Differential Factors: Improved Attacks on SERPENT

Cihangir Tezcan[*] and Ferruh Özbudak

Department of Mathematics and Institute of Applied Mathematics
Middle East Technical University, Ankara, Turkey

Abstract. A differential attack tries to capture the round keys corresponding to the S-boxes activated by a differential. In this work, we show that for a fixed output difference of an S-box, it may not be possible to distinguish the guessed keys that have a specific difference. We introduce these differences as differential factors. Existence of differential factors can reduce the time complexity of differential attacks and as an example we show that the 10, 11, and 12-round differential-linear attacks of Dunkelman *et al.* on SERPENT can actually be performed with time complexities reduced by a factor of 4, 4, and 8, respectively.

Keywords : S-box, differential factor, SERPENT, differential-linear attack

1 Introduction

Confusion layer of cryptographic algorithms mostly consists of substitution boxes (S-boxes) and in order to provide better security against known attacks, S-boxes are selected depending on their cryptographic properties: Low non-linear and differential uniformity [30] provide resistance against linear and differential cryptanalysis, respectively; high algebraic degree and branch number provides resistance against algebraic [14] and cube [16] attacks; lack of undisturbed bits [36] provides resistance against truncated [20], impossible [2], and improbable [11] differential cryptanalysis. Moreover, recently it was shown in [9] that additive shares can be used in threshold implementations to provide resistance against side-channel attacks like differential power analysis [21] and the number of shares affects the performance.

In this work, we show that a fixed S-box output difference μ may remain invariant when the possible input pairs are XORed with some λ. We define such λ as a differential factor for the output difference μ and we show that such a property of an S-box can significantly reduce the attacked key space in a differential attack which results in an attack with reduced time complexity. Our analysis of S-boxes that are used in cryptographic algorithms show that differential factors are observed mostly in small S-boxes. We observed that 73% of all possible

bijective 3×3 S-boxes contain differential factors. Moreover, 4×4 S-boxes of DES [28], GOST [38], LBLOCK [40], LED [18], LUFFA [12], NOEKEON [15], *Piccolo* [35], PRESENT [11], RECTANGLE [41], SARMAL [39], SERPENT [1], SPONGENT [10] and Twofish [33] contain differential factors.

Lightweight cryptography has become very vital with the emerging needs in sensitive applications like RFID (Radio-frequency identification) systems and sensor networks. For these types of special purposes, there is a strong demand in designing secure lightweight cryptographic modules. Since most of such lightweight algorithms have a hardware oriented design, they use small S-boxes. Thus, differential factors pose a threat to lightweight block ciphers. We indicate the importance of this new S-box criteria on cryptanalysis by reducing the time complexities of the 10, 11, and 12-round differential-linear attacks of Dunkelman *et al.* on SERPENT by a factor of 4, 4, and 8, respectively. By changing the differential, we further modify these attacks to marginally reduce the data complexity. We compare our improved attacks on SERPENT with the previous ones in Table 1.

2   S-box Evaluation

S-boxes are commonly used as non-linear components for symmetric cryptosystems and hash functions. Properties of S-boxes provide resistance against many cryptanalytic techniques.

Differential Uniformity

Definition 1. *For a mapping $S : F_2^n \to \mathbb{F}_2^m$, and all $\Delta_i \in \mathbb{F}_2^n$ and $\Delta_o \in \mathbb{F}_2^m$, let t be the number of elements x that satisfy $S(x \oplus \Delta_i) = S(x) \oplus \Delta_o$. Then $t/|2^n|$ is the differential probability of the characteristic $S(\Delta_i \to \Delta_o)$. The table that lists all t values for every $i, o \in X$ is called the* Difference Distribution Table (DDT).

The maximum value in a DDT, excluding the zero difference case, is called differential uniformity. S-box designers aim to minimize differential uniformity since differential cryptanalysis [8] uses characteristics with high differential probability.

Non-linear Uniformity

Definition 2. *For a mapping $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$, and all $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2^m$, let the numbers $L_f(a, b)$ be defined as*

$$L_f(a, b) := |\#\{x \in F_2^n | a \cdot x = b \cdot S(x)\} - 2^{n-1}|$$

where $a \cdot b$ denotes the parity of the bit-wise product of a and b. Then S is called non-linearly l-uniform if $L_f(a, b) \leq l$ for all a and b with $b \neq 0$.

S-box designers aim to minimize the non-linear uniformity l since linear cryptanalysis [26] uses linear approximations with high bias.

Table 1: Summary of attacks on SERPENT. Note that it is claimed in [27] that the multidimensional linear attacks of [29] may not work as claimed depending on the linear hull effect. If the claims are correct, then our use of differential factors in the attacks of [17] becomes the best attacks for this cipher.
En - Encryptions, MA - Memory Accesses, B - bytes, CP - Chosen Plaintexts, KP - Known Plaintexts.

