Construction of Lightweight S-Boxes Using Feistel and MISTY Structures

Anne Canteaut[✉], Sébastien Duval, and Gaëtan Leurent

Inria, project-team SECRET, Rocquencourt, France
{Anne.Canteaut,Sebastien.Duval,Gaetan.Leurent}@inria.fr

Abstract. The aim of this work is to find large S-Boxes, typically operating on 8 bits, having both good cryptographic properties and a low implementation cost. Such S-Boxes are suitable building-blocks in many lightweight block ciphers since they may achieve a better security level than designs based directly on smaller S-Boxes. We focus on S-Boxes corresponding to three rounds of a balanced Feistel and of a balanced MISTY structure, and generalize the recent results by Li and Wang on the best differential uniformity and linearity offered by such a construction. Most notably, we prove that Feistel networks supersede MISTY networks for the construction of 8-bit permutations. Based on these results, we also provide a particular instantiation of an 8-bit permutation with better properties than the S-Boxes used in several ciphers, including Robin, Fantomas or CRYPTON.

Keywords: S-Box · Feistel network · MISTY network · Lightweight block-cipher

1 Introduction

A secure block cipher must follow Shannon's criteria and provide confusion and diffusion [42]. In most cases, confusion is achieved with small substitution boxes (S-Boxes) operating on parts of the state (usually bytes) in parallel, and diffusion is achieved with linear operations mixing the state. The security of the cipher is then strongly dependent on the cryptographic properties of the S-Boxes. For instance, the AES uses an 8-bit S-Box based on the inversion in the finite field with 2^8 elements. This S-Box has the smallest known differential probability and linear correlation, and then allows the AES to be secure with a small number of rounds, and to reach very good performances. However, it is not always the best option for constrained environments. In software, an S-Box can be implemented with a look-up table in memory, but this takes 256 bytes for the AES S-Box, and there might be issues with cache-timing attacks [7]. In hardware, the best known implementation of the AES S-Box requires 115 gates [13]; this hardware description can also be used for a bit-sliced software implementation [25].

For some constrained environments, this cost might be too high. Therefore, the field of lightweight cryptography has produced many alternatives with a smaller footprint, such as TEA [47], CRYPTON [29,30], NOEKEON [16], PRESENT [11], KATAN [17], LBLOCK [48], PRINCE [12], TWINE [44], the LS-Designs [23], or PRIDE [2]. In particular, many of those lightweight ciphers use S-Boxes operating on 4-bit words, or even on a smaller alphabet like in [1]. But, reducing the number of variables increases the values of the optimal differential probability and linear correlation. Therefore, more rounds are required in order to achieve the same resistance against differential and linear attacks.

An alternative approach when constructing a lightweight cipher consists in using larger S-Boxes, typically operating on 8 bits like in the AES, but with a lower implementation cost. Then, we search for S-Boxes with better implementations than the AES S-Box, at the cost of suboptimal cryptographic properties. Finding 8-bit S-Boxes which offer such an interesting trade-off is a difficult problem: they cannot be classified like in the 4-bit case [18,27], and randomly chosen S-Boxes have a high implementation cost [46]. Therefore, we focus on constructions based on smaller S-Boxes and linear operations. This general approach has been used in several previous constructions: CRYPTON v0.5 [29] (3-round Feistel), CRYPTON v1.0 [30] (2-round SPN), WHIRLPOOL [5] (using five small S-Boxes), KHAZAD [4] (3-round SPN), ICEBERG [43] (3-round SPN), ZORRO [21] (4-round Feistel with mixing layer), and the LS-Designs [23] (3-round Feistel and MISTY network). As in [23], we here focus on constructions with a 3-round Feistel network, or a 3-round balanced MISTY network, because they use only 3 smaller S-Boxes, but can still provide good large S-Boxes. And we study the respective merits of these two constructions, since this comparison is raised as an open question in [23].

The Feistel and MISTY structures have been intensively studied in the context of block cipher design, and bounds are known for the maximum expected differential probability (MEDP) [3,32,38] and maximum expected linear potential (MELP) [3,37]. However, those results are not relevant for the construction of S-Boxes, because they only consider the average value over all the keys, while an S-Box is unkeyed. Therefore, the differential and linear properties of the Feistel and MISTY constructions need to be analyzed in the unkeyed setting. Such a study has been initiated recently by Li and Wang [28] in the case of 3 rounds of a Feistel network. In this work, we expand the results of Li and Wang, by giving some more general theoretical results for unkeyed Feistel and MISTY structures, with a particular focus on the construction of 8-bit permutations. Due to the page limitation, some of the results are not detailed here and are presented in the full version of this paper [15] only.

Our contributions. We first explain why the usual MEDP and MELP notions are meaningless in the unkeyed setting. In particular we exhibit a 3-round MISTY network where there exists a differential with probability higher than the MEDP *for any fixed key*, but this optimal differential depends on the key. Then, Sect. 3 gives some lower bounds on the differential uniformity and linearity of any 3-round balanced MISTY structure, which involve the properties of the three

inner S-Boxes. Similar results on 3-round Feistel networks are also detailed in Sect. 5 of [15], which generalize the previous results from [28]. Section 4 then focuses on the construction of 8-bit permutations. Most notably, we show that 3 rounds of a Feistel network with appropriate inner S-Boxes provide better cryptographic properties than any 3-round MISTY network, explaining some experimental results reported in [23]. Section 5 then gives an instantiation of such an 8-bit permutation, which offers a very good trade-off between the cryptographic properties and the implementation cost. It can be implemented efficiently in hardware and for bit-sliced software, and has good properties for side-channel resistant implementations with masking. In particular, this S-Box supersedes the S-Boxes considered in many lightweight ciphers including CRYPTON, Robin and Fantomas.

2   From Keyed Constructions to Unkeyed S-Boxes

2.1   Main Cryptographic Properties for an S-Box

In this paper, we focus on S-Boxes having the same number of input and output bits. The resistance offered by an S-Box against differential [10] and linear [31] cryptanalysis is quantified by the highest value in its difference table (resp. table of linear biases, aka linear-approximation table). More precisely, these two major security parameters are defined as follows.

Definition 1 (Differential uniformity [36]**).** *Let F be a function from $\mathbb{F}_2^n$ into $\mathbb{F}_2^n$. For any pair of differences (a, b) in $\mathbb{F}_2^n$, we define the set*

$$D_F(a \to b) = \{x \in \mathbb{F}_2^n \mid F(x \oplus a) \oplus F(x) = b\}.$$

The entry at position (a, b) in the difference table of F then corresponds to the cardinality of $D_F(a \to b)$ and will be denoted by $\delta_F(a, b)$.

Moreover, the differential uniformity of F is

$$\delta(F) = \max_{a \neq 0, b} \delta_F(a, b).$$

Obviously, the differential uniformity of an S-Box is always even, implying that, for any F, $\delta(F) \geq 2$. The functions F for which equality holds arze named *almost perfect nonlinear (APN) functions*.

Similarly, the bias of the best linear approximation of an S-Box is measured by its linearity.

Definition 2 (Walsh transform of an S-Box). *Let F be a function from $\mathbb{F}_2^n$ into $\mathbb{F}_2^n$. The Walsh transform of F is the function*

$$\begin{aligned}\mathbb{F}_2^n \times \mathbb{F}_2^n &\to \mathbb{Z} \\ (a, b) &\mapsto \lambda_F(a, b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{b \cdot F(x) + a \cdot x}.\end{aligned}$$

Moreover, the linearity *of F is*

$$\mathcal{L}(F) = \max_{a, b \in \mathbb{F}_2^n, b \neq 0} |\lambda_F(a, b)|.$$

Indeed, up to a factor 2^n, the linearity corresponds to the bias of the best linear relation between the input and output of F:

$$\Pr_X[b \cdot F(X) + a \cdot X = 1] = \frac{1}{2^n}\left(2^{n-1} - \frac{1}{2}\sum_{x \in \mathbb{F}_2^n}(-1)^{b \cdot F(x) + a \cdot x}\right) = \frac{1}{2}\left(1 - \frac{\lambda_F(a,b)}{2^n}\right).$$

It is worth noticing that, for any fixed output mask $b \in \mathbb{F}_2^n$, the function $a \mapsto \lambda_F(a,b)$, corresponds to the Walsh transform of the n-variable Boolean component of F: $x \mapsto b \cdot F(x)$. In particular, it enjoys all properties of a discrete Fourier transform, for instance the Parseval relation.

