NLAT: the NonLinear Approximation Table of Vectorial Boolean Mappings

Jorge Nakahara Jr

Abstract. This paper studies an extension of the Linear Approximation Table (LAT) of vectorial Boolean mappings (also known as Substitution boxes) used in Linear Cryptanalysis (LC). This extended table was called NonLinear Approximation Table (NLAT).
Similar concepts and parameters of a LAT are associated to the NLAT as well such as the nonlinear uniformity, the nonlinear spectrum and the nonlinear bias/correlation.

Keywords: linear cryptanalysis, Linear Approximation Table (LAT), S-boxes, vectorial Boolean mappings.

1 Introduction

Linear Cryptanalysis (LC) is a statistical method to distinguish a given block cipher from a (family) of random permutations with the same block size.

LC was first described by Matsui and Yamagishi to attack the FEAL block cipher [9] in 1992. Later, in 1993-1994, Matsui [8, 7] applied LC to analyze the DES block cipher [11].

The analyses on DES were initially focused on the 6×4 Substitution boxes (S-boxes) which are the only nonlinear components in DES. One of the tools Matsui built to study systematically the propagation of linear relations across the DES S-boxes is the Linear Approximation Table (LAT), that allows to identify for each 6-bit input linear relation, the most probable 4-bit output linear relation for each S-box.

Further studying the LATs combined with the diffusion properties of the P permutation in DES allowed Matsui to multi-round linear relations.

Therefore, the analysis approach was bottom-up, from small nonlinear components (the S-boxes and the P permutation) to larger cipher components such as a full round, and then covering multiple-rounds.

This report is organized as follows: Section 2 contains notations and definitions. Section 2.1 describes the NonLinear Approximation Table (NLAT) of a vectorial Boolean mapping (or S-box). Section 3 describes an application of nonlinear approximations. Section 4 presents the conclusions.

2 Notation

Substitution boxes (S-boxes) are vectorial Boolean mappings [4] defined as $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$ with $n, m \in \mathbb{N}^+$.

The inner-product or dot-product of two Boolean vectors $u, x \in \mathbb{F}_2^n$ is defined as

$$< u, x >= u.x^T =< x, u >= x.u^T = \bigoplus_{0 \le i \le n-1} u_i.x_i, \tag{1}$$

where u^T is the transpose vector of $u = (u_{n-1}, \ldots, u_0)$.

In the LC setting, u is a fixed n-bit mask, while x is a n-bit word in the cipher state. The meaning of $< u, x >$ is that the bits in u select specific bits of x when $u_i = 1$. The result of $< u, x >$ is the parity of the bits of x selected by the bits of u.

A systematic analysis of all linear approximations involving all possible input and output bits of a given S-box S is summarized in its Linear Approximation Table or LAT.

Definition 1. *Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an S-box, and let $u \in \mathbb{F}_2^n$, $v \in \mathbb{F}_2^m$. The* Linear Approximation Table (LAT) *of S is the $2^n \times 2^m$ matrix whose entries are defined as*

$$LAT_S(u, v) = \gamma_S(u, v)$$

where

$$\gamma_S(u, v) = |\{x \in \mathbb{F}_2^n | < u, x >=< v, S[x] >\}| - 2^{n-1}$$

The linear approximation (or linear relation) $< u, x >=< v, S[x] >$ implies a 0/1 parity value $< u, x > \oplus < v, S[x] >$. Therefore, $\gamma_S(u, v)$ counts how often this parity value deviates from a balanced parity distribution.

From the definition of LAT_S, if S is a bijective and involutory S-box, then LAT_S is a symmetric matrix ie. the LAT of S^{-1} is the transpose of the LAT of S.

The strength of a linear approximation of S with bit-masks u, v is measured by its bias:

$$\epsilon_S(u, v) = |\gamma_S(u, v)/2^n| \in [0, 1/2]. \tag{2}$$

The closer the bias $\epsilon_S(u, v)$ is to $1/2$, the better the approximation of S is to a linear function given by the masks (u, v).