#Rounds	Attack Type	Key Size	Data	Time	Memory	Advantage	Success	Reference
6	Meet-in-the-middle	256	512 KP	2^{247} En	2^{246} B	-	-	[22]
6	Differential	All	2^{83} CP	2^{90} En	2^{40} B	-	-	[22]
6	Differential	All	2^{71} CP	2^{103} En	2^{75} B	-	-	[22]
6	Differential	192, 256	2^{41} CP	2^{163} En	2^{45} B	124	-	[22]
7	Differential	256	2^{122} CP	2^{248} En	2^{126} B	128	-	[22]
7	Improbable	All	$2^{116.85}$ CP	$2^{117.57}$ En	2^{113} B	112	99.9%	[37]
7	Differential	All	2^{84} CP	2^{85} MA	2^{56} B	-	-	[4]
10	Rectangle	192, 256	$2^{126.3}$ CP	$2^{173.8}$ MA	$2^{131.8}$ B	80	-	[6]
10	Boomerang	192, 256	$2^{126.3}$ AC	$2^{173.8}$ MA	2^{89} B	80	-	[6]
10	Differential-Linear	All	$2^{101.2}$ CP	$2^{115.2}$ En	2^{40} B	40	84%	[17]
10	**Differential-Linear**	**All**	$2^{101.2}$ **CP**	$2^{113.2}$ **En**	2^{40} **B**	**38**	**84%**	**Sect. 4.4**
10	**Differential-Linear**	**All**	$2^{100.55}$ **CP**	$2^{116.55}$ **En**	2^{40} **B**	**42**	**84%**	**Appx. B**
11	Linear	256	2^{118} KP	2^{214} MA	2^{85} B	140	78.5%	[3]
11	Multidimensional Linear [a]	All	2^{116} KP	$2^{107.5}$ En	2^{108} B	48	78.5%	[29]
11	Multidimensional Linear [b]	All	2^{118} KP	$2^{109.5}$ En	2^{104} B	44	78.5%	[29]
11	Nonlinear	192, 256	$2^{120.36}$ KP	$2^{139.63}$ MA	$2^{133.17}$ B	118	78.5%	[27]
11	Filtered Nonlinear	192, 256	$2^{114.55}$ KP	$2^{155.76}$ MA	$2^{146.59}$ B	132	78.5%	[27]
11	Differential-Linear	192, 256	$2^{121.8}$ CP	$2^{135.7}$ MA	2^{76} B	48	84%	[17]
11	**Differential-Linear**	**192, 256**	$2^{121.8}$ **CP**	$2^{133.7}$ **MA**	2^{76} **B**	**46**	**84%**	**Sect. 4.4**
11	**Differential-Linear**	**192, 256**	$2^{121.15}$ **CP**	$2^{137.05}$ **MA**	2^{76} **B**	**50**	**84%**	**Appx. B**
12	Multidimensional Linear [c]	256	2^{116} KP	$2^{237.5}$ En	2^{125} B	174	78.5%	[29]
12	Differential-Linear	256	$2^{123.5}$ CP	$2^{249.4}$ En	$2^{128.5}$ B	160	84%	[17]
12	**Differential-Linear**	**256**	$2^{123.5}$ **CP**	$2^{246.4}$ **En**	$2^{128.5}$ **B**	**157**	**84%**	**Sect. 4.4**

[a] In [27], it is claimed that the correct data complexity of this attack is $2^{125.81}$ KP and the time complexity is $2^{101.44}$ En $+2^{114.13}$ MA.

[b] In [27], it is claimed that the correct data complexity of this attack is $2^{127.78}$ KP and the time complexity is $2^{97.41}$ En $+2^{110.10}$ MA.

[c] In [27], it is claimed that the correct data complexity of this attack is $\geq 2^{125.81}$ KP and the time complexity is $2^{229.44}$ En $+2^{242.13}$ MA.

Branch Number

Definition 3. *[32] The branch number of an $n \times n$ S-box is*

$$BN = \min_{a,b \neq a} (wt(a \oplus b) + wt(S(a) \oplus S(b))),$$

where $a, b \in X$ and $wt(a)$ is the Hamming weight of the bit vector a.

For a bijective S-box, the branch number is at least 2 and this property of S-boxes is closely related to algebraic [14] and cube attacks [16].

Number of Shares S-boxes are also studied for their security against side-channel attacks. Side-channel attacks are based on the information leakage dur-

ing the computation of the hardware implementation of a cryptographic algorithm. For instance, differential power analysis (DPA) [21] exploits the correlation between the instantaneous power consumption of a device and the intermediate results of a cryptographic algorithm. One countermeasure against side-channel attacks is threshold implementation in which a variable is split into additive shares. Bilgin *et al.* analyzed the number of shares of S-boxes by categorizing all 3×3 and 4×4 S-boxes using affine equivalence classes and investigated the cost of this kind of protection in [9].

Undisturbed Bits Recently in [36], undisturbed bits are introduced as probability 1 truncated differentials for S-boxes. A 13-round improbable differential attack on PRESENT that uses undisturbed bits is provided in [36] and it was shown that the attack reduces to 7 rounds when the S-box is replaced with a similar one that lacks undisturbed bits. Moreover, it is shown that every bijective 3×3 S-box contains undisturbed bits and a list of ciphers were provided in [36] whose 4×4 S-boxes contain undisturbed bits. S-boxes with undisturbed bits should be avoided to increase security against truncated, impossible, and improbable differential cryptanalysis.

3 Differential Factors

A differential attack on block ciphers tries to capture the round keys corresponding to the S-boxes activated by a differential. However, output difference of the S-box operation may be invariant when the round key is XORed with some specific value. Such a case would prevent the attacker from fully capturing the round key. This observation is similar to the *linear factors* of block ciphers but here we are focusing on the S-box instead of some rounds of the cipher and we focus on key differences.

Definition 4 ([13]). *A block cipher is said to have a* linear factor *if, for all plaintexts and keys, there is a fixed non-empty set of key bits whose simultaneous complementation leaves the XOR sum of a fixed non-empty set of ciphertext bits unchanged.*

In order to have a similar property for S-boxes in the concept of differential cryptanalysis, we define the differential factors as follows:

Definition 5. *Let S be a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$. For all $x, y \in \mathbb{F}_2^n$ that satisfy $S(x) \oplus S(y) = \mu$, if we also have $S(x \oplus \lambda) \oplus S(y \oplus \lambda) = \mu$, then we say that the S-box has a* differential factor λ *for the output difference μ. (i.e. μ remains invariant for λ).*

When undisturbed bits are introduced in [36], the undisturbed bits of S-boxes and their inverses are considered together because in substitution permutation networks (SPNs), the inverse of an S-box is used for decryption. For instance, a 6-round impossible differential for PRESENT is obtained in [36] by using both

undisturbed bits of its S-box and the inverse of it. In the following theorem, we prove that the number of differential factors of an S-box is the same with the number of differential factors of its inverse. Moreover, it also provides the differential factors of the inverse S-box when we know the differential factors of the S-box. Hence, there is no need to check the differential factors of the inverse of S-boxes.