2.2 Constructing S-Boxes from Smaller Ones

If this paper we focus on the construction of S-Boxes using several smaller S-Boxes. Indeed small S-Boxes are much cheaper to implement that large S-Boxes:

– for table-based software implementations, the tables are smaller;
– for hardware implementations, the gate count is lower;
– for bit-sliced software implementation, the instructions count is lower;
– for vectorized implementation, small S-Boxes can use vector permutations.

In many cases, implementing several small S-Boxes requires less resources than implementing a large one. Therefore, constructing S-Boxes from smaller ones can reduce the implementation cost.

The Feistel construction is a well-known construction to build a $2n$-bit permutation from smaller n-bit functions, introduced in 1971 for the design of Lucifer (which later became DES [34]). It is a good candidate for constructing large S-Boxes from smaller ones at a reasonable implementation cost. In particular, this construction has been used for the S-Boxes of Crypton v0.5 [29], ZUC [20] (for S_0), Robin [23] and iScream [22]. The MISTY construction introduced by Matsui [32] uses a different structure, but offers a similar level of security. The main advantage of the MISTY network is that it can offer a reduced latency because the first two S-Boxes can be evaluated in parallel. Therefore it is a natural alternative to Feistel networks for the construction of lightweight S-Boxes, and it has been used in the design of Fantomas [23] and Scream [22]. In order to reduce the number of gates used for implementing the construction, we focus on *balanced* MISTY networks, while the MISTY block cipher proposed in [33] is unbalanced and combines an $(n-1)$-bit S-Box and an $(n+1)$-bit S-Box.

The two structures we study are depicted in Figs. 1 and 2. It is worth noticing a major difference between the two: the function resulting from the Feistel construction is always invertible (since one round is an involution, up to a permutation of the outputs), while the function resulting from the MISTY construction is invertible if and only if all the inner S-Boxes are invertible.
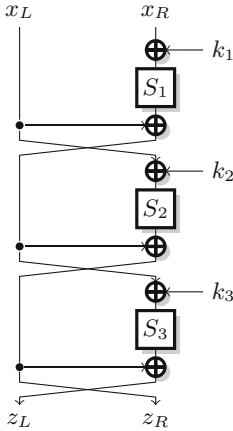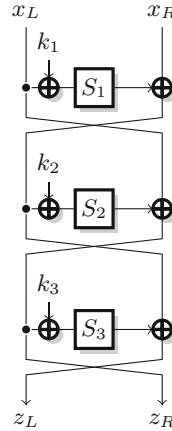
Fig. 1. 3-round MISTY network

Fig. 2. 3-round Feistel network

Analysis of Feistel and MISTY Structures. Since these two constructions have been used for the design of many block ciphers (in particular the DES [34] and MISTY [33], respectively), their security properties have been intensively studied. A natural way to measure the resistance of the resulting block cipher against differential and linear cryptanalysis is to study the probabilities of the differentials (respectively the potentials of the linear approximations) averaged over all keys.

Definition 3 (MEDP and MELP)**.** *Let F_K be a family of function from $\mathbb{F}_2^n$ into $\mathbb{F}_2^n$. The MEDP is the maximum probability of a differential, averaged over all keys:*

$$\mathsf{MEDP}(F_K) = \max_{a \neq 0, b} \frac{1}{2^k} \sum_{K \in \mathbb{F}_2^k} \frac{\delta_{F_K}(a, b)}{2^n}.$$

The MELP is the maximum potential of a linear approximation, averaged over all keys:

$$\mathsf{MELP}(F_K) = \max_{a, b \neq 0} \frac{1}{2^k} \sum_{K \in \mathbb{F}_2^k} \left(\frac{\lambda_{F_K}(a, b)}{2^n} \right)^2.$$

The following theorem shows that the MEDP and MELP of a Feistel or MISTY network can be bounded.

Theorem 1 (Feistel or MISTY, averaged over all keys, [3,32,38]). *Given S_1, S_2 and S_3 three n-bit permutations, let $p = \max_i \delta(S_i)/2^n$ and $q = \max_i (\mathcal{L}(S_i)/2^n)^2$. Then the family of functions $(F_K)_{K=(K_1,K_2,K_3)\in\mathbb{F}_2^{3k}}$ defined by 3 rounds of a Feistel or of a MISTY network with S_i as inner functions verifies*

$$\mathsf{MEDP}(F_K) \leq p^2 \text{ and } \mathsf{MELP}(F_K) \leq q^2.$$

This theorem is very powerful for the construction of iterated block ciphers: it shows that a big function with strong cryptographic properties can be built from small functions with strong cryptographic properties. However, for the design of an S-Box from smaller S-Boxes, it is of little use. Indeed, we are interested in the properties of a single fixed S-Box, rather than the average properties of a family of S-Boxes. For a fixed (a, b) the theorem proves that the average values of $\delta_{F_K}(a, b)$ and $\lambda_{F_K}(a, b)$ are bounded, therefore there exists at least one key for which the value is smaller than or equal to the average. However, it might be that the values a, b where the maximum is reached are not the same for every key. Therefore if we select a key so that $\delta_{F_K}(a, b)$ is small for an a, b maximizing the average probability, the maximum can be reached for another entry of the differential table.

More strikingly, we discovered some choices of S_1, S_2, S_3 such that the maximum differential probability of the functions in the corresponding family is always higher than the MEDP.

Example 1. We consider a MISTY structure with three identical S-Boxes:

$$S_i = [\mathtt{A}, 7, 9, 6, 0, 1, 5, \mathtt{B}, 3, \mathtt{E}, 8, 2, \mathtt{C}, \mathtt{D}, 4, \mathtt{F}].$$

We have $\mathsf{MEDP}(F_K) \leq 16/256$ according to Theorem 1, because $\delta(S_i) = 4$. However, for *any* function in this family, there exists a differential with probability $32/256$. This is not a contradiction, because the differential reaching the maximum depends on the key.

The relevant property for the construction of an S-Box is the maximum differential probability (respectively maximum linear potential). Therefore, we could derive some information on this quantity for F_K for some fixed keys from the knowledge of the average value of the maximal differential probability, i.e., the EMDP (resp. EMLP), which may significantly differ from the MEDP (resp. MELP). We would like to point out that there is a confusion between the two notions in [33]: the definition corresponds to the expected maximum differential probability (respectively expected maximum linear potential), while the theorems apply to the MEDP and MELP.

Analysis of Feistel and MISTY Structures with Fixed Key. In order to study the properties of Feistel and MISTY structures for the construction of lightweight S-Boxes, we must study these structures with a fixed key. Equivalently, we can consider the structures without any key, because a structure with a fixed key is equivalent to an unkeyed one with different S-Boxes. Indeed, using an S-Box S_i with round key k_i is equivalent to using $S_i' : x \mapsto S_i(x + k_i)$ as an S-Box without any key. In the following, we always consider a key-less variant.

In a recent analysis of the fixed-key Feistel structure [28], Li and Wang derive the best differential uniformity and linearity which can be achieved by a 3-round Feistel cipher with a fixed key, and give examples reaching this bound. Their main results are as follows:

Theorem 2 (Feistel unkeyed, [28]). *Let S_1, S_2 and S_3 be three n-bit S-Boxes and F be the 2n-bit function defined by the corresponding 3-round Feistel network. Then, $\delta(F) \geq 2\delta(S_2)$. Moreover, if S_2 is not a permutation, $\delta(F) \geq 2^{n+1}$. If $n = 4$, F satisfies $\delta(F) \geq 8$. If equality holds, then $\mathcal{L}(F) \geq 64$.*

3 S-Boxes Obtained from 3 Rounds of MISTY or Feistel

3.1 Our Results

In this paper, we generalize the bounds of Li and Wang [28] on Feistel structures, and derive bounds for MISTY structures. The results are very similar for the two structures, but for a MISTY structure, optimal results are only achieved with non-invertible inner functions. Therefore, our work shows that Feistel structures allow better results than MISTY structures for the design of invertible 8-bit S-Boxes.