If the bias is nonzero, then the input linear relation represented by the n-bit mask 'u' is said to propagate across S to the output linear relation represented by the m-bit mask 'v', and the S-box is said to be (linearly) active. This linear approximation is denoted $u \xrightarrow{S} v$.

If we consider the mapping $l_u(x) =< u, x >$ then the correlation between the mappings $l_u(x)$ and $l_v(S[x])$ is:

$$\mathrm{corr}(l_u(x), l_v(S[x])) = 2 * \epsilon_s(u, v), \tag{3}$$

which normalizes the range to the interval $[0, 1]$.

The signed values of $\gamma_S(u, v)$ tabulated in the LAT_S constitute **linear spectrum** [4] of S.

Outstanding values in the LAT_S refer to the pairs of input/output bit-masks (u, v) with the highest bias which is denoted as Γ_S:

$$\Gamma_S = \max_{u \in (\mathbb{F}_2^n)^*, v \in (\mathbb{F}_2^m)^*} |\gamma_S(u, v)|, \tag{4}$$

where $(\mathbb{F}_2^n)^*$ means $\mathbb{F}_2^n \setminus \mathbf{0}$, and $\mathbf{0}$ is the zero vector.

The S-box S is called a linearly Γ_S-uniform mapping [12], which means that the absolute value of all LAT_S entries are smaller than or equal to Γ_S.

Table 1 lists the linear uniformity of the 6×4 DES S-boxes [8, 7] and that of the AES S-box [5]. In AES, Γ_S is used along with the diffusion components ShiftRows and MixColumns to upperbound the bias of linear relations across multiple rounds, and therefore demonstrate AES's resistance against conventional LC.

Table 1. Examples of S-boxes and their linear uniformity.

S-box	linear uniformity	maximum bias
DES S_1	18	$18/2^6 \approx 2^{-1.83}$
DES S_2	16	$16/2^6 = 2^{-2}$
DES S_3	16	$16/2^6 = 2^{-2}$
DES S_4	16	$16/2^6 = 2^{-2}$
DES S_5	20	$20/2^6 \approx 2^{-1.67}$
DES S_6	14	$14/2^6 \approx 2^{-2.19}$
DES S_7	18	$18/2^6 \approx 2^{-1.83}$
DES S_8	16	$16/2^6 = 2^{-2}$
AES S-box	16	$16/2^8 = 2^{-4}$

Also, in [8], Matsui exploited the fact that $\Gamma_{S_5} = 20$ to develop linear relations across up to 20 rounds of DES. According to Table 1, Γ_{S_5} is the largest linear uniformity among all eight DES S-boxes.

As an example of a LAT, consider the 4×4 S-box $S_1 : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ used in the Serpent block cipher [2]. The truth table of S_1 is in Table 4 in the Appendix, where the subscript x denotes hexadecimal notation.

The LAT of S_1 is in Table 2. Therefore, S_1 is a linearly 4-uniform mapping.

Table 2. LAT of S-box S_1 of the Serpent cipher with input bit-mask 'u' (row) and output bit-mask 'v' (column). For instance, $\gamma_{S_1}(6_x, 1_x) = -2$.