Theorem 1. *If a bijective S-box S has a differential factor λ for an output difference μ, then S^{-1} has a differential factor μ for the output difference λ.*

Proof. Let us assume that S has a differential factor λ for an output difference μ. If $S^{-1}(c_1) \oplus S^{-1}(c_2) = \lambda$ for some c_1 and c_2, then we need to show that $S^{-1}(c_1 \oplus \mu) \oplus S^{-1}(c_2 \oplus \mu) = \lambda$.

Let $c_1 \oplus \mu = S(p_1)$ for some p_1, then we have $S(S^{-1}(c_1) \oplus \lambda) \oplus S(p_1 \oplus \lambda) = \mu$ since λ is a differential factor of S for μ. Thus, we have

$$\begin{aligned} S^{-1}(c_1 \oplus \mu) \oplus S^{-1}(c_2 \oplus \mu) &= S^{-1}(S(p_1)) \oplus S^{-1}(S(S^{-1}(c_1) \oplus \lambda) \oplus \mu) \\ &= p_1 \oplus S^{-1}(S(p_1 \oplus \lambda)) \\ &= p_1 \oplus p_1 \oplus \lambda \\ &= \lambda \end{aligned}$$

$\square$

Theorem 2. *If λ_1 and λ_2 are differential factors for an output difference μ, then $\lambda_1 \oplus \lambda_2$ is also a differential factor for the output difference μ. i.e. All differential factors λ_i for μ form a vector space.*

Proof. We are going to use the following change of variables: $x' = x \oplus \lambda_1$ and $y' = y \oplus \lambda_1$. For all (x, y) pairs satisfying $S(x) \oplus S(y) = \mu$, we have $S(x \oplus \lambda_1) \oplus S(y \oplus \lambda_1) = \mu$ and $S(x \oplus \lambda_2) \oplus S(y \oplus \lambda_2) = \mu$. Thus, we have

$$S(x \oplus \lambda_1 \oplus \lambda_2) \oplus S(y \oplus \lambda_1 \oplus \lambda_2) = S(x' \oplus \lambda_2) \oplus S(y' \oplus \lambda_2) = \mu$$

$\square$

In this section we used two variables x and y since they are directly linked to the input pairs in differential cryptanalysis. However, same definition and theorems can be given using a single variable for bijective S-boxes and we provide them in Appendix A.

3.1 Differential Factors and Cryptanalysis

We start by recalling the definition of advantage.

Definition 6 ([34]). *If an attack on an m-bit key gets the correct value ranked among the top r out of 2^m possible candidates, we say the attack obtained an $(m - log(r))$-bit* advantage *over exhaustive search.*

Theorem 3. *In a block cipher let an S-box S contain a differential factor λ for an output difference μ and the partial round key k is XORed with the input of S. If an input pair provides the output difference μ under a partial subkey k, then the same output difference is observed under the partial subkey $k \oplus \lambda$. Therefore, during a differential attack involving the guess of a partial subkey corresponding to the output difference μ, the advantage of the cryptanalyst is reduced by 1 bit and the time complexity of this key guess step is halved.*

Proof. In a differential attack for any key k, k and $k \oplus \lambda$ would get the same number of hits since λ is a differential factor. Hence the attacker cannot distinguish half of the guessed keys with the other half. Therefore during the key guessing step, the attacker does not need to guess half of the keys. Thus, the time complexity of this step is halved. $\square$

Corollary 1. *During a differential attack involving the guess of a partial subkey corresponding to the output difference μ of an S-box that has a vector space of differential factors of dimension r for μ, the advantage of the cryptanalyst is reduced by r bits and the time complexity of the key guess step is reduced by a factor of 2^r.*

Proof. Follows directly from Theorem 2 and Theorem 3. $\square$

3.2   Relating Differential Factors to Other Properties of S-boxes

Since we are considering non-zero μ and λ, a 3×3 S-box can contain at most $7 \cdot 7 = 49$ differential factors. In such a case, an S-box provides no security at all. In [36], it was shown that every bijective 3×3 S-box contains undisturbed bits. However, this is not the case for differential factors. Among the $8! = 40320$ different bijective 3×3 S-boxes, we observed that 10752 of them do not contain any differential factor. Moreover, 18816 of them contain 9, 9408 of them contain 25, and 1344 of them contain 49 differential factors.

We further observed that the 3×3 S-boxes that do not have any differential factor also have 6 undisturbed bits, which is the smallest number of undisturbed bits a 3×3 S-box can have. Thus, for the case of 3×3 S-boxes, it is enough to check differential factors.

In our literature search we found 102 unique 4×4 S-boxes that are used in block ciphers and hash functions and observed that 40 of them have 74 differential factors in total, without counting the differential factors of their inverses. These are the S-boxes of DES, GOST, LBLOCK, LED, LUFFA, NOEKEON, *Piccolo*, PRESENT, RECTANGLE, SARMAL, SERPENT, SPONGENT and Twofish and they are provided in Table 2.

During our analysis, we observed that the existence of differential factors for an S-box is closely related to the number of nonzero entries in the columns of the DDT table. For instance, for a differentially 4-uniform 4×4 S-box, which is the best case for S-boxes of this size, we observed the following phenomenon:

Conjecture 1. A differential 4-uniform 4×4 S-box S has a differential factor for the output difference μ if and only if the μ-th column of the DDT table of S consists of only zeros and fours.

The only 8×8 S-boxes we found with differential factors are the two S-boxes of the initial version of the CRYPTON cipher [24]. They contain 15 differential factors each and they are provided in Table 2. These S-boxes are replaced in the revised version of the CRYPTON cipher [25] and the new S-boxes do not contain any differential factors.

4 Improved Differential-Linear Attacks on SERPENT

4.1 SERPENT

SERPENT was designed by Anderson, Biham and Knudsen in 1998. It was submitted to the AES contest and came second after Rijndael. It has a block size of 128 bits and accepts any key size of length 0 to 256 bits. It is a 32-round SPN, where each round consists of key mixing, a layer of S-boxes and a linear transformation.

The 128-bit input value before round i is denoted by $\hat{B}_i$, $i \in \{0, \dots, 31\}$. Each $\hat{B}_i$ is composed of four 32-bit words X_0, X_1, X_2, X_3 where X_0 is the leftmost word.