More precisely, we introduce two new S-Box properties $\delta_{\min}$ and $\mathcal{L}_{\min}$ in order to derive our bounds: $\mathcal{L}_{\min}$ is the smallest linearity we can have for a non-trivial component of the S-Box. Similarly, $\delta_{\min}$ is the smallest value we can have for the maximum $\max_b \delta(a, b)$ within a row in the difference table. In particular, for any 4-bit function S, $\delta_{\min}(S) \geq 2$ and $\mathcal{L}_{\min}(S) \geq 4$. Moreover, if S is a 4-bit permutation, then $\delta_{\min}(S) \geq 4$ and $\mathcal{L}_{\min}(S) \geq 8$.

We first present the general lower bounds we obtain on the differential uniformity and linearity of 3 rounds of a Feistel and of a MISTY construction.

1. For a Feistel network with inner S-Boxes S_1, S_2 and S_3:
   - $\delta(F) \geq \delta(S_2) \max(\delta_{\min}(S_1), \delta_{\min}(S_3))$
   - if S_2 is not a permutation, $\delta(F) \geq 2^{n+1}$.
   - if S_2 is a permutation, $\delta(F) \geq \max\limits_{i \neq 2, j \neq i, 2}\left(\delta(S_i)\delta_{\min}(S_j),\ \delta(S_i)\delta_{\min}(S_2^{-1})\right)$.
   - $\mathcal{L}(F) \geq \mathcal{L}(S_2) \max(\mathcal{L}_{\min}(S_1),\ \mathcal{L}_{\min}(S_3))$
   - if S_2 is a permutation, $\mathcal{L}(F) \geq \max\limits_{i \neq 2, j \neq i, 2}\left(\mathcal{L}(S_i)\mathcal{L}_{\min}(S_j),\ \mathcal{L}(S_i)\mathcal{L}_{\min}(S_2^{-1})\right)$.
2. For a MISTY network with inner S-Boxes S_1, S_2 and S_3:
   - $\delta(F) \geq \delta(S_1) \max(\delta_{\min}(S_2), \delta_{\min}(S_3))$
   - if S_1 is not a permutation, $\delta(F) \geq 2^{n+1}$.
   - if S_1 is a permutation, $\delta(F) \geq \max\limits_{i \neq 1, j \neq 1, i}\left(\delta(S_i)\delta_{\min}(S_j),\ \delta(S_i)\delta_{\min}(S_1^{-1})\right)$;
   - $\mathcal{L}(F) \geq \max\left(\mathcal{L}(S_1)\mathcal{L}_{\min}(S_2), \mathcal{L}(S_2)\mathcal{L}_{\min}(S_1),\ \mathcal{L}(S_3)\mathcal{L}_{\min}(S_1)\right)$;
   - if S_3 is a permutation, $\mathcal{L}(F) \geq \mathcal{L}(S_1)\mathcal{L}_{\min}(S_3^{-1})$.
   - if S_1 is a permutation, $\mathcal{L}(F) \geq \mathcal{L}(S_3)\mathcal{L}_{\min}(S_2)$.
   - if S_1 and S_3 are permutations, $\mathcal{L}(F) \geq \mathcal{L}(S_2)\mathcal{L}_{\min}(S_3^{-1})$.

If $n = 4$ this yields for both constructions:

$$\delta(F) \geq 8 \text{ and } \mathcal{L}(F) \geq 48.$$

Moreover, $\mathcal{L}(F) \geq 64$ unless $\delta(F) \geq 32$.

For the MISTY construction with $n = 4$, if F is a permutation, we obtain tighter bounds: $\delta(F) \geq 16$ and $\mathcal{L}(F) \geq 64$. This implies that the Feistel construction is more appropriate for constructing 8-bit permutations. We will also show that all these bounds for $n = 4$ are tight. We now detail the results in the case of the MISTY construction, while the results for the Feistel construction are presented in the full version [15].

3.2 Differential Uniformity of 3 Rounds of MISTY

Our lower bound on the differential uniformity of the 3-round MISTY relies on the evaluation of the number of solutions of some differentials for which the input difference of one of the 3 S-Boxes is canceled (see Fig. 3 in [15]).

Proposition 1. *Let S_1, S_2 and S_3 be three n-bit S-Boxes and F be the $2n$-bit function defined by the corresponding 3-round MISTY network. Then, for all a, b and c in $\mathbb{F}_2^n$, we have:*

(i)
$$\delta_F(0\|a, b\|c) = \delta_{S_1}(a, c) \times \delta_{S_3}(c, b \oplus c);$$

(ii) *If S_1 is bijective,*

$$\delta_F(a\|0, b\|c) = \delta_{S_2}(a, a \oplus c) \times \delta_{S_3}(a, b \oplus c);$$

(iii) $\delta_{S_1}(a, b) \times \delta_{S_2}(b, c) \leq \delta_F(b\|a, c\|c) \leq \sum_{d \in \mathbb{F}_2^n} \delta_{S_1}(a, b \oplus d) \times \delta_{S_2}(b, c \oplus d) \times \gamma_{S_3}(d)$

where $\gamma_{S_3}(d)$ is 0 if $\delta_{S_3}(d, 0) = 0$ and 1 otherwise. Most notably, if S_3 is bijective,
$$\delta_F(b\|a, c\|c) = \delta_{S_1}(a, b) \times \delta_{S_2}(b, c).$$

Proof. Let x be the input of the MISTY network, and let x_L and x_R be its left and right parts respectively.

(i) $x = (x_L, x_R)$ satisfies $F(x_L\|x_R) \oplus F(x_L\|(x_R \oplus a)) = b\|c$ if and only if

$$\begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L) = b \oplus c, \\ S_2(x_L) \oplus S_1(x_R) \oplus x_L \oplus S_2(x_L) \oplus S_1(x_R \oplus a) \oplus x_L = c \end{cases}$$
$$\Leftrightarrow \begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L) = b \oplus c, \\ S_1(x_R) \oplus S_1(x_R \oplus a) = c \end{cases}$$

or equivalently

$$x_R \in D_{S_1}(a \to c) \text{ and } x_L \in S_1(x_R) \oplus D_{S_3}(c \to b \oplus c).$$

Hence, we deduce that there are exactly $\delta_{S_1}(a, c)$ values of x_R, and for each of those, $\delta_{S_3}(c, b \oplus c)$ values of x_L, such that x verifies the differential.

(ii) $x = (x_L, x_R)$ satisfies $F(x_L\|x_R) \oplus F((x_L \oplus a)\|x_R) = b\|c$ if and only if

$$\begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R) \oplus x_L \oplus a) = b \oplus c, \\ S_2(x_L) \oplus S_1(x_R) \oplus x_L \oplus S_2(x_L \oplus a) \oplus S_1(x_R) \oplus x_L \oplus a = c \end{cases}$$

$$\Leftrightarrow \begin{cases} S_1(x_R) \oplus x_L \in D_{S_3}(a \to b \oplus c), \\ S_2(x_L) \oplus S_2(x_L \oplus a) = a \oplus c \end{cases}$$

or equivalently,

$$x_L \in D_{S_2}(a \to a \oplus c) \text{ and } S_1(x_R) \in x_L \oplus D_{S_3}(a \to b \oplus c).$$

If S_1 is invertible, for any fixed x_L, each one of the $\delta_{S_3}(a, b \oplus c)$ values defined by the second condition determines a unique value of x_R. Therefore, the number of (x_L, x_R) satisfying the differential is exactly $\delta_{S_2}(a, a \oplus c) \times \delta_{S_3}(a, b \oplus c)$.