u\v	0_x	1_x	2_x	3_x	4_x	5_x	6_x	7_x	8_x	9_x	A_x	B_x	C_x	D_x	E_x	F_x
0_x	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1_x	0	-2	-2	-4	0	-2	2	0	2	0	0	-2	2	0	-4	2
2_x	0	-2	2	0	0	-2	2	0	-2	4	0	-2	2	0	4	2
3_x	0	0	0	0	0	0	4	-4	0	0	0	0	-4	-4	0	0
4_x	0	0	-2	2	0	0	2	-2	-2	-2	0	-4	2	2	0	-4
5_x	0	2	0	-2	0	2	0	-2	0	2	4	2	4	-2	0	-2
6_x	0	-2	0	2	0	-2	-4	-2	0	-2	4	-2	0	-2	0	2
7_x	0	-4	2	2	0	4	2	2	2	-2	0	0	2	-2	0	0
8_x	0	0	0	0	0	4	0	-4	0	0	0	0	0	4	0	4
9_x	0	-2	2	0	0	2	-2	0	-2	4	0	-2	-2	0	-4	-2
A_x	0	2	2	4	0	-2	2	0	-2	0	0	2	2	0	-4	2
B_x	0	-4	-4	0	0	0	0	0	-4	0	0	4	0	0	0	0
C_x	0	0	-2	2	-4	0	-2	-2	2	2	-4	0	2	-2	0	0
D_x	0	2	-4	2	4	2	0	2	0	2	0	-2	0	-2	0	2
E_x	0	2	0	-2	-4	2	0	2	-4	-2	0	-2	0	-2	0	2
F_x	0	0	-2	2	-4	0	2	2	2	2	4	0	-2	2	0	0

2.1 Nonlinear Approximation Table (NLAT)

A natural extension of the LAT of an S-box is to consider nonlinear bit-masks instead of linear ones.

Consider an S-box $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$, with $n, m \in \mathbb{N}^+$. In LC, the LAT of S lists all $2^n \times 2^m$ linear combinations of the input and output bits of S.

As an example, consider a 3-bit linear mask operating on an input $x = (x_2, x_1, x_0)$. The 2^3 linear masks are typically ordered as $\{0,\ x_0,\ x_1,\ x_0 + x_1,\ x_2,\ x_0 + x_2,\ x_1 + x_2,\ x_0 + x_1 + x_2\}$, where '+' denotes bitwise exclusive-or.

The nonlinear masks for this same 3-bit word would include all possible combinations of the $2^3 - 1 = 7$ nonzero monomials $\{x_0,\ x_1,\ x_0.x_1,\ x_2,\ x_0.x_2,\ x_1.x_2,\ x_0.x_1.x_2\}$, where '.' means bitwise-AND. There are $2^7 - 1$ possible combinations of nonzero monomials.

In general, for n-bit data, while linear masks are n bits long, the nonlinear masks are $2^n - 1$ bits long.

The use of nonlinear masks requires the dot-product (1) to be redefined. We call this new concept a nonlinear dot-product.

Due to the asymmetry in size, the order of the parameters matter in the nonlinear dot-product which is the following binary operator:

$$\lhd u, x \rhd : \mathbb{F}_2^{2^n-1} \times \mathbb{F}_2^n \to \mathbb{F}_2,$$

because the nonlinear mask u is $2^n - 1$ bits long, while the second parameter (on which the mask is applied) is only n bits long.

The nonlinear dot-product is defined as

$$\lhd u, x \rhd = \bigoplus_{\forall x_i} f_u(x), \tag{5}$$

where $f_u(x)$ is a nonlinear Boolean function of the bits of $x = (x_{n-1}, \ldots, x_0)$.

For instance, consider a 3-bit input $x = (x_2, x_1, x_0)$. Nonlinear masks applied to x may include any of the $2^3 - 1 = 7$ monomials derived from the 3 bits (x_2, x_1, x_0). These monomials, in reverse lexicographic order, are: $(x_0.x_1.x_2,\ x_1.x_2,\ x_0.x_2,\ x_2,\ x_0.x_1,\ x_1,\ x_0)$.

For instance, the 7-bit nonlinear mask $u = 1111000_2$ stands for $f_u(x) = x_2 + x_0.x_2 + x_1.x_2 + x_0.x_1.x_2$, and $\lhd u, x \rhd = \bigoplus_{x_i; 0 \le i \le 2}(x_2 + x_0.x_2 + x_1.x_2 + x_0.x_1.x_2)$. Table 5 in the Appendix shows how to translate a binary representation of a nonlinear mask to its 'symbolic' representation.

A systematic analysis of all nonlinear approximations involving all the input and output bits of a given S-box S is summarized in its NonLinear Approximation Table or NLAT.