Three round operations are specified as follows:

1. Key Mixing: At each round R_i, a 128-bit subkey K_i is XORed with the current intermediate data $\hat{B}_i$.
2. S-boxes: At each round, R_i uses a single S-box S_j, where $i \equiv j \pmod 8$ and $i \in \{0, \dots, 31\}$, 32 times in parallel. In this paper, we use the bitsliced version of SERPENT. For example, in the first round the first copy of S_0 takes the least significant bits from X_0, X_1, X_2, X_3 and returns the output to the same bits. Thus, we obtain 32 4-bit slices referred as b_i's, where $i \in \{0, \dots, 31\}$ and b_0 is the right most slice.
3. Linear Transformation: The four 32-bit words X_0, X_1, X_2, X_3 are linearly mixed by the following linear operations:

$$X_0 := X_0 \lll 13$$
$$X_2 := X_2 \lll 3$$
$$X_1 := X_1 \oplus X_0 \oplus X_2$$
$$X_3 := X_3 \oplus X_2 \oplus (X_0 \ll 3)$$
$$X_1 := X_1 \lll 1$$
$$X_3 := X_3 \lll 7$$
$$X_0 := X_0 \oplus X_1 \oplus X_3$$
$$X_2 := X_2 \oplus X_3 \oplus (X_1 \ll 7)$$
$$X_0 := X_0 \lll 5$$
$$X_2 := X_2 \lll 22$$
$$\hat{B}_{i+1} := X_0, X_1, X_2, X_3$$

where $\lll$ denotes the left rotation operation and $\ll$ denotes the left shift operation.

Table 2: Differential Factors of Cryptographic Algorithms

S-box	λ	μ		S-box	λ	μ
CRYPTON S_0, S_1	10_x	10_x		DES7 Row4	1_x	C_x
CRYPTON S_0, S_1	20_x	20_x		DES7 Row4	4_x	C_x
CRYPTON S_0, S_1	30_x	30_x		DES7 Row4	5_x	C_x
CRYPTON S_0, S_1	40_x	40_x		DES7 Row4	4_x	F_x
CRYPTON S_0, S_1	50_x	50_x		DES8 Row2	6_x	7_x
CRYPTON S_0, S_1	60_x	60_x		DES8 Row2	B_x	8_x
CRYPTON S_0, S_1	70_x	70_x		GOST S_1	5_x	3_x
CRYPTON S_0, S_1	80_x	80_x		GOST S_4	D_x	5_x
CRYPTON S_0, S_1	90_x	90_x		GOST S_6	9_x	B_x
CRYPTON S_0, S_1	$A0_x$	$A0_x$		GOST S_8	7_x	5_x
CRYPTON S_0, S_1	$B0_x$	$B0_x$		GOST S_8	E_x	6_x
CRYPTON S_0, S_1	$C0_x$	$C0_x$		LBLOCK S_0, S_8	B_x	1_x
CRYPTON S_0, S_1	$D0_x$	$D0_x$		LBLOCK S_0, S_8	3_x	4_x
CRYPTON S_0, S_1	$E0_x$	$E0_x$		LBLOCK S_1, S_6, S_7, S_9	B_x	2_x
CRYPTON S_0, S_1	$F0_x$	$F0_x$		LBLOCK S_1, S_6, S_7, S_9	3_x	4_x
DES1 Row3	F_x	2_x		LBLOCK S_2	3_x	1_x
DES1 Row3	F_x	8_x		LBLOCK S_2	B_x	2_x
DES1 Row3	F_x	A_x		LBLOCK S_3	B_x	1_x
DES2 Row1	6_x	A_x		LBLOCK S_3	3_x	8_x
DES2 Row2	2_x	7_x		LBLOCK S_4, S_5	B_x	1_x
DES2 Row2	4_x	7_x		LBLOCK S_4, S_5	3_x	2_x
DES2 Row2	6_x	7_x		LUFFA	4_x	1_x
DES2 Row3	1_x	A_x		LUFFA	2_x	2_x
DES2 Row3	6_x	A_x		NOEKEON	1_x	1_x
DES2 Row3	7_x	A_x		NOEKEON	B_x	B_x
DES3 Row3	2_x	6_x		Piccolo	1_x	2_x
DES3 Row3	8_x	6_x		Piccolo	2_x	5_x
DES3 Row3	A_x	6_x		PRESENT, LED	1_x	5_x
DES3 Row4	3_x	1_x		PRESENT, LED	F_x	F_x
DES3 Row4	3_x	6_x		RECTANGLE	2_x	4_x
DES3 Row4	3_x	7_x		RECTANGLE	E_x	C_x
DES3 Row4	3_x	8_x		SARMAL S_2	F_x	4_x
DES3 Row4	3_x	9_x		SARMAL S_2	A_x	9_x
DES3 Row4	1_x	E_x		SERPENT S_0	4_x	4_x
DES3 Row4	2_x	E_x		SERPENT S_0	D_x	F_x
DES3 Row4	3_x	E_x		SERPENT S_1	4_x	4_x
DES3 Row4	3_x	F_x		SERPENT S_1	F_x	E_x
DES5 Row4	2_x	F_x		SERPENT S_2	2_x	1_x
DES6 Row1	9_x	D_x		SERPENT S_2	4_x	D_x
DES6 Row2	B_x	4_x		SERPENT S_6	6_x	2_x
DES6 Row4	6_x	6_x		SERPENT S_6	F_x	F_x
DES7 Row2	4_x	D_x		SPONGENT	F_x	9_x
DES7 Row2	9_x	D_x		SPONGENT	1_x	F_x
DES7 Row2	D_x	D_x		Twofish q0 t1	6_x	9_x
DES7 Row4	4_x	3_x		Twofish q1 t2	5_x	B_x

32-round SERPENT cipher may be described by the following equations:

$$\hat{B}_0 := P \qquad \hat{B}_{i+1} := R_i(\hat{B}_i),\ i \in \{0, \ldots, 31\} \qquad C := \hat{B}_{32}$$

where

$$R_i(X) = LT(\hat{S}_i(X \oplus K_i)),\ i \in \{0, \ldots, 30\}$$
$$R_{31}(X) = \hat{S}_{31}(X \oplus K_{31}) \oplus K_{32}$$

and $\hat{S}_i$ is the application of the S-box $S_{(i\ (mod\ 8))}$ 32 times in parallel, and LT is the linear transformation.

