Boomerang Uniformity of Popular S-box Constructions

Shizhu Tian, Christina Boura, Léo Perrin

Boomerang Uniformity of Popular S-box Constructions

Shizhu Tian[1,3,4], Christina Boura[1,2], and Léo Perrin[1]

[1] Inria, Paris, France.
leo.perrin@inria.fr
[2] Université de Versailles, Versailles, France
christina.boura@uvsq.fr
[3] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[4] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
tianshizhu@iie.ac.cn

Abstract. In order to study the resistance of a block cipher against boomerang attacks, a tool called the Boomerang Connectivity Table (BCT) for S-boxes was recently introduced. Very little is known today about the properties of this table especially for bijective S-boxes defined for n variables with $n \equiv 0 \mod 4$. In this work we study the boomerang uniformity of some popular constructions used for building large S-boxes, e.g. for 8 variables, from smaller ones. We show that the BCTs of all the studied constructions have abnormally high values in some positions. This remark permits us in some cases to link the boomerang properties of an S-box with other well-known cryptanalytic techniques on such constructions while in other cases it leads to the discovery of new ones. A surprising outcome concerns notably the Feistel and MISTY networks. While these two structures are very similar, their boomerang uniformity can be very different. In a second time, we investigate the boomerang uniformity under EA-equivalence for Gold and the inverse function (as used respectively in MPC-friendly ciphers and the AES) and we prove that the boomerang uniformity is EA-invariant in these cases. Finally, we present an algorithm for inverting a given BCT and provide experimental results on the size of the BCT-equivalence classes for some 4 and 8-bit S-boxes.
Keywords: BCT, S-box, Feistel, MISTY, Lai-Massey, Gold

1 Introduction

To evaluate the security level of a block cipher, cryptanalysts have designed multiple techniques that aim at searching for undesirable patterns. One such technique is the so-called *differential cryptanalysis* [3] which looks for pairs (a,b) of input and output differences such that $E_k(x \oplus a) \oplus E_k(x) = b$, where E_k is (a round-reduced version of) the studied block cipher.

Block ciphers—but also other symmetric primitives such as hash functions for example—are often built using the so-called *S-boxes* as the source of their non-linearity. These are small functions mapping n bits to m, specified by their lookup tables. Typical values of n and m would be $n = m = 8$ and $n = m = 4$. In order to study the resilience of a block cipher against differential attacks, we first study the differential properties of its S-boxes. To this end, a key tool is the so-called *Difference Distribution Table* (DDT). The DDT is a table defined for any function $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$, of size $2^n \times 2^m$ and such that

$$\delta_S(a,b) = \#\{x \in \mathbb{F}_2^n, S(x \oplus a) \oplus S(x) = b\}$$

for all $a \neq 0$. The maximal coefficient in the DDT of a function is called its *differential uniformity* and is denoted by δ_S. If the S-boxes of a block cipher have a low differential uniformity, we can expect—and, depending on the properties of its diffusion layer, prove—that it does not have any high probability differential pattern.

Since its inception, many variants of the differential cryptanalysis have been designed. In particular, the *boomerang attack* [25] considers both the encryption and the decryption functions at the same time by looking for pairs (a,b) such that

$$\begin{cases} E_k(x) & = y \\ E_k(x \oplus a) & = y' \end{cases} \text{ and } \begin{cases} E_k^{-1}(y \oplus b) & = z \\ E_k^{-1}(y' \oplus b) & = z \oplus a \end{cases},$$

where E_k is the block cipher under consideration. In order to estimate the probability of such an event in an S-box-based block cipher, we need to use the DDT of its S-box. However, in order to properly take into account the coupling between the different encryptions, we also need to look at another table called the *Boomerang Connectivity Table* (BCT) introduced in [11]. For a permutation $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$, it is a $2^n \times 2^n$ table of integers $\beta_S(a,b)$ defined as

$$\beta_S(a,b) = \#\left\{x \in \mathbb{F}_2^n, S^{-1}\left(S(x) \oplus b\right) \oplus S^{-1}\left(S(x \oplus a) \oplus b\right) = a\right\} .$$

The BCT coefficients are always equal to 2^n (the maximum) when $a = 0$ or $b = 0$. Thus, we define the *boomerang uniformity* of S, denoted β_S, to be the maximum value of $\beta_S(a,b)$ for $a \neq 0$ and $b \neq 0$. As with the differential uniformity, the lower the boomerang uniformity the better. Before going any further, we also recall some basic properties of the BCT which were established in [8, 11].

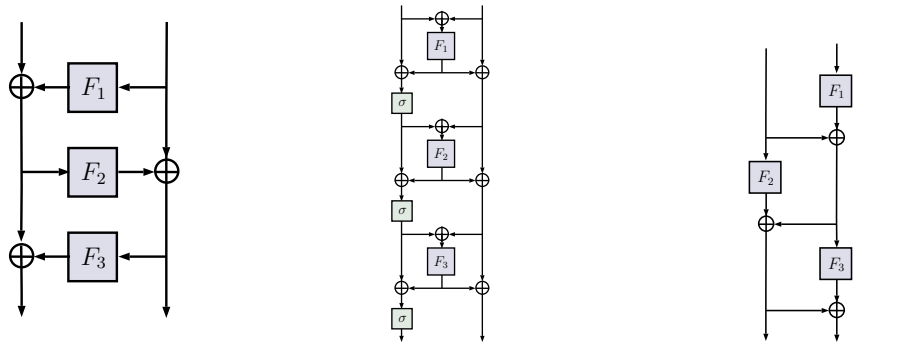
Proposition 1. *Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a permutation. Then S has the same boomerang uniformity as S^{-1} and $B \circ S \circ A$, where A and B are affine permutations, and $\beta_S \geq \delta_S$. Furthermore, if we let $S_b : x \mapsto S^{-1}(S(x) + b)$ be a permutation, then*

$$\beta_S(a,b) = \#\{x \in \mathbb{F}_2^n, S_b(x) \oplus S_b(x \oplus a) = a\} . \tag{1}$$

Finally, it holds that $\beta_{S^{-1}}(a,b) = \beta_S(b,a)$ for all a,b in $\mathbb{F}_2^n$.

Unlike the DDT, little is known about the BCT of even the most common cryptographic components. It is easy for example to prove (see e.g. [11]) that for permutations S providing an optimal resistance to differential cryptanalysis, called *Almost Perfect Nonlinear (APN)* permutations, $\beta_S = \delta_S = 2$. Then, Boura and Canteaut studied in [8] the boomerang properties of the inverse mapping and quadratic power permutations of $\mathbb{F}_{2^n}$. They showed that both families have an optimal boomerang uniformity when $n \equiv 2 \mod 4$, where optimal means that the boomerang uniformity equals the differential uniformity and is 4 in both cases. Since these first results, permutations with boomerang uniformity 4 were further investigated and some new families of permutations with optimal boomerang uniformity were presented [18, 22]. However, besides these first advances, determining the boomerang uniformity of other cryptographically relevant families of permutations or constructions remains an open problem.

In order to ease the implementation of their ciphers on constrained platforms, cryptographers often use specific block-cipher-like structures for their S-boxes. This technique permits the construction of large S-boxes from smaller, much cheaper ones. In this paper, we investigate the BCT of several such lightweight S-box structures, notably the 3-round Feistel, Lai-Massey and (unbalanced) MISTY structures. These three classical constructions are depicted in Figures 1a, 1b and 1c respectively. We then look at two 1-round structures: the 1-round SPN and the specific structure used in the FLY block cipher [17].



(A) 3-round Feistel. **(B)** 3-round Lai-Massey. **(C)** 3-round MISTY.

FIG. 1: The structures investigated in this paper.

For each of those structures, we derive a lower bound on the boomerang uniformity of S-boxes built using it. Table 1 presents the differential and boomerang uniformity of several S-boxes from the literature. It also contains the lower bounds derived in this paper.

Rounds	S-box Struct.	Cipher	Ref.	δ_S	β_S	Lower bound
3	Feistel	Scream	[10, 16]	8	**256**	256
	MISTY-like	Fantomas	[15]	16	**160**	64 (∗)
	Lai-Massey	Fox	[24]	16	**256**	256
1	SPN	Midori	[2]	64	**256**	256
	Lai-Massey-like	FLY	[17]	16	**256**	256 (∗)

TABLE 1: The boomerang and differential uniformity of various 8-bit S-boxes. (∗) The bound depends on the inner components of these constructions.

For building S-boxes for symmetric-key primitives, apart from the block-cipher-like structure, the most popular S-box structure is the finite field monomial. In particular, the AES S-box is based on the multiplicative inverse ($x \mapsto x^{2^n-2}$), but other monomials are also used, such as Gold functions ($x \mapsto x^{2^i+1}$). The latter class has the lowest possible non-linear algebraic degree and this makes them ideal for use cases where such a low degree is a requirement i.e. for building Multi-Party Computation-friendly block ciphers like MiMC [1].

Furthermore, the study of such functions could help us answer the following more general question: Are there 8-bit S-boxes with boomerang uniformity equal to 4? In the absence of APN permutations for $n = 8$, such S-boxes would have the lowest possible boomerang uniformity in this dimension. However, finding such permutations in $\mathbb{F}_2^n$ with $n \equiv 0 \mod 4$ is considered as an open problem. On the other hand, when $n \equiv 2 \mod 4$ the situation is much simpler and it was shown in [8] that both the inverse mapping $x \mapsto x^{2^n-2}$ and the Gold power permutations $x \mapsto x^{2^i+1}$ with $\gcd(i,n) = 2$ over $\mathbb{F}_{2^n}$, have a boomerang uniformity equal to 4. One strategy then for finding permutations S of $\mathbb{F}_2^n$ with $n \equiv 0 \mod 4$ for which $\beta_S = 4$ would be to search inside the extended-affine (EA) classes of the Gold family of functions. Indeed, almost all boomerang 4-uniformed permutations that are known today are, in fact, EA-equivalent to the Gold function (see notably the examples given in [18] and [22]). We study this problem and provide results about the EA-equivalence class of Gold functions for $n \equiv 0 \mod 4$. Then, we analyze the EA-equivalence class of the inverse permutation and show that even if in general the boomerang uniformity is not preserved inside an EA-equivalence class, this is so for the inverse function.

Finally, we investigate a third problem concerning BCTs. Notably, we are interested in the problem of finding all permutations sharing a given BCT and provide an algorithm for doing such a computation. This is a similar problem to the one studied in [9] for DDTs. We are in particular interested in the number of permutations that can share a same BCT and analyse different S-boxes from the literature with respect to this notion.

The rest of the paper is organized all follows. In Section 2 we provide our results about the boomerang uniformity of 3-round Feistel and Lai-Massey constructions. In Sections 3 and 4 we do the same for 3-round MISTY constructions and some non-iterative S-boxes respectively. In Section 5, we formally analyze the significance of the properties discovered in the previous sections. Section 6 is dedicated to the analysis of the boomerang uniformity of functions that are EA-equivalent to Gold and the inverse function. Finally, in Section 7 we present an algorithm for inverting a given BCT.

2 3-round Feistel and Lai-Massey Networks

We study in this section the properties of the BCT of 3-round Feistel and Lai-Massey constructions and show that the boomerang uniformity of both is the worst one possible. In both cases, the results

we derive are an inherent property of the structures used: even if the subcomponents are chosen so as to have excellent properties, the boomerang uniformity will be the worst possible.

2.1 The Feistel Case

A popular way to construct large S-boxes from smaller ones is by using the Feistel construction. This scheme was used for constructing the 8-bit S-boxes of multiple ciphers including ZUC [14], and Scream [16]. All use 3 rounds. This number of rounds is the smallest allowing the overall structure to have good properties. In particular, if only 2 rounds are used, a part of the output depends linearly on a part of the input.

A 3-round Feistel network has the worst possible boomerang uniformity, no matter the choice of the inner functions. Proposition 2 formalizes this statement.

Proposition 2. *Let $S : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n \times \mathbb{F}_2^m$ be a 3-round Feistel network and let F_1, F_2 and F_3 be its inner functions as depicted in Fig. 1a, with $F_1, F_3 : \mathbb{F}_2^m \to \mathbb{F}_2^n$ and $F_2 : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Then, for any $b \in \mathbb{F}_2^n \times \mathbb{F}_2^m$,*

$$\beta_S(b,b) = \beta_S = 2^{n+m}.$$

Proof. For any $b \in \mathbb{F}_2^n \times \mathbb{F}_2^m$ define

$$S_b : \begin{cases} \mathbb{F}_2^n \times \mathbb{F}_2^m & \to \ \mathbb{F}_2^n \times \mathbb{F}_2^m \\ (x,y) & \mapsto \ S^{-1}(S(x,y)+b) \ . \end{cases}$$

Consider now a constant b of the form $b = (b_1, 0) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$. In this case, the function S_b, depicted in the upper part of Fig. 2, reduces to the structure depicted to the lower part of Fig. 2.

