A FAMILY OF 6-TO-4-BIT S-BOXES WITH LARGE LINEAR BRANCH NUMBER

Daniel Loebenberger and Michael Nüsken

2 April 2013

Abstract. We propose a family of 6-to-4-bit S-boxes with linear branch number 3. Since they also fulfill various further desirable properties such S-boxes can serve as a building block for various block ciphers.

Keywords. DES, S-box, branch number

1. Introduction

During all the time of analysis and improvement proposals to DES it seemed like there might be no 6-to-4-bit S-box with linear branch number 3. While almost all S-boxes used so far have differential branch number 2 and only fail at differentials with zero output differences to achieve even differential branch number 3, they only have linear branch number 2. Since Matsui (1994) discovered linear cryptanalysis, several teams have tried to describe a set of properties that makes DES invulnerable to differential and linear cryptanalysis. Kim, Lee, Park & Lee (1994, 1995) use some conditions asking several specific 1-to-1-bit biases to be zero. In contrast, linear branch number 3 (or higher) means that *all* 1-to-1-bit biases are zero. Neither the DES S-boxes nor later replacements like the s^5DES S-boxes found in Kim *et al.* (1994, 1995) have linear branch number 3. DESL, the lightweight variant of DES proposed by Leander, Paar, Poschmann & Schramm (2007), employs a single S-box in place of the eight different ones in DES. Also the DESL S-box has linear branch number 2 only.

Here, we propose the S-box U:

$efgh$	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$U(0efgh0)$	0	9	7	2	B	E	C	5	3	F	D	8	4	1	A	6
$U(0efgh1)$	B	6	8	F	2	1	5	C	D	A	E	3	7	4	0	9
$U(1efgh0)$	E	4	8	D	2	7	1	B	5	A	6	3	9	C	F	0
$U(1efgh1)$	1	D	4	2	F	8	A	7	6	0	9	5	C	B	3	E

Note that we represent four bit strings $efgh$ as hexadecimal digits as usual.

It does have linear branch number 3 and enjoys the properties summarized in Figure 1.1, that ensure good resistance against differential and linear cryptanalysis. Actually, we started scanning for an S-box with the legendary conditions given by Coppersmith (1994) and augmented them with conditions from Leander *et al.* (2007).

S-7 $\mathrm{diff}_S(\Delta x \to \Delta y) \leq \frac{16}{64}$ for $\Delta x \neq 0$.

S-3 $\mathrm{diff}_S(0{*}{*}{*}{*}0 \to 0000) = 0$.

Q1' $\mathrm{diff}_S({*}{*}{*}{*}00 \to 0000) = 0$.

S-4 $\mathrm{diff}_S(\underbrace{\quad}_{\mathrm{wt}\,1} \to \underbrace{\quad}_{\mathrm{wt}\leq 1}) = 0$.

S-5 $\mathrm{diff}_S(001100 \to \underbrace{\quad}_{\mathrm{wt}\leq 1}) = 0$.

Q2$^+$ $|\mathrm{bias}_S(a,b)| \leq \frac{24}{64}$ for $a \neq 0$.

Q3$^+$ $|\mathrm{bias}_S(\underbrace{\quad}_{\mathrm{wt}\,1}, \underbrace{\quad}_{\mathrm{wt}\,1})| = 0$.

Q4$^+$ $|\mathrm{bias}_S(\underbrace{\quad}_{\mathrm{wt}\,k}, \underbrace{\quad}_{\mathrm{wt}\,\ell})| \leq \frac{16}{64}$ when $0 < k + \ell \leq 4$.

Figure 1.1: Summary of conditions for 6-to-4-bit S-boxes.
The properties **S-?** are from Coppersmith (1994), the properties **Q?**$^+$ are stronger forms of the conditions in Leander *et al.* (2007). The property **Q1'** is equivalent to Leander *et al.*'s Condition 1 under **S-3**.

