $x_i \in \mathbb{F}_2^m$. In general, t not necessarily identical bijective S-boxes are used in each round, though commonly these are same. We will denote the space of m-variable Boolean functions $f : \mathbb{F}_2^m \to \mathbb{F}_2$ with $\mathcal{B}_m$.

The main reason for avoiding the use of S-boxes that admit invariants of small algebraic degree, is the possibility to extend these invariants (in certain cases) to the whole encryption round as demonstrated in [YWP19], see also [BCLR17].

Theorem 1. *[YWP19] Assume that the round function of an SPN-based block cipher uses LS design rationale and that a binary representation of the linear layer $\mathcal{L}$ is an orthogonal matrix $M \in \mathbb{F}_2^{n\times n}$. If there is a quadratic generalized nonlinear invariant $g \in \mathcal{U}(S, a_1, a_2) = \{g \in \mathcal{B}_m : g(x_i + a_1) = g(S(x_i) + a_2) + c, c \in \mathbb{F}_2, x_i \in \mathbb{F}_2^m\}$, then the function*

$$G(x_1, \ldots, x_t) = \sum_{i=1}^{t} g(x_i)$$

is also a generalized nonlinear invariant for the round function $\mathcal{L} \circ \mathcal{S}$.

2.1 Specifying standard invariants for arbitrary bijective S-boxes

The possibility of applying a distinguishing attack, based on the existence of standard invariants, to many lightweight block ciphers [BCLR17] is a serious security concern. A natural question regarding the existence of standard nonlinear invariants, their cardinality $\#g$ and structural properties has been considered in [TLS16]. The main conclusion is that there always exist standard invariants for any given bijective S-box and their number is closely related to the cycle structure of the S-box. More precisely, Todo *et al.* [TLS16] showed that $\#g = 2^{(\#\text{ cycles of } F)}$ when F (representing a bijective S-box) has at least one cycle of odd length; alternatively $\#g = 2^{(\#\text{ cycles of } F)+1}$ when F only has cycles of even length.

An important consequence of the above result is that full-cycle permutations admit the smallest cardinality of standard invariants, thus reducing the probability of finding suitable invariants, e.g. of low algebraic degree or containing a few terms in their algebraic representation. Assuming that identical S-boxes are used in encryption rounds, it was furthermore proved that such a set of invariants, say $\{g_i : i = 1, \ldots, t\}$, can be used to construct an invariant for the whole substitution layer $\mathcal{S} : \mathbb{F}_2^{tm} \to \mathbb{F}_2^{tm}$ of the form $G = \sum_i^t \alpha_i g_i$, for any nonzero choice of the binary coefficients α_i [TLS16, Proposition 1]. Nevertheless, one can also deduce other standard invariants that do not stem from this simple technique [TLS16].

3 Closed loop invariants for S-boxes

In this section, we derive exact cardinality of CLIs of a given bijective mapping $S(x)$ over $\mathbb{F}_{2^m}$ that entirely depends of the cycle structure of S.

One of the main conclusions, using the framework of generalized nonlinear invariants [YWP19], was that the choice of round constants is not only related to the linear layer of a block cipher, but there is a rather strong dependency on the substitution layer as well. In other words, a large dimension of the parameter $W_L(D)$ (see [BCLR17] for more details), suggested as a design rationale by Beierle *et al.* [BCLR17], is not a sufficient condition to protect block ciphers against generalized nonlinear invariant attacks.

In addition, a (different) concept of closed-loop invariants was used in [YWP19] for mounting a distinguishing attack on a variant of the Midori64 block cipher [BBI+15]. For any bijective S-box $S : \mathbb{F}_2^m \longrightarrow \mathbb{F}_2^m$, its *closed-loop invariant* was defined as

$$CLI(S) = \{(g_1, g_2) \mid g_1(x) + g_2(S(x)) = c_1, g_2(x) + g_1(S(x)) = c_2, c_i \in \mathbb{F}_2\}.$$

Remark 1 *Given an S-box and taking any function g_1 one can compute $g_1(S(x))$, which then uniquely specifies $g_2(x)$ so that $g_2(x) + g_1(S(x)) = c_2$ is satisfied. In general, there is*

no guarantee that such g_1 and g_2 satisfy $g_1(x)+g_2(S(x)) = c_1$, for a fixed $c_1 \in \mathbb{F}_2$. However, there are cases when one can assure it, for instance, when S is an involution. In this case, we have that $g_2(x)+g_1(S(x)) = c_2$ for every $x \in \mathbb{F}_2^m$, implies that $g_2(S(x))+g_1(x) = c_2$ for every $x \in \mathbb{F}_2^m$, thus $(g_1, g_2) \in CLI(S)$.

To characterize robust round constants, the authors in [YWP19] proposed a new design criterion which requires that for each closed loop invariant in $CLI(S)$ the round constants should not be contained in $\mathsf{LS}(g_i)$, $i \in \{1, 2\}$. Here, $\mathsf{LS}(g_i)$ denotes the subspace of *linear structures* given by $\mathsf{LS}(g_i) = \{a \in \mathbb{F}_2^m : g_i(x) + g_i(x + a) = c\}$, where $c \in \mathbb{F}_2$.

3.1 Counting closed loop invariants

Similarly to the case of standard/generalized invariants, the set $CLI(S)$ forms a subspace of $\mathcal{B}_m \times \mathcal{B}_m$.

Lemma 1. *For every permutation S on $\mathbb{F}_2^m$, $CLI(S)$ is a subspace of $\mathcal{B}_m \times \mathcal{B}_m$.*

Proof. Clearly, $(0, 0) \in CLI(S)$. For different tuples $(g_1, g_2), (g_1', g_2') \in CLI(S)$ we have

$$g_1(x) + g_2(S(x)) = c_1; \;\; g_2(x) + g_1(S(x)) = c_2$$

and

$$g_1'(x) + g_2'(S(x)) = c_1'; \;\; g_2'(x) + g_1'(S(x)) = c_2',$$

for some constants $c_1, c_2, c_1', c_2' \in \mathbb{F}_2$. Then, $g_1(x)+g_1'(x)+g_2(S(x))+g_2'(S(x)) = c_1+c_1'$ and similarly $g_2(x) + g_2'(x) + g_1(S(x)) + g_1'(S(x)) = c_2 + c_2'$. Hence, $(g_1 + g_1', g_2 + g_2') \in CLI(S)$ and thus $CLI(S)$ is a subspace of $\mathbb{F}_2^m$. □

Note that for every standard invariant g of S, we have $(g, g) \in CLI(S)$. So $CLI(S)$ has dimension at least $k + \varepsilon$, where k denotes the number of cycles and ε is equal to zero or one depending whether there are odd cycles or not. Throughout this article a cycle of length l containing an element $x_0 \in \mathbb{F}_2^m$, with respect to S, is the set of distinct values $C_{x_0} = \{S^j(x_0) : 1 \leq j \leq l\}$, where S^j denotes the composition of S with itself j times and $S^l(x_0) = x_0$. We obtain a better bound if we observe that the complement of every element induces a new pair, i.e., $(g, g + 1) \in CLI(S)$. So $\dim CLI(S) \geq k + \varepsilon + 1$. Actually, we can deduce the exact cardinality of $CLI(S)$.

Theorem 2. *Let S be a permutation on $\mathbb{F}_2^m$ and k_e and k_o be the number of cycles of even and odd length of S, respectively. The cardinality of $CLI(S)$ is given by:*

- 4^{k_e+1} *if there are only cycles of even length greater than two.*
- $4^{k_e}2^{k_o+1}$ *otherwise.*

Proof. Suppose first that $k_o = 0$ and there are no transpositions. Let us fix c_1 and c_2 and consider a fixed $x_0 \in \mathbb{F}_2^m$. We consider the images of g_j, $j \in \{1, 2\}$, evaluated at the cycle $C_{x_0} := \{ S^i(x_0) \;\; : \;\; i \geq 0\}$. Notice that every fixed value for $g_1(x_0)$ completely determines the value of $g_2(S(x_0))$. Now, once $g_2(S(x_0))$ is determined, $g_1(S^2(x_0))$ can be computed using $g_2(S(x_0)) + g_1(S^2(x_0)) = c_2$. We see that all values in the sequence $g_2(S(x_0)), g_1(S^2(x_0)), \ldots, g_2(S^{|C_{x_0}|-1}(x_0))$ are determined. Similarly, any fixed value of $g_2(x_0)$ completely determines the value of $g_1(S(x_0)), g_2(S^2(x_0)), \ldots, g_1(S^{|C_{x_0}|-1}(x_0))$.
In total, we have four possible independent choices for $g_1(x_0)$ and $g_2(x_0)$ and considering

fixed elements in different cycles there are 4^{k_e} pairs (g_1, g_2) in $CLI(S)$, for any fixed c_1 and c_2. Noticing that different choices of constants give rise to distinct pairs, we deduce that there are 4^{k_e+1} elements in $CLI(S)$.