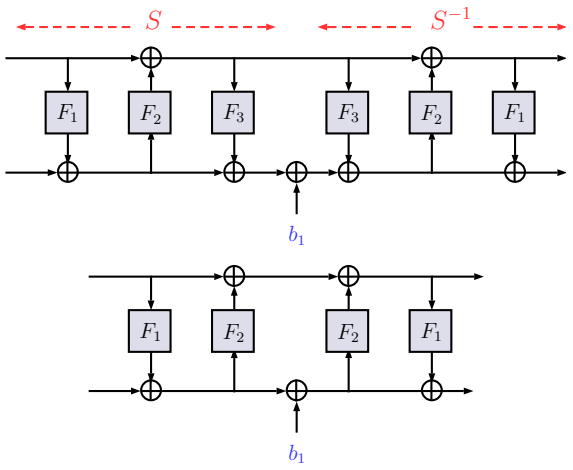


FIG. 2: The function $S_b(x,y) = S^{-1}(S(x,y)+b)$, for $(x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$ and b of the form $(b_1, 0)$, for S a 3-round Feistel structure.

The central observation here is that for any $b_1 \in \mathbb{F}_2^n$

$$S_b(x,y) + S_b((x,y)+(b_1,0)) = (b_1,0), \text{ for all } (x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m. \tag{2}$$

The statement of Eq. (2) can be verified by simply developing S_b and performing the necessary computation:

$$\begin{aligned} S_b(x,y) = (&b_1 + x + F_1(x) + F_1(y + F_2(x + F_1(y)) + F_2(b_1 + x + F_1(y))), \\ &y + F_2(x + F_1(y)) + F_2(b_1 + x + F_1(y))). \end{aligned}$$

By rewriting Eq. (2) as

$$S^{-1}(S(x,y)+b)+S^{-1}(S(x+b,y)+b)=(b_1,0)\ ,$$

we get that $\beta_S(b,b)=2^{n+m}$. This proves that $\beta_S=2^{n+m}$. □

The property corresponding to Proposition 2 was known before. In fact, this property was recently used by Biryukov et al. for performing guess and determine attacks against Feistel networks [4]. We deduce that the knowledge of the BCT can help the cryptanalyst in contexts different from boomerang attacks.

2.2 The Lai-Massey Case

We consider a variant of the structure depicted in Fig. 1b where the last application of the linear orthomorphism σ is removed for two reasons. First, the S-box of the cipher Fox is built in this way. Second, this version yields a simpler proof but, since σ is linear and since the boomerang uniformity is constant in an affine-equivalence class, the result still holds if the last σ is present.

Proposition 3. *Let $S:(\mathbb{F}_2^n)^2\to(\mathbb{F}_2^n)^2$ be a 3-round Lai-Massey structure. Let $F_1,F_2,F_3:\mathbb{F}_2^n\to\mathbb{F}_2^n$ be its inner functions and let σ be a linear permutation. For any nonzero $a\in\mathbb{F}_2^n$, consider $a_1=(a,a)$ and $b=(\sigma(a),\sigma(a))\in(\mathbb{F}_2^n)^2$. Then,*

$$\beta_S(a_1,b)=\beta_S=2^{2n}.$$

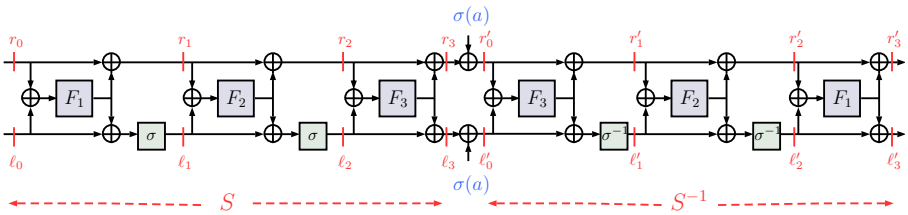


FIG. 3: The function $S_b(x)=S^{-1}(S(x,y)+b)$, where S is a 3-round Lai-Massey construction for which the last application of σ is omitted.

Proof. Let $b\in(\mathbb{F}_2^n)^2$ be of the form $b=(\sigma(a),\sigma(a))\in(\mathbb{F}_2^n)^2$ where a is nonzero. Furthermore, let S_b be as defined in Eq. (1) (see Fig. 3). Our aim is to prove that

$$S_b(x,y)+S_b((x,y)+(a,a))=(a,a),\ \text{for all}\ (x,y)\in(\mathbb{F}_2^n)^2. \tag{3}$$

The remainder of the proof consists in checking Eq. (3) by developing the expression of S_b.

Begin with the input $(\ell_0,r_0)=(x,y)\in\mathbb{F}_2^{2n}$ and denote

$$\begin{aligned}
A_1(x,y)&=A_1=x+y,\ \ A_2(x,y)=A_2=\sigma(x)+y+F_1(A_1)+\sigma(F_1(A_1)),\\
A_3(x,y)&=A_3=\sigma(\sigma(x+F_1(A_1))+F_2(A_2))+y+F_1(A_1)+F_2(A_2),\\
T(x,y)&=T=F_2(A_2)+F_2(A_2+\sigma(a)+a),\\
B_1(x,y)&=B_1=A_1+\sigma^{-1}(a)+\sigma(a)+T+\sigma^{-1}(T).
\end{aligned}$$

The output (ℓ_i,r_i) of the i-th round, $i=1,2,3$ of S, is detailed below:

$$\begin{aligned}
(\ell_1,r_1)&=(\sigma(x+F_1(A_1)),y+F_1(A_1))\,,\\
(\ell_2,r_2)&=(\sigma(\sigma(x+F_1(A_1))+F_2(A_2)),y+F_1(A_1)+F_2(A_2))\,,\\
(\ell_3,r_3)&=(\sigma(\sigma(x+F_1(A_1))+F_2(A_2))+F_3(A_3),y+F_1(A_1)+F_2(A_2)+F_3(A_3))\,.
\end{aligned}$$

Denote by $(\ell'_0, r'_0) = (\ell_3, r_3) + (\sigma(a), \sigma(a))$ the input of S^{-1}. Similarly, we have

$$\begin{aligned}
(\ell'_1, r'_1) &= (a + \sigma(x + F_1(A_1)) + F_2(A_2), \sigma(a) + y + F_1(A_1) + F_2(A_2)), \\
(\ell'_2, r'_2) &= \left(\sigma^{-1}(a) + x + F_1(A_1) + \sigma^{-1}(T), \sigma(a) + y + F_1(A_1) + T\right), \\
(\ell'_3, r'_3) &= ((\sigma^{-1}(a) + x + F_1(A_1) + F_1(B_1) + \sigma^{-1}(T), \\
&\quad \sigma(a) + y + F_1(A_1) + F_1(B_1) + T).
\end{aligned}$$

As $A_1(x+a, y+a) = A_1(x,y)$ and $A_2(x+a, y+a) = A_2(x,y) + \sigma(a) + a$, it holds that $T(x+a, y+a) = T(x,y)$ and $B_1(x+a, y+a) = B_1(x,y)$. Therefore, Eq. (3) holds and we deduce the proposition. □

3 BCTs of 3-round MISTY Networks

The MISTY network mimicks a structure used in the MISTY cipher [21]. While it might ressemble a Feistel network it requires all three inner functions to be bijective in order for the whole function to be a permutation. As we will show, the MISTY structure also differs from the Feistel one via its BCT: the boomerang uniformity of a 3-round MISTY structure *depends* on the specifics of the subfunctions used.

Though the results are similar in both cases, the case where the branches are of the same size (balanced) and of different sizes (unbalanced) require different analyses which we provide in Sections 3.1 and 3.2 respectively.

3.1 3-round Balanced MISTY Networks

Proposition 4. *Let $S : (\mathbb{F}_2^n)^2 \to (\mathbb{F}_2^n)^2$ be a 3-round balanced MISTY network and let F_1, F_2 and F_3 be its inner functions as depicted in Fig. 1c, with $F_1, F_2, F_3 : \mathbb{F}_2^n \to \mathbb{F}_2^n$ bijective. Then,*

$$\beta_S \geq 2^n \beta_{F_2}.$$

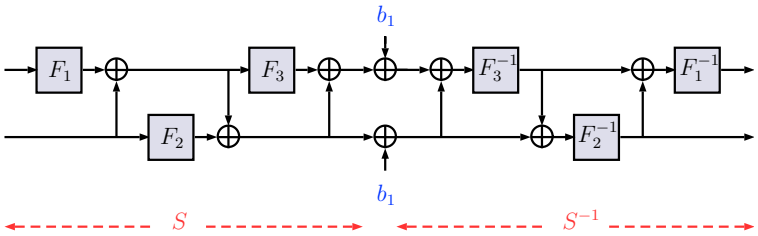


FIG. 4: The function $S_b(x) = S^{-1}(S(x) + b)$, for $x \in \mathbb{F}_2^n \times \mathbb{F}_2^n$ and $b = (b_1, b_1)$ where S is a 3-round balanced MISTY structure.

Proof. The boomerang uniformity of F_2 is β_{F_2} meaning that there exists $(a_1, b_1) \in (\mathbb{F}_2^n)^2$ such that $\beta_{F_2}(a_1, b_1) = \beta_{F_2}$. If $a = (a_1, 0), b = (b_1, b_1)$ then we deduce from Proposition 1 that

$$\beta_S(a,b) = \#\{(x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n \mid S_b(x,y) + S_b(x + a_1, y) = (a_1, 0)\},$$

where S_b is as depicted in Fig. 4. Since $b = (b_1, b_1)$, it can be simplified into

$$S_b(x,y) = \left(F_2^{-1}(F_2(x) + b_1),\ F_1^{-1}(F_2^{-1}(F_2(x) + b_1) + F_1(y) + x)\right)$$

so we deduce

$$\begin{aligned}
S_b(x,y) + S_b(x + a_1, y) = \Big(&F_2^{-1}(F_2(x) + b_1) + F_2^{-1}(F_2(x + a_1) + b_1), \\
&F_1^{-1}\left(F_2^{-1}(F_2(x) + b_1) + F_1(y) + x\right) + \\
&F_1^{-1}\left(F_2^{-1}(F_2(x + a_1) + b_1) + F_1(y) + x + a_1\right)\Big).
\end{aligned} \tag{4}$$

We now show that $\beta_S(a,b) = 2^n|A|$, where

$$A \;=\; \left\{x \in \mathbb{F}_2^n \mid F_2^{-1}(F_2(x)+b_1)+F_2^{-1}(F_2(x+a_1)+b_1) = a_1\right\}\ .$$

For any $x \in A$, the right hand side of $S_b(x,y)+S_b(x+a_1,y)$ can be simplified:

$$\begin{aligned}
&F_1^{-1}\big(F_2^{-1}(F_2(x)+b_1)+F_1(y)+x\big)\\
&\quad+F_1^{-1}\big(F_2^{-1}(F_2(x+a_1)+b_1)+F_1(y)+x+a_1\big)\\
=\;& F_1^{-1}\big(F_2^{-1}(F_2(x)+b_1)+F_1(y)+x\big)+F_1^{-1}\big(F_2^{-1}(F_2(x)+b_1)+F_1(y)+x\big)\\
=\;& 0\ .
\end{aligned}$$

As $|A| = \beta_{F_2}(a_1,b_1) = \beta_{F_2}$, we conclude that the boomerang uniformity of a 3-round balanced MISTY network is lower bounded by $2^n\beta_{F_2}$. □

Remark 1. We give here the minimal bounds for a 3-round MISTY network, for some popular choices of n. All functions F_1, F_2 and F_3 are supposed bijective.

n = 4 As proved in [8], the minimal boomerang uniformity for a permutation of $\mathbb{F}_2^4$ is 6. Therefore, by choosing F_2, with $\beta_{F_2} = 6$, we get that $\beta_S \geq 96$. As we will argue in Section 5.1, this is a very high value for an 8-bit S-box. Note here that the bound of Proposition 4 is tight. Indeed, the 8-bit MISTY network with $F_1 = F_2 = F_3 = [8,0,1,12,15,5,6,7,4,3,10,11,9,13,14,2]$, where $\beta_{F_2} = 6$, has boomerang uniformity 96 and reaches thus this lower bound.

n = 3 APN permutations exist for $n = 3$. Using one as F_2, we obtain $\beta_S \geq 16$.

Proposition 5. *Let $S : (\mathbb{F}_2^n)^2 \to (\mathbb{F}_2^n)^2$ be a 3-round MISTY network with inner functions F_1, F_2 and F_3. If F_1 or F_3 is an affine permutation, then $\beta_S = 2^{2n}$.*

Proof. Suppose that F_1 is affine. In this case, we focus on the BCT value at the point $(a,b) = \big((0,a_2),(b_1,b_1)\big)$, where $a_2, b_1 \in \mathbb{F}_2^n \setminus \{0\}$. Then

$$\beta_S(a,b) = \#\{(x,y) \in (\mathbb{F}_2^n)^2 | S_b(x,y)+S_b(x,y+a_2) = (0,a_2)\}$$

and, again, we have

$$\begin{aligned}
S_b(x,y)+S_b(x,y+a_2) = \Big(&0, F_1^{-1}\big(F_2^{-1}(F_2(x)+b_1)+F_1(y)+x\big)\\
&+F_1^{-1}\big(F_2^{-1}(F_2(x)+b_1)+F_1(y+a_2)+x\big)\Big).
\end{aligned} \tag{5}$$

For $x \in \mathbb{F}_2^n$, denote $f(x) = F_2^{-1}(F_2(x)+b_1)+x$. Then, $\beta_S(a,b) = \sum_{x\in\mathbb{F}_2^n} \beta_{F_1}(a_2,f(x))$. Suppose $f(x) = c$, then $F_2(x+c)+F_2(x) = b_1$. This means that there are $\delta_{F_2}(c,b_1)$ different x corresponding to the same $f(x)$. We deduce that

$$\sum_{x\in\mathbb{F}_2^n} \beta_{F_1}(a_2,f(x)) = \sum_{c\in\mathbb{F}_2^n} \beta_{F_1}(a_2,c)\delta_{F_2}(c,b_1) = 2^n \sum_{c\in\mathbb{F}_2^n} \delta_{F_2}(c,b_1) = 2^{2n}$$

since when F_1 is affine, we always have $\beta_{F_1}(a_2,c) = 2^n$. Therefore, $\beta_S(a,b) = 2^{2n}$ which is the worst possible case.