(iii) (x_L, x_R) satisfies $F(x_L\|x_R) \oplus F((x_L \oplus b)\|(x_R \oplus a)) = c\|c$ if and only if

$$\begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L \oplus b) = 0, \\ S_2(x_L) \oplus S_1(x_R) \oplus x_L \oplus S_2(x_L \oplus b) \oplus S_1(x_R \oplus a) \oplus x_L \oplus b = c \end{cases}$$

$$\Leftrightarrow \begin{cases} S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L \oplus b) = 0, \\ S_2(x_L) \oplus S_1(x_R) \oplus S_2(x_L \oplus b) \oplus S_1(x_R \oplus a) = b \oplus c \end{cases}$$

This equivalently means that there exists some $d \in \mathbb{F}_2^n$ such that

$$\begin{cases} x_R \in D_{S_1}(a \to b \oplus d), \ x_L \in D_{S_2}(b \to c \oplus d), \\ S_3(S_1(x_R) \oplus x_L) \oplus S_3(S_1(x_R \oplus a) \oplus x_L \oplus b) = 0, \end{cases}$$

i.e.,

$$x_R \in D_{S_1}(a \to b \oplus d), \ x_L \in D_{S_2}(b \to c \oplus d) \text{ and } S_1(x_R) \oplus x_L \in D_{S_3}(d \to 0).$$

Then, for any fixed $d \in \mathbb{F}_2^n$ such that $\delta_{S_3}(d, 0) = 0$, no pair (x_L, x_R) satisfies the third condition. If $\delta_{S_3}(d, 0) > 0$, then some of the values (x_L, x_R) defined by the first two conditions may also satisfy the third one, and if $d = 0$, the third condition is always satisfied. It then follows that

$$\delta_{S_1}(a, b) \times \delta_{S_2}(b, c) \leq \delta_F(b\|a, c\|c) \leq \sum_{d \in \mathbb{F}_2^n} \delta_{S_1}(a, b \oplus d) \times \delta_{S_2}(b, c \oplus d) \times \gamma_{S_3}(d)$$

where $\gamma_{S_3}(d)$ is 0 if $\delta_{S_3}(d, 0) = 0$ and 1 otherwise. Moreover, if S_3 is bijective, $\delta_{S_3}(d, 0) > 0$ if and only if $d = 0$, implying that the two previous bounds are equal, i.e.,

$$\delta_F(b\|a, c\|c) = \delta_{S_1}(a, b) \times \delta_{S_2}(b, c).$$

$\square$

These three particular types of differentials provide us with the following lower bound on the differential uniformity of any 3-round MISTY network.

Theorem 3. *Let S_1, S_2 and S_3 be three n-bit S-Boxes and let F be the 2n-bit function defined by the corresponding 3-round MISTY network. Then,*

$$\delta(F) \geq \delta(S_1) \max\left(\delta_{\min}(S_2), \delta_{\min}(S_3)\right) \text{ where } \delta_{\min}(S) = \min_{a \neq 0} \max_b \delta_S(a,b).$$

Moreover,

– *if S_1 is a permutation,*

$$\delta(F) \geq \max_{i \neq 1, j \neq 1, i} \max\left(\delta(S_i)\delta_{\min}(S_j), \; \delta(S_i)\delta_{\min}(S_1^{-1})\right),$$

– *if S_1 is not a permutation, $\delta(F) \geq 2^{n+1}$.*

Proof. The result is a direct consequence of Proposition 1. We here derive the bounds from the first item in Proposition 1; the other cases can be similarly deduced from the other two items. Let us first consider a pair of differences (α, β) which achieves the differential uniformity of S_1, i.e., $\delta(S_1) = \delta_{S_1}(\alpha, \beta)$. Then, we choose $a = \alpha$ and $c = \beta$, and get that, for any $b \in \mathbb{F}_2^n$,

$$\delta_F(0\|\alpha, b\|\beta) = \delta(S_1) \times \delta_{S_3}(\beta, \beta \oplus b).$$

Then, we can choose for b the value which maximizes $\delta_{S_3}(\beta, \beta \oplus b)$. This value is always greater than or equal to $\delta_{\min}(S_3)$. Similarly, we can now consider a pair of differences (α, β) which achieves the differential uniformity of S_3, i.e., $\delta(S_3) = \delta_{S_3}(\alpha, \beta)$. In this case, we choose $c = \alpha$ and $b = \alpha \oplus \beta$, and get that, for any $a \in \mathbb{F}_2^n$,

$$\delta_F(0\|a, (\alpha \oplus \beta)\|\alpha) = \delta_{S_1}(a, \alpha) \times \delta(S_3).$$

We then choose for a the value which maximizes $\delta_{S_1}(a, \alpha)$ which is always greater than or equal to $\delta_{\min}(S_1^{-1})$ when S_1 is a permutation.

Let us now assume that S_1 is not bijective. This means that there exists some nonzero $a \in \mathbb{F}_2^n$ such that $\delta_{S_1}(a, 0) \geq 2$. Then, we deduce from the first item in Proposition 1, with $b = c = 0$, that $F(x_L\|x_R) \oplus F_K(x_L \oplus a\|x_R) = (0, 0)$ has $\delta_{S_1}(a, 0) \times \delta_{S_3}(0, 0) \geq 2 \times 2^n = 2^{n+1}$ solutions in $\mathbb{F}_2^{2n}$.   □

3.3   Linearity of 3 Rounds of MISTY

The lower bound on the linearity of a three-round MISTY structure can be derived in a similar way. The proofs of the following results are given in [15].

Proposition 2. *Let S_1, S_2 and S_3 be three n-bit S-Boxes and F the 2n-bit function defined by the corresponding 3-round MISTY network. Then, for all a, b and c in $\mathbb{F}_2^n$, we have:*

(i) $\lambda_F(a\|b, 0\|c) = \lambda_{S_1}(b, c)\lambda_{S_2}(a \oplus c, c)$
(ii) $\lambda_F(a\|b, c\|c) = \lambda_{S_1}(b, a)\lambda_{S_3}(a, c)$
(iii) *If S_1 is bijective,* $\lambda_F(a\|0, b\|c) = \lambda_{S_2}(a, b \oplus c)\lambda_{S_3}(b \oplus c, b)$

As in the differential case, the previous three linear approximations provide us with a lower bound on the linearity of any 3-round MISTY network. This bound involves both the linearity of the constituent S-Boxes and another quantity denoted by $\mathcal{L}_{\min}$ computed from the table of linear biases as follows.

Definition 4 ($\mathcal{L}_{\min}$). *Let F be an n-bit S-Box. We define*

$$\mathcal{L}_{\min}(F) = \min_{b \in \mathbb{F}_2^n, b \neq 0} \max_{a \in \mathbb{F}_2^n} |\lambda_F(a, b)|.$$

Most notably, $\mathcal{L}_{\min}(F) \geq 2^{\frac{n}{2}}$ and this bound is not tight when F is bijective.

Proof. By definition, $\mathcal{L}_{\min}(F)$ is the smallest linearity achieved by a component $F_b : x \mapsto b \cdot F(x)$ of F, when b varies in $\mathbb{F}_2^n \backslash \{0\}$. Since any F_b is an n-variable Boolean function, its linearity is at least $2^{\frac{n}{2}}$ with equality if and only if F_b is bent [41]. Since bent functions are not balanced, none of the components of a permutation is bent, implying that $\mathcal{L}_{\min}(F) > 2^{\frac{n}{2}}$ when F is a permutation. $\square$

We then derive the following lower bound on the linearity of any 3-round MISTY network. The proof is given in [15].