Now suppose $k_o > 0$ or there are transpositions. In the latter case, consider $x_0 \in \mathbb{F}_2^m$ such that $S^2(x_0) = x_0$ which implies $g_1(x_0)+c_1+c_2 = g_2(S(x_0))+c_2 = g_1(x_0)$; thus $c_1 = c_2$. When $k_0 > 0$, we take x_0 that belongs to a cycle of odd length and then $g_1(x_0)+c_1+c_2+\ldots+c_1 = g_2(x_0)$, where c_1 and c_2 appear an even and an odd number of times, respectively. Hence, $g_1(x_0) + c_2 = g_2(x_0)$. In a similar fashion $g_2(x_0) + c_1 = g_1(x_0)$, thus $c_1 = c_2$. In both cases we have $c_1 = c_2$, and therefore for these types of cycles there are only two possible choices for the constants c_1, c_2. In addition, for fixed c_1, c_2 and a given x_0 belonging to a cycle of odd length, the value of $g_1(x_0)$ determines the values of g_1 and g_2 in this cycle. This gives only the two possibilities when a cycle has odd length and four when it is of even length (as shown above). Therefore, $|CLI(S)| = 2 \cdot 2^{k_o} \cdot 4^{k_e}$. □

We observe that k_o is always even, since adding the lengths of each disjoint cycle must equal 2^m which is impossible when k_o is odd.

Remark 2 *Note that if g admits a non-trivial linear structure, then it must be the case that the weight of g is even: $g(x) = g(x + \alpha) + 1$ implies g is balanced and $g(x) = g(x + \alpha)$ implies we always have pairs with the same value. Therefore, when S is a product of two odd cycles its two non-constant invariants will not admit any linear structure (since they have odd weight). According to Theorem 2, such a permutation S will only admit eight CLIs which are just distinct pairs of standard invariants and none of them will admit linear structures.*

Corollary 1. *Let S be a permutation on $\mathbb{F}_2^m$. If k_s is the dimension of the space of standard invariants of $S^2 = S \circ S$, then $k_e + 1 \leq k_s$. Furthermore, if S has a cycle of odd length or a transposition then $2k_e + k_o + 1 \leq 2k_s$.*

Proof. Taking $(g_1, g_2) \in CLI(S)$, we deduce that $g_1(x) + g_1(S^2(x)) = c_1 + c_2$ and $g_2(x) + g_2(S^2(x)) = c_1 + c_2$. So g_1 and g_2 are invariants of S^2. This means that $\dim(CLI(S)) \leq 2k_s$. Applying Theorem 2, we obtain the desired result. □

For instance, if S^2 is a full-length cycle or a product of two cycles of odd length $k_s = 2$, then S has to be a full-length cycle or a product of two cycles of odd length.[3]

The existence of quadratic CLI-s, assuming the use of an orthogonal matrix as a linear layer, can induce immediate weaknesses in the design [YWP19] whose resistance to distinguishing attacks then entirely depends on the choice of round constants. Below, we give a detailed analysis related to the existence of quadratic CLI-s for the inverse S-box of AES of variable size. It is shown that there are exactly 63 quadratic CLI-s for the inverse S-box of size 4×4 out of which exactly 35 CLI-s admit linear structures of dimension two (see Theorem 8). On the other hand, there provably do not exist quadratic CLI-s for the inverse S-box of size $m \times m$ when $m > 4$.

3.2 CLIs of the inverse AES box and its small-scaled variant

Here, we describe an efficient approach for specifying CLIs of the inverse S-box using its compact univariate representation. This method can be easily applied to arbitrary bijective S-boxes (of reasonable size) using their univariate representation ober the finite field $\mathbb{F}_{2^m}$.

[3] Moreover, S^2 splits just the cycles of even length in the cycle decomposition of S in halves. S^3 will split only the cycles with length a multiple of 3, and so on. In particular, if we start with two cycles of odd length then $S, S^2, S^4, \ldots, S^{2^m} = 1$ are all conjugates.

We recall that the univariate representation of a mapping $F : \mathbb{F}_{2^m} \to \mathbb{F}_{2^m}$ is given as $F(x) = \sum_{i=0}^{2^m-1} a_i x^i$, where $a_i \in \mathbb{F}_{2^m}$. In particular, if the coefficients a_i satisfy the following (Boolean conditions): $a_0, a_{2^m-1} \in \mathbb{F}_2$ and $a_{2i \pmod{2^m-1}} = a_i^2$ for $i = 1, \ldots, 2^m - 2$, then it turns out that F is essentially a Boolean mapping thus $F : \mathbb{F}_{2^m} \to \mathbb{F}_2$ instead. This is due to the condition that for any Boolean mapping $f(x)^2 \equiv f(x) \pmod{x^{2^m} - x}$. This representation is convenient when analysing the composition $g(S(x)) = g(y)$ since both S and g can be described as univariate polynomials. We use this approach to demonstrate that the inverse S-box defined as a permutation $S(x) = x^{-1}$ over $\mathbb{F}_{2^4}$ not only admit a large set of quadratic CLIs but furthermore a majority of these invariants have linear structures of dimension two.

Theorem 3. *There are exactly 63 quadratic closed loop invariants for the inverse S-box $S(x) = x^{-1}$ over $\mathbb{F}_{2^4}$ and 35 of these invariants admit linear structures of dimension two.*

Proof. To determine CLIs for the inverse S-box $S(x) = x^{-1}$, we write g_1 and g_2 as $g_1(x) = \sum_{i=0}^{2^4-1} a_i x^i$ and $g_2(x) = \sum_{i=0}^{2^4-1} b_i x^i$, where $a_i, b_i \in \mathbb{F}_2^4$. Since we are looking for quadratic g_1 and g_2 then all a_i, b_i whose Hamming weight is greater than two equal zero. Thus, neglecting the constant terms,

$$g_1(x) = (a_1 x + a_2 x^2 + a_4 x^4 + a_8 x^8) + a_3 x^3 + a_5 x^5 + a_6 x^6 + a_9 x^9 + a_{10} x^{10} + a_{12} x^{12} \text{ and}$$

$$g_2(x) = (b_1 x + b_2 x^2 + b_4 x^4 + b_8 x^8) + b_3 x^3 + b_5 x^5 + b_6 x^6 + b_9 x^9 + b_{10} x^{10} + b_{12} x^{12}.$$

Now the condition that $g_1(x) + g_2(y) = g_1(x) + g_2(x^{-1}) = c$ implies the following. First notice that, after cutting exponents modulo 15,

$$g_2(x^{-1}) = g_2(x^{14}) = (b_1 x^{14} + b_2 x^{13} + b_4 x^{11} + b_8 x^7) + b_3 x^{12} + b_5 x^{10} + b_6 x^9 + b_9 x^6 + b_{10} x^5 + b_{12} x^3,$$

which then implies that $b_1 = b_2 = b_4 = b_8 = 0$ since these terms are cubic or quartic. Similarly, to satisfy $g_1(x) + g_2(y) = c$ we have $a_1 = a_2 = a_4 = a_8 = 0$. Furthermore, we have $a_3 = b_{12}, a_5 = b_{10}, a_6 = b_9, a_9 = b_6, a_{10} = b_5, a_{12} = b_3$. Then using the Boolean conditions on the coefficients a_i, b_i ($a_6 = a_3^2; a_{10} = a_5^2; a_6^2 = a_{12}; a_{12}^2 = a_9; a_9^2 = a_3$ and similarly for b_i) we get:

$$\begin{aligned} g_1(x) &= a_3 x^3 + a_5 x^5 + a_3^2 x^6 + a_3^8 x^9 + a_5^2 x^{10} + a_3^4 x^{12} \\ g_2(x^{-1}) &= a_3 x^3 + a_5 x^5 + a_6 x^6 + a_9 x^9 + a_{10} x^{10} + a_{12} x^{12} = g_1(x). \end{aligned}$$