Definition 2. *Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an S-box and $u \in \mathbb{F}_2^{2^n-1}$ and $v \in \mathbb{F}_2^{2^m-1}$ be nonlinear masks. The* NonLinear Approximation Table (NLAT) *of S is the $2^{2^n-1} \times 2^{2^m-1}$ matrix whose entries are defined as*

$$NLAT_S(u, v) = \gamma_S^*(u, v)$$

where

$$\gamma_S^*(u, v) = |\{x \in \mathbb{F}_2^n | \lhd u, x \rhd = \lhd v, S[x] \rhd\}| - 2^{n-1}$$

Corollary 1. *If S is a bijective and involutory S-box then the NLAT of S is symmetric, that is, the NLAT of S^{-1} is the transpose the NLAT of S.*

Similar to linear approximations, the nonlinear approximations $\lhd u, x \rhd = \lhd v, S[x] \rhd$ imply a parity value $\lhd u, x \rhd + \lhd v, S[x] \rhd$. Therefore, $\gamma^*_S(u, v)$ counts how often the parity of the nonlinear approximation deviates from that of a balanced or uniform distribution.

The strength of a nonlinear approximation of S with bit-masks u, v is measured by its bias:

$$\epsilon^*_S(u, v) = |\gamma^*_S(u, v)/2^n| \in [0, 1/2]. \tag{6}$$

The closer the bias is to $1/2$, the better the nonlinear approximation is.

If we consider the mapping $l^*_u(x) = \lhd u, x \rhd$ then the correlation between the mappings $l^*_u(x)$ and $l^*_v(S[x])$ can be defined as:

$$\text{corr}(l^*_u(x), l^*_v(S[x])) = 2 * \epsilon^*_S(u, v), \tag{7}$$

which normalizes the range to the interval $[0, 1]$.

Definition 3. *The values $\gamma^*_S(u, v)$ in the NLAT of S constitute its* **nonlinear spectrum**.

Outstanding values in the NLAT_S refer to the pairs of input/output nonlinear bit-masks (u, v) with the highest bias. Masks that reach the maximum bias in the NLAT_S are associated to the nonlinear uniformity of S.

Let the highest bias in the NLAT of S be denoted:

$$\Gamma^*_S = \max_{u \in (\mathbb{F}_2^{2^n-1})^*, v \in (\mathbb{F}_2^{2^m-1})^*} |\gamma^*_S(u, v)|, \tag{8}$$

where $(\mathbb{F}_2^{2^n-1})^*$ means $\mathbb{F}_2^{2^n-1} \setminus \mathbf{0}$, and $\mathbf{0}$ is the zero vector.

Definition 4. *The S-box S is called a* nonlinearly Γ^*_S-uniform *mapping, which means that the absolute value of all $NLAT_S$ entries are smaller than or equal to Γ^*_S.*

As an example of an NLAT, consider the S_1 S-box in Table 4. Its $2^{2^4-1} \times 2^{2^4-1}$ NLAT is too large to be displayed in this report but S_1's nonlinear uniformity can be found to be 8. Comparatively, S_1's linear uniformity is 4.

The NLAT of a $n \times m$ S-box S can be partitioned into four non-overlapping sub-tables denoted: (L-L), (L-NL), (NL-L) and (NL-NL) to indicate the nature of the masks used at the input and output of S, respectively.

(L-L) a $2^n \times 2^m$ sub-table with linear masks (u, v), which is the LAT of S. This fact shows that the NLAT is a natural extension of the LAT.
(L-NL) a $2^n \times (2^{2^m-1} - 2^m)$ sub-table with masks (u, v) where u is linear but v is nonlinear.
(NL-L) a $(2^{2^n-1} - 2^n) \times 2^m$ sub-table with masks (u, v) where u is nonlinear but v is linear.
(NL-NL) a $(2^{2^n-1} - 2^n) \times (2^{2^m-1} - 2^m)$ sub-table with masks (u, v) where both u and v are nonlinear.

For instance, for $n = m = 3$, (L-L) is an 8×8 matrix, while (L-NL) is an 8×120 matrix, (NL-L) is a 120×8 matrix, and (NL-NL) is a 120×120 matrix.