The case where F_3 is affine reduces to where F_1 is. The reduction uses that the boomerang uniformity is preserved by both functional inversion and affine equivalence (Proposition 1). Suppose now that F_3 is an affine permutation. This is obviously also the case for F_3^{-1}. Therefore, the 3-round MISTY network with inner functions $(F_3^{-1}, F_2^{-1}, F_1^{-1})$ has a boomerang uniformity equal to 2^{2n} because of the previous proposition. However, this network is a simple affine equivalent of the inverse of the MISTY network with inner functions (F_1,F_2,F_3), as this can be seen from Fig. 4. The result follows from the fact that the boomerang uniformity is preserved under affine equivalence. □

3.2 3-round Unbalanced MISTY Networks

In practice, all known MISTY constructions are unbalanced. More specifically, they use branches of $n-1$ and $n+1$ bits. This is the case for the original MISTY1 cipher [21], where some inner components follow a 3-round construction with one 7-bit and one 9-bit branch. It is also the case of the 8-bit S-box of Fantomas [15], where a 3-bit and a 5-bit branch are used. One reason for this is that an unbalanced 3-round MISTY structure can potentially achieve a better differential uniformity than a balanced one. For example, Canteaut et al. showed in [10] that for $n = 4$, a $2n$-bit permutation S following a balanced MISTY structure has $\delta_S \geq 16$, while by taking a 3-bit and a 5-bit branch it is possible to find a permutation S with $\delta_S = 8$. Another reason is that taking branches of an odd length of bits permits to use APN permutations for the inner components—as was indeed done in MISTY1.

We studied the boomerang uniformity of 3-round unbalanced MISTY networks. Our arguments differ slightly depending on whether the widest branch is on the left or on the right. Let m be the size of the smallest branch and n the size of the biggest one. If S_1 is the construction on the left part of Fig. 5 and S_2 is the construction on the right part of the same figure then we get the following proposition:

Proposition 6. *The boomerang uniformities of the unbalanced MISTY structures in Fig. 5 are bounded as follows:*

$$\beta_{S_1} \geq 2^n \beta_{F_2} \quad \text{and} \quad \beta_{S_2} \geq 2^m \beta_{F_2|_{\mathbb{F}_2^m}} .$$

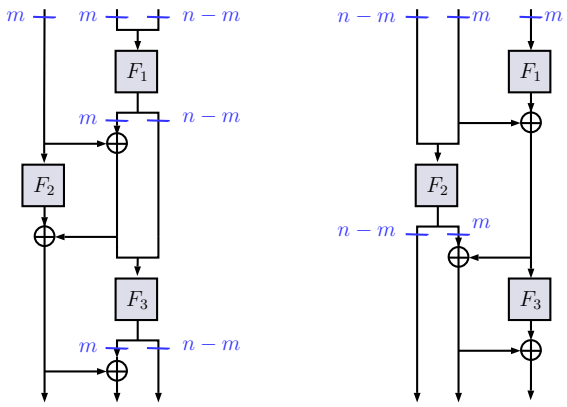


FIG. 5: The two different types of unbalanced 3-round MISTY networks (S_1 is on the left, S_2 is on the right).

The proof of the above proposition is provided in Appendix A.

4 Non-Iterative Constructions

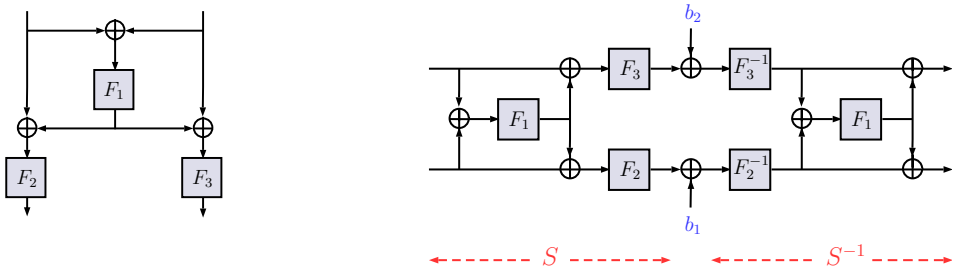
Until now, we have only considered 3-round constructions. However, S-box designers sometimes choose non-iterative structures. In this section, we look at 1-round SPN (as used e.g. in Midori [2]) and at the *ad hoc* Lai-Massey-like structure used by the *Littlun* S-Box of the block cipher FLY [17]. It is composed of a single Lai-Massey round followed by an S-box layer (see Fig. 6a). While Littlun is such that $F_1 = F_2 = F_3$, we do not make this assumption.

The following straightforward proposition deals with the properties of a 1-round SPN.

Proposition 7. *Let F_1, F_2 be n-bit permutations and let $S : (\mathbb{F}_2^n)^2 \to (\mathbb{F}_2^n)^2$ be such that $S(x_1, x_2) = (F_1(x_1), F_2(x_2))$. Then we have*

$$\beta_S((a_1,a_2),(b_1,b_2)) \;=\; \beta_{F_1}(a_1,b_1) \times \beta_{F_2}(a_2,b_2) ,$$

so that in particular $\beta_S = \beta_S((a,0),(0,b)) = 2^{2n}$. □



(A) Littlun S-box.

(B) The function $S_b(x) = S^{-1}(S(x,y)+b)$, where S is the Littlun S-box.

FIG. 6: Analysis of the Littlun structure

Let us now consider the Littlun construction. If F_1 is an affine permutation, then the corresponding Littlun-like S-box is a 1-round SPN structure which has thus the worst possible boomerang uniformity. If not, the following proposition relates its boomerang uniformity to that of its subcomponents.

Proposition 8. *Let S be a generalised-Littlun structure with n-bit permutations F_1, F_2 and F_3 (see Fig. 6a). Then $\beta_S \geq 2^n \max\{\beta_{F_2}, \beta_{F_3}\}$.*

Proof. We define the two following functions H and G over $(\mathbb{F}_2^n)^2$:

$$H(x,y) = (x+F_1(x+y), y+F_1(x+y)) \quad \text{and} \quad G(x,y) = (F_2(x), F_3(y)) \ .$$

Both H and G are permutations of $(\mathbb{F}_2^n)^2$, H is an involution, and $S(x,y) = (G \circ H)(x,y)$. we claim that $\beta_S(a,b) \geq \beta_G(a,b)$, for $a = (a_1,a_1)$, $a_1 \neq 0$ and $b = (b_1,b_2)$. Indeed,

$$\begin{aligned}
\beta_S(a,b) &= \#\left\{X \in (\mathbb{F}_2^n)^2 \mid H(G^{-1}(G(H(X)+a)+b)) + H(G^{-1}(G(H(X)+b))) = a\right\} \\
&= \#\left\{Y \in (\mathbb{F}_2^n)^2 \mid H(G^{-1}(G(Y+a)+b)) + H(G^{-1}(G(Y)+b)) = a\right\} \\
&= \#\left\{Y \in (\mathbb{F}_2^n)^2 \mid H(G^{-1}(G(Y)+b) + G^{-1}(G(Y)+b)) \right. \\
&\qquad \left. + G^{-1}(G(Y+a)+b)) + H(G^{-1}(G(Y)+b)) = a\right\} \\
&\geq \#\left\{Y \in (\mathbb{F}_2^n)^2 \mid G^{-1}(G(Y)+b) + G^{-1}(G(Y+a)+b) = a\right\} = \beta_G(a,b) \ .
\end{aligned}$$

By applying Proposition 7 to $\beta_G(a,b)$ with $a = (a_1,a_1), b = (b_1,b_2)$, we obtain that $\beta_S(a,b) = 2^n \beta_{F_2}(a_1,b_1)$ when $b_2 = 0$ and $\beta_S(a,b) = 2^n \beta_{F_2}(a_1,b_2)$ when $b_1 = 0$. We deduce that $\beta_S \geq 2^n \max(\beta_{F_2}, \beta_{F_3})$. □

5 Using the Maximum BCT Coefficients

The S-boxes we study in this paper are highly structured, it is then not surprising that we find specific artefacts in their BCT. In this section, we explore the significance of these properties. First, we recall the expected behaviour of the BCT of a random permutation (Section 5.1) and deduce the expected value of the boomerang uniformity of a random n-permutation. We then identify a specific pattern that exists whenever the boomerang uniformity is maximum and compare it with the patterns that we have identified in Feistel and Lai-Massey networks (Section 5.2).

5.1 A High Boomerang Uniformity is Unlikely

We first recall the following result.

Proposition 9 ([7]). *If S is a permutation of $\mathbb{F}_2^n$ picked uniformly at random, then its BCT coefficients $\beta_S(a,b)$ with $a,b \neq 0$ can be modelled like independent and identically distributed random variables with the following distribution:*

$$\Pr[\beta_S(a,b) = c] \ = \sum_{2i_1+4i_2=c} P_1(i_1) P_2(i_2) \ ,$$

where P_1 and P_2 are stochastic variables following binomial distributions:

$$P_1(i) = \text{Binomial}\left(i; 2^{n-1}, \frac{1}{2^n-1}\right) ,$$
$$P_2(i) = \text{Binomial}\left(i; 2^{2n-2}-2^{n-1}, \frac{1}{(2^n-1)^2}\right) .$$

Applying the proposition above, we can derive a formula to compute the expected boomerang uniformity for an n-bit random permutation, namely:

$$E(\beta_s) \;=\; \sum_{c=0}^{2^n} c \left(\left(\sum_{i=0}^{c} \Pr[\beta_S(a,b)=i] \right)^t - \left(\sum_{i=0}^{c-2} \Pr[\beta_S(a,b)=i] \right)^t \right) ,$$

where $t = (2^n-1)^2$. We used this formula to build a table containing the expected boomerang uniformity for random permutations from $n = 4$ to 14 (see Table 2).

n	4	5	6	7	8	9	10	11	12	13	14
$E(\beta_s)$	11.6	14.2	16.3	18.3	20.2	22.1	23.9	25.7	27.4	29.1	30.8
$E(\beta_s)/2^n$	0.7250	0.4437	0.2547	0.1430	0.0789	0.0432	0.0233	0.0125	0.0067	0.0036	0.0019

TABLE 2: The expected boomerang uniformity for n-bit random permutations.

As we can see, the expected boomerang uniformity increases slowly with n and it is significantly smaller than 2^n. This shows that a maximal boomerang uniformity is an extremely rare event indicative of a very strong structure.

Furthermore, the non-maximal but still very high boomerang uniformity of the 3-round MISTY and Littlun structures can also be leveraged to identify such potentially hidden structures. Indeed, for $n = 8$, it holds that the boomerang uniformity of a balanced MISTY structure is at least 96 (see Sec. 3.1), which is much higher than the expected 20.2.

5.2 Patterns in the Maximum BCT Coefficients

The coordinates of the BCT coefficients equal to the maximum possible value (namely 2^n) always have particular structures captured by the following proposition.

Proposition 10. *Let S be a permutation of $\mathbb{F}_2^n$. For every $x \in \mathbb{F}_2^n$, the following sets are vector spaces:*

$$\left\{y \in \mathbb{F}_2^n, \beta_S(y,x) = 2^n\right\} \;\; \textit{and} \;\; \left\{y \in \mathbb{F}_2^n, \beta_S(x,y) = 2^n\right\} .$$

Proof. Suppose that $\beta_S(a,b) = \beta_S(a',b) = 2^n$. This is equivalent to saying that, for all $x \in \mathbb{F}_2^n$, the following two equations hold: $S_b(x \oplus a) \oplus S_b(x) = a$ and $S_b(x \oplus a') \oplus S_b(x) = a'$. Summing them and replacing x with $y = x \oplus a$, we get that

$$S_b(y) \oplus S_b(y \oplus a \oplus a') = a \oplus a'$$

for all $y \in \mathbb{F}_2^n$, which is equivalent to $\beta_S(a \oplus a', b) = 2^n$. We deduce that the first set in the proposition is a vector space.

Since the BCT of S^{-1} is the transpose of that of S, applying what we just proved to the BCT of S^{-1} yields that if $\beta_S(a,b) = \beta_S(a,b') = 2^n$, then $\beta_S(a, b \oplus b') = 2^n$. As a consequence, the second set in the proposition is also a vector space. □

These properties are reminiscent of those of *linear structures*. We first recall their definition.

Definition 1 (Linear Structure). *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be a Boolean function. A* linear structure *of f is an element $a \in \mathbb{F}_2^n$ such that $f(x \oplus a) \oplus f(x) = \varepsilon$, where $\varepsilon \in \mathbb{F}_2$ is a constant. All such elements form a vector space called the* linear space of *f which is denoted $\mathsf{LS}(f)$.*

The linear structures of a Boolean function always form a vector space, much like how the input differences yielding a maximum BCT coefficient for a given output difference yield a vector space (Proposition 10). This likeness is not a coincidence. Indeed, for a fixed output difference b, the maximum BCT coefficients correspond to a such that $S_b(x \oplus a) \oplus S_b(x) = a$ for all $x \in \mathbb{F}_2^n$, i.e. essentially to linear structures that are shared by all the coordinates of S_b.