Hence, we can specify g_2 as

$$g_2(x) = b_3 x^3 + b_5 x^5 + b_6 x^6 + b_9 x^9 + b_{10} x^{10} + b_{12} x^{12} = a_{12} x^3 + a_{10} x^5 + a_9 x^6 + a_6 x^9 + a_5 x^{10} + a_3 x^{12}.$$

Notice that the conditions $a_6 = a_3^2; a_{12} = a_6^2 = a_3^4; a_9 = a_{12}^2 = a_3^8; a_5 = a_{10}^2$ implies that all the other coefficients can be expressed in terms of a_3 and additionally $a_5^4 = a_5$ must be satisfied. There are exactly four elements x in $\mathbb{F}_{2^4}$ such that $x^4 = x$, thus there are $16 \cdot 4$ possible choices for $(a_3, a_5) \in \mathbb{F}_{2^4} \times \mathbb{F}_{2^4}$. Excluding $a_3 = a_5 = 0$ gives 63 pairs of quadratic invariants out of which 35 admit linear structures of dimension two. □

Using computer based simulations one can verify that out of the 63 closed loop invariants there are 35 functions g such that its space of linear structures has dimension 2, i.e. $|LS(g)| = 4$, and 28 functions g with a trivial space of linear structures. Partial specification of these invariants along with their corresponding linear structures is given in Table 1.

Pair (a_3, a_5)	Linear structures
(α, α^5)	$0, \alpha, \alpha^8, \alpha^{10}$
(α, α^{10})	$0, 1, \alpha^6, \alpha^{13}$
$(\alpha, 1)$	$0, \alpha^3, \alpha^5\alpha^{11}$
(α^2, α^5)	$0, 1, \alpha^{11}, \alpha^{12}$
(α^2, α^{10})	$0, \alpha, \alpha^2, \alpha^5$
$(\alpha^2, 1)$	$0, \alpha^6, \alpha^7, \alpha^{10}$
$(\alpha^3, 0)$	$0, \alpha^4, \alpha^9, \alpha^{14}$
(α^5, α^5)	$0, \alpha^5, \alpha^6, \alpha^9$
(α^5, α^{10})	$0, \alpha^{10}, \alpha^{11}, \alpha^{14}$
$(\alpha^5, 1)$	$0, 1, \alpha, \alpha^4$
$(\alpha^6, 0)$	$0, \alpha^3, \alpha^8, \alpha^{13}$
(α^7, α^5)	$0, \alpha^4, \alpha^{11}, \alpha^{13}$
(α^7, α^{10})	$0, \alpha, \alpha^3, \alpha^9$
$(\alpha^7, 1)$	$0, \alpha^6, \alpha^8, \alpha^{14}$
(α^{10}, α^5)	$0, \alpha^5, \alpha^7, \alpha^{13}$
$(\alpha^{10}, \alpha^{10})$	$0, \alpha^3, \alpha^{10}, \alpha^{12}$
$(\alpha^{10}, 1)$	$0, 1, \alpha^2, \alpha^8$
(α^{11}, α^5)	$0, \alpha^8, \alpha^9, \alpha^{12}$
$(\alpha^{11}, \alpha^{10})$	$0, \alpha^2, \alpha^{13}, \alpha^{13}$
$(\alpha^{11}, 1)$	$0, \alpha^3, \alpha^4, \alpha^7$
$(1, 0)$	$0, \alpha^5, \alpha^{10}, \alpha^{15}$

Table 1. Specifying CLIs of $S(x) = x^{-1}$ over $\mathbb{F}_{2^4}$ of the form $g_1(x) = a_3x^3 + a_5x^5 + a_3^2x^6 + a_3^8x^9 + a_5^2x^{10} + a_3^4x^{12}$ and $g_2(x) = a_3^4x^3 + a_5^2x^5 + a_3^8x^6 + a_3^2x^9 + a_5x^{10} + a_3x^{12}$ determined by the pair (a_3, a_5), where $a_5 \in \{0, 1, \alpha^5, \alpha^{10}\}$ with α being a generator of $\mathbb{F}_{2^4}^*$.

Remark 3 *A design whose substitution layer consists of identical 4-bit inverse S-boxes along with the use of an orthogonal binary matrix to provide diffusion, might be vulnerable to CLI-based attacks if the round constants are not carefully chosen.*

The same analysis for $m > 4$, representing both g_1 and g_2 as univariate polynomials and keeping only linear and quadratic terms, gives:

$$g_1(x) = \sum_{0 \le i \le j < m} a_{2^i+2^j} x^{2^i+2^j}; \quad g_2(x) = \sum_{0 \le i \le j < m} b_{2^i+2^j} x^{2^i+2^j}.$$

As before, the condition that $g_1(x) + g_2(y) = g_1(x) + g_2(x^{-1}) = c$ this time implies that the coefficients $b_{2^i+2^j}$ of $g_2(x)$ must satisfy the following equation,

$$b_l \neq 0 \iff -l \equiv k \bmod 2^m - 1, \tag{2}$$

where $1 \le l, k \le 2^{m-1} + 2^{m-2}$ and both have weight less than or equal to two. Notice that $2 \le k + l \le 2(2^{m-1}) + 2(2^{m-2}) = 2^m + 2^{m-1}$. Therefore the only possible solutions to (2) is $k + l = 2^m - 1$ which would imply that (for quadratic coefficients) we would have $2^m - 1 = 2^a + 2^b + 2^c + 2^d$ for some integers $a, \ldots, d$ of weight one. This is however impossible whenever $m > 4$ and therefore there are no quadratic CLIs for the inverse function $S(x) = x^{-1}$ over $\mathbb{F}_{2^m}$ when $m > 4$.

4 Cycle structure of generalized nonlinear invariants

Structurally, generalized nonlinear invariants offer more diversity due to the involvement of two constants in their definition which may be useful for eliminating the impact of round constants. Most notably, the distinguishing attack on iSCREAM [GLSV14] block cipher described in [YWP19] uses completely different family of weak keys than those related to a distinguishing attack based on standard invariants.

A natural question, concerning the cardinality and structure of generalized invariants, is whether similar results can be deduced using the cycle structure of a given bijective S-box. More specifically, having $g(x + a_1) + g(S(x) + a_2)) = c$ satisfied for all $x \in \mathbb{F}_2^m$, which is equivalent to $g(x) = g(S(x + a_1) + a_2))$ when $c = 0$, means that we look for cycles of the form

$$x^{(1)} \to S(x^{(1)} + a_1) + a_2 x^{(2)} \cdots \to S(x^{(2)} + a_1) + a_2 x^{(r-1)} \to S(x^{(r-1)} + a_1) + a_2 x^{(1)}.$$

Since $S(x+a_1)+a_2$ is a permutation, the same reasoning applies as for standard invariants. However, the cycle structure depends on the choice of a_1, a_2 and it is not necessarily identical to the structure of S.

A trivial connection relating cryptographically (affine) equivalent S-boxes on $\mathbb{F}_2^m$ can be easily deduced. Namely, defining affinely equivalent S-box of S as $S'(x) = S(x + a_1) + a_2$, which is also a permutation, a generalized invariant $g(x + a_1) + g(S(x) + a_2)) = c$ gives rise to a standard invariant of S' since now we have $g(x_i) + g(S'(x_i)) = c$. Notice that S' is a unique S-box derived from S for a given a_1, a_2, indicating the equivalence of S and S' with respect to a fixed g and a_1, a_2. More formally, defining the translates $\pi_a(x) = x + a$ where $x, a \in \mathbb{F}_2^m$, this affine transform can be compactly written as $S'(x) = \pi_{a_2} \circ S \circ \pi_{a_1}(x)$. This equivalence can be easily extended to the whole substitution layer $\mathcal{S}$ so that

$$\mathcal{S} \overset{\mathbf{a}_1,\mathbf{a}_2}{\simeq} \mathcal{S}' \Leftrightarrow \mathcal{S}' = S_1' \times S_2' \times \cdots \times S_t',$$

where $S_i'(x) = S(x + a_1) + a_2$, for $i = 1, \ldots, t$, and $\mathbf{a}_1 = (a_1, \ldots, a_1)$, $\mathbf{a}_2 = (a_2, \ldots, a_2)$.