Theorem 4. *Let S_1, S_2 and S_3 be three n-bit S-Boxes and let F be the $2n$-bit function defined by the corresponding 3-round MISTY network. Then,*

$$\mathcal{L}(F) \geq \max \left(\mathcal{L}(S_1)\mathcal{L}_{\min}(S_2), \ \mathcal{L}(S_2)\mathcal{L}_{\min}(S_1), \ \mathcal{L}(S_3)\mathcal{L}_{\min}(S_1) \right).$$

Moreover, if S_1 is a permutation, $\mathcal{L}(F) \geq \mathcal{L}(S_3)\mathcal{L}_{\min}(S_2)$; if S_3 is a permutation, $\mathcal{L}(F) \geq \mathcal{L}(S_1)\mathcal{L}_{\min}(S_3^{-1})$, and if both S_1 and S_3 are permutations, then $\mathcal{L}(F) \geq \mathcal{L}(S_2)\mathcal{L}_{\min}(S_3^{-1})$.

4 Application to 8-Bit S-Boxes

In this section, we investigate the cryptographic properties of 8-bit S-Boxes corresponding to a 3-round MISTY structure with 4-bit inner S-Boxes, with a particular focus on the case where the three inner S-Boxes are bijective, since it corresponds to the case where the resulting function is a permutation.

4.1 Differential Uniformity

The following bound on the differential uniformity of any 3-round MISTY network over $\mathbb{F}_2^8$ is a direct consequence of Theorem 3.

Corollary 1. *Any 8-bit function F corresponding to a 3-round MISTY network satisfies $\delta(F) \geq 8$.*

Proof. The bound clearly holds when S_1 is not bijective, since we known from Theorem 3 that $\delta(F) \geq 32$ in this case. If S_1 is bijective, then $\delta(S_1) \geq 4$ since APN permutations over $\mathbb{F}_2^4$ do not exist, as proved in [24, Theorem 2.3]. Obviously, any 4-bit S-Box S satisfies $\delta_{\min}(S) \geq 2$, implying that

$$\delta(F) \geq \delta(S_1)\delta_{\min}(S_2) \geq 8.$$

$\square$

Besides this general result, we can provide some necessary conditions on the constituent S-Boxes to achieve the previous lower bound.

Theorem 5. *Let S_1, S_2 and S_3 be three 4-bit S-Boxes and let F be the 8-bit function defined by the corresponding 3-round MISTY network. Then, $\delta(F) = 8$ implies that S_1 is a permutation with $\delta(S_1) = 4$ and S_2 and S_3 are two APN functions. Otherwise, $\delta(F) \geq 12$.*

Proof. Since $\delta(F) \geq 32$ when S_1 is not bijective, we only need to focus on the case where S_1 is a permutation. If any of the constituent S-Boxes S_i has differential uniformity strictly greater than 4, i.e., $\delta(S_i) \geq 6$, we deduce from Theorem 3 that $\delta(F) \geq \delta(S_i)\delta_{\min}(S_j) \geq 12$. Therefore, $\delta(F) = 8$ can be achieved only if $\delta(S_1) = 4$, $\delta(S_2) \leq 4$, and $\delta(S_3) \leq 4$. The fact that $\delta(F) \geq 16$ when at least one of the S-Boxes S_2 or S_3 has differential uniformity 4 is proved in Lemma 1 in the full version [15]. □

We can then prove that the lower bound in Corollary 1 is tight by exhibiting three 4-bit S-Boxes satisfying the previous conditions which lead to a 3-round MISTY network with differential uniformity 8.

Example 2. The following 4-bit S-Boxes yield an 8-bit S-Box with differential uniformity 8 and linearity 64 when used in a MISTY structure:

$$S_1 = [4, 0, 1, \mathsf{f}, 2, \mathsf{b}, 6, 7, 3, 9, \mathsf{a}, 5, \mathsf{c}, \mathsf{d}, \mathsf{e}, 8]$$
$$S_2 = [0, 0, 0, 1, 0, \mathsf{a}, 8, 3, 0, 8, 2, \mathsf{b}, 4, 6, \mathsf{e}, \mathsf{d}]$$
$$S_3 = [0, 7, \mathsf{b}, \mathsf{d}, 4, 1, \mathsf{b}, \mathsf{f}, 1, 2, \mathsf{c}, \mathsf{e}, \mathsf{d}, \mathsf{c}, 5, 5]$$

With Bijective Inner S-Boxes. We now focus on the case where the three inner S-Boxes are permutations since this guarantees that the resulting MISTY network is a permutation. We have proved that, in this case, the lowest possible differential uniformity we can obtain is 12. Here, we refine this result and show that the differential uniformity cannot be lower than 16. This improved bound exploits the following lemma on the difference tables of 4-bit permutations.

Lemma 1. *Let S_1, S_2 and S_3 be 4-bit permutations. Then, there exists a nonzero difference $\gamma \in \mathbb{F}_2^4 \backslash \{0\}$ such that at least one of the following statements holds:*

- *The difference table of S_1 has at least one value greater than or equal to 4 in Column γ and the difference table of S_2 has at least one value greater than or equal to 4 in Row γ;*
- *The difference table of S_1 has at least one value greater than or equal to 4 in Column γ and the difference table of S_3 has at least one value greater than or equal to 4 in Row γ,*
- *The difference table of S_2 has at least one value greater than or equal to 4 in Row γ and the difference table of S_3 has at least one value greater than or equal to 4 in Row γ.*

Proof. This result relies on an exhaustive search over the equivalence classes defined by composition on the left and on the right by an affine transformation, exactly as in the classification of optimal 4-bit S-Boxes in [18,27]. There are 302 equivalence classes for 4-bit permutations. From each of the classes we picked a representative, and checked that its difference table has at least six rows defined by some nonzero input difference a which contain a value greater than or equal to 4. Let $\mathcal{R}(S)$ denote the corresponding set (of size at least six):

$$\mathcal{R}(S) = \left\{ a \in \mathbb{F}_2^4 \backslash \{0\} \ : \exists b \in \mathbb{F}_2^4 \backslash \{0\}, \delta_S(a,b) \geq 4 \right\}.$$

Therefore, if there exists no difference $\gamma \in \mathbb{F}_2^4 \backslash \{0\}$ satisfying one of the three statements in the lemma, then this would mean that the three sets $\mathcal{R}(S_2)$, $\mathcal{R}(S_3)$ and $\mathcal{R}(S_1^{-1})$ are disjoint. In other words, we could find 18 distinct values amongst the 15 nonzero elements in $\mathbb{F}_2^4$, which is impossible. □

We then deduce the following refined lower bound on the differential uniformity of a 3-round MISTY network over $\mathbb{F}_2^8$ with inner permutations.

Theorem 6. *Let S_1, S_2 and S_3 be three 4-bit permutations and let F be the 8-bit function defined by the corresponding 3-round MISTY network. Then, $\delta(F) \geq 16$.*

Proof. The result is a direct consequence of Proposition 1 combined with the previous lemma. Indeed, Lemma 1 guarantees the existence of a, b and c such that at least one of the three following properties holds:

– $\delta_{S_1}(a,c) \geq 4$ and $\delta_{S_3}(c, b \oplus c) \geq 4$,
– $\delta_{S_2}(a, a \oplus c) \geq 4$ and $\delta_{S_3}(a, b \oplus c) \geq 4$,
– $\delta_{S_1}(a,b) \geq 4$ and $\delta_{S_2}(b,c) \geq 4$.

In each of these three situations, Proposition 1 exhibits a differential (α, β) for F with $\delta_F(\alpha, \beta) = 16$. □

4.2   Linearity

In order to apply Theorem 4 to the case of 8-bit MISTY network, we need to estimate the best linearity (and $\mathcal{L}_{\min}$) for 4-bit S-Boxes. It is well-known that the lowest linearity for a 4-bit permutation is 8. But, this result still holds if the S-Box is not bijective.