As we have established, 3-round Feistel and Lai-Massey networks have maximum boomerang uniformity. Furthermore, we have showed that the coordinates of the maximum coefficients have a specific structure that is even stronger than the one implied by Proposition 10: in both cases, these coordinates form a vector space of a specific dimension. The coordinates are both involved at the same time in the construction of the vector space, unlike in Proposition 10.

- For the 3-round Feistel network operating on $\mathbb{F}_2^{n+m}$, the maximum coefficients have coordinates (b,b) for $b \in \mathbb{F}_2^{n+m}$. It is a vector space of dimension $n+m$ corresponding to the span of all vectors of the form (e_i, e_i) where the vectors e_i form the canonical basis of $\mathbb{F}_2^{n+m}$.
- For the 3-round Lai-Massey structure operating on $\mathbb{F}_2^{2n}$, the maximum coefficients are at positions $((a,a),((\sigma(a),\sigma(a)))$ for all $a \in \mathbb{F}_2^n$. Since σ is a linear permutation of $\mathbb{F}_2^n$, we deduce that these coordinates form a vector space of dimension n.

In [7], the authors introduced a bases extraction algorithm returning a basis for each of the vector spaces of a given dimension contained in a specific set. We can then use this algorithm to figure out if a permutation is affine-equivalent to a 3-round Feistel or Lai-Massey network. Indeed, let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a permutation. If running this algorithm on the set of the coordinates of all the coefficients of β_S equal to the maximum returns a space of dimension n, then it is a strong indication that S is affine-equivalent to a 3-round Feistel network. If it instead finds a space of dimension $n/2$ then it is probably affine-equivalent to a 3-round Lai-Massey network.

6 Boomerang Uniformity for some Specific EA-Equivalence Classes

In this section, we investigate the boomerang properties of n-bit monomial permutations, namely those with exponents $2^n - 2$ (multiplicative inverse) and $2^i + 1$ (Gold functions). As mentioned before, these S-boxes are common in symmetric cryptography because of their excellent differential properties. Due to the simplicity of their mathematical structure, they are also interesting targets for a more general study of boomerang properties and in particular for the search of permutations with optimal boomerang uniformity.

Before going further, we recall some notions of finite fields arithmetics. We denote with $\mathbb{F}_q$ the finite field with q elements. If $q = p^n$ and if $n = k \times m$ for some integers k and m, then we define the *trace* from $\mathbb{F}_{p^n}$ to $\mathbb{F}_{p^k}$ as

$$\mathrm{Tr}_k^n(x) = \sum_{i=0}^{m-1} x^{p^{i \times k}} .$$

If $k = 1$ then we simply write Tr. The vector space $(\mathbb{F}_p)^n$ is isomorphic to the finite field $\mathbb{F}_{p^n}$. Thus, our results over $\mathbb{F}_2^n$ can easily be re-written over $\mathbb{F}_{2^n}$—and vice-versa.

The inverse and the Gold functions have been investigated in [8]. In fact, almost all permutations with optimal boomerang uniformity known today are *extended-affine equivalent* to the Gold function. We recall the definition of this equivalence below.

Definition 2 (EA-equivalence). *Two functions $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $G : \mathbb{F}_2^n \to \mathbb{F}_2^m$ are extended-affine (EA)-equivalent if there exist two affine permutations $A : \mathbb{F}_2^n \to \mathbb{F}_2^n$, $B : \mathbb{F}_2^m \to \mathbb{F}_2^m$ and an affine function $C : \mathbb{F}_2^n \to \mathbb{F}_2^m$ such that, for all $x \in \mathbb{F}_2^n$,*

$$F(x) = (B \circ G \circ A)(x) + C(x) .$$

This investigation has two aspects. First, we need to find permutations in the EA-equivalence class of the functions studied. Second, we need to verify that the boomerang uniformity is preserved under EA-equivalence.

6.1 EA-Equivalence for Gold Functions

We investigate first the EA-equivalence class of some Gold functions. As we are only interested in permutations in our analysis, we focus on the problem of finding all linear polynomials L such that $F(x) = x^{2^i+1} + L(x)$ is a permutation over $\mathbb{F}_{2^n}$. Note that this problem was investigated by Yongqiang Li et al. in [19] for the case $\gcd(i,n) = 1$. In the following, we focus on the case $\gcd(i,n) = k > 1$.

Proposition 11. *Suppose* $\gcd(i,n) = k > 1$ *with* m *even and* $n = k \times m$*. Then, there are no permutations EA-equivalent (expect for affine equivalent) to* x^{2^i+1} *over* $\mathbb{F}_{2^n}$*.*

Proof. First, notice that

$$\gcd(2^{2i-1}, 2^n - 1) = 2^{2k} - 1 > 2^k - 1 = \gcd(2^i - 1, 2^n - 1)$$

when m is even. Let S_1, S_2 be the image sets of x^{2^i-1} and $x^{2^{2i}-1}$ respectively, then

$$|S_1| = |\{x^{2^i-1} | x \in \mathbb{F}_{2^n}\}| > |\{x^{2^{2i}-1} | x \in \mathbb{F}_{2^n}\}| = |S_2|$$

implying that $S_1 \backslash S_2$ is not empty. By choosing $b \in \mathbb{F}_{2^n}^*$ we have that $(b^{-1})^{2^i-1} \notin S_2$. Then, $b^{2^i} x^{2^{2i}} + bx$ is a linear permutation over $\mathbb{F}_{2^n}$.

Assume that there exists a linear polynomial L such that $F(x) = x^{2^i+1} + L(x)$ is a permutation. Then, for any $b' \in \mathbb{F}_{2^n}$, the Boolean function $\mathrm{Tr}(b'F(x))$ is balanced. However, $g(x) = \mathrm{Tr}(bF(x))$ is a bent function. Indeed, for any $a \in \mathbb{F}_{2^n}$, we have

$$\begin{aligned}\left(\sum_{x\in\mathbb{F}_{2^n}} (-1)^{g(x)+\mathrm{Tr}(ax)}\right)^2 &= \sum_{x,u\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(b(x^{2^i+1}+(x+u)^{2^i+1})+au+L(u))} \\ &= \sum_{u\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}(bu^{2^i+1}+au+L(u))} \sum_{x\in\mathbb{F}_{2^n}} (-1)^{\mathrm{Tr}\left((bu^{2^i}+(bu)^{2^{n-i}})x\right)} \\ &= 2^n.\end{aligned}$$

This indicates that $g(x)$ is never balanced, which contradicts the assumption. Therefore, in this case, there are no permutations EA-equivalent (expect for affine equivalent) to Gold. □

Remark 2. Gold permutations (i.e., $k = 2$) have an optimal boomerang uniformity over $\mathbb{F}_{2^{2m}}$ when m is odd. Unfortunately, these functions are not permutations anymore when m is even. In order to find permutations with a good boomerang uniformity for n even, an idea would be to search for permutations in the EA-equivalence class of a Gold function and hope that it has good boomerang properties. Unfortunately, Proposition 11 shows that there exist no permutations EA-equivalent to Gold over $\mathbb{F}_{2^{2m}}$ when m is even.

Next we consider the case of $n/k = m$ odd. We start by recalling a useful lemma from [6].

Lemma 1 ([6]). *Suppose* F *is finite and* $\lambda \in F^*$*. There is* $y \in F$ *such that* $y^q - y = \lambda$ *if and only if* $\mathrm{Tr}_{F\backslash\mathbb{F}_Q}(\lambda) = 0$*, where* q *is a power of prime* p *and* $\mathbb{F}_Q = F \cap \mathbb{F}_Q(q)$*.*

If we set $F = \mathbb{F}_{2^n}$ and $q = 2^i$ with $\gcd(i,n) = k$, then $\mathbb{F}_Q = \mathbb{F}_Q(2^k)$. The following result follows.

Corollary 1. *Suppose* $\gcd(i,n) = k$ *and* $\lambda \in \mathbb{F}_{2^n}^*$*. Then, there exists* $y \in \mathbb{F}_{2^n}$ *such that* $y^{2^i} + y = \lambda$ *if and only if* $\mathrm{Tr}_k^n(\lambda) = 0$*.*

Proposition 12. *Suppose* $\gcd(i,n) = k > 1$ *with* $m = n/k$ *odd and* L *a linear polynomial. Then* $F(x) = x^{2^i+1} + L(x)$ *is a permutation over* $\mathbb{F}_{2^n}$ *if and only if, for all* $u \in \mathbb{F}_{2^n}^*$,

$$\mathrm{Tr}_k^n\left(\frac{L(u)}{u^{2^i+1}}\right) \neq 1 .$$

Proof. Suppose L is a linear polynomial. Then $F(x) = x^{2^i+1} + L(x)$ is a permutation if and only if, for any $u \in \mathbb{F}_{2^n}^*$, the following equation

$$x^{2^i+1} + (x+u)^{2^i+1} = L(u)$$

which is equivalent to

$$\frac{x}{u} + \left(\frac{x}{u}\right)^{2^i} = \frac{L(u)}{u^{2^i+1}} + 1$$

has no solutions in $\mathbb{F}_{2^n}$. Thus, F is a permutation if and only if

$$\frac{L(u)}{u^{2^i+1}} + 1 \notin \{x + x^{2^i} \,|\, x \in \mathbb{F}_{2^n}\}.$$

Here, we claim that

$$\{x + x^{2^i} \,|\, x \in \mathbb{F}_{2^n}\} = \{\mathrm{Tr}_k^n(x) = 0 \,|\, x \in \mathbb{F}_{2^n}\}.$$

Indeed, on one hand, if $b = x_0 + x_0^{2^i}$, it is obvious that $\mathrm{Tr}_k^n(b) = \mathrm{Tr}_k^n(x_0) + \mathrm{Tr}_k^n(x_0^{2^i}) = 0$. On the other hand, for any $b \in \mathbb{F}_{2^n}^*$ with $\mathrm{Tr}_k^n(b) = 0$, the equation $x^{2^i} + x = b$ always has solutions according to Corollary 1 and for the case of $b = 0$, $x = 0$ is the corresponding solution. Note that m is odd, then $\mathrm{Tr}_k^n(1) = 1$ and therefore F is a permutation if and only if $\mathrm{Tr}_k^n\left(\frac{L(u)}{u^{2^i+1}} + 1\right) \neq 0$, i.e., $\mathrm{Tr}_k^n\left(\frac{L(u)}{u^{2^i+1}}\right) \neq 1$ for any $u \in \mathbb{F}_{2^n}^*$. □

Remark 3. Note that it is easy to find a linear function L satisfying the conditions of Proposition 12. Indeed, the most trivial form for L is $L(x) = ax^{2^i} + a^{2^i}x$. However, characterizing all linear functions L that satisfy this condition is hard. By doing experiments in $\mathbb{F}_{2^6}$ we were able to find, besides the trivial case, other examples. For instance, when $i = 2$, $L(x) = gx^8 + g^{62}x^2$ and $L(x) = g^{36}x^{32} + g^{18}x^8$, where g is a primitive element, are another examples of admissible functions.

6.2 Boomerang Uniformity under EA-Equivalence of the Gold Functions

Proposition 12 shows that when $\gcd(i,n) = k > 1$ and $n/k = m$ is odd it is possible to find permutations inside the EA-equivalence class of the Gold function x^{2^i+1}. By doing experiments, we generated many such permutations and observed that all of them have a boomerang uniformity equal to 4. This seems not to be a coincidence. Although in general the boomerang uniformity is not preserved under EA-equivalence, the above observation shows that the boomerang uniformity might be EA-invariant for some families of functions. Indeed, as we will prove in Theorem 2, this is so for the Gold family of functions.

We begin with a useful theorem from [18], which we will frequently mention in the proofs. This theorem shows an alternative way for computing the boomerang uniformity of a function.

Theorem 1 ([18]). *Let* F *be a permutation over* $\mathbb{F}_2^n$. *Then the boomerang uniformity of* F *is the maximum number of solutions of the following equation system*

$$\begin{cases} F(x+b) + F(y+b) = a, \\ F(x) + F(y) = a \end{cases} \tag{6}$$

for any $a, b \in \mathbb{F}_2^n$.

To determine the boomerang uniformity of members of the EA-equivalence class of a permutation F, we need to compute β_F and β_G, where $G(x) = F(x) + L(x)$ for L a linear function. According to Theorem 1 this is equivalent to finding the number of solutions of the following systems of equations:

$$\begin{cases} F(x)+F(y) = a \\ F(x+b)+F(x)+F(y)+F(y+b) = 0 \end{cases} \tag{7}$$

and

$$\begin{cases} F(x)+F(y)+L(x+y) = a \\ F(x+b)+F(x)+F(y)+F(y+b) = 0 \end{cases} \tag{8}$$

Note that Systems (7) and (8) share the same equation:

$$F(x+b)+F(x)+F(y)+F(y+b) = 0. \tag{9}$$

Therefore, for solving these systems, one can first determine the number of solutions to Eq. (9) and then consider the difference between Systems (7) and (8).