Given a permutation P on $\mathbb{F}_2^m$, we will denote the set of standard invariants by $\mathcal{U}(P) := \{g \in \mathcal{B}_m : g(x) + g(P(x)) = c, \text{ for all } x \in \mathbb{F}_2^m\}$. Note that for $a, b \in \mathbb{F}_2^m$ we have $\mathcal{U}(P, a, b) = \mathcal{U}(\pi_b \circ P \circ \pi_a)$. The set of generalized invariants of P will be denoted by $\mathsf{GI}(P)$, thus

$$\mathsf{GI}(P) = \bigcup_{a,b \in \mathbb{F}_2^m} \mathcal{U}(\pi_b \circ P \circ \pi_a).$$

We will usually refer to elements of $\mathsf{GI}(P)$ simply as invariants.

A further refinement of structural properties of generalized invariants can be deduced by specifying a subset of permutations for which a fixed Boolean function g is a standard invariant,

$$\mathcal{P}_g := \{P \in Sym(\mathbb{F}_2^m) : g \in \mathcal{U}(P)\}, \tag{3}$$

where $Sym(\mathbb{F}_2^m)$ denotes the set of all permutations on $\mathbb{F}_2^m$. The following properties related to $\mathcal{P}_g$ are then easily established.

Theorem 4. *For a fixed Boolean function $g \in \mathcal{B}_m$ the set $\mathcal{P}_g$ is a group under composition.*

Proof. Firstly, $id \in \mathcal{P}_g$ since $g(x) + g(x) = 0$ for every $x \in \mathbb{F}_2^m$. For $P \in \mathcal{P}_g$ and $Q \in \mathcal{P}_g$ define $y = Q^{-1}(x)$, where $x \in \mathbb{F}_2^m$. Then

$$g(x) + g(P \circ Q^{-1}(x)) = g(Q(y)) + g(P(y)) = g(Q(y)) + g(y) + g(y) + g(P(y)) = c.$$

This proves that $P \circ Q^{-1} \in \mathcal{P}_g$, thus $\mathcal{P}_g$ is a group under composition. □

Corollary 2. *Let P and Q be two permutations on $\mathbb{F}_2^m$. Consider $g \in \mathcal{B}_m$ such that $g \in \mathcal{U}(Q \circ P)$. Then, it holds that $g \in \mathcal{U}(P) \iff g \in \mathcal{U}(Q)$.*

It should also be noted that standard nonlinear invariants always give rise to generalized invariants whenever an S-box admits nonzero linear structures. More precisely, assuming that there exist $a \neq 0 \in \mathbb{F}_2^m$ and $b \in \mathbb{F}_2^m$ such that $S(x + a) + S(x) = b$ for all $x \in \mathbb{F}_2^m$, then the existence of a non-constant function $g \in \mathcal{B}_m$ so that $g(x) + g(S(x)) = 0$, for all $x \in \mathbb{F}_2^m$, implies that $g(x + a) + g(S(x + a)) = 0$ which consequently gives $g(x + a) + g(S(x) + b) = 0$. Thus, assuming the existence of linear structures of S, a standard invariant induces a generalized nonlinear invariant as:

$$g(x) + g(S(x)) = 0 \Longleftrightarrow g(x + a) + g(S(x + a)) = 0 \Longleftrightarrow g(x + a) + g(S(x) + b) = 0.$$

Using these results we have the following connection between standard and generalized invariants. We recall that linear structures of a Boolean function $g \in \mathcal{B}_m$ build a linear subspace which is denoted by $\mathsf{LS}(g) = \{a \in \mathbb{F}_2^m : g(x) + g(x + a) = c, \forall x \in \mathbb{F}_2^m\}$.

Theorem 5. *Let S be a bijective S-box on $\mathbb{F}_2^m$ and $g \in \mathcal{B}_m$. If $a, b \in \mathsf{LS}(g)$, then $g \in \mathcal{U}(S) \iff g \in \mathcal{U}(S, a, b)$. Also, if $S(x) + S(x + a) = b$ then $g \in \mathcal{U}(S) \iff g \in \mathcal{U}(S, a, b)$.*

Proof. Suppose that $g \in \mathcal{U}(S)$. Since $a \in \mathsf{LS}(g)$, we have $g(x)+g(x+a) = c_1$ and furthermore, as g is an invariant, $g(x + a) + g(S(x + a)) = c_2$. Now, because $b \in \mathsf{LS}(g)$ it holds that $g(S(x + a)) + g(S(x + a) + b) = c_3$. Putting this together gives

$$g(x)+g(S(x+a)+b) = g(x)+g(x+a)+g(x+a)+g(S(x+a))+g(S(x+a))+g(S(x+a)+b)$$

$$= c_1 + c_2 + c_3,$$

thus $\mathcal{U}(S, a, b)$. Conversely, suppose that $g \in \mathcal{U}(S, a, b)$. We simply observe that

$$g(x) + g(S(x)) = g(x) + g(x + a) + g(x + a) + g(S(x) + b) + g(S(x) + b) + g(S(x)) = c.$$

If $S(x) = S(x+a)+b$, then the result comes from $g(x)+g(S(x)) = g(x)+g(S(x+a)+b)$. □

Corollary 3. *Let S be a permutation on $\mathbb{F}_2^m$. Suppose that $g \in \mathcal{B}_m$ is a standard invariant of S, then $\alpha \in \mathsf{LS}(g)$ if and only if g is a standard invariant of $S \circ \pi_\alpha$.*

Proof. The sufficiency is given by the first part of Theorem 5 with $b = 0$. Now suppose $g(x)+g(S(x+\alpha)) = c_1$, for every $x \in F_2^m$. By hypothesis, we know that $g(x+\alpha)+g(S(x+\alpha)) = c_2$ for all $x \in \mathbb{F}_2^m$. Adding together the two equations we obtain that $g(x) + g(x + \alpha) = c_1 + c_2$ for every $x \in F_2^m$, which proves the claim. □

Another important and useful property regarding invariants is that translated permutations have the same set of invariants, which is the content of the following theorem.

Theorem 6. *Let S and S' be two permutations on $\mathbb{F}_2^m$. If there exist $a, b \in \mathbb{F}_2^m$ such that $S = \pi_b \circ S' \circ \pi_a$, then $\mathsf{GI}(S) = \mathsf{GI}(S')$.*

Proof. Consider an arbitrary invariant $g \in \mathsf{GI}(S)$, by definition there exist α and β such that $g \in \mathcal{U}(\pi_\beta \circ S \circ \pi_\alpha)$. The fact that $S = \pi_b \circ S' \circ \pi_a$ implies that for every element $x \in \mathbb{F}_2^m$, we have that

$$g(x) + g(S'(x + \alpha + a) + b + \beta) = g(x) + g(S(x + \alpha) + \beta).$$

Since $g \in \mathcal{U}(\pi_\beta \circ S \circ \pi_\alpha)$ the latter is equal to a constant c for every $x \in \mathbb{F}_2^m$. This means that $g \in \mathcal{U}(\pi_{b+\beta} \circ S' \circ \pi_{\alpha+a})$ therefore $g \in \mathsf{GI}(S')$. The other inclusion follows by symmetry. □

Example 1 *We consider a permutation S on $\mathbb{F}_2^4$ specified as:*

$$\{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\} \overset{S}{\mapsto} \{0\ 3\ C\ F\ 7\ 8\ A\ E\ 1\ 5\ 6\ B\ 4\ 2\ D\ 9\}.$$

Taking $a_1 = a_2 = 0001$, the equivalent mapping $S'(x) = S(x + a_1) + a_1$ is given by

$$\{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\} \overset{S'}{\mapsto} \{2\ 1\ E\ D\ 9\ 6\ F\ B\ 4\ 0\ A\ 7\ 3\ 5\ 8\ C\},$$

and its cycle structure is given by:

$$(0, 2, E, 8, 4, 9), (1), (3, D, 5, 6, F, C), (7, B), (A),$$

whose cycle structure remains unchanged compared to S. This can be deduced by noting that π_{a_1} has order two in $Sym(\mathbb{F}_2^4)$ and S' is the composition $S' = \pi_{a_1} \circ S \circ \pi_{a_1}$. Then, S' is a conjugate of S having the same cycle structure as S. Both S-boxes have $2^5 = 32$ standard and 2146 generalized invariants.