Lemma 2. *Any 4-bit S-Box S satisfies $\mathcal{L}(S) \geq 8$.*

Proof. Assume that there exists some S from $\mathbb{F}_2^4$ into $\mathbb{F}_2^4$ with $\mathcal{L}(S) < 8$, i.e., with $\mathcal{L}(S) \leq 6$. Then, all nonzero Boolean components of S, $S_c : x \mapsto c \cdot S(x)$ with $c \neq 0$, satisfy $\mathcal{L}(S_c) \leq 6$. From the classification of all Boolean functions of at most 5 variables by Berlekamp and Welch [6], we deduce that any S_c, $c \neq 0$, is affine equivalent either to $x_1 x_2 x_3 x_4 + x_1 x_2 + x_3 x_4$ or to $x_1 x_2 + x_3 x_4$, because these are the only classes of Boolean functions with linearity at most 6. Let L_1

(resp. L_2) denote the set of all nonzero $c \in \mathbb{F}_2^4$ such that S_c belongs to the first (resp. second) class. Since the degree is invariant under affine transformations, L_1 (resp. L_2) corresponds to the components with degree 4 (resp. with degree at most 2). The sum of two components of degree at most 2 has degree at most 2, implying that $L_2 \cup \{0\}$ is a linear subspace V of $\mathbb{F}_2^4$. It follows that the projection of S on V can be seen as a function from $\mathbb{F}_2^4$ into $\mathbb{F}_2^{\dim V}$ with linearity 4, i.e., a bent function. It has been shown by Nyberg [35] that, if a function F from $\mathbb{F}_2^n$ into $\mathbb{F}_2^m$ is bent, then $m \leq n/2$. Therefore, $\dim V \leq 2$. But, the sum of any two components S_c of degree 4 cannot have degree 4 since there is a single monomial of degree 4 of 4 variables. We deduce that, if L_1 contains t words of weight 1 (i.e., if S has t coordinates with linearity 6), then

$$\#L_2 \geq \binom{t}{2} + 2^{4-t} - 1 > 3,$$

for all $0 \leq t \leq 4$, a contradiction.   □

Combined with the previous lemma and with Definition 4, Theorem 4 provides the following lower bound on the linearity of a 3-round MISTY network over $\mathbb{F}_2^8$.

Corollary 2. *Any 8-bit function F corresponding to a 3-round MISTY network satisfies $\mathcal{L}(F) \geq 32$.*

This bound is of marginal interest since, up to our best knowledge, $\mathcal{L}(S) = 32$ is the lowest known linearity for an 8-bit S-Box. But, once again, the previous lower bound can be improved when focusing on permutations. Indeed, we can exploit that $\mathcal{L}_{\min}(S) \geq 8$ for any 4-bit permutation:

Lemma 3. *For any 4-bit permutation S, the table of linear biases of S has at least one value greater than or equal to 8 on every row and column.*

Proof. This result is obtained by an exhaustive search over all affine equivalence classes. The 302 representatives have been examined, and we could check the result for each of them.   □

Using that any 4-bit permutation S satisfies $\mathcal{L}(S) \geq 8$ and $\mathcal{L}_{\min}(S) \geq 8$, we directly deduce from Theorem 4 the following improved lower bound.

Proposition 3. *Let S_1, S_2 and S_3 be three 4-bit S-Boxes and let F be the 8-bit function defined by the corresponding 3-round MISTY network. If any of the three inner S-Boxes is a permutation, then $\mathcal{L}(F) \geq 64$. Most notably, if $\mathcal{L}(F) < 64$, then $\delta(F) \geq 32$.*

The last statement in the previous theorem is deduced from the first item in Theorem 3. While it shows that 3-round MISTY with $\mathcal{L}(F) < 64$ would be of little interest, we show in [15] the following theorem proving that their linearity is at least 48.

Theorem 7. *Let S_1, S_2 and S_3 be three 4-bit S-Boxes and let F be the 8-bit function defined by the corresponding 3-round MISTY network. Then $\mathcal{L}(F) \geq 48$.*

We conjecture that any MISTY network with 4-bit inner functions actually satisfies $\mathcal{L}(F) \geq 64$, but it seems hard to prove without a full classification of the 4-bit functions.

5   Constructions

We now use the previous results to design concrete 8-bit invertible S-Boxes optimized for lightweight implementations. We use Feistel and MISTY networks, and select 4-bit S-Boxes S_i's with a low-cost implementation that provide good cryptographic properties of the resulting 8-bit S-Box. Such S-Boxes have been considered as good candidates for many lightweight constructions (*e.g.* the LS-designs [23]), but their respective merits and their cryptographic properties remained open.

We focus on implementing functions with a low gate count for hardware implementations, and a low instruction count for bit-sliced implementations (for table-based implementations, the table size is independent of the concrete S-Boxes). Moreover, we focus on implementations with a small number of non-linear gates, because non-linear gates are much harder to implement than linear gates in some dedicated settings such as masking [40], multi-party computation, or homomorphic encryption [1]. Bit-slicing can be used as an implementation technique to take advantage of some platform characteristics (for instance, it yields the fastest known implementation of AES on some Intel processors [25]), but it can also be a design criterion. Indeed, using a bit-sliced S-Box allows compact implementations without tables, and good performances both in software and hardware. In addition, S-Boxes implemented in this way are easier to protect against side-channel attacks with masking. Therefore, this approach is used by many lightweight designs such as SERPENT [9], NOEKEON [16], KECCAK [8], ROBIN and FANTOMAS [23], PRIDE [2], PRØST [26], or ASCON [19]. This makes the construction of S-Boxes with a low gate count particularly relevant for lightweight cryptography.

Following the previous sections, the best results we can achieve for an 8-bit invertible S-Box are:

With a MISTY network: $\delta(F) = 16$ and $\mathcal{L}(F) = 64$.
With a Feistel network: $\delta(F) = 8$ and $\mathcal{L}(F) = 64$.

We can provide some examples fulfilling these bounds: Example 2 is optimal for the MISTY construction, and an example for the Feistel construction is now exhibited. It is worth noticing that these results explain the compared properties of the S-Boxes obtained by the simulations reported in [23]. Since Feistel networks can reach a better security, we will mostly consider this construction. In this case, the optimal differential uniformity can be reached only if S_1, S_3 are APN, and S_2 is a permutation with $\delta(S_2) = 4$, as proved in Th. 9 in the full

version [15]. Note that, in some other contexts, the MISTY construction presents some advantages since it offers better performance in terms of throughput and latency because the first two S-Boxes can be evaluated in parallel.

### 5.1   Feistel Network with Low Gate Count and Instruction Count

Rather than choosing S-Boxes S_1, S_2 and S_3 with good properties first, and then searching for an efficient implementation of these S-Boxes (as in [13,39] for instance), we take the opposite approach, following Ullrich *et al.* [45]. We build gate descriptions of S-Boxes, and we test their cryptographic properties until we find a good candidate. Indeed, we do not have to specify in advance the 4-bit S-Boxes S_1, S_2, S_3. Instead, we look for a good implementation of a permutation with $\delta(S) = 4$ for S_2, a good implementation of an APN function for S_1 and S_3, and we test the properties of the resulting Feistel structure. With good probability, this results in a Feistel network F with $\delta(F) = 8$ and $\mathcal{L}(F) = 64$.

Following Ullrich *et al.*, we run a search oriented towards bit-sliced implementations. We consider sequences of software instructions, with instructions AND, OR, XOR, NOT, and MOV, using at most 5 registers. This directly translates to a hardware representation: the MOV instruction becomes a branch while the other instructions represent the corresponding gates. There are 85 choices of instructions at each step, but we use an equivalence relation to restrict the search. For S_2, we can directly reuse the results of [45]: they give an optimal implementation of a 4-bit permutation with $\delta(S) = 4$. For S_1 and S_3, we implemented a version of their algorithm, and searched for APN functions. We found that there is no construction of an APN function with 9 or fewer instructions. There are solutions with 10 instructions, but they have at least 6 non-linear instructions (AND, OR), which is not efficient for a masked implementation. Finally, with 11 instructions, there are constructions of APN functions with 4 non-linear instructions, 5 XOR instructions, and 2 MOV (copy) instructions. This search requires about 6000 core-hours of computation. The branching factor of our search is close to 10, while Ullrich *et al.* report a branching factor of less than 7; this is because we do not restrict the search to permutations (indeed, 4-bit APN functions are not permutations). This results in a very efficient 8-bit S-Box with good cryptographic properties, using 12 nonlinear gates, and 26 XORs. According to Theorem 9 in [15] and to the following lemma, this is the optimal number of non-linear gates.