The key scheduling algorithm of SERPENT takes a 256-bit key as an input. If the key is shorter, then it is padded by a single bit of 1 and the remaining part is padded by bits of 0 up to 256 bits. By using an affine recurrence, the 256-bit key is used to construct 132 *prekeys* having length of 32 bits. The S-boxes are used to produce 32-bit keywords from prekeys. The round keys are obtained by combining these keywords.

4.2 Differential-Linear Cryptanalysis

In 1994, Langford and Hellman combined differential cryptanalysis with linear cryptanalysis and introduced differential-linear cryptanalysis [23]. They suggested using a truncated differential with probability 1 and concatenating a linear approximation with bias q (i.e. probability $1/2 + q$) where the output difference of the differential should contain zero differences in the places where input bits masked in the linear approximation. This way one can construct differential-linear distinguishers and the data complexity of the distinguisher is $O(q^{-4})$ chosen plaintexts. The exact number depends on the success probability and the number of possible subkeys.

Moreover, Biham, Dunkelman and Keller showed that it is possible to construct a differential-linear distinguisher where the differential holds with probability $p < 1$ and introduced enhanced differential-linear cryptanalysis [5]. They also showed that the attack is still applicable if the XOR of the masked bits of the differential is 1. In the enhanced method, the data complexity becomes $O(p^{-2}q^{-4})$ chosen plaintexts.

4.3 Differential-Linear Attacks on SERPENT

In [7] a differential-linear attack on 11-round SERPENT-192 and SERPENT-256 is presented. The attack combines the 3-round differential

Δ : 00000000000000000000000040050000 → 0??00?000?000000000?00?0??0??0?0

that has a probability of $p = 2^{-7}$ with the 6-round linear approximation

Λ : 20060040000001001000000000000000 → 00001000000000005000010000100001

of [3] that has bias $q = 2^{-27}$.

The first attack on 10-round Serpent-128 is also presented in [7] which is obtained by removing the last round of this linear approximation. The data and time complexities of these attacks are reduced in [17] by using the following improvements:

1. Better analysis of the bias of the differential-linear approximation,
2. Better analysis of the success probability,
3. Changing the output mask.

Moreover in [17], these reduced complexities are used to extend the 11-round attack and obtain the first 12-round attack on Serpent-256. In the following section we further improve these differential-linear attacks by using the differential factors of Serpent's S-boxes S_0 and S_1.

4.4 Improved Differential-Linear Attacks using Differential Factors

The differential-linear attacks of [7, 17] start at round 1 and the 3-round differential activates 5 S-boxes in this round. Two of the output differences of these activated S-boxes are 4_x and E_x which have differential factors as shown in Table 2. The authors guess every possible 20 subkey bits corresponding to these five S-boxes but the attacker can only obtain 18-bit advantage for this subkey due to Theorem 3 and there is no need to try half of the subkeys corresponding to these two S-boxes having differential factors. Thus, the advantage of the differential-linear attacks on 10, 11, and 12 rounds of Serpent are actually 38, 46, and 158 bits instead of 40, 48, and 160 bits, respectively. And again by Theorem 3, the same attacks can be performed with time complexities reduced by a factor of 4.

Moreover, the 12-round attack of [17] adds one more round to the top of the differential which affects every S-box at round 0 except the S-boxes 2, 3, 19, and 23 and guesses the 112 bits of the subkey corresponding to these active S-boxes. However, by using the undisturbed bits of Serpent, we observed that the output difference of the S-box 8 is exactly 4_x. Since $\mu = 4_x$ also has a differential factor for S_0, the attacker's advantage reduces to 157 bits and the time complexity of the attack further reduces by a factor of 2. Table 3 summarizes this 12-round attack and highlights the differential factors and the undisturbed bits that are used to reduce the time complexity.

We also observed that by replacing the 3-round differential with a more probable one, we can perform these attacks with less data complexity and capture four more subkey bits with a time complexity increased by a factor of $2^{3.35}$. These modified attacks are provided in Appendix B.

5 Conclusion

In this paper, we introduced a new S-box evaluation criteria that we call differential factors. Differential factors are mostly observed in small S-boxes like 3×3 and 4×4 which are preferred in hardware oriented lightweight block ciphers.

Table 3: 12-round differential-linear attack of [17]. Output differences μ that contain differential factors, which are 4_x and E_x for S_1 and 4_x for S_0, are shown in bold. Undisturbed bits are shown in italic.

Input	X_0: ???? ???? 0??? 0??? ???? ???? ???? 00??
	X_1: ???? ???? 0??? 0??? ???? ???? ???? 00??
	X_2: ???? ???? 0??? 0??? ???? ???*1* ???? 00??
	X_3: ???? ???? 0??? 0??? ???? ???? ???? 00??
S_0	X_0: ??0? 00?0 0000 0?00 00?0 000**0** 00?? 00??
	X_1: ??0? ???? 00?0 0??? 0??? ???**0** 0?00 0000
	X_2: 000? 00?? 0??0 0?00 ??00 ?00**1** 0?00 0000
	X_3: ?0?? ?0?? 00?? 0??? ??0? 0??**0** ?001 0000
LT	X_0: ?000 0000 0000 0??0 0?00 ?000 0000 0000
	X_1: ?000 0000 0000 0??0 0?00 ?000 0000 0000
	X_2: ?000 0000 0000 0??0 0?00 ?000 0000 0000
	X_3: ?000 0000 0000 0*1*?0 0?00 *1*000 0000 0000
S_1	X_0: 0000 0000 0000 0100 0**000** **0**000 0000 0000
	X_1: 1000 0000 0000 0010 0**100** **0**000 0000 0000
	X_2: 0000 0000 0000 0000 0**100** **1**000 0000 0000
	X_3: 0000 0000 0000 0010 0**100** **0**000 0000 0000
LT	X_0: 0000 0000 0000 0000 0000 0000 0001 0000
	X_1: 0000 0000 0000 0000 0000 0000 0000 0000
	X_2: 0000 0000 0000 0000 0000 0000 1001 0000
	X_3: 0000 0000 0000 0000 0000 0000 0000 0000
9-Round Differential-Linear Characteristic $\Delta \circ \Lambda$	
Last Round	