Now, taking for instance a permutation T on $\mathbb{F}_2^4$ to be a cyclic shift of its inputs

$$\{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\} \overset{T}{\mapsto} \{1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\ 0\},$$

the degree distribution of $\mathsf{GI}(T)$ is given by: 2 constant, 2 linear, 104 quadratic, 2976 cubic and 4680 quartic Boolean functions.

Remark 4 *It is important to notice that none of the 106 quadratic invariants of T in Example 1 is a standard invariant. Moreover, all of these quadratic invariants admit non-trivial linear structures, in fact, for every such quadratic invariant g it holds that $|\mathsf{LS}(g)| = 4$.*

4.1 Specifying cardinality of generalized invariants

We now address the problem of deriving the exact cardinality of triples (g, a, b) such that $g \in \mathcal{U}(\pi_b \circ S \circ \pi_a)$ based on the cycle structure of a given bijective permutation S on $\mathbb{F}_2^m$.

Lemma 2. *Let S and T be two permutations on $\mathbb{F}_2^m$. The set*

$$C_T := \{\ (a, b) \in \mathbb{F}_2^m \times \mathbb{F}_2^m\ :\ \pi_b \circ S \circ \pi_a \text{ has the same cycle structure as } T\ \}$$

has $2^m l$ elements, where l is the size of the set $B := \{\ b \in \mathbb{F}_2^m : \pi_b \circ S$ is a conjugate of $T\ \}$.

Proof. Since conjugacy is a transitive relation we have

$$\pi_b \circ S \circ \pi_a \text{ is a conjugate of } T \iff \pi_a \circ \pi_b \circ S \circ \pi_a \circ \pi_a = \pi_{a+b} \circ S \text{ is a conjugate of } T.$$

This in turn implies that

$$(a,b) \in C_T \iff a+b \in B \tag{4}$$

In particular, $C_T = \emptyset$ if and only if $B = \emptyset$, therefore if $B = \emptyset$ the result follows. Thus, we may suppose that $B \neq \emptyset$.

Consider the function $F: C_T \to \mathbb{F}_2^m \times B$ defined by $(a,b) \mapsto (a, a+b)$. Using (4), we see that F is well-defined. Now we will prove that F is bijective. First, we show that F is injective. Indeed, taking $(a,b) \neq (a',b') \in C_T$ such that $a \neq a'$ we obviously have $F((a,b)) \neq F((a',b'))$. Now, assuming $a = a'$ and $b \neq b'$ we clearly have $a+b \neq a+b'$ hence $F((a,b)) \neq F((a',b'))$, proving that F is injective. To show that F is surjective, let $y = (a,d) \in F_2^m \times B$ and use (4) to deduce that $x \in C_T$ for which $F(x) = y$. Thus, F is a bijection and consequently $|C_T| = 2^m l$. □

Remark 5 *The previous lemma also gives an efficient method for finding all the elements of the set C_T provided that $l < 2^m$: compute the elements of the set B, then for every $b \in B$ and $a \in \mathbb{F}_2^m$ obtain the element $(a, a+b_i) \in C_T$.*

We extend this approach for the purpose of an exact specification of the cardinality of generalized invariants. Consider a relation $\backsim$ over $\mathbb{F}_2^m \times \mathbb{F}_2^m$ defined by

$$(a,b) \backsim (a',b') \iff \pi_a \circ S \circ \pi_b \text{ has the same cycle structure of } \pi_{a'} \circ S \circ \pi_{b'}.$$

It can be readily seen verified that this is an equivalence relation. Let us denote with s the number of elements in $\mathbb{F}_2^m \times \mathbb{F}_2^m / \backsim$. To specify a set of representatives, let R denote a subset of $\mathbb{F}_2^m$ which is maximal with respect to the following property,

$$\text{for every } b, b' \in R \text{ it holds that } \pi_b \circ S \nsim \pi_{b'} \circ S. \tag{5}$$

Then a set of representatives for this equivalence relation is given by $\{\ (0,b)\ \}_{b \in R}$. Let $b_1, \ldots, b_s$ denote the distinct elements in R. In the light of the previous lemma, we note that each equivalence class has $2^m l_i$ elements, where

$$l_i = |\{b \in \mathbb{F}_2^m : \pi_b \circ S \text{ is a conjugate of } \pi_{b_i} \circ S\}|.$$

Thereby, we deduce the following formula to count the number of triples (g,a,b) where $g \in \mathcal{U}(\pi_b \circ S \circ \pi_a)$.

Theorem 7. *Let S be a permutation on $\mathbb{F}_2^m$ and let k_i denote the number of disjoint cycles of $\pi_{b_i} \circ S$, where $\{b_i\}_{1 \leq i \leq s}$ are elements of $R \subset \mathbb{F}_2^m$ which is maximal w.r.t. the property given in (5). Let us also define $\varepsilon_i = 1$ if there are no cycles of odd length and $\varepsilon_i = 0$ otherwise. The cardinality of distinct non-trivial (g being a non-constant function) triples (g,a,b) satisfying $g(x) + g(S(x+a)+b) = c$ for all $x \in \mathbb{F}_2^m$, equals*

$$2^m \sum_{i=1}^{s} l_i (2^{k_i + \varepsilon_i} - 2).$$

Proof. For every choice of (a,b), g has to be a standard invariant of $\pi_b \circ S \circ \pi_a$. As there exists i such that $\pi_b \circ S \circ \pi_a$ has the same cycle structure as $\pi_{b_i} \circ S$, then the number of non-constant standard invariants of $\pi_b \circ S \circ \pi_a$ is $2^{k_i+\varepsilon_i} - 2$. Since there are $2^m l_i$ possible pairs with the same cycle structure, we have $2^m l_i (2^{k_i+\varepsilon_i} - 2)$ invariants coming from the elements in the class of $(0, b_i)$. The same argument applies to every equivalence class, so there are in total $\sum_{i=1}^{s} 2^m l_i (2^{k_i+\varepsilon_i} - 2)$ distinct triples. □

Corollary 4. *The size of the subspace of linear structures of S is less than or equal to l_i, where l_i is defined as before and $(0, b_i) \sim (0, 0)$.*

Proof. Every linear structure a of S satisfies $\pi_b \circ S \circ \pi_a = S$ so (a, b) is related to $(0, 0)$ thus with $(0, b_i)$, but also this equation uniquely determines a for a given b. Therefore, there can be at most l_i linear structures of S. □

Example 2 *We consider the same permutation S on $\mathbb{F}_2^4$ as in Example 1 specified as:*

$$\{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\} \overset{S}{\mapsto} \{2\ 1\ E\ D\ 9\ 6\ F\ B\ 4\ 0\ A\ 7\ 3\ 5\ 8\ C\},$$

By letting R be the set $\{0, 1, 2, 5, 6, 8, 9, B, C, E, F\}$, computer simulations verify that R indeed satisfies (5), therefore $s = 11$ and the parameters l_i and k_i can be determined.

Rep.	*Cycle structure*	k_i	l_i
S	$(0,2,E,8,4,9),(1),(3,D,5,6,F,C),(7,B),(A)$	5	2
$\pi_1 \circ S$	$(0,1,2,D,3,E,C,5,9,4,6,B,A,7,F,8)$	1	4
$\pi_2 \circ S$	$(0,2,E,F,B,9,7,C,6,8,3,D),(1),(4,5,A)$	3	2
$\pi_5 \circ S$	$(0,5,D,7,B,E,8,4,2,9),(1,6,F,C),(3,A)$	3	1
$\pi_6 \circ S$	$(0,6,C,2,A),(1,5,E,B,D,4),(3,9),(7,8),(F)$	5	1
$\pi_8 \circ S$	$(0,8,9,D,A,E,5),(1,B,3,7,6,2,4,F),(C)$	3	1
$\pi_9 \circ S$	$(0,9,C,D,B,2,5,1,A,F),(3,6),(4,E),(7),(8)$	5	1
$\pi_B \circ S$	$(0,B),(1,8,A,D,9,E,6),(2,7,5,3,4,C,F)$	3	1
$\pi_C \circ S$	$(0,C,8,D,E,1,F,5,4,B,7,2),(3),(6),(9),(A)$	5	1
$\pi_E \circ S$	$(0,E,3,1,D,C,A,8,F,7),(2),(4,9,B,5,6)$	3	1
$\pi_F \circ S$	$(0,F,6,5,7,1,C,B,4,8,E,2,3),(9,A),(D)$	3	1