Theorem 2. *Let n be odd and $\gcd(i,2n) = 2$. Then, all permutations EA-equivalent to x^{2^i+1} over $\mathbb{F}_{2^{2n}}$ have boomerang uniformity 4.*

Proof. Applying the method discussed above to compute $\beta_F(a,b)$, we plug $F(x) = x^{2^i+1}$ into Eq. (9) and get

$$b(x+y)^{2^i} + b^{2^i}(x+y) = 0\,.$$

That is, $(x+y)^{2^i-1} = b^{2^i-1}$. Then, $y \in \{x+w, x+w^2, x+1\}$, where w is the primitive element of $\mathbb{F}_{2^{2n}} \bigcap \mathbb{F}_{2^i} = \mathbb{F}_{2^2}$. To make this clear, we denote by $y = x+\alpha$, $\alpha \in \mathbb{F}_{2^2}^*$ and plug it into the first equation of System (8). We have,

$$x^{2^i+1} + (x+\alpha)^{2^i+1} + L(\alpha) + a = 0$$

which can be simplified as

$$\left(\frac{x}{\alpha}\right)^{2^i} + \frac{x}{\alpha} + 1 + \frac{a+L(\alpha)}{\alpha^{2^i+1}} = 0. \tag{10}$$

From Proposition 12, we know that $\mathrm{Tr}_2^{2n}\left(\frac{L(\alpha)}{\alpha^{2^i+1}}\right) \neq 1$. Meanwhile, according to Corollary 1, Eq. (10) has solutions if and only if $\mathrm{Tr}_2^{2n}\left(1+\frac{a+L(\alpha)}{\alpha^{2^i+1}}\right) = 0$. Now, we focus on the number of $\alpha \in \mathbb{F}_{2^2}^*$ such that Eq. (10) has solutions. Note that, $\alpha^{2^i} = \alpha$ and α makes Eq. (10) have solutions means that

$$\mathrm{Tr}_2^{2n}\left(1+\frac{a+L(\alpha)}{\alpha^{2^i+1}}\right) = 1 + \frac{\mathrm{Tr}_2^{2n}(a+L(\alpha))}{\alpha^2} = 0$$

due to n being odd. We thus need $\mathrm{Tr}_2^{2n}(a+L(\alpha)) = \alpha^2$. We claim that there is at most one $\alpha \in \mathbb{F}_{2^2}^*$ that makes the above equation hold. Indeed, we consider the solutions over $\mathbb{F}_{2^2}$ of the equation

$$x^2 + \mathrm{Tr}_2^{2n}(L(x)) + \mathrm{Tr}_2^{2n}(a) = 0, \tag{11}$$

which is affine. Therefore, whether it has no solutions or it has the same number of solutions over $\mathbb{F}_{2^2}$ as

$$x^2 + \mathrm{Tr}_2^{2n}(L(x)) = 0. \tag{12}$$

Since $\mathrm{Tr}_2^{2n}\left(\frac{L(x)}{x^{2^i+1}}\right) = \frac{\mathrm{Tr}_2^{2n}(L(x))}{x^2} \neq 1$ (Proposition 12), we deduce that Eq. (12) has a unique solution $x_0 = 0$ and therefore Eq. (11) has at most one solution over $\mathbb{F}_{2^2}$. Consequently, there is at most one a such that Eq. (10) holds and thus System (8) has at most 4 solutions. □

Remark 4. Theorem 2 offers us an easy way to understand why the permutation proved in [18] and the 960 examples given in Table 1 of [22] have boomerang uniformity 4. To find new boomerang 4-uniform permutations, we first need to verify whether they are in the EA-equivalence class of Gold.

6.3 Boomerang Uniformity under EA-Equivalence of the Inverse

Now, we investigate the boomerang uniformity for another optimal class, namely for the inverse permutation.

Theorem 3. *All permutations EA-equivalent to x^{2^n-2} over $\mathbb{F}_{2^{2n}}$ have boomerang uniformity at most 6.*

Proof. Similarly to the proof of Theorem 2, we plug $F(x) = x^{2^n-2}$ into Eq. (9) and get

$$(x+1)^{2^n-2} + (y+1)^{2^n-2} + x^{2^n-2} + y^{2^n-2} = 0. \tag{13}$$

We investigate different cases.

Case 1: $x, y \notin \{0,1\}$ and $y = x+1$. Consider System (8), we have

$$x^{2^n-2} + (x+1)^{2^n-2} + L(1) + a = 0,$$

i.e.,

$$(L(1)+a)(x^2+x) + 1 = 0.$$

If $L(1)+a = 0$, it has no solutions. Otherwise, it has no solution when $\mathrm{Tr}(\frac{1}{L(1)+a}) = 1$ and has 2 solutions when $\mathrm{Tr}(\frac{1}{L(1)+a}) = 0$. In general, we have no solutions if

$$a = L(1) \text{ or } \begin{cases} a \neq L(1) \\ \mathrm{Tr}\left(\frac{1}{L(1)+a}\right) = 1 \end{cases}$$

and we have 2 solutions otherwise, i.e. if

$$a \neq L(1) \text{ and } \mathrm{Tr}\left(\frac{1}{L(1)+a}\right) = 0\,,$$

in which case (x_0, x_0+1) and (x_0+1, x_0) are the solutions.

Case 2: $x = 0$. It implies $y \neq 0$ since $a \neq 0$ in System (8). Eq. (13) means $y^{2^n-2} + (y+1)^{2^n-2} + 1 = 0$. We then have $y \in \{1, w, w^2\}$, where w is a primitive element of $\mathbb{F}_{2^2}$. We plug it back into System (8) and obtain

$$\begin{aligned} a = L(1)+1 &\implies (0,1) \text{ is a solution}\,, \\ a = L(w)+w^2 &\implies (0,w) \text{ is a solution}\,, \\ a = L(w^2)+w &\implies (0,w^2) \text{ is a solution}\,, \\ a \notin \{L(1)+1, L(w)+w^2, L(w^2)+w &\implies \text{there is no solution}\,. \end{aligned}$$

Case 3: $x = 1$. As above, we deduce $y \neq 1$, because Eq. (13) implies $y^{2^n-2} + (y+1)^{2^n-2} + 1 = 0$. Then $y = \{0, w, w^2\}$, where w is a primitive element of $\mathbb{F}_{2^2}$. We deduce that the situation is similar to the one in Case 2:

$$\begin{aligned} a = L(1)+1 &\implies (1,0) \text{ is a solution}\,, \\ a = L(w)+w^2 &\implies (1,w) \text{ is a solution}\,, \\ a = L(w^2)+w &\implies (1,w^2) \text{ is a solution}\,, \\ a \notin \{L(1)+1, L(w)+w^2, L(w^2)+w &\implies \text{there is no solution}\,. \end{aligned}$$

The last two cases are similar to the previous ones, thus we omit their details and just display the results.

Case 4: $y = 0$, which implies $x \neq 0$. Using the same method, we obtain that

$$\begin{aligned} a = L(1)+1 &\implies (1,0) \text{ is a solution ,} \\ a = L(w)+w^2 &\implies (w,0) \text{ is a solution ,} \\ a = L(w^2)+w &\implies (w^2,0) \text{ is a solution ,} \\ a \notin \{L(1)+1, L(w)+w^2, L(w^2)+w &\implies \text{there is no solution .} \end{aligned}$$

Case 5: $y = 1$, in which case $x \neq 1$. We deduce:

$$\begin{aligned} a = L(1)+1 &\implies (0,1) \text{ is a solution ,} \\ a = L(w)+w^2 &\implies (w,1) \text{ is a solution ,} \\ a = L(w^2)+w &\implies (w^2,1) \text{ is a solution ,} \\ a \notin \{L(1)+1, L(w)+w^2, L(w^2)+w &\implies \text{there is no solution .} \end{aligned}$$

Let us use the knowledge of these cases to deduce the boomerang uniformity of the permutation $x \mapsto x^{2^n-2} + L(x)$. First, note that $L(1)+1, L(w^2)+w$ and $L(w)+w^2$ are different elements, since $L(x)+x^{2^n-2}$ permutes $\mathbb{F}_{2^{2n}}$. Using all the observations above, we consider the three cases that depend on the value of a.

- If $a = L(1)+1$ then $a \neq L(1)$ and $\mathrm{Tr}(\frac{1}{L(1)+a}) = \mathrm{Tr}(1) = 0$. Then System (8) has 4 solutions: $\{(0,1),(1,0),(x_0,x_0+1),(x_0+1,x_0)\}$.
- If $a = L(w^2)+w$ (respectively $a = L(w)+w^2$) then System (8) has at most 6 solutions, namely: $\{(1,w),(w,1),(0,w^2),(w^2,0),(x_0,x_0+1),(x_0+1,x_0)\}$ (respectively $\{(1,w^2),(w^2,1),(0,w),(w,0),(x_0,x_0+1),(x_0+1,x_0)\}$).

Therefore, for all $a \in \mathbb{F}^*_{2^{2n}}$, System (8) has at most 6 solutions. We deduce the theorem. □

Remark 5. From the above theorem, we may come up with the idea of finding optimal permutations among the EA-equivalence class of the inverse permutation. However, [20] shows that there are no EA-equivalent permutations (except for the affine equivalent) to the inverse over $\mathbb{F}_{2^m}$ when $m \geq 5$. Furthermore, when $m = 4$, the only class of $L(x)$ that makes $x^{2^n-2} + L(x)$ a permutation is $L(x) = \alpha x^2 + (\alpha x)^8$, where $\alpha \in \mathbb{F}^*_{2^4}$ and $\alpha^5 = 1$. In fact, the boomerang uniformity of this specified EA-equivalence class is exactly 6.

Corollary 2. *All permutations EA-equivalent to x^{2^n-2} over $\mathbb{F}_{2^{2n}}$ have boomerang uniformity* 4 *when n is odd and* 6 *when n is even.*

Proof. The proof is straightforward from the above remark and the fact that $\beta_{x^{2^n-2}} = 4$ when n is odd and $\beta_{x^{2^n-2}} = 6$ when n is even (see [8]). □

7 An algorithm for inverting a given BCT

We are interested now in a different but related problem concerning Boomerang Connectivity Tables. This problem is stated as follows: "Given a table B, find all permutations that have B as their Boomerang Connectivity Table." This same problem has been recently investigated by Boura et al. in [9] for the Difference Distribution Table (DDT). Reconstructing an S-box from its DDT or BCT and finding how many permutations share the same table is a theoretically interesting problem that can however also be useful to designers or cryptanalysts of block ciphers. For studying this problem, authors in [9] introduced the notion of DDT-equivalence, applying to functions sharing the same DDT. Similarly, we introduce here the notions of *BCT-equivalence* and *BCT-equivalence class*.

Definition 3. *Two permutations F and G are said* BCT-equivalent *if they have the same Boomerang Connectivity Table. Furthermore, for a given permutation F, all permutations sharing the same BCT as F form its* BCT-equivalence class*, $\mathscr{C}_{\mathsf{BCT}}(F)$.*

In [9], an algorithm for computing the DDT-equivalence class of a given function was provided. However, finding a similar algorithm for computing the BCT-equivalence class of a permutation was stated as an open problem in [8] and [13]. The main goal of this section is to present such an algorithm that permitted us to reconstruct the BCT-equivalence class of many known permutations. Before this, we start by presenting some simple observations concerning the nature of the BCT-equivalence classes.

The first natural question is whether there exist permutations sharing the same BCT and if yes then what are the possible sizes of a BCT-equivalence class. The answer to this question is rather trivial and is given by the following proposition, directly adapted from Proposition 2 in [9].

Proposition 13. *Let F be a permutation of $\mathbb{F}_2^n$ and let ℓ denote the dimension of its linear space, i.e., of the space formed by all linear structures of F. Then, the BCT-equivalence class of F contains the $2^{2n-\ell}$ distinct permutations of the form*

$$x \mapsto F(x \oplus c) \oplus d, \qquad c,d \in \mathbb{F}_2^n.$$

Proof. From Theorem 1, for $a,b \in \mathbb{F}_2^n \setminus \{0\}$, $\beta_F(a,b)$ equals the number of solutions $(x,y) \in \mathbb{F}_2^n \times \mathbb{F}_2^n$ of the system of equations

$$\begin{cases} F(x \oplus a) \oplus F(y \oplus a) = b, \\ F(x) \oplus F(y) = b. \end{cases}$$

Then, for G a permutation of $\mathbb{F}_2^n$ with $G(x) = F(x \oplus c) \oplus d$, the system of equations

$$\begin{cases} G(x \oplus a) \oplus G(y \oplus a) = b, \\ G(x) \oplus G(y) = b. \end{cases} \Leftrightarrow \begin{cases} F(x \oplus c \oplus a) \oplus F(y \oplus c \oplus a) = b, \\ F(x \oplus c) \oplus F(y \oplus c) = b. \end{cases}$$

has the same number of solutions as the first one, by setting $x' = x \oplus c$ and $y' = y \oplus c$. This means that for any $a,b \in \mathbb{F}_2^n \setminus \{0\}$, $\beta_G(a,b) = \beta_F(a,b)$. The rest of the proof on the minimal size of a class is identical to the proof of Proposition 2 in [9] and is therefore omitted here. □

Let F, G be permutations of $\mathbb{F}_2^n$ with $G(x) = F(x \oplus c) \oplus d$, where $c,d \in \mathbb{F}_2^n$. We will say that F and G are *trivially BCT-equivalent*. Proposition 13, when compared to Proposition 2 of [9] shows that the lower bound on the size of equivalence classes is the same for both DDT and BCT equivalences. Note however that while the lower bounds for the two equivalence notions are the same the two classes for a given permutation can be very different. Indeed, for $n = 6$, the DDT-equivalence class of the quadratic permutation $x \mapsto x^5$ contains 2^{12} trivial permutations, while its BCT-equivalence class is much larger.