We show that differential factors may reduce the attacked key space in differential cryptanalysis and its variants which results in an attack with reduced time complexity. As an example, we show that the differential factors of Serpent's S-boxes are overlooked in Dunkelman *et al.*'s differential-linear attacks on Serpent and the attacked round keys cannot be fully recovered in these attacks. We reduce the time complexities of these attacks by using the differential factors and provide the best differential-linear attacks on this cipher.

References

1. Biham, E., Anderson, R.J., Knudsen, L.R.: Serpent: A new block cipher proposal. In: Vaudenay, S. (ed.) FSE. Lecture Notes in Computer Science, vol. 1372, pp. 222–238. Springer (1998)
2. Biham, E., Biryukov, A., Shamir, A.: Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. J. Cryptology 18(4), 291–311 (2005)

3. Biham, E., Dunkelman, O., Keller, N.: Linear cryptanalysis of reduced round Serpent. In: Matsui, M. (ed.) FSE. Lecture Notes in Computer Science, vol. 2355, pp. 16–27. Springer (2001)
4. Biham, E., Dunkelman, O., Keller, N.: The rectangle attack - rectangling the Serpent. In: Pfitzmann, B. (ed.) EUROCRYPT. Lecture Notes in Computer Science, vol. 2045, pp. 340–357. Springer (2001)
5. Biham, E., Dunkelman, O., Keller, N.: Enhancing differential-linear cryptanalysis. In: Zheng [42], pp. 254–266
6. Biham, E., Dunkelman, O., Keller, N.: New results on boomerang and rectangle attacks. In: Daemen, J., Rijmen, V. (eds.) FSE. Lecture Notes in Computer Science, vol. 2365, pp. 1–16. Springer (2002)
7. Biham, E., Dunkelman, O., Keller, N.: Differential-linear cryptanalysis of Serpent. In: Johansson, T. (ed.) FSE. Lecture Notes in Computer Science, vol. 2887, pp. 9–21. Springer (2003)
8. Biham, E., Shamir, A.: Differential Cryptanalysis of DES-like Cryptosystems. J. Cryptology 4(1), 3–72 (1991)
9. Bilgin, B., Nikova, S., Nikov, V., Rijmen, V., Stütz, G.: Threshold implementations of all 3x3 and 4x4 S-boxes. In: Prouff, E., Schaumont, P. (eds.) CHES. Lecture Notes in Computer Science, vol. 7428, pp. 76–91. Springer (2012)
10. Bogdanov, A., Knezevic, M., Leander, G., Toz, D., Varici, K., Verbauwhede, I.: spongent: A lightweight hash function. In: Preneel and Takagi [31], pp. 312–325
11. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: An ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) CHES. Lecture Notes in Computer Science, vol. 4727, pp. 450–466. Springer (2007)
12. Canniere, C.D., Sato, H., Watanabe, D.: Hash function Luffa: Specification. Submission to NIST (Round 2) (2009)
13. Chaum, D., Evertse, J.H.: Crytanalysis of DES with a reduced number of rounds: Sequences of linear factors in block ciphers. In: Williams, H.C. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 218, pp. 192–211. Springer (1985)
14. Courtois, N., Pieprzyk, J.: Cryptanalysis of block ciphers with overdefined systems of equations. In: Zheng [42], pp. 267–287
15. Daemen, J., Peeters, M., Assche, G.V., Rijmen, V.: Nessie proposal: NOEKEON. NESSIE proposal, 27 October 2000.
16. Dinur, I., Shamir, A.: Cube attacks on tweakable black box polynomials. In: Joux, A. (ed.) EUROCRYPT. Lecture Notes in Computer Science, vol. 5479, pp. 278–299. Springer (2009)
17. Dunkelman, O., Indesteege, S., Keller, N.: A differential-linear attack on 12-round Serpent. In: Chowdhury, D.R., Rijmen, V., Das, A. (eds.) INDOCRYPT. Lecture Notes in Computer Science, vol. 5365, pp. 308–321. Springer (2008)
18. Guo, J., Peyrin, T., Poschmann, A., Robshaw, M.J.B.: The led block cipher. In: Preneel and Takagi [31], pp. 326–341
19. Helleseth, T. (ed.): Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings, Lecture Notes in Computer Science, vol. 765. Springer (1994)
20. Knudsen, L.R.: Truncated and higher order differentials. In: Preneel, B. (ed.) FSE. Lecture Notes in Computer Science, vol. 1008, pp. 196–211. Springer (1994)
21. Kocher, P.C., Jaffe, J., Jun, B.: Differential power analysis. In: Wiener, M.J. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 1666, pp. 388–397. Springer (1999)