In the NLAT of an S-box S, some nonlinear masks are expected to reach the maximum possible bias $1/2$. Examples of such nonlinear expressions are the Algebraic Normal Form (ANF) [4] of the output coordinates y_i of S because equality in the ANFs always hold. If S is bijective then the ANF of the output coordinates x_i of S^{-1} are also present in the NLAT_S.

Therefore, the nonlinear uniformity of an S-box $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is 2^{n-1} and is achieved by the nonlinear masks associated to the ANF of the S-box output coordinates (among other nonlinear relations).

For instance, let $(y_3, y_2, y_1, y_0) = S_1[x_3, x_2, x_1, x_0]$. Then, the (NL-L) sub-table of the NLAT of S_1 contains the nonlinear relations corresponding to its ANF coordinates [13]:

(a) $y_3 = 1 + x_1 + x_2.x_0 + x_3 + x_3.x_0 + x_3.x_1.x_0 + x_3.x_2.x_0 + x_3.x_2.x_1$,
(b) $y_2 = 1 + x_1 + x_1.x_0 + x_2 + x_3$,
(c) $y_1 = 1 + x_0 + x_1.x_0 + x_2 + x_2.x_0 + x_3 + x_3.x_1 + x_3.x_1.x_0 + x_3.x_2.x_0 + x_3.x_2.x_1$,
(d) $y_0 = 1 + x_0 + x_1 + x_2.x_1 + x_3.x_0 + x_3.x_2 + x_3.x_2.x_0 + x_3.x_2.x_1$

and likewise the (L-NL) sub-table of the NLAT_{S_1} contains the following ANFs [13] of S_1^{-1}:

(e) $x_3 = y_0 + y_2 + y_3 + y_3.y_1$,
(f) $x_2 = 1 + y_0 + y_1 + y_2.y_0 + y_2.y_1 + y_2.y_1.y_0 + y_3 + y_3.y_2.y_0$,
(g) $x_1 = y_1 + y_2 + y_2.y_1.y_0 + y_3 + y_3.y_0 + y_3.y_1 + y_3.y_2.y_0 + y_3.y_2.y_1$,
(h) $x_0 = 1 + y_0 + y_1 + y_1.y_0 + y_2.y_1.y_0 + y_3.y_1 + y_3.y_2.y_0 + y_3.y_2.y_1$.

In the NLAT_{S_1}, we have also found the 'negation' of (h): $x_0 = y_0 + y_1 + y_1.y_0 + y_2.y_1.y_0 + y_3.y_1 + y_3.y_2.y_0 + y_3.y_2.y_1$, with bias 1/2, and (g) whose equality always holds (bias is 1/2).

Actually, these nonlinear approximations with maximum possible bias are not approximations but rather deterministic nonlinear relations, in contrast to the majority of the entries in the NLAT_{S_1}, which are probabilistic.

Since the ANFs of every output coordinate of an $n \times m$ S-box S will be in the NLAT (as well as the ANF of the input coordinates), the nonlinear uniformity of S will be maximum (2^{n-1}), and their biases will be the maximum possible (1/2).

Moreover, there are other nonlinear relations in the NLAT_{S_1} (holding with maximum possible bias 1/2), that are not the ANFs of S_1's input or output coordinates, such as:

$$x_0.x_1 = y_2 + y_0.y_2 + y_1.y_2 + y_3 + y_0.y_3 + y_1.y_2.y_3,$$
$$x_0.x_2 = y_0 + y_1 + y_0.y_1 + y_3 + y_0.y_3 + y_0.y_2.y_3 + y_1.y_2.y_3,$$
$$x_1.x_2 = y_0.y_1 + y_2 + y_1.y_2 + y_0.y_1.y_2 + y_1.y_3 + y_2.y_3,$$
$$x_0.x_1.x_2 = y_2 + y_0.y_2 + y_1.y_2 + y_0.y_1.y_2 + y_1.y_3 + y_0.y_1.y_3 + y_2.y_3 + y_0.y_2.y_3,$$
$$x_0.x_3 = y_2 + y_0.y_2 + y_1.y_2 + y_0.y_1.y_2 + y_3 + y_0.y_3 + y_1.y_3 + y_0.y_2.y_3$$

These deterministic nonlinear relations may be of independent interest for algebraic cryptanalysis [1].