For instance, we used Condition 5 from Leander *et al.* (2007):

$$\textbf{Q5}\ \ |\mathrm{bias}_S(a,b_1)\cdot\mathrm{bias}_S(a,b_2)| \leq \tfrac{240}{64^2} \quad \text{for all } a \in \mathbb{F}_2^6,\ b_1,b_2 \in \mathbb{F}_2^4 \text{ with } \mathrm{wt}(b_1+b_2)=1.$$

Our program, however, also reported S-boxes that violate this particular condition. Among those we found the S-box U. It only has $|\mathrm{bias}_S(a,b_1)\cdot\mathrm{bias}_S(a,b_2)| \leq \frac{384}{64^2}$ for all $a \in \mathbb{F}_2^6$, $b_1,b_2 \in \mathbb{F}_2^4$ with $\mathrm{wt}(b_1+b_2)=1$.[1]

It turns out that this new S-box has good differential and linear properties, including linear branch number 3 implied by **Q3**$^+$. To our knowledge this is the first known 6-to-4-bit S-box with these properties.

Finally, note that all these properties are invariant under many transformations of S-boxes. Thus we are actually talking about a family of $2^6 6! \cdot 2^4 4!$ S-boxes when looking at the high level conditions listed in Section 2 or of a smaller family of $2^6 2!^2 \cdot 2^4 4!$ S-boxes when considering all properties from Figure 1.1.

In the following, we consider properties of the proposed family of S-boxes in the spirit of Saarinen (2012). That work continues many other investigations including Biryukov, De Cannière, Braeken & Preneel (2003); Courtois & Bard (2007); Daemen & Rijmen (2002); Kim *et al.* (1994, 1995); Leander *et al.* (2007). In this text, we do not treat properties like most of the DES design properties in Coppersmith (1994), the conditions in Leander *et al.* (2007), or the ones listed in Figure 1.1 any further.

2. S-box properties

To formulate all conditions we fix the following standard notions. Let $k, \ell \in \mathbb{N}_{>0}$ and consider a candidate k-to-ℓ-bit S-box

$$S\colon \mathbb{F}_2^k \longrightarrow \mathbb{F}_2^\ell.$$

[1] For the considerations in Leander *et al.* (2007) this weakened condition would have still been sufficient for their reasonings. The issue of so far mostly omitted cases involving neighbored active S-boxes is treated by us in a future paper.

2.1. Differential cryptanalysis. For differential cryptanalysis the following notion is central.

DEFINITION 2.1 (Differential probabilities). *Given an input difference $\Delta x \in \mathbb{F}_2^k$ and an output difference $\Delta y \in \mathbb{F}_2^\ell$ we define*

$$\begin{aligned}\operatorname{diff}_S(\Delta x \to \Delta y) &= \operatorname{prob}\left(S(X) \oplus S(X \oplus \Delta x) = \Delta y \,\middle|\, X \xleftarrow{\text{🎲}} \mathbb{F}_2^k\right)\\ &= \frac{1}{2^k}\#\left\{x \in \mathbb{F}_2^k \,\middle|\, S(x) \oplus S(x \oplus \Delta x) = \Delta y\right\}\\ &\in [0,1].\end{aligned}$$

Here, X denotes a uniform random variable with values $\mathbb{F}_2^k$.

This definition matches Definition 1 in Saarinen (2012).

2.2. Linear cryptanalysis. For linear cryptanalysis the bias of a linear expression in inputs and outputs is essential: We write $\langle a \,|\, x\rangle = \bigoplus_i a_i x_i$ for applying a linear form $a \in \mathbb{F}_2^{k\vee}$ to a vector $x \in \mathbb{F}_2^k$. Actually this way, we identify the dual space $\mathbb{F}_2^{k\vee} := \left\{\mathbb{F}_2^k \to \mathbb{F}_2\right\}$ with $\mathbb{F}_2^k$, matching the bilinear form $\langle \cdot \,|\, \cdot \rangle$. As it turns out to be important not to mix vectors and dual vectors, we keep the notation, at least as a reminder.

DEFINITION 2.2. *Given linear forms $a \in \mathbb{F}_2^{k\vee}$ and $b \in \mathbb{F}_2^{\ell\vee}$ we define the bias*

$$\begin{aligned}\operatorname{bias}_S(a,b) &= \operatorname{prob}\left(\langle a \,|\, X\rangle = \langle b \,|\, S(X)\rangle\right) - \operatorname{prob}\left(\langle a \,|\, X\rangle \neq \