22. Kohno, T., Kelsey, J., Schneier, B.: Preliminary cryptanalysis of reduced-round Serpent. In: AES Candidate Conference. pp. 195–211 (2000)
23. Langford, S.K., Hellman, M.E.: Differential-linear cryptanalysis. In: Desmedt, Y. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 839, pp. 17–25. Springer (1994)
24. Lim, C.H.: Crypton: A new 128-bit block cipher - specification and analysis (1998)
25. Lim, C.H.: A revised version of crypton - crypton v1.0. In: Knudsen, L.R. (ed.) FSE. Lecture Notes in Computer Science, vol. 1636, pp. 31–45. Springer (1999)
26. Matsui, M.: Linear cryptoanalysis method for DES cipher. In: Helleseth [19], pp. 386–397
27. McLaughlin, J., Clark, J.A.: Filtered nonlinear cryptanalysis of reduced-round serpent, and the wrong-key randomization hypothesis. In: Stam, M. (ed.) IMA Int. Conf. Lecture Notes in Computer Science, vol. 8308, pp. 120–140. Springer (2013)
28. National Bureau of Standards: Data Encryption Standard. FIPS PUB 46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C., (15 January 1977)
29. Nguyen, P.H., Wu, H., Wang, H.: Improving the algorithm 2 in multidimensional linear cryptanalysis. In: Parampalli, U., Hawkes, P. (eds.) ACISP. Lecture Notes in Computer Science, vol. 6812, pp. 61–74. Springer (2011)
30. Nyberg, K.: Differentially uniform mappings for cryptography. In: Helleseth [19], pp. 55–64
31. Preneel, B., Takagi, T. (eds.): Cryptographic Hardware and Embedded Systems - CHES 2011 - 13th International Workshop, Nara, Japan, September 28 - October 1, 2011. Proceedings, Lecture Notes in Computer Science, vol. 6917. Springer (2011)
32. Saarinen, M.J.O.: Cryptographic analysis of all 4×4 s-boxes. In: Miri, A., Vaudenay, S. (eds.) Selected Areas in Cryptography. Lecture Notes in Computer Science, vol. 7118, pp. 118–133. Springer (2011)
33. Schneier, B., Kelsey, J., Whiting, D., Wagner, D., Hall, C., Ferguson, N.: Twofish: A 128-bit block cipher. In: First Advanced Encryption Standard (AES) Conference (1998)
34. Selçuk, A.A.: On probability of success in linear and differential cryptanalysis. J. Cryptology 21(1), 131–147 (2008)
35. Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T., Shirai, T.: Piccolo: An ultra-lightweight blockcipher. In: Preneel and Takagi [31], pp. 342–357
36. Tezcan, C.: Improbable differential attacks on PRESENT using undisturbed bits. Journal of Computational and Applied Mathematics 259, Part B(0), 503 – 511 (2014)
37. Tezcan, C., Takn, H.K., Demirciolu, M.: Improbable differential attacks on SERPENT using undisturbed bits. In: SIN (2014)
38. V. Dolmatov (Ed.): GOST 28147-89: Encryption, decryption, and message authentication code (MAC) algorithms. In: Internet Engineering Task Force RFC 5830 (March 2010)
39. Varici, K., Özen, O., Çelebi Kocair: Sarmal: Sha-3 proposal. Submission to NIST (2008)
40. Wu, W., Zhang, L.: LBlock: A lightweight block cipher. In: Lopez, J., Tsudik, G. (eds.) ACNS. Lecture Notes in Computer Science, vol. 6715, pp. 327–344 (2011)
41. Zhang, W., Bao, Z., Lin, D., Rijmen, V., Yang, B., Verbauwhede, I.: Rectangle: A bit-slice ultra-lightweight block cipher suitable for multiple platforms. IACR Cryptology ePrint Archive 2014, 84 (2014)

42. Zheng, Y. (ed.): Advances in Cryptology - ASIACRYPT 2002, 8th International Conference on the Theory and Application of Cryptology and Information Security, Queenstown, New Zealand, December 1-5, 2002, Proceedings, Lecture Notes in Computer Science, vol. 2501. Springer (2002)

A   Equivalent Definitions with only One Variable

When defining differential factors in Section 3, we used two variables x and y since they are directly linked to the input pairs in differential cryptanalysis. One can observe that the same definition and theorems of Section 3 for bijective S-boxes can be given by using a single variable. We provide them as follows.

Definition 7. *S has a* differential factor λ *for the output difference* μ *if*

$$S^{-1}(S(x) \oplus \mu) \oplus \lambda = S^{-1}(S(x \oplus \lambda) \oplus \mu)$$

for all x.

Proposition 1. *Definition 5 is equivalent to Definition 7.*

Proof. Since $S(x) \oplus S(y) = \mu$, we have $y = S^{-1}(S(x) \oplus \mu)$. Similarly, $y \oplus \lambda = S^{-1}(S(x \oplus \lambda) \oplus \mu)$ since $S(x \oplus \lambda) \oplus S(y \oplus \lambda) = \mu$. XORing both equations gives $\lambda = S^{-1}(S(x) \oplus \mu) \oplus S^{-1}(S(x \oplus \lambda) \oplus \mu)$ and we are done. $\square$

Definition 8. *S has a* differential factor λ *for the output difference* μ *if*

$$S(S^{-1}(x) \oplus \lambda) \oplus \mu = S(S^{-1}(x \oplus \mu) \oplus \lambda)$$

for all x.

Proposition 2. *Definition 5 is equivalent to Definition 8.*

Proof. Let $y = S(x)$. Then the Definition 7 becomes

$$S^{-1}(y \oplus \mu) \oplus \lambda = S^{-1}(S(S^{-1}(y) \oplus \lambda) \oplus \mu)$$

for all y. Applying the S operation on both sides of the equation gives

$$S(S^{-1}(y \oplus \mu) \oplus \lambda) = S(S^{-1}(y) \oplus \lambda) \oplus \mu$$

for all y and we are done. $\square$

Thus, Propositions 1 and 2 prove the Theorem 1.