7.1 Invariance of the Sizes of BCT-Equivalence Classes Under Affine Equivalence

Another natural question is the relation of the BCT-equivalence with other classical equivalences. We show here that the size of the BCT-equivalence classes is preserved under affine-equivalence.

Proposition 14 (adapted from [9]). *Let F and G be two functions which are affine-equivalent, i.e., there exist two affine permutations A_1, A_2 such that $G = A_2 \circ F \circ A_1$. Then, the BCT-equivalence classes of F and of G have the same size. Moreover, the class of G is composed of all $A_2 \circ F' \circ A_1$ where F' varies in the class of F.*

Proof. Let L_1 and L_2 denote the linear parts of the affine functions A_1 and A_2. It was shown in [8] that the BCT coefficients of F and G are related by

$$\beta_G(a,b) = \beta_F(L_1(a), L_2^{-1}(b)), \text{ for all } (a,b)\,. \tag{14}$$

Let $F' \in \mathscr{C}_{\mathsf{BCT}}(F)$ be an element in the BCT-equivalence class of F and let us consider $G' = A_2 \circ F' \circ A_1$. Then, the BCT of F' and G' satisfy: for all (a,b)

$$\beta_{G'}(a,b) = \beta_{F'}(L_1(a), L_2^{-1}(b)) = \beta_F(L_1(a), L_2^{-1}(b))),$$

where the last equality comes from the fact that F and F' have the same BCT. Then, we deduce from (14) that $\beta_{G'}(a,b) = \beta_G(a,b)$ for all (a,b). It follows that

$$\left\{(A_2 \circ F' \circ A_1), F' \in \mathscr{C}_{\mathsf{BCT}}(F)\right\} \subseteq \mathscr{C}_{\mathsf{BCT}}(G).$$

By exchanging the roles of F and G, we deduce that both sets coincide. □

This last result is useful for computations as it shows that is sufficient to compute the size of a BCT-equivalence class for only one representative of the affine-equivalence class.

7.2 Relation between DDT and BCT-Equivalence Classes

To understand the way that the DDT and BCT-equivalence classes are related we have computed and analyzed all DDT and BCT-equivalence classes for 3-bit permutations. Our experiments showed that for $n = 3$ there are 924 DDT-equivalence classes and 512 BCT-equivalence classes. All DDT-equivalence classes are trivial while the BCT-equivalence classes are partitioned as follows:

- 294 trivial classes containing 32 permutations each. Each class corresponds to a different DDT-equivalence class.
- 168 trivial classes containing 64 permutations each. Each class corresponds to a different DDT-equivalence class.
- 1 class containing 1344 permutations. This class corresponds to the BCT where all entries are 8. It encloses 168 DDT-equivalence classes of 8 permutations each, that correspond to all linear permutations for $n = 3$.
- 49 classes containing 384 members each. Each class encapsulates 6 different DDT-equivalence classes containing 64 permutations. For example the BCT-equivalence class shown in Fig. 7 contains all permutations corresponding to the 6 DDTs shown in Fig. 8.

	0	1	2	3	4	5	6	7
0	8	8	8	8	8	8	8	8
1	8	2	.	2	2	.	2	.
2	8	.	8	.	.	8	.	8
3	8	2	.	2	2	.	2	.
4	8	2	.	2	2	.	2	.
5	8	.	8	.	.	8	.	8
6	8	2	.	2	2	.	2	.
7	8	.	8	.	.	8	.	8

FIG. 7: A BCT-equivalence class containing 384 permutations with 6 different DDTs.

As shown by the above computations, two permutations F and G that are BCT-equivalent are not necessarily DDT-equivalent. Moreover, for a permutation F of $\mathbb{F}_2^3$ we always have $\mathscr{C}_{\mathsf{DDT}}(F) \subseteq \mathscr{C}_{\mathsf{BCT}}(F)$, with equality occurring for some classes. However, for higher dimensions this inclusion does not always hold. For example, the two functions of Table 3 are (non-trivially) DDT-equivalent but have different BCTs. More precisely, the DDT-equivalence class of these two permutations has $7 * 2^{11}$ members inside. These permutations can be partitioned into 28 groups of 2^9 permutations each, each group belonging a different BCT class. Furthermore, each of these BCT classes has also $7 * 2^{11}$ permutations inside, coming here also from 28 different DDT-equivalence classes. This last example shows that the relation between DDT and BCT-equivalences is not trivial and needs further investigation.

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	4	.	.	4	.	.
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	4	.	.	.	.	4
6	.	2	.	2	2	.	2	.
7	.	.	.	.	.	4	.	4

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	4	.	.	4	.	.
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	.	.	.	4	.	4
6	.	2	.	2	2	.	2	.
7	.	.	4	.	.	.	.	4

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	4	.	.	.	.	4
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	4	.	.	4	.	.
6	.	2	.	2	2	.	2	.
7	.	.	.	.	.	4	.	4

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	.	.	.	4	.	4
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	4	.	.	.	.	4
6	.	2	.	2	2	.	2	.
7	.	.	4	.	.	4	.	.

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	.	.	.	4	.	4
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	4	.	.	4	.	.
6	.	2	.	2	2	.	2	.
7	.	.	4	.	.	.	.	4

	0	1	2	3	4	5	6	7
0	8	.	.	.	.	.	.	.
1	.	2	.	2	2	.	2	.
2	.	.	.	.	.	4	.	4
3	.	2	.	2	2	.	2	.
4	.	2	.	2	2	.	2	.
5	.	.	4	.	.	4	.	.
6	.	2	.	2	2	.	2	.
7	.	.	4	.	.	.	.	4

FIG. 8: The 6 DDT-equivalence classes contained in the BCT-equivalence classe shown in Fig. 7.

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$F(x)$	0	1	2	3	4	5	6	7	8	9	10	11	13	12	15	14
$F'(x)$	0	1	2	3	4	5	6	7	8	9	10	11	13	12	15	14

x	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$F(x)$	16	17	19	18	20	21	23	22	25	24	26	27	28	29	31	30
$F'(x)$	16	17	19	18	21	20	22	23	24	25	27	26	28	29	31	30

TABLE 3: Two non-trivially DDT-equivalent permutations with different BCTs.

7.3 An Algorithm for Computing the Size of a BCT-Equivalence Class

We present in this section an algorithm that given a table B of size $2^n \times 2^n$ returns all permutations S whose BCT is the same as B. In other words, the presented algorithm computes and returns the BCT-equivalence class of permutations of a given table B. For this, it implements a tree-traversal search where each level i of the tree corresponds to the possible values of $S(i)$. Therefore, the tree depth is 2^n. The general structure recalls the algorithm given in [9] for recovering the DDT-equivalence class corresponding to a given input table but the inner details remain quite different, because of the differences in the nature of these two tables. More precisely, the main difference in the two algorithms is the `ComputePossibleValues` subroutine (described in detail below) that at each level i computes and stores in a list all possible values for $S(i)$ by taking into account values for S computed at a higher level of the search (notably the values $S(0),\ldots,S(i-1)$) and the constraints imposed by the input table B.

At first sight, the difficulty for providing an algorithm for reconstructing an S-box from its BCT similar to the one given in [9] is that in the original definition of a BCT entry, both S and its inverse S^{-1} appear. However, by using the equivalent definition of $\beta(a,b)$ provided in Theorem 1, that sees this quantity as the number of solutions of a system of equations involving only S and not its inverse, the above problem can be bypassed.

The `ComputePossibleValues` *subroutine* At line 1, en empty list $\mathscr{L}$ is initialised that will contain at the end of the algorithm all possible values for $S(i)$. For this, all possible values $0 \leq x < 2^n$ are checked. A flag for each value x, $flag_x$ is initialised at line 3. This flag equals 1 if the value is possible and 0 otherwise. The first test in lines 4-6 verifies if the value x has not already been taken by a

Algorithm 1 `Main`

Input: A table B of size $2^n \times 2^n$
Output: A list $\mathscr{F}$ of all permutations S of $\mathbb{F}_2^n$ whose BCT equals B

1: $\mathscr{F} \leftarrow \{\emptyset\}$ ▷ Globaly defined
2: $S \leftarrow [0,1,2,\ldots,2^n-1]$ ▷ $\texttt{len}(S) = 2^n$
3: `RecursiveSearch`$(S,1)$
4: **return** $\mathscr{F}$

Algorithm 2 `RecursiveSearch`

Input: A table S of size 2^n, an integer i

1: **if** $i < 2^n$ **then**
2: `ComputePossibleValues`(S,i)
3: **else**
4: **if** $\text{BCT}_S = B$ **then**
5: Append S to $\mathscr{F}$
6: **return**
7: **if** $\mathscr{L} \neq \emptyset$ **then**
8: **for all** $x \in \mathscr{L}$ **do**
9: $S[i] \leftarrow x$
10: `RecursiveSearch`$(S,i+1)$
11: **else**
12: **return**

value $S(j)$ for $0 \leq j < i$. If this is the case, then x is not possible for $S(i)$, as S is a permutation. The following central test will be performed for all possible differences δ. The idea is to verify if for some δ and some value $k \leq i$ the equation $x \oplus S(k) = S(i \oplus \delta) \oplus S(k \oplus \delta)$ is satisfied. If this equation has a solution and at the same time $B[\delta][x \oplus S[k]] = 0$, then x cannot be a solution for $S(i)$. The condition at line 10 ensures that the tested values have already been computed and are not at a lower level of the tree.

7.4 Experiments

Using the above algorithm we were able to compute the BCT-equivalence classes of many known permutations. In our experiments we are not interested in APN permutations, as their BCT-equivalence class equals their DDT-equivalence class and this problem was studied in [9]. We provide here a summary of the results:

S-boxes for $n = 4$ We have computed the BCT-equivalence classes for all 4-bit S-boxes. According to the classification of De Cannière [12] there are 302 classes up to affine equivalence and further to Proposition 14 it is enough to compute the size of the BCT-equivalence class for one representative of the class only. Our algorithm shows that among the 302 classes, 280 are trivial, meaning that they are only composed of permutations of the form $F(x \oplus c) \oplus d$, for $c,d \in \mathbb{F}_2^n$. The other 22 non-trivial classes are summarized in Table 4. The class numbering follows the one given in Table 5.8 of [12]. A first observation is that all S-boxes that appear in Table 4 have differential uniformity $\delta_S \geq 8$. This means that normally, all S-boxes used in practice (i.e. having $\delta_S = 4,6$) have trivial BCT-equivalence classes. Another remark concerns quadratic permutations. These permutations are particularly interesting as many different classes of quadratic permutations were shown to have an optimal boomerang uniformity [8, 18, 22]. There are 6 affine-equivalent quadratic classes for $n = 4$ and all of them have non-trivial BCT-equivalence classes. Finally, as it can be seen, all permutations of Table 4 have BCT-equivalence classes larger or equal than the corresponding DDT-equivalence classes. The symbol † in the table below means that we were not able to finish the computation for the corresponding classes. The reason is that our algorithm does not work well for very dense BCTs (i.e. with only a few zeros). However, we were able to verify that the concerned classes are not trivial, as we found non trivial members inside.

Algorithm 3 `ComputePossibleValues`

Input: A table S of size 2^n, an integer i

$\mathscr{V} \leftarrow \{\emptyset\}$
for all $x \in \{0,1,\dots 2^n-1\}$ **do**
  $flag_x \leftarrow 1$
  for all $j \in \{0,\dots,i-1\}$ **do** ▷ S is a permutation.
    if $S(j) = x$ **then**
      $flag_x \leftarrow 0$
  if $flag_x = 1$ **then**
    for all $\delta \in \{1,\dots,2^n-1\}$ **do**
      for all $k \in \{0,\dots,i\}$ **do**
        if $i \oplus \delta \leq i$ and $k \oplus a \leq i$ **then**
          if $x \oplus S(k) = S(i \oplus \delta) \oplus S(k \oplus \delta)$ and $B[\delta][x \oplus S[k]] = 0$ **then**
            $flag_x \leftarrow 0$
  if $flag_x = 1$ **then**
    Append x to $\mathscr{V}$.