Lemma 4. *Let S be a 4-bit permutation with $\delta(S) \leq 4$ or a 4-bit APN function. Any implementation of S requires at least 4 non-linear gates.*

Proof. If S can be implemented with 3 non-linear gates or less, then the algebraic expression of the output variables is a linear combination of the input variables, and of the 3 polynomials corresponding to the output of the 3 non-linear gates. Therefore, there exists a linear combination of the input and output variables that sums to a constant, *i.e.* $\mathcal{L}(S) = 16$. According to the classification of 4-bit

permutation in [18], any permutation with $\delta(S) = 4$ satisfies $\mathcal{L}(S) \leq 12$. Furthermore, the classification of 4-bit APN functions [14] shows that they satisfy $\mathcal{L}(S) = 8$, which proves the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$
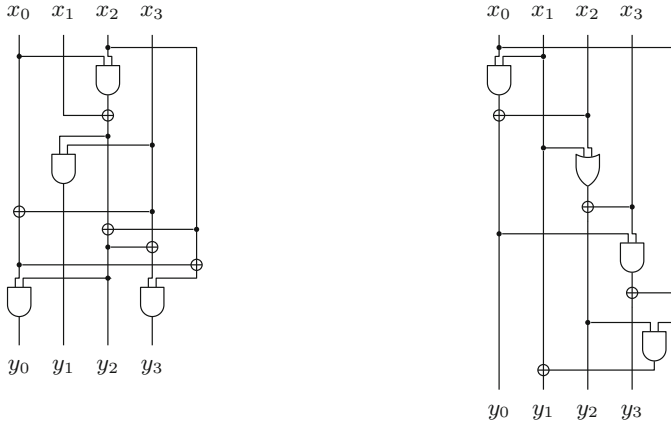
3.1. S_1, APN function with $\delta(S_1) = 2$.
$S_1 = [0, 0, 4, \mathsf{d}, \mathsf{c}, 0, 0, 5, 8, 0, 7, 6, 5, \mathsf{a}, 2, 4]$

3.2. S_2, permutation with $\delta(S_2) = 4$.
$S_2 = [0, 8, 6, \mathsf{d}, 5, \mathsf{f}, 7, \mathsf{c}, 4, \mathsf{e}, 2, 3, 9, 1, \mathsf{b}, \mathsf{a}]$

Fig. 3. Construction of a lightweight S-Box S with a three-round Feistel (S_1, S_2, S_1) satisfying $\delta(S) = 8$ and $\mathcal{L}(S) = 64$.

We give an example of such an implementation in Fig. 3, and we compare our results with previous designs in Table 1. In particular, we reach a better differential uniformity than the S-Boxes used in Robin and Fantomas [23], for a small number of extra gates. For comparing the respective merits of the S-Boxes considered in Table 1, we use the fact that, as a simple approximation, the number of rounds needed to reach a fixed security level against differential attacks is proportional to $1/\log(\delta(S)/256)$, and the implementation cost per round is proportional to the number of non-linear gates (for a bit-sliced software implementation with masking). This allows to derive a simple implementation cost metric for the S-Boxes presented in the last column, taking 1 for the AES, and considering only security against differential attacks.

5.2   Unbalanced MISTY Structure

Finally, we consider an alternative to MISTY structures as studied in this paper. Instead of dividing the input into two halves of equal size, we consider unbalanced networks. The idea is to split the 8 input bits in two unequal parts of 3 and 5 bits. Thus, the MISTY network will use only 3- and 5-bit S-Boxes. The advantage of 3- and 5-bit S-Boxes is that invertible S-Boxes with $\delta = 2$ exist, contrarily to the

case of 4-bit S-Boxes. We managed to obtain 8-bit S-Boxes S with $\delta(S) = 8$ using unbalanced MISTY networks, which is better than the lower bound $\delta(S) \geq 16$ proved for balanced MISTY networks. However, this method uses 5-bit S-Boxes, which are more complicated to implement than 4-bit S-Boxes.

Example 3. We consider a 3-round unbalanced MISTY structure, with 5-bit permutations S_1, S_3 and a 3-bit permutation S_2. After S_1 and S_3, the 3-bit x_L is xored in the 3 MSB of x_R; after S_2 the 3 MSB of the 5-bit x_L are xored into x_L. The following S-Boxes define an 8-bit S-Box with $\delta = 8$ and $\mathcal{L} = 64$:

$$S_1 = [00, 01, 02, 04, 03, 08, 0d, 10, 05, 11, 1c, 1b, 1e, 0e, 18, 0a,$$
$$06, 13, 0b, 14, 1f, 1d, 0c, 15, 12, 1a, 0f, 19, 07, 16, 17, 09]$$
$$S_2 = [2, 5, 6, 4, 0, 1, 3, 7]$$
$$S_3 = [00, 01, 02, 04, 03, 08, 10, 1c, 05, 0a, 1a, 12, 11, 14, 1f, 1d,$$
$$06, 15, 18, 0c, 16, 0f, 19, 07, 0e, 13, 0d, 17, 09, 1e, 1b, 0b]$$

This shows that generalizing our results to the unbalanced case, especially for the MISTY construction, may be of interest.

Table 1. Comparison of some 8-bit S-Boxes. δ and $\mathcal{L}$ respectively denote the differential uniformity and the linearity of the S-Box (see Sect. 2.1 for the definitions), the last column presents the relative overall implementation cost (taking 1 for the AES).

S-Box	Construction	Implementation		Properties		
		AND/OR	XOR	$\mathcal{L}$	δ	cost
AES [13]	Inversion in $\mathbb{F}_{2^8}$ + affine	32	83	32	4	1
Whirlpool [5]	Lai-Massey	36	58	56	8	1.35
CRYPTON [29]	3-round Feistel	49	12	64	8	1.83
Robin [23]	3-round Feistel	12	24	64	16	0.56
Fantomas [23]	3-round MISTY (3/5 bits)	11	25	64	16	0.51
Unnamed [23]	Whirlpool-like	16	41	64	10	0.64
New	3-round Feistel	12	26	64	8	0.45

6   Conclusion

Our results give a better understanding of the cryptographic properties of lightweight S-Boxes built from smaller S-Boxes. We give a precise description of the best security achievable with a 3-round balanced Feistel or MISTY structure for an 8-bit S-Box, and necessary conditions to reach the bound. Interestingly, the MISTY network cannot offer the same security as the Feistel network for

constructing an invertible 8-bit S-Box. Using those results, we describe an 8-bit S-Box S using only 12 non-linear gates and 26 XOR gates, with $\delta(S) = 8$ and $\mathcal{L}(S) = 64$. This is the best security that can be achieved with a 3-round Feistel or MISTY structure, and our construction uses the minimal number of non-linear gates to reach this security. This is an improvement over previous proposals, including the S-Boxes used in CRYPTON, Fantomas and Robin, but further work is required to determine whether different structures can provide better S-Boxes.