By Theorem 7, the number of non-constant triples (g, a, b) where $g \in \mathcal{U}(\pi_b \circ S \circ \pi_a)$ is exactly

$$2^4(2 \cdot 30 + 4 \cdot 0 + 2 \cdot 6 + 1 \cdot 6 + 1 \cdot 30 + 1 \cdot 6 + 1 \cdot 30 + 1 \cdot 6 + 1 \cdot 30 + 1 \cdot 6 + 1 \cdot 6) = 3072.$$

The degree distribution of $\mathsf{GI}(S)$ consists of 2 *constant,* 0 *linear,* 32 *quadratic,* 1088 *cubic and* 1024 *quartic invariants, thus giving* 2146 *invariants altogether. Once again, none of the quadratic ones being a standard invariant. By Theorem 5, every affine translate of S will have the same set of generalized invariants hence the same degree distribution.*

Remark 6 *Notice that the number of standard invariants for the S-box in Example 2 is given as $\#g = 2^{(\#\text{ cycles of } F)} = 2^5 = 32$ which is negligible compared to the number of generalized invariants!*

As already demonstrated, every permutation admits a large number of generalized invariants which entirely depends on its cycle structure. Nevertheless some permutations do not admit quadratic invariants at all. Two such examples are the 8-bit AES S-box and the 4-bit PRESENT S-box [BKL+07].

Example 3 *The PRESENT cipher uses as S-box the following permutation:*

$$\{0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\} \overset{S}{\mapsto} \{C\ 5\ 6\ B\ 9\ 0\ A\ D\ 3\ E\ F\ 8\ 4\ 7\ 1\ 2\}.$$

Based on computer simulations its set of generalized invariants $\mathsf{GI}(S)$ is of size $|\mathsf{GI}(S)| =$ 2178 and none of these is a quadratic one. The inverse S-box (Rijndael S-box) of AES, as a permutation on $\mathbb{F}_2^8$, only admit generalized invariants of degree at least 6.

4.2 Finding generalized invariants - an application to AES

We have already remarked that there are no quadratic generalized invariants for the S-box of AES, neglecting the fact that their exact specification is computationally infeasible due to a large search space for the triples (g, a_1, a_2), when $g \in \mathcal{B}_8$ and $a_1, a_2 \in \mathbb{F}_2^8$. In what follows, we describe an efficient search algorithm which exploits certain regularities in the spectrum of cycle decompositions of the translates of a given S-box. The succeeding analysis is performed on the S-box of AES but it might be applied to other S-boxes which exhibit similar regularities.

For a given bijective S-box S on $\mathbb{F}_2^m$ let $\mathcal{T}$ denote the set of all translates of S, i.e, elements of the form $\pi_b \circ S \circ \pi_a$ for some $a, b \in \mathbb{F}_2^n$. Define the sets P_δ consisting of all elements of $\mathcal{T}$ whose cycle decomposition has exactly δ disjoint cycles. Let Δ be a subset of the natural numbers such that $\{P_\delta\}_{\delta \in \Delta}$ is a partition of $\mathcal{T}$.

As an invariant must also be a standard invariant of some element in $\mathcal{T}$, we are searching for standard invariants through all the elements in $\mathcal{T}$. An initial approach might be simply to compute $\mathcal{U}(T)$ for every $T \in \mathcal{T}$ using the cycle decomposition of T and then calculating the algebraic degree of every element $g \in \mathcal{U}(T)$. However, this procedure has a time complexity of $\mathcal{O}(m2^{6m})$ operations. We propose a slightly modified version of this approach which induces a time memory trade-off.

Let us identify $\mathbb{F}_2^m$ with the set $\{1, \ldots, 2^m\}$ via the lexicographical ordering of its elements and suppose that $\Delta = \{\delta_1, \ldots, \delta_{|\Delta|}\}$. In order to find the invariants of a permutation S on $\mathbb{F}_2^m$ we carry out the steps detailed in Algorithm 1.

Algorithm 1: Finding invariants and their degree distribution of bijective S-box

Input: A permutation S on $\mathbb{F}_2^m$.
Output: The set $\mathsf{GI}(S)$ and its degree distribution

1. Initialize a $2^m \times 2^m-$matrix M containing all the information regarding the cycle structure of elements in $\mathcal{T}$, namely, M_{ij} is the cycle decomposition of $\pi_j \circ S \circ \pi_i$.
2. Create a second matrix I with $|\Delta|$ columns whose k-th column contains a list of positions (i, j) with the property that the translate $\pi_j \circ S \circ \pi_i$ lies in P_{δ_k}.
3. For every k such that $1 \leq k \leq |\Delta|$ and for every $(i, j) \in I_k$ compute each element $g \in \mathcal{U}(\pi_i \circ S \circ \pi_j)$ using the information in M_{ij} and then calculate its algebraic degree.

The outcome of the algorithm will be a set of generalized invariants of S and their degree distribution. The running time of the algorithm depends upon how large the spaces of standard invariants $\mathcal{U}(T)$ of elements $T \in \mathcal{T}$ are, which in turn depends on the permutation S. Running the algorithm in $|\Delta|$ parallel processes is a plausible option for the purpose of saving both time and memory. Additionally, assuming that for every $\delta \in \Delta$ the sizes $|P_\delta|2^{\delta+1}$ are *approximately* equal to 2^{2m}, we can estimate the number of steps of the algorithm for a fixed $\delta \in \Delta$ to be $\mathcal{O}(m2^{4m})$.

In the case of AES, the algorithm can be efficiently applied since $\Delta = \{1, 3, 5, 7, 9, 11\}$, moreover, the following facts are easily obtained.

- $S \in P_5$;
- $|P_1| = 512$; $|P_3| = 7936$; $|P_5| = 26112$; $|P_7| = 18432$; $|P_9| = 10496$; $|P_{11}| = 2048$.

Table 2 shows the running time of the algorithm in the case of AES for every $\delta \in \Delta$.

δ	1	3	5	7	9	11
$\log_2(ns(\delta))$	22.4594	27.4136	32.1319	32.6294	33.817	33.4594

Table 2. Binary logarithm of the time complexity of Algorithm 1 for finding the generalized invariants and their degree spectrum of the Rijndael S-box for a fixed $\delta \in \{1, 3, 5, 7, 9, 11\}$.

Remark 7 *Using the cycle decomposition of translates of a given permutation S on $\mathbb{F}_2^8$ one can compute the degree distribution $\mathsf{GI}(S)$ much faster than using a brute-force approach.*

5 Nonlinear invariant through active cycle sets

In this section, we introduce the concept of an active cycle set which gives us the possibility to specify nonlinear invariants of the round function for AES-like ciphers.

5.1 The concept of active cycle set and induced invariants

To investigate the cycle structure of two concatenated S-boxes, we introduce the concept of an active cycle set as follows.

Definition 1. *Let S_1 and S_2 be two bijective S-boxes defined on the spaces $\mathbb{F}_2^{n_1}$ and $\mathbb{F}_2^{n_2}$, respectively. For a fixed element $x_0 \in \mathbb{F}_2^{n_2}$, let $C_{x_0} := \{ S_2^j(x_0) \ : \ 1 \leq j \leq d\}$ be its corresponding cycle of length d. We call the set $\Gamma = \{(x_1, x_2), x_1 \in \mathbb{F}_2^{n_1}, x_2 \in C_{x_0}\}$ an* active cycle set *with respect to $S^*(x_1, x_2) = (S_1(x_1), S_2(x_2))$.*

Remark 8 *If Γ is an active cycle set then $|\Gamma| = 2^{n_1}d$. Notice also that for every $\alpha = (\alpha_1, \mathbf{0}) \in \mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2}$ and $\beta \in \Gamma$, we have $\alpha + \beta \in \Gamma$. Additionally, an active cycle set can be regarded as a union $\Gamma = \bigcup_{x \in \mathbb{F}_2^{n_1}} (C_x \times C_{x_0})$.*

It was pointed out in [TLS18] that the characteristic functions of cycles form a basis of the space of standard invariants. We make this statement more precise and establish the relation between active cycle sets and nonlinear invariants of two concatenated S-boxes.