	Class	Representative	Degree	δ_S	β_S	$\#\mathscr{C}_{DDT}$	$\#\mathscr{C}_{BCT}$
1.	32	[12,0,10,2,3,5,4,7,6,9,1,11,8,13,14,15]	3	8	16	2^8	2^9
2.	33	[13,0,10,2,3,5,4,7,6,9,1,11,12,8,14,15]	3	8	16	2^8	2^9
3.	253	[4,0,1,2,3,5,6,7,12,9,8,11,10,13,14,15]	3	8	12	2^8	2^9
4.	254	[4,0,1,2,3,5,6,7,13,9,8,11,12,10,14,15]	3	8	16	2^8	2^9
5.	255	[4,0,1,3,2,5,6,7,10,13,9,11,12,8,14,15]	3	8	16	2^8	2^9
6.	256	[6,0,1,2,3,5,4,7,12,9,10,11,8,13,14,15]	3	8	16	2^8	2^9
7.	257	[6,0,1,2,3,5,4,7,13,9,10,11,12,8,14,15]	3	8	16	2^8	2^9
8.	258	[6,5,1,2,3,0,4,7,8,9,10,11,12,13,14,15]	2	8	16	2^8	63×2^{14}
9.	267	[4,0,1,2,3,5,6,7,12,9,10,11,8,13,14,15]	3	8	16	2^8	2^9
10.	268	[4,0,1,2,3,5,6,7,15,9,10,11,12,13,14,8]	3	8	16	2^8	2^9
11.	277	[4,0,1,6,2,5,3,7,8,9,10,11,12,13,14,15]	3	10	16	2^9	2^9
12.	292	[2,0,1,3,4,5,6,7,8,9,10,11,12,13,14,15]	3	12	16	2^8	†
13.	293	[8,0,10,2,3,5,1,7,4,9,6,11,12,13,14,15]	3	16	16	2^7	2^{14}
14.	294	[4,0,1,3,2,5,6,7,12,8,9,11,10,13,14,15]	2	16	16	2^7	2^{14}
15.	295	[4,0,6,2,3,5,1,7,10,9,8,11,12,13,14,15]	2	16	16	2^7	2^{12}
16.	296	[4,0,6,2,3,5,1,7,8,9,10,11,12,13,14,15]	3	16	16	2^7	2^{13}
17.	297	[2,0,1,3,6,4,5,7,8,9,10,11,12,13,14,15]	2	16	16	2^7	1.5×2^{14}
18.	298	[3,0,1,2,6,5,4,7,8,9,10,11,12,13,14,15]	2	16	16	2^7	2^{10}
19.	299	[3,0,1,2,4,5,6,7,8,9,10,11,12,13,14,15]	3	16	16	2^7	2^{10}
20.	300	[1,0,2,3,4,5,6,7,8,9,10,11,12,13,14,15]	3	16	16	2^7	†
21.	301	[1,0,3,2,4,5,6,7,8,9,10,11,12,13,14,15]	2	16	16	2^6	9×2^9
22.	302	[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15]	1	16	16	2^4	†

TABLE 4: Size of the DDT and the BCT equivalence classes of all permutations for $n = 4$ having a non-trivial BCT-equivalence class. The second column corresponds to the class numbering corresponding to the classification of [12] .

Some popular S-boxes for $n = 8$ We have also computed the BCT and DDT-equivalence classes for some 8-bit S-boxes from the literature. For all of them we have found both DDT and BCT-equivalence classes to be trivial. We give below the list of the S-boxes that we analyzed and we divide them according to their structure, in the same way as in Table 9.7 of [23].

Mathematical : AES, BelT, E2, MAGENTA, Safer, Snow-3G
SPN : CLEFIA S_0, Crypton 0.5, Enocoro, Iceberg, Khazad, Twofish ρ_0, Twofish ρ_1
Feistel : Crypton 1.0, Zorro, Scream, iScream, Zuc s_0, CS-cipher
Lai-Massey : Fox, Fly, Whirlpool
Hill-climbing : Anubis, Kalyna π_0, Kalyna π_1, Kalyna π_2, Kalyna π_3

Pseudo-random : MD2, newDES, Turing
Unknown : Skipjack, Kuznyechik

8 Conclusion

Our results on the BCT and on the boomerang uniformity of permutations with various structures have several consequences. First, by the fact that $\beta_S = 2$ if and only if S is APN, we can immediately see that 3-round Feistel, Lai-Massey and MISTY structures can never be APN.

As mentioned in Section 2.1, the consequences in terms of cryptanalysis can also extend further than boomerang attacks. The guess and determine of Biryukov et al. [4] uses a property equivalent to the fact that the boomerang uniformity of a 3-round Feistel network is always maximal. We can therefore expect the same attack to work against Lai-Massey structures. Interestingly, our results show that it is possible to construct a 3-round MISTY structure immune against the existence of such probability 1 patterns, meaning that they seem to offer some inherent resilience against these attacks. Not only are our results regarding Feistel and Lai-Massey structures very similar, the arguments we used to derive them are also very close—independently of the choice of the linear mapping σ. While the similarity between these two structures makes intuitive sense, we find it interesting to see it displayed in such a clear manner.

We have encountered several cases: for 3-round Lai-Massey, 3-round Feistel and 1-round SPN, the boomerang uniformity is maximal regardless of the subcomponents used. In the 3-round MISTY case, the boomerang uniformity is bounded by the *differential* uniformity of its subfunction and, in the Littlun case, it is bounded by the *boomerang* uniformity of its subfunction.

Another application of our results lies in S-box reverse-engineering [5]. In this context, the aim is to recover the hidden structure of an S-box using only its lookup table. If an S-box has non-trivial differential and linear properties but a boomerang uniformity equal to 2^n then we can suspect that it is a 3-round Lai-Massey or Feistel structure. Since the boomerang uniformity is preserved under the composition with an affine permutation, this test would work even if the S-box structure is obfuscated by such permutations—as is the case for instance in the S-box of ZUC [14].

Finally, the initial analysis of the BCT-equivalence classes problem that we provided, gives rise to many open questions. For example, for a permutation F, is the cardinality of a BCT-equivalence class always higher or equal to the size of the corresponding DDT-equivalence class? Further, can we derive any bounds on the size of the BCT-equivalence classes for quadratic permutations? Finally, an interesting direction is to further investigate the relation between the two equivalence notions.

References

1. Albrecht, M.R., Grassi, L., Rechberger, C., Roy, A., Tiessen, T.: MiMC: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In: J.H. Cheon, T. Takagi (eds.) ASIACRYPT 2016, Part I, *LNCS*, vol. 10031, pp. 191–219. Springer, Heidelberg (2016). DOI 10.1007/978-3-662-53887-6_7
2. Banik, S., Bogdanov, A., Isobe, T., Shibutani, K., Hiwatari, H., Akishita, T., Regazzoni, F.: Midori: A block cipher for low energy. In: T. Iwata, J.H. Cheon (eds.) ASIACRYPT 2015, Part II, *LNCS*, vol. 9453, pp. 411–436. Springer, Heidelberg (2015). DOI 10.1007/978-3-662-48800-3_17
3. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of Cryptology **4**(1), 3–72 (1991). DOI 10.1007/BF00630563
4. Biryukov, A., Leurent, G., Perrin, L.: Cryptanalysis of Feistel networks with secret round functions. In: O. Dunkelman, L. Keliher (eds.) SAC 2015, *LNCS*, vol. 9566, pp. 102–121. Springer, Heidelberg (2016). DOI 10.1007/978-3-319-31301-6_6
5. Biryukov, A., Perrin, L.: On reverse-engineering S-boxes with hidden design criteria or structure. In: R. Gennaro, M.J.B. Robshaw (eds.) CRYPTO 2015, Part I, *LNCS*, vol. 9215, pp. 116–140. Springer, Heidelberg (2015). DOI 10.1007/978-3-662-47989-6_6
6. Bluher, A.W.: On x^{q+1}+ax+b. Finite Fields and Their Applications **10**(3), 285–305 (2004). DOI 10.1016/j.ffa.2003.08.004. URL https://doi.org/10.1016/j.ffa.2003.08.004
7. Bonnetain, X., Perrin, L., Tian, S.: Anomalies and Vector Space Search: Tools for S-Box Reverse-Engineering. Cryptology ePrint Archive, Report 2019/528 (2019). URL https://eprint.iacr.org/2019/528

8. Boura, C., Canteaut, A.: On the boomerang uniformity of cryptographic sboxes. IACR Trans. Symm. Cryptol. **2018**(3), 290–310 (2018). DOI 10.13154/tosc.v2018.i3.290-310
9. Boura, C., Canteaut, A., Jean, J., Suder, V.: Two notions of differential equivalence on Sboxes. Des. Codes Cryptogr. **87**(2-3), 185–202 (2019). DOI 10.1007/s10623-018-0496-z. URL https://doi.org/10.1007/s10623-018-0496-z
10. Canteaut, A., Duval, S., Leurent, G.: Construction of lightweight S-boxes using Feistel and MISTY structures. In: O. Dunkelman, L. Keliher (eds.) SAC 2015, *LNCS*, vol. 9566, pp. 373–393. Springer, Heidelberg (2016). DOI 10.1007/978-3-319-31301-6_22
11. Cid, C., Huang, T., Peyrin, T., Sasaki, Y., Song, L.: Boomerang connectivity table: A new cryptanalysis tool. In: J.B. Nielsen, V. Rijmen (eds.) EUROCRYPT 2018, Part II, *LNCS*, vol. 10821, pp. 683–714. Springer, Heidelberg (2018). DOI 10.1007/978-3-319-78375-8_22
12. De Cannière, C.: Analysis and Design of Symmetric Encryption Algorithms. PhD thesis, Katholieke Universiteit Leuven (2007)
13. Dunkelman, O., Huang, S.: Reconstructing an S-box from its Difference Distribution Table. IACR Trans. Symmetric Cryptol. **2019**(2), 193–217 (2019). DOI 10.13154/tosc.v2019.i2.193-217. URL https://doi.org/10.13154/tosc.v2019.i2.193-217
14. ETSI/Sage: Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3. Document 4 : Design and Evaluation Report. Tech. rep., ETSI/Sage (2011). Available at http://www.gsma.com/aboutus/wp-content/uploads/2014/12/EEA3_EIA3_Design_Evaluation_v2_0.pdf
15. Grosso, V., Leurent, G., Standaert, F.X., Varici, K.: LS-designs: Bitslice encryption for efficient masked software implementations. In: C. Cid, C. Rechberger (eds.) FSE 2014, *LNCS*, vol. 8540, pp. 18–37. Springer, Heidelberg (2015). DOI 10.1007/978-3-662-46706-0_2
16. Grosso, V., Leurent, G., Standaert, F.X., Varici, K., Anthony Journault, F.D., Gaspar, L., Kerckhof, S.: SCREAM & iSCREAM Side-Channel Resistant Authenticated Encryption with Masking. Candidate for the CAESAR Competition. See also http://perso.uclouvain.be/fstandae/SCREAM/ (2014)
17. Karpman, P., Grégoire, B.: The LITTLUN S-box and the FLY block cipher. In: Lightweight Cryptography Workshop 2016, October 17-18 (informal proceedings). National Institute of Standards and Technology (2016)
18. Li, K., Qu, L., Sun, B., Li, C.: New results about the boomerang uniformity of permutation polynomials. Cryptology ePrint Archive, Report 2019/079 (2019). https://eprint.iacr.org/2019/079
19. Li, Y., Wang, M.: On EA-equivalence of certain permutations to power mappings. Des. Codes Cryptogr. **58**(3), 259–269 (2011). DOI 10.1007/s10623-010-9406-8. URL https://doi.org/10.1007/s10623-010-9406-8
20. Li, Y., Wang, M.: Permutation polynomials EA-equivalent to the inverse function over GF (2^n). Cryptography and Communications **3**(3), 175–186 (2011). DOI 10.1007/s12095-011-0045-3. URL https://doi.org/10.1007/s12095-011-0045-3
21. Matsui, M.: New block encryption algorithm MISTY. In: E. Biham (ed.) FSE'97, *LNCS*, vol. 1267, pp. 54–68. Springer, Heidelberg (1997). DOI 10.1007/BFb0052334
22. Mesnager, S., Tang, C., Xiong, M.: On the boomerang uniformity of (quadratic) permutations over $\mathrm{f}_{2^{\mathrm{n}}}$. CoRR **abs/1903.00501** (2019). URL http://arxiv.org/abs/1903.00501
23. Perrin, L.: Cryptanalysis, Reverse-Engineering and Design of Symmetric Cryptographic Algorithms. Ph.D. thesis, University of Luxembourg (2017). URL http://orbilu.uni.lu/handle/10993/31195
24. Vaudenay, S., Junod, P.: Device and method for encrypting and decrypting a block of data. United States Patent (20040247117), see also "Fox, a New Family of Block Ciphers" http://crypto.junod.info/sac04a.pdf (2004)
25. Wagner, D.: The boomerang attack. In: L.R. Knudsen (ed.) FSE'99, *LNCS*, vol. 1636, pp. 156–170. Springer, Heidelberg (1999). DOI 10.1007/3-540-48519-8_12

A Proof of Proposition 6

We prove here the bounds provided in Proposition 6 for the two types of unbalanced MISTY network depicted in Fig. 5. We start by the network on the left.