Proposition 3. *If* λ_1 *and* λ_2 *are differential factors for an output difference* μ, *then* $\lambda_1 \oplus \lambda_2$ *is also differential factor for the output difference* μ. *i.e. All differential factors* λ_i *for* μ *forms a vector space.*

Proof. We have

$$S^{-1}(S(x) \oplus \mu) \oplus \lambda_1 = S^{-1}(S(x \oplus \lambda_1) \oplus \mu)$$

for all x, by Definition 7. And we have

$$S^{-1}(S(x \oplus \lambda_1) \oplus \mu) \oplus \lambda_2 = S^{-1}(S(x \oplus \lambda_1 + \lambda_2) \oplus \mu)$$

since λ_2 is a differential factor. Thus, we get

$$S^{-1}(S(x) \oplus \mu) \oplus \lambda_2 \oplus \lambda_2 = S^{-1}(S(x \oplus \lambda_1 \oplus \lambda_2) \oplus \mu)$$

for all x and we are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

B 3-Round Differentials with Higher Probability

The rounds of the 3-round differential used in the differential-linear attacks of [7, 17] have probabilities 2^{-5}, 2^{-1}, and 1 but the authors observed experimentally that this differential has probability 2^{-7} instead of 2^{-6}. We observed that there are 3-round differentials of SERPENT with probability 2^{-5} that can be combined with the same linear approximations. The rounds of these differential have probabilities 2^{-5}, 1, and 1 and for this reason, the theoretical and practical probabilities of these differentials are the same. However, these differentials activate six S-boxes at the first round of the attack instead of five. So replacing the original differential with one of them results in capturing four more subkey bits but time complexity of the attacks also increases by a factor of 2^4.

Since the data complexity of a differential-linear attack is of $O(p^{-2}q^{-4})$ and replacing the differential result in $p = 2^{-5}$ instead of 2^{-7}, one would expect the modified attacks to have data and time complexities reduced by a factor of 2^4. However, experiment results show that the gain in the modified attacks is at most a factor of $(2^{-0.32})^2$. This is because the transition between the original differential and the linear approximation is far better than expected. For instance, when the original 3-round differential is combined with a 1-round linear approximation of bias 2^{-5}, Dunkelman *et al.* experimentally verified that the 4-round differential-linear path has bias $2^{-13.75}$, instead of $2 \cdot 2^{-7} \cdot (2^{-5})^2 = 2^{-16}$. We performed similar experiments on five different 3-round differentials with probability 2^{-5} using 2^{34} pairs and the results are summarized in Table 4.

We replace the original differential with the second one from Table 4 and obtain new 10, and 11 round differential-linear attacks. This change provides a 4-round bias of $2^{-13.43}$ instead of $2^{-13.75}$. Thus the data and time complexity gain in the modified attack is a factor of $(2^{-0.32})^2$. This differential activates six S-boxes instead of five so we capture four more subkey bits and the time complexity is multiplied by 2^4. We summarize this modified attack in Table 5. Note that there are two differential factors for this differential, too. Since the rest of our modified attacks are almost identical to the attacks of [17], we refer the interested reader to [17].

Table 4: 4-Round biases for 3-round differentials with probability 2^{-5} and 1-round linear approximation with bias 2^{-5}.

#	Input Difference X_0 X_1 X_2 X_3 (in Hexadecimal)	#Active S-boxes	Bias	Standard Deviation
1	40000000 00000000 40000002 00000000	6	$2^{-13,49}$	$2^{-18.03}$
2	00000000 40000000 40000002 00000000	6	$2^{-13,43}$	$2^{-18.11}$
3	00000000 40000000 00000002 40000000	6	$2^{-13,56}$	$2^{-18.07}$
4	00000000 40000000 40000002 00000002	6	$2^{-13,43}$	$2^{-18.19}$
5	00000002 00000000 00000012 00000000	6	$2^{-14,65}$	$2^{-18.00}$

Table 5: 11-Round differential-linear attack with a 3-round differential of probability 2^{-5}. Output differences $\mu = 4_x$ and $\mu = E_x$ that contain differential factors for S_1 are shown in bold. Undisturbed bits are shown in italic.

Input	X_0: 0??0 0000 0000 00?0 0000 ?00? 00?0 0000
	X_1: 0??0 0000 0000 00?0 0000 ?00? 00?0 0000
	X_2: 0??0 0000 0000 00?0 0000 ?00? 00?0 0000
	X_3: 0??0 0000 0000 00*10* 0000 ?00? 00*10* 0000
S_1	X_0: 0000 0000 0000 0010 0000 000**0** 0000 0000
	X_1: 0110 0000 0000 0000 0000 100**1** 00**0**0 0000
	X_2: 0000 0000 0000 0000 0000 0001 00**1**0 0000
	X_3: 0000 0000 0000 0000 0000 100**1** 00**0**0 0000
LT	X_0: 0000 0000 0000 0000 0000 0000 0000 0000
	X_1: 0100 0000 0000 0000 0000 0000 0000 0000
	X_2: 0100 0000 0000 0000 0000 0000 0000 0010
	X_3: 0000 0000 0000 0000 0000 0000 0000 0000
S_2	X_0: 0000 0000 0000 0000 0000 0000 0000 0000
	X_1: 0000 0000 0000 0000 0000 0000 0000 0010
	X_2: 0100 0000 0000 0000 0000 0000 0000 0000
	X_3: 0000 0000 0000 0000 0000 0000 0000 0010
LT	X_0: 0000 0000 0000 0000 0000 0000 0000 0000
	X_1: 0000 0000 0000 0000 0000 0000 0000 0000
	X_2: 0000 0000 1000 0000 0000 0000 0000 0000
	X_3: 0000 0000 0000 0000 0000 0000 0000 0000
S_3	X_0: 0000 0000 ?000 0000 0000 0000 0000 0000
	X_1: 0000 0000 ?000 0000 0000 0000 0000 0000
	X_2: 0000 0000 ?000 0000 0000 0000 0000 0000
	X_3: 0000 0000 ?000 0000 0000 0000 0000 0000
LT	X_0: 00?0 0000 0000 ?000 0000 0??0 0?00 ?00?
	X_1: 0000 ?00? 0000 0000 0000 0000 00?0 0000
	X_2: 0000 0000 ?0?? 000? 0000 0000 000? 0?00
	X_3: 0?00 0000 0000 0000 0?00 0000 0000 00?0
S_4	X_0: 0??0 ?00? ?0?? ?00? 0?00 0??0 0??? ????
	X_1: 0??0 ?00? ?0?? ?00? 0?00 0??0 0??? ????
	X_2: 0??0 ?00? ?0?? ?00? 0?00 0??0 0??? ????
	X_3: 0??0 ?00? ?0?? ?00? 0?00 0??0 0??? ????
6-Round Linear Approximation Λ	
Last Round	

$p = 2^{-5}$