While the absolute value of all entries in the LAT of S_1 are even integers, in the NLAT of S_1 there are odd integer values.

These odd-valued entries in the NLAT are due the monomial $\prod_{0 \le i \le 3} y_i$ that never happens in the ANF of bijective S-boxes.

For instance, for S_1, the monomial $y_3.y_2.y_1.y_0$ never happens in S_1's ANF and the nonlinear relations that contain $y_3.y_2.y_1.y_0$, the NLAT_{S_1} entries are odd valued.

3 Applications

In [6], Knudsen and Robshaw had the idea of using nonlinear approximations (along with linear ones) to improve the linear analysis of block ciphers.

For instance, they noticed that for the DES cipher, there are nonlinear approximations with larger bias than linear approximations could provide.

But, there are clear difficulties of: (i) combining nonlinear relations across successive rounds, and (ii) linking nonlinear relations of S-boxes with those of linear components (such as the diffusion layers) in a single round.

In this section, we exploit their idea of using nonlinear relations only in the beginning and at the end of an existing linear relation.

The most promising targets to test this idea of combining linear and nonlinear relations are ciphers designed for a hardware environment and that use bit-based diffusion components.

In particular, we will experiment with the PRESENT block cipher [3].

PRESENT is a 64-bit block cipher with an SPN design, key sizes of 80 or 128 bits, and iterating 31 rounds. PRESENT uses only one 4×4 S-box S_P whose truth table is listed in Table 6 in the Appendix.

Each full round in PRESENT consists of the following operation (in order):

(i) bitwise-xor of a 64-bit round key with the cipher state;
(ii) an S-box layer where each nibble is input to S_P;
(iii) a diffusion layer called pLayer, which consists of bit permutation involving the 64 bits in the state.

In [10], linear distinguishers for a variable number of rounds of the PRESENT block cipher were described.

All the linear distinguishers in [10] use the one-round iterative linear relation (9):

$$0000000000200000_x \xrightarrow{1r} 0000000000200000_x, \tag{9}$$

where each hexadecimal digit stands for a 4-bit piece of a linear mask applied to a single nibble (4-bit word) in the 64-bit state of PRESENT.

The 11th nibble (from left to right) is the only active nibble at both the input and the output of (9). This iterative linear relation is based on the $(u, v) = (2, 2)$ linear mask across S_P: $2 \xrightarrow{S_P} 2$, with bias 2^{-3}. It is a fix-point linear relation across S_P. The LAT of S_P can be found in the appendix of [10].

Moreover, (9) exploits a fix-point of the pLayer of PRESENT, which is just a bit permutation. This fix-point connects the second least significant output bit of S_P to the second input bit of S_P in the next round. For further details about the pLayer, see [10].

Concatenating (9) with itself allows to extend this linear relation across several rounds with a fixed decrease in the bias due to the Piling-Up Lemma [7].

Further, exploiting the linear hull effect [12], several linear trails can be counted when (9) covers multiple rounds. For instance, after 20 rounds of (9), there are 20466576 linear trails inside it, and the estimated bias of 2^{-41} (without considering the linear hull effect) becomes $2^{-28.85}$ when all the internal linear trails are accounted for.