Proof. Suppose that $m < n$ and define the function

$$S_b : \begin{cases} \mathbb{F}_2^m \times \mathbb{F}_2^m \times \mathbb{F}_2^{n-m} & \to \mathbb{F}_2^m \times \mathbb{F}_2^m \times \mathbb{F}_2^{n-m} \\ (x,y,z) & \mapsto S^{-1}(S(x,y,z)+b) \end{cases}$$

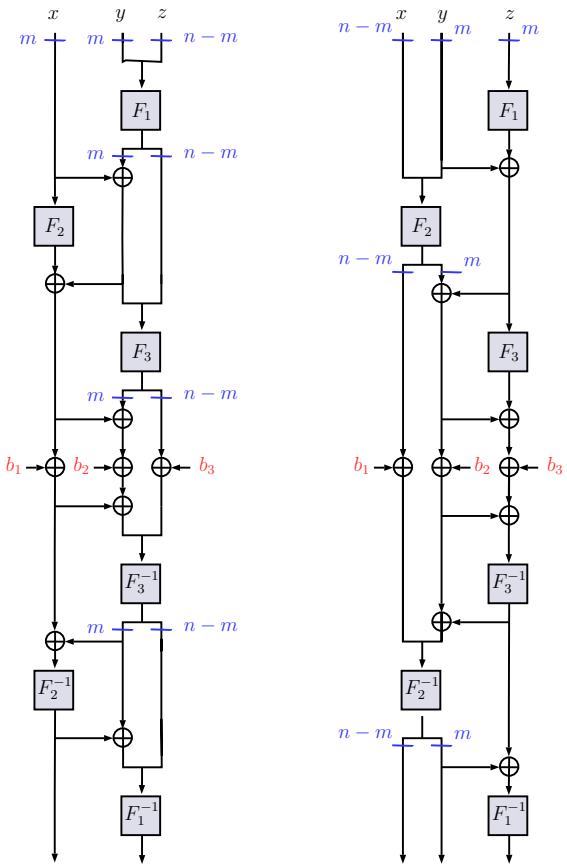


FIG. 9: The functions $S_b(x) = S^{-1}(S(x)+b)$, for $x \in \mathbb{F}_2^m \times \mathbb{F}_2^n$ (on the left), and for $x \in \mathbb{F}_2^n \times \mathbb{F}_2^m$ (on the right), where S is a 3-round unbalanced MISTY structure.

where F_1 and F_3 are permutations over $\mathbb{F}_2^n$ and F_2 is a permutation over $\mathbb{F}_2^m$. This function is displayed at the left of Fig. 9. We denote the values of (x,y,z) after i rounds for both cases as (ℓ_i, m_i, r_i).

Begin with $(l_0, m_0, r_0) = (x,y,z) \in \mathbb{F}_2^m \times \mathbb{F}_2^m \times \mathbb{F}_2^{n-m}$. Then

$$\begin{aligned}
(l_1,m_1,r_1) &= (x, F_1(y||z)_l + x, F_1(y||z)_r)\,;\\
(l_2,m_2,r_2) &= (F_2(x) + F_1(y||z)_l + x, F_1(y||z)_l + x, F_1(y||z)_r)\,;\\
(l_3,m_3,r_3) &= \big(F_2(x) + F_1(y||z)_l + x, F_2(x) + F_1(y||z)_l + x + F_3(F_1(y||z) + x||0)_l,\\
&\qquad F_3(F_1(y||z) + x||0)_r\big).
\end{aligned}$$

Consider $b = (b_1, b_2, b_3)$ with $b_1 = b_2 \neq 0$ and $b_3 = 0$, with b_1 and b_2 being m-bit values and b_3 being an $(n-m)$-bit value. By starting with $(l_3 + b_1, m_3 + b_1, r_3)$, we get

$$\begin{aligned}
(l_4,m_4,r_4) &= (F_2(x) + F_1(y||z)_l + x + b_1, F_1(y||z)_l + x, F_1(y||z)_r)\,;\\
(l_5,m_5,r_5) &= \big(F_2^{-1}(F_2(x) + b_1), F_1(y||z)_l + x, F_1(y||z)_r\big)\,;\\
(l_6,m_6,r_6) &= \big(F_2^{-1}(F_2(x) + b_1), F_1^{-1}\left(F_1(y||z) + (F_2^{-1}(F_2(x) + b_1) + x)||0\right)_l,\\
&\qquad F_1^{-1}\left(F_1(y||z) + (F_2^{-1}(F_2(x) + b_1) + x)||0\right)_r\big).
\end{aligned}$$

That is, for $b = (b_1, b_1, 0)$ with $b_1 \neq 0 \in \mathbb{F}_2^m$,

$$\begin{aligned}
S_b(x,y,z) = &\big(F_2^{-1}(F_2(x) + b_1), F_1^{-1}\left(F_1(y||z) + (F_2^{-1}(F_2(x) + b_1) + x)||0\right)_l,\\
&F_1^{-1}\left(F_1(y||z) + (F_2^{-1}(F_2(x) + b_1) + x)||0\right)_r\big).
\end{aligned}$$

Then, for $a=(a_1,0,0), b=(b_1,b_1,0)\in\mathbb{F}_2^m\times\mathbb{F}_2^m\times\mathbb{F}_2^{n-m}$,

$$\begin{aligned}
&S^{-1}(S(x,y,z)+(b_1,b_1,0))+S^{-1}(S(x+a_1,y,z)+(b_1,b_1,0))\\
&=S_b(x,y,z)+S_b(x+a_1,y,z)\\
&=\left(F_2^{-1}(F_2(x)+b_1)+F_2^{-1}(F_2(x+a_1)+b_1),F_1^{-1}\left(F_1(y||z)+(F_2^{-1}(F_2(x)+b_1)+x)||0\right)_l\right.\\
&\quad+F_1^{-1}\left(F_1(y||z)+(F_2^{-1}(F_2(x+a_1)+b_1)+x+a_1)||0\right)_l,F_1^{-1}\left((F_2^{-1}(F_2(x)+b_1)+x)||0\right)\\
&\quad\left.+F_1(y||z)\right)_r+F_1^{-1}\left(F_1(y||z)+(F_2^{-1}(F_2(x+a_1)+b_1)+x+a_1)||0\right)_r\Big).
\end{aligned}\tag{15}$$

By definition, we deduce that $\beta_S(a,b)=2^n\beta_{F_2}(a_1,b_1)$. Indeed, since for any $x\in T$ where

$$T=\{x\in\mathbb{F}_2^n|F_2^{-1}(F_2(x)+b_1)+F_2^{-1}(F_2(x+a_1)+b_1)=a_1\}$$

we have $F_2^{-1}(F_2(x+a_1)+b_1)=F_2^{-1}(F_2(x)+b_1)+a_1$, Eq. (15) becomes

$$S^{-1}(S(x,y,z)+(b_1,b_1,0))+S^{-1}(S(x+a_1,y,z)+(b_1,b_1,0))=(a_1,0,0).$$

Now, by choosing a_1,b_1 such that $\beta_{F_2}(a_1,b_1)=\beta_{F_2}$, it follows that $\beta_S\geq 2^n\beta_{F_2}$. □

Lemma 2. *If F_1 is an affine permutation then the unbalanced MISTY network on the left of Fig. 5 has the worst possible BCT.*

Proof. Since F_1 is affine, we can simplify the function $S_b(x,y,z)$ as

$$\begin{aligned}
S_b(x,y,z)=&\left(F_2^{-1}(F_2(x)+b_1),y+F_1^{-1}(0)_l+F_1^{-1}\left((F_2^{-1}(F_2(x)+b_1)+x)||0\right)_l,\right.\\
&\left.z+F_1^{-1}(0)_r+F_1^{-1}\left((F_2^{-1}(F_2(x)+b_1)+x)||0\right)_r\right).
\end{aligned}$$

By chosing $a=(0,a_2,a_3)\in\mathbb{F}_2^m\times\mathbb{F}_2^m\times\mathbb{F}_2^{n-m}$ with $a_2,a_3\neq 0$, we get

$$\begin{aligned}
&S^{-1}(S(x,y,z)+(b_1,b_1,0))+S^{-1}(S(x,y+a_2,z+a_3)+(b_1,b_1,0))\\
&=S_b(x,y,z)+S_b(x,y+a_2,z+a_3)\\
&=(0,a_2,a_3)
\end{aligned}$$

which holds for any $(x,y,z)\in\mathbb{F}_2^m\times\mathbb{F}_2^m\times\mathbb{F}_2^{n-m}$. Therefore, $\beta_S=2^{n+m}$. □

We provide now the proof for the network depicted on the right of Fig. 5. The proof for this case is very similar to the previous one and we only provide it here for the sake of completeness.

Proof. Suppose that $m<n$ and define the function

$$S_b:\begin{cases}\mathbb{F}_2^{n-m}\times\mathbb{F}_2^m\times\mathbb{F}_2^m & \to\mathbb{F}_2^{n-m}\times\mathbb{F}_2^m\times\mathbb{F}_2^m\\ (x,y,z) & \mapsto S^{-1}(S(x,y,z)+b)\end{cases}$$

when F_1 and F_3 are permutations over $\mathbb{F}_2^m$ and F_2 is a permutation over $\mathbb{F}_2^n$. This function is displayed on the rightside of Fig. 9. We denote the values of (x,y,z) after i rounds for both cases as (ℓ_i,m_i,r_i).

Begin with $(l_0,m_0,r_0)=(x,y,z)\in\mathbb{F}_2^{n-m}\times\mathbb{F}_2^m\times\mathbb{F}_2^m$. Then,

$$\begin{aligned}
&(l_1,m_1,r_1)=(x,y,y+F_1(z));\\
&(l_2,m_2,r_2)=(F_2(x||y)_l,y+F_1(z)+F_2(x||y)_r,y+F_1(z));\\
&(l_3,m_3,r_3)=\left(F_2(x||y)_l,y+F_1(z)+F_2(x||y)_r,y+F_1(z)+F_2(x||y)_r+F_3\left(y+F_1(z)\right)\right).
\end{aligned}$$

Consider $b=(b_1,b_2,b_3)$ with $b_1=0$ and $b_2=b_3\neq 0$, with b_1 being an $(n-m)$-bit value and b_2,b_3 being m-bit values. Then, by beginning with (l_3,m_3+b_2,r_3+b_2) we get

$$\begin{aligned}
&(l_4,m_4,r_4)=\left(F_2(x||y)_l,b_2+y+F_1(z)+F_2(x||y)_r,y+F_1(z)\right);\\
&(l_5,m_5,r_5)=\left(F_2^{-1}\left(F_2(x||y)+(0||b_2)\right)_l,F_2^{-1}\left(F_2(x||y)+(0||b_2)\right)_r,y+F_1(z)\right);\\
&(l_6,m_6,r_6)=\left(F_2^{-1}\left(F_2(x||y)+(0||b_2)\right)_l,F_2^{-1}\left(F_2(x||y)+(0||b_2)\right)_r,\right.\\
&\qquad\left.F_1^{-1}\left(y+F_1(z)+F_2^{-1}\left(F_2(x||y)+(0||b_2)\right)_r\right)\right).
\end{aligned}$$

That is, for $b = (0, b_2, b_2)$ with $b_2 \neq 0 \in \mathbb{F}_2^m$,

$$\begin{aligned} S_b(x,y,z) = \big(&F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_l, F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_r, \\ &F_1^{-1}\left(y + F_1(z) + F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_r\right)\big). \end{aligned}$$

Then, for $a = (0, a_2, 0)$, $b = (0, b_2, b_2) \in \mathbb{F}_2^{n-m} \times \mathbb{F}_2^m \times \mathbb{F}_2^m$ with $a_2, b_2 \neq 0$ we have

$$\begin{aligned} &S^{-1}(S(x,y,z) + (0,b_2,b_2)) + S^{-1}(S(x,y+a_2,z) + (0,b_2,b_2)) \\ &= S_b(x,y,z) + S_b(x,y+a_2,z) \\ &= \big(F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_l + F_2^{-1}\left(F_2(x||y+a_2) + (0||b_2)\right)_l, F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_r \\ &\quad + F_2^{-1}\left(F_2(x||y+a_2) + (0||b_2)\right)_r, F_1^{-1}\left(y + F_1(z) + F_2^{-1}\left(F_2(x||y) + (0||b_2)\right)_r\right) \\ &\quad + F_1^{-1}\left(y + F_1(z) + F_2^{-1}\left(F_2(x||y+a_2) + (0||b_2)\right)_r\right)\big). \end{aligned} \tag{16}$$

By definition, we deduce that $\beta_S(a,b) = 2^m \beta_{F_2}(0||a_2, 0||b_2)$. Indeed, since for any $x||y \in T$ where

$$T = \{x||y \in \mathbb{F}_2^n | F_2^{-1}\left(F_2(x||y) + (0||b_2)\right) + F_2^{-1}\left(F_2(x||y+a_2) + (0||b_2)\right) = 0||a_2\},$$

$F_2^{-1}\left(F_2(x||y+a_2) + (0||b_2)\right) = F_2^{-1}\left(F_2(x||y) + (0||b_2)\right) + 0||a_2$, Eq. (16) becomes

$$S^{-1}(S(x,y,z) + (b_1,b_1,0)) + S^{-1}(S(x,y,z) + (b_1,b_1,0)) = (0,a_2,0).$$

Now, we choose a_2, b_2 such that $\beta_{F_2}(0||a_2, 0||b_2) = \beta_{F_2|_{\mathbb{F}_2^m}}$. It follows that $\beta_S \geq 2^m \beta_{F_2|_{\mathbb{F}_2^m}}$. □

From the above lower bound, we may come up with the idea of constructing an S-box with boomerang uniformity equal to 4 by setting $m = 1$ and by choosing F_2 to be an APN permutation. However, in this case, F_1 would have to be an affine permutation as the only 1-bit permutations are $x \mapsto x$ and $x \mapsto x \oplus 1$. Lemma 2 can be adapted to this situation and yields the same conclusion: the boomerang uniformity of such a 3-round MISTY network is maximal. Hence, this approach cannot work.