Observation 1 *Let x_0 be a fixed element in $\mathbb{F}_2^{n_2}$ and $\Gamma = \{(x_1, x_2), x_1 \in \mathbb{F}_2^{n_1}, x_2 \in C_{x_0}\}$ be an active cycle set with respect to $S^*(x_1, x_2) = (S_1(x_1), S_2(x_2))$. Let $g \in \mathcal{B}_{n_1+n_2}$ be the characteristic function of Γ so that $g(x_1, x_2) = 1$ if $(x_1, x_2) \in \Gamma$ and $g(x_1, x_2) = 0$ otherwise. Then, g is a standard nonlinear invariant of S^*.*

Proof. Suppose that $x = (x_1, x_2) \in \Gamma$ so that $g(x) = 1$. It easily follows that $y = (y_1, y_2) = (S_1(x_1), S_2(x_2)) \in \Gamma$ implying that $g(y) = 1$. Similarly, for any $x \notin \Gamma$ we have that $g(x) = 0$. The fact that $x \notin \Gamma$ implies that $x_2 \notin C_{x_0}$, which means $S_2(x_2) \notin C_{x_0}$ and consequently $g(y) = 0$. Therefore, $g(x)$ is a nonlinear invariant of S^*. □

In a similar fashion, one can construct other invariants of $S^*(x_1, x_2) = (S(x_1), S(x_2))$. Indeed, instead of using an active cycle set containing a single cycle one can consider a union of disjoint cycles $\mathcal{C}$ so that $\Gamma^* = \{(x_1, x_2), x_1 \in \mathbb{F}_2^{n_1}, x_2 \in \mathcal{C}\}$, and accordingly define h to be the characteristic function of Γ^*. It can be readily verified that h is also a nonlinear invariant of S^*. In particular, h is a balanced Boolean function on $\mathbb{F}_2^{n_1+n_2}$ when $|\Gamma^*| = 2^{n_1+n_2-1}$.

The above discussion implies the possibility of deriving a general method of designing nonlinear invariants by combining the active bytes and cycles rather than the traditional usage of algebraic normal form (ANF) of Boolean function (or only the cycle structure of a given S-box). Moreover, it also indicates that the support of a nonlinear invariant is closely related to the active cycle set with respect to S^*.

5.2 Cycle structure of the entire S-box layer for AES-like ciphers

In this section, we propose an algorithm for determining the exact number of cycles (and their lengths) of two concatenated bijective S-boxes which can be used iteratively for specifying the cycle structure of the entire S-box layer.

Observation 2 *Let S_1 and S_2 be two permutations defined over $\mathbb{F}_2^m$. Let x_1 and x_2 be two fixed elements in $\mathbb{F}_2^m$ and consider the cycles $C_{x_1} = \{ S_1^i(x_1) : 1 \leq i \leq d_1\}$ and $C_{x_2} = \{ S_2^i(x_2) : 1 \leq i \leq d_2\}$ of lengths d_1 and d_2, respectively. Consider the permutation $S : \mathbb{F}_2^{2m} \to \mathbb{F}_2^{2m}$ given as a parallel application (or concatenation) of S_1 and S_2, i.e., $S(x_1, \ldots, x_{2m}) = (S_1(x_1, \ldots, x_m), S_2(x_{m+1} \ldots, x_{2m}))$. Then, there are exactly $\gcd(d_1, d_2)$ cycles of S induced by the elements in C_{x_1} and C_{x_2}. In particular, if $\gcd(d_1, d_2) = 1$, then there is only one cycle stemming from C_{x_1} and C_{x_2} in the cycle decomposition of S.*

Proof. We claim that the length l of the cycle $C_z = \{S^i(z) : 1 \leq i \leq l\}$, for $z := (x, y) \in C_{x_1} \times C_{x_2}$, is equal to $lcm(d_1, d_2)$. Indeed, it is obvious that $l \leq lcm(d_1, d_2)$. Now, as $S^l(x, y) = (S_1^l(x), S_2^l(y))$ we infer that $S_1^l(x) = x$ and $S_2^l(y) = y$. Therefore, $d_1 | \ l$ and $d_2 | \ l$ and thus $lcm(d_1, d_2) \leq l$. Hence, $l = lcm(d_1, d_2)$. Since there are $d_1 d_2$ elements in $C_{x_1} \times C_{x_2}$, we observe that there will be $\frac{d_1 d_2}{lcm(d_1, d_2)} = \gcd(d_1, d_2)$ different cycles in the cycle structure of S, coming from elements in $C_{x_1} \times C_{x_2}$. The case $\gcd(d_1, d_2) = 1$ easily follows. □

Remark 9 *In general, the above result is also valid when bijective S-boxes S_1 and S_2 are not defined on the same variable space.*

Example 4 *Let $A := \{a_1, a_2, a_3, a_4\}$ and $B := \{b_1, b_2, \ldots, b_6\}$ be two cycles in the cycle structure of the S-boxes S_1 and S_2, respectively. From Observation 2, we conclude there are $\gcd(4, 6) = 2$ cycles for $S(x, y) = (S_1(x), S_2(y))$, induced by A and B, given by:*

$$D_0 = \{(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), (a_1, b_5), (a_2, b_6), (a_3, b_1), (a_4, b_2),$$

$$(a_1, b_3), (a_2, b_4), (a_3, b_5), (a_4, b_6)\},$$

$$D_1 = \{(a_2, b_1), (a_3, b_2), (a_4, b_3), (a_1, b_4), (a_2, b_5), (a_3, b_6), (a_4, b_1), (a_1, b_2),$$

$$(a_2, b_3), (a_3, b_4), (a_4, b_5), (a_1, b_6)\}.$$

Notice that the length of each cycle is $lcm(4, 6) = 12$.

Observation 2 induces the following algorithm for determine the exact number of cycles in the cycle structure of two concatenated S-boxes.

Algorithm 2: Finding the exact number of cycles in the cycle structure of two concatenated S-boxes

Input: Two permutations S_1 and S_2.

Output: The exact number of cycles for the concatenation $S = (S_1, S_2)$

1. Calculate the lengths of each cycle in the cycle decompositions of both S_1 and S_2.
2. For each pair of these lengths, compute the corresponding length of the cycle in the cycle decomposition of S via Observation 2.
3. Return the sum of the numbers obtained in the previous step.

Example 5 *Let S_1 be the 4-bit S-box used in PRINCE [BCG+12] block cipher. It is easily checked that its cycle structure consists of four cycles. More precisely,*

- $A_1 = \{0, 11\}$,
- $A_2 = \{1, 15, 4, 10, 8, 6, 9, 7\}$,
- $A_3 = \{2, 3\}$,
- $A_4 = \{5, 12, 14, 13\}$.

The cardinalities of these cycles are $|A_1| = 2, |A_2| = 8, |A_3| = 2$ and $|A_4| = 4$. Using Algorithm 2, we can determine the exact number of cycles in the cycle structure of two concatenated Prince S-boxes $S = (S_1, S_1)$, i.e., $4 \times 2 + 2 + 8 + 2 + 4 + 4 \times 2 + 2 + 4 + 2 + 4 = 44$. Moreover, we can also determine the specific lengths of these cycles. We can further apply Algorithm 2 to other concatenated S-boxes, for instance, the number of cycles in the cycle decomposition of two concatenated Skinny [BKL+16] S-boxes is 30 (Skinny's cycle decomposition contains 4 cycles) and the number of cycles in the cycle structure of two concatenated AES S-boxes is 336.

In fact, using Algorithm 2 iteratively, one can in general calculate the exact number of cycles in the decomposition of the permutation consisting of l parallel applications of (not necessarily) identical bijective S-boxes. We illustrate this in Table 3 below, by considering S-boxes of LED cipher [GPPRM] of size 4×4.

Number l of parallel applications of S	2	4	8
Number of cycles	30	2928	82695528

Table 3. Number of cycles for l parallel applications of LED S-box

Remark 10 *The complexity of Algorithm 2 entirely depends on the cycle structure of involved S-boxes. In the worst case scenario, when the number of cycles is very large, the advantage of this approach may be insignificant compared to the basic method of considering $S = (S_1, S_2)$ as a bijective superbox and finding the cycles of S directly. Nevertheless, the case of having many small cycles seems not to be common which is also confirmed when analyzing the cycle structure of the S-boxes of AES and LED cipher for instance. Moreover, large cycles can also be identified in the $MC(SB(x))$ operation typical for AES-like ciphers, where MC and SB stand for the mix column and substitute byte operation, respectively.*

One can calculate the length of the cycles corresponding to $MC(SB(x))$ operation via Algorithm 2. For instance, when considering the LED cipher, the number of cycles referring to one column of $MC(SB(x))$ nonlinear operation (corresponding to a superbox which maps $\mathbb{F}_2^{4\times 4}$ to $\mathbb{F}_2^{4\times 4}$) is very small, namely it equals 10. The number of cycles for two columns of $MC(SB(x))$ operation, viewing these two columns as a superbox consisting of eight concatenated S-boxes and thus mapping $\mathbb{F}_2^{4\times 8}$ to $\mathbb{F}_2^{4\times 8}$, for LED cipher is 65740. These cycles could be obtained on a standard PC in less than 1 minute. The exact number of cycles can also be deduced using Algorithm 2 for the full permutation $MC(SB(x))$ consisting of 16 concatenated S-boxes, which can be performed efficiently (due to a small number of cycles for $MC(SB(x))$ columns) on dedicated platforms.