Let $(y_3, y_2, y_1, y_0) = S_P[x_3, x_2, x_1, x_0]$. The ANF of the output coordinates of S_P are:

$$\begin{aligned}
&y_3 = 1 + x_0 + x_1 + x_1.x_2 + x_0.x_1.x_2 + x_3 + x_0.x_1.x_3 + x_0.x_2.x_3,\\
&y_2 = 1 + x_0.x_1 + x_2 + x_3 + x_0.x_3 + x_1.x_3 + x_0.x_1.x_3 + x_0.x_2.x_3,\\
(*)\quad &y_1 = 1 + x_0.x_1.x_2 + x_3 + x_1.x_3 + x_0.x_1.x_3 + x_2.x_3 + x_0.x_2.x_3,\\
&y_0 = x_0 + x_2 + x_1.x_2 + x_3.
\end{aligned}$$

and the ANF of the output coordinates of S_P^{-1} are:

$$
\begin{aligned}
& x_3 = y_0 + y_1 + y_0.y_1 + y_2 + y_0.y_1.y_2 + y_3 + y_0.y_2.y_3, \\
& x_2 = 1 + y_0.y_1 + y_0.y_2 + y_1.y_2 + y_0.y_1.y_2 + y_3 + y_0.y_3 + y_1.y_3 + y_0.y_1.y_3 + y_0.y_2.y_3, \\
(**)\quad & x_1 = y_0 + y_1 + y_0.y_2 + y_0.y_1.y_2 + y_3 + y_1.y_3 + y_0.y_1.y_3 + y_2.y_3 + y_0.y_2.y_3, \\
& x_0 = 1 + y_0 + y_2 + y_1.y_3.
\end{aligned}
$$

The ANF of S_P and those of its inverse are interesting nonlinear relations to combine with (9) because:

(i) they holds with maximum possible bias;
(ii) one end of the equality has only one bit, which makes them attractive to combine with linear relations in (9). Here it becomes clear the advantage of attacking ciphers that employ bit permutations as diffusion components.

Now, we can extend the 20-round relation using (9) by adding the ANF of S_P at the top and the ANF of S_P^{-1} at the end.

At the top-end of 20 rounds of (9), the second least significant input bit to the 11th nibble in (9) is connected to the second output bit of S_P one round above. That is bit y_1, whose nonlinear expression is described in (*). It is an exact nonlinear expression from the ANF of S_P with maximum bias 1/2.

At the bottom-end of 20 rounds of (9), the second least significant output bit of the 11th nibble in (9) is connected to the second input bit of S_P one round below. That is bit x_1, whose nonlinear expression is described in (**). It is an exact nonlinear expression from the ANF of S_P^{-1} with maximum bias 1/2.

Combining both nonlinear expression, the result is a 22-round nonlinear relation with the same bias as the original 20-round linear relation (since the ANF correspond to nonlinear relations with maximum bias).

4 Conclusions

In this report, an extension of the Linear Approximation Table (LAT) of an S-box was investigated using nonlinear bit masks. The extemded table was called NonLinear Approximation Table (NLAT).

Table 3 compares some features of the LAT and NLAT of an $n \times m$-bit S-box.

Table 3. Some features of the LAT and the NLAT of an $n \times m$-bit S-box $S : \mathbb{F}_2^n \to \mathbb{F}_2^m$.

LAT	NLAT
n-bit linear input mask	$(2^n - 1)$-bit nonlinear input mask
m-bit linear output mask	$(2^m - 1)$-bit nonlinear output mask
linear uniformity	nonlinear uniformity
$2^n \times 2^m$ matrix	$2^{2^n-1} \times 2^{2^m-1}$ matrix
absolute entry values are even integers	absolute entry values can be even or odd integers