References

1. Albrecht, M.R., Rechberger, C., Schneider, T., Tiessen, T., Zohner, M.: Ciphers for MPC and FHE. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9056, pp. 430–454. Springer, Heidelberg (2015)
2. Albrecht, M.R., Driessen, B., Kavun, E.B., Leander, G., Paar, C., Yalçın, T.: Block ciphers – focus on the linear layer (feat. PRIDE). In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 57–76. Springer, Heidelberg (2014)
3. Aoki, K., Ohta, K.: Strict evaluation of the maximum average of differential probability and the maximum average of linear probability. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E80A**(1), 2–8 (1997)
4. Barreto, P.S., Rijmen, V.: The KHAZAD Legacy-Level Block Cipher. NESSIE submission
5. Barreto, P.S., Rijmen, V.: The WHIRLPOOL Hashing Function. NESSIE submission
6. Berlekamp, E.R., Welch, L.R.: Weight distributions of the cosets of the (32, 6) Reed-Muller code. IEEE Trans. Inf. Theor. **18**(1), 203–207 (1972)
7. Bernstein, D.J.: Cache-timing on AES. In: Symmetric-Key Encryption Workshop - SKEW 2005 (2005). http://cr.yp.to/antiforgery/cachetiming-20050414.pdf
8. Bertoni, G., Daemen, J., Peeters, M., Assche, G.V.: The Keccak reference, January 2011. http://keccak.noekeon.org/
9. Biham, E., Anderson, R., Knudsen, L.R.: Serpent: a new block cipher proposal. In: Vaudenay, S. (ed.) FSE 1998. LNCS, vol. 1372, pp. 222–238. Springer, Heidelberg (1998)
10. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. J. Crypt. **4**(1), 3–72 (1991)
11. Bogdanov, A.A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) CHES 2007. LNCS, vol. 4727, pp. 450–466. Springer, Heidelberg (2007)
12. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçın, T.: PRINCE – a low-latency block cipher for pervasive computing applications. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225. Springer, Heidelberg (2012)
13. Boyar, J., Peralta, R.: A new combinational logic minimization technique with applications to cryptology. In: Festa, P. (ed.) SEA 2010. LNCS, vol. 6049, pp. 178–189. Springer, Heidelberg (2010)

14. Brinkmann, M., Leander, G.: On the classification of APN functions up to dimension five. Des. Codes Crypt. **49**(1–3), 273–288 (2008)
15. Canteaut, A., Duval, S., Leurent, G.: Construction of Lightweight S-Boxes using Feistel and MISTY structures (Full Version). IACR eprint report 2015/711, Jul 2015. http://eprint.iacr.org/2015/711
16. Daemen, J., Peeters, M., Assche, G.V., Rijmen, V.: Nessie proposal. In: NOEKEON (2000)
17. De Cannière, C., Dunkelman, O., Knežević, M.: KATAN and KTANTAN — a family of small and efficient hardware-oriented block ciphers. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 272–288. Springer, Heidelberg (2009)
18. De Cannière, C.: Analysis and Design of Symmetric Encryption Algorithms. Ph.D. thesis, KU Leuven (2007)
19. Dobraunig, C., Eichlseder, M., Mendel, F., Schläffer, M.: Ascon v1. In: CAESAR Competition (2014). http://competitions.cr.yp.to/round1/asconv1.pdf
20. ETSI/SAGE: specification of the 3GPP confidentiality and integrity algorithms 128-EEA3 & 128-EIA3. document 4: design and evaluation report. Technical report (2011)
21. Gérard, B., Grosso, V., Naya-Plasencia, M., Standaert, F.-X.: Block ciphers that are easier to mask: how far can we go? In: Bertoni, G., Coron, J.-S. (eds.) CHES 2013. LNCS, vol. 8086, pp. 383–399. Springer, Heidelberg (2013)
22. Grosso, V., Leurent, G., Standaert, F.X., Varici, K., Durvaux, F., Gaspar, L., Kerckhof, S.: SCREAM & iSCREAM side-channel resistant authenticated encryption with masking. In: CAESAR Competition (2014). http://competitions.cr.yp.to/round1/screamv1.pdf
23. Grosso, V., Leurent, G., Standaert, F.-X., Varıcı, K.: LS-designs: bitslice encryption for efficient masked software implementations. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 18–37. Springer, Heidelberg (2015)
24. Hou, X.D.: Affinity of permutations of $\mathbb{F}_2^n$. Discrete Appl. Math. **154**(2), 313–325 (2006)
25. Käsper, E., Schwabe, P.: Faster and timing-attack resistant AES-GCM. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 1–17. Springer, Heidelberg (2009)
26. Kavun, E.B., Lauridsen, M.M., Leander, G., Rechberger, C., Schwabe, P., Yalçın, T.: Prøst. CAESAR Proposal (2014). http://proest.compute.dtu.dk
27. Leander, G., Poschmann, A.: On the classification of 4 bit S-Boxes. In: Carlet, C., Sunar, B. (eds.) WAIFI 2007. LNCS, vol. 4547, pp. 159–176. Springer, Heidelberg (2007)
28. Li, Y., Wang, M.: Constructing S-Boxes for lightweight cryptography with Feistel structure. In: Batina, L., Robshaw, M. (eds.) CHES 2014. LNCS, vol. 8731, pp. 127–146. Springer, Heidelberg (2014)
29. Lim, C.H.: CRYPTON: A new 128-bit block cipher. AES submission (1998)
30. Lim, C.H.: A revised version of CRYPTON - CRYPTON V1.0. In: Knudsen, L.R. (ed.) FSE 1999. LNCS, vol. 1636, pp. 31–45. Springer, Heidelberg (1999)
31. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer, Heidelberg (1994)
32. Matsui, M.: New structure of block ciphers with provable security against differential. In: Gollmann, Dieter (ed.) FSE 1996. LNCS, vol. 1039, pp. 205–218. Springer, Heidelberg (1996)
33. Matsui, M.: New block encryption algorithm MISTY. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 54–68. Springer, Heidelberg (1997)

34. National Institute of Standards and Technology: Data Encryption Standard, FIPS Publication 46-2, December 1993
35. Nyberg, K.: Perfect nonlinear S-Boxes. In: Davies, D.W. (ed.) EUROCRYPT 1991. LNCS, vol. 547, pp. 378–386. Springer, Heidelberg (1991)
36. Nyberg, K.: Differentially uniform mappings for cryptography. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 55–64. Springer, Heidelberg (1994)
37. Nyberg, K.: Linear approximation of block ciphers. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 439–444. Springer, Heidelberg (1995)
38. Nyberg, K., Knudsen, L.R.: Provable security against a differential attack. J. Crypt. **8**(1), 27–37 (1995)
39. Osvik, D.A.: Speeding up Serpent. In: AES Candidate Conference, pp. 317–329 (2000)
40. Rivain, M., Prouff, E.: Provably secure higher-order masking of AES. In: Mangard, S., Standaert, F.-X. (eds.) CHES 2010. LNCS, vol. 6225, pp. 413–427. Springer, Heidelberg (2010)
41. Rothaus, O.S.: On "bent" functions. J. Comb. Theor. Ser. A **20**(3), 300–305 (1976)
42. Shannon, C.E.: Communication theory of secrecy systems. Bell Syst. Tech. J. **28**(4), 656–715 (1949)
43. Standaert, F.-X., Piret, G., Rouvroy, G., Quisquater, J.-J., Legat, J.-D.: ICEBERG: an involutional cipher efficient for block encryption in reconfigurable hardware. In: Roy, B., Meier, W. (eds.) FSE 2004. LNCS, vol. 3017, pp. 279–299. Springer, Heidelberg (2004)
44. Suzaki, T., Minematsu, K., Morioka, S., Kobayashi, E.: Twine: a lightweight block cipher for multiple platforms. In: Knudsen, L.R., Wu, H. (eds.) SAC 2012. LNCS, vol. 7707, pp. 339–354. Springer, Heidelberg (2013)
45. Ullrich, M., De Cannière, C., Indesteege, S., Küçük, Ö., Mouha, N., Preneel, B.: Finding optimal bitsliced implementations of 4×4-bit S-Boxes. In: SKEW 2011 Symmetric Key Encryption Workshop, Copenhagen, Denmark, pp. 16–17 (2011)
46. Wegener, I.: The Complexity of Boolean Functions. Wiley-Teubner, New York (1987)
47. Wheeler, D.J., Needham, R.M.: TEA, a Tiny Encryption Algorithm. In: Preneel, B. (ed.) FSE 1994. LNCS, vol. 1008, pp. 363–366. Springer, Heidelberg (1994)
48. Wu, W., Zhang, L.: LBlock: a lightweight block cipher. In: Lopez, J., Tsudik, G. (eds.) ACNS 2011. LNCS, vol. 6715, pp. 327–344. Springer, Heidelberg (2011)