6 Conclusions

A detailed theoretical analysis of generalized and closed loop invariants, the concepts introduced in [YWP19], in terms of their cardinality and structure has been provided. The generalized concept of nonlinear invariants seems to be fully justified since in many cases real-life lightweight block ciphers do not admit quadratic invariants while having many quadratic generalized ones. A novel concept of active cycle set also appears to be useful when defining nonlinear invariants of concatenated S-boxes. Practical applications of this theory (in particular the use of active cycle set for specifying invariants) are currently investigated and these look quite promising.

7 Acknowledgements

Yongzhuang Wei is supported in part by the National Natural Science Foundation of China (61872103), in part by Guangxi Science and Technology Foundation (Guike AB18281019), in part by Guangxi Natural Science Foundation (2019GXNSFGA245004). Enes Pasalic is partly supported by the Slovenian Research Agency (research program P1-0404 and research projects J1-1694, J1-9108 and N1-0159).

References

[BBI+15] S. Banik, A. Bogdanov, T. Isobe, K. Shibutani, H. Hiwatari, T. Akishita & F. Regazzoni. Midori: A Block Cipher for Low Energy. In *Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part II*, pp. 411–436, 2015.

[BGL19] Z. Bao, J. Guo & E. List. Extended Truncated-differential Distinguishers on Round-reduced AES. Cryptology ePrint Archive, Report 2019/622, 2019, https://eprint.iacr.org/2019/622

[BCLR17] C. Beierle, A. Canteaut, G. Leander & Y. Rotella. Proving Resistance Against Invariant Attacks: How to Choose the Round Constants. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part II*, pp. 647–678, 2017.

[BCL18] C. Beierle, A. Canteaut & G. Leander. Nonlinear Approximations in Cryptanalysis Revisited. In *IACR Trans. Symmetric Cryptol.*, 2018(4), pp. 80–101, 2018.

[BKL+16] C. Beierle, J. Jean, S. Kölbl, G. Leander, A. Moradi, T. Peyrin, Y. Sasaki, P. Sasdrich & S.M. Sim. The SKINNY family of block ciphers and its low-latency variant MANTIS. In *CRYPTO 2016*, Robshaw, M., Katz, J. (eds.), Part II. LNCS, vol. 9815, pp. 123–153. Springer (2016)

[Beyne2018] T. Beyne. Block Cipher Invariants as Eigenvectors of Correlation Matrices. In *Advances in Cryptology - ASIACRYPT 2018 - 24th Annual International Cryptology Conference, Brisbane, Australia, December 02-06, 2018, Proceedings, Part I*, pp. 3–31, 2018.

[BBG+09] R. Benadjila, O. Billet, H. Gilbert G. Macario-Rat, T. Peyrin, M. Robshaw & Y. Seurin SHA-3 Proposal: ECHO. Submission to NIST (updated), 2009. http://crypto.rd.francetelecom.com/echo/doc/echo_description_1-5.pdf

[BKL+07] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin & C. Vikkelsoe. PRESENT: An Ultra-lightweight Block Cipher. In *Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings*, pp. 450–466, 2007.

[BS90] E. Biham & A. Shamir. Differential Cryptanalysis of DES-like Cryptosystems. In *Advances in Cryptology - CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings*, pp. 2–21, 1990.

[BCG+12] J. Borghoff, A. Canteaut, T. Güneysu, E. B. Kavun, M. Kneževič, L. R. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S. S. Thomsen & T. Yalcin. PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In *ASIACRYPT 2012,* Wang, X., Sako, K. (eds.) LNCS, vol. 7658, pp. 208–225. Springer (2012)

[DR02] J. Daemen, V. Rijmen. The Design of Rijndael: AES - The Advanced Encryption Standard. Information Security and Cryptography, Springer 2002, ISBN 3-540-42580-2

[GKM+11] P. Gauravaram, L. R. Knudsen, K. Matusiewicz, F. Mendel, C. Rechberger, M. Schläffer & S. S. Thomsen. Grøstl - a SHA-3 candidate. 2011-03-02.

[GLRTW19] L. Grassi, G. Leander, C. Rechberger, C. Tezcan & F. Wiemer. Weak-Key Subspace Trails and Applications to AES. Available at *IACR Cryptology ePrint Archive 2019*, https://eprint.iacr.org/2019/852.pdf.

[GLSV14] V. Grosso, G. Leurent, F. Standaert & K. Varici. SCREAM v2. Submission to CAESAR competition. 2014.

[HKM95] C. Harpes, G. G. Kramer & J. L. Massey. A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma. In *Advances in Cryptology - EUROCRYPT '95, International Conference on the Theory and Application of Cryptographic Techniques, Saint-Malo, France, May 21-25, 1995, Proceeding*, pp. 24–38, 1995.

[HTW15] T. Huang, I. Tjuawinata & H. Wu. Differential-Linear Cryptanalysis of ICEPOLE. In *Fast Software Encryption - 22nd International Workshop, FSE 2015, Istanbul, Turkey, March 8-11, 2015, Revised Selected Papers*, pp. 243–263, 2015.

[IP09] S. Indesteege & B. Preneel. SHA-3 Proposal: Lane. Submission to NIST (updated), 2009. http://www.cosic.esat.kuleuven.be/lane/

[KR96] L. R. Knudsen & M. J. B. Robshaw. Non-Linear Approximations in Linear Cryptanalysis. In *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, pp. 224–236, 1996.

[LAAZ11] G. Leander, M. A. Abdelraheem, H. AlKhzaimi & E. Zenner. A Cryptanalysis of PRINTcipher: The Invariant Subspace Attack. In *Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings*, pp. 206–221, 2011.

[LH94] S. K. Langford & M. E. Hellman. Differential-Linear Cryptanalysis. In *Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings*, pp. 17–25, 1994.

[MS2001] I. Mantin & A. Shamir. A Practical Attack on Broadcast RC4. In *Fast Software Encryption - FSE* '2001, LNCS 2355, pp. 152—64, 2001.

[Mat93] M. Matsui. Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, pp. 386–397, 1993.

[GPPRM] J. Guo, T. Peyrin, A. Poschmann & M. J. B. Robshaw. The LED block cipher. In *CHES 2011*, Preneel, B., Takagi, T. (eds.) LNCS, vol. 6917, pp. 326–341. Springer (2011)

[RGH17] S. Rønjom, N. G. Bardeh & T. Helleseth. Yoyo Tricks with AES, In *Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I*, Lecture Notes in Computer Science vol. 10624, pp. 217–243, Springer, 2017.

[TLS16] Y. Todo, G. Leander & Y. Sasaki. Nonlinear Invariant Attack - Practical Attack on Full SCREAM, iSCREAM, and Midori64. In *Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part II*, pp. 3–33, 2016.

[AAS08] A. A. Selcuk. On Probability of Success in Linear and Differential Cryptanalysis. *Journal of Cryptology*, vol. 21 (1), pp. 131-147, January 2008.

[TLS18] Y. Todo, G. Leander & Y. Sasaki. Nonlinear Invariant Attack: Practical Attack on Full SCREAM, iSCREAM, and Midori64. *Journal of Cryptology*, pp. 1432–1378, Apr 2018.

[YWP19] Y. Wei, T. Ye, W. Wu & E. Pasalic. Generalized Nonlinear Invariant Attack and a New Design Criterion for Round Constants. In *Fast Software Encryption, FSE '2019, Paris, France, March 24-26, 2019, Proceedings*, IACR Trans. Symmetric Cryptology, vol. 2018(4), pp. 62–79, 2019.