References

1. G. V. Bard. *Algebraic Cryptanalysis*. Springer, 2009.
2. E. Biham, R. J. Anderson, and L. R. Knudsen. Serpent: a New Block Cipher Proposal. In S. Vaudenay, editor, *Fast Software Encryption (FSE)*, LNCS 1372, pages 222–238. Springer, 1998.
3. A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, and C. Vikkelsoe. Present: an Utra-Lightweight Block Cipher. In P. Paillier and I. Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems (CHES)*, LNCS 4727, pages 450–466. Springer, 2007.
4. A. Canteaut. Lecture Notes on Cryptographic Boolean Functions. https://www.rocq.inria.fr/secret/Anne.Canteaut, 2016.
5. J. Daemen and V. Rijmen. AES Proposal: Rijndael. First AES Conference, California, USA, http://www.nist.gov/aes, 1998.
6. L.R. Knudsen and M.J.B. Robshaw. Non-Linear Approximations in Linear Cryptanalysis. In *Advances in Cryptology, Eurocrypt*, LNCS 1070, pages 224–236. Springer, 1996.
7. M. Matsui. Linear Cryptanalysis of DES Cipher (I), version 1.03.
8. M. Matsui. Linear Cryptanalysis Method for DES Cipher. In T. Helleseth, editor, *Advances in Cryptology, Eurocrypt*, LNCS 765, pages 386–397. Springer, 1993.
9. M. Matsui and A. Yamagishi. A New Method for Known-Plaintext Attack of FEAL Cipher. In R. A. Rueppel, editor, *Advances in Cryptology, Eurocrypt*, LNCS 658, pages 81–91. Springer, 1993.
10. J. Nakahara, Jr, P. Sepehrdad, B. Zhang, and M. Wang. Linear (Hull) and Algebraic Cryptanalysis of the Block Cipher PRESENT. In J. A. Garay and A. Otsuka, editors, *8th International Conference on Cryptology and Network Security (CANS)*, LNCS 5888, pages 58–75. Springer, 2009.
11. NBS. Data Encryption Standard (DES). FIPS PUB 46, Federal Information Processing Standards Publication 46, U.S. Department of Commerce, Jan 1977.
12. K. Nyberg. Linear Approximation of Block Ciphers. In A. De, Santis, editor, *Advances in Cryptology, Eurocrypt*, LNCS 950, pages 439–444. Springer, 1995.
13. B. Singh, L. Alexander, and S. Burman. On Algebraic Relations of Serpent S-boxes. IACR ePrint 2009-038, 2009.

A Tables

Table 4. Truth Table (TT) of S-box S_1 used in the Serpent block cipher.

i	0_x	1_x	2_x	3_x	4_x	5_x	6_x	7_x	8_x	9_x	A_x	B_x	C_x	D_x	E_x	F_x
$S_1[i]$	F_x	C_x	2_x	7_x	9_x	0_x	5_x	A_x	1_x	B_x	E_x	8_x	6_x	D_x	3_x	4_x

Table 5. Notation for 7-bit nonlinear masks based on reverse lexicographic order of monomials.

nonlinear mask		Monomials						
binary	symbolic	$x_0.x_1.x_2$	$x_1.x_2$	$x_0.x_2$	x_2	$x_0.x_1$	x_1	x_0
0000001_2	x_0	0	0	0	0	0	0	1
0000010_2	x_1	0	0	0	0	0	1	0
0000011_2	$x_0 + x_1$	0	0	0	0	0	1	1
0000100_2	$x_0.x_1$	0	0	0	0	1	0	0
0000101_2	$x_0.x_1 + x_0$	0	0	0	0	1	0	1
0000110_2	$x_0.x_1 + x_1$	0	0	0	0	1	1	0
0000111_2	$x_0.x_1 + x_1 + x_0$	0	0	0	0	1	1	1
...	...	...	...	...	...	...	...	...
1111111_2	$x_0.x_1.x_2 + x_1.x_2 + ... + x_0$	1	1	1	1	1	1	1

Table 6. Truth Table (TT) of S-box S_P and S_P^{-1} used in the PRESENT block cipher.

i	0_x	1_x	2_x	3_x	4_x	5_x	6_x	7_x	8_x	9_x	A_x	B_x	C_x	D_x	E_x	F_x
$S_P[i]$	C_x	5_x	6_x	B_x	9_x	0_x	A_x	D_x	3_x	E_x	F_x	8_x	4_x	7_x	1_x	2_x
$S_P^{-1}[i]$	5_x	E_x	F_x	8_x	C_x	1_x	2_x	D_x	B_x	4_x	6_x	3_x	0_x	7_x	9_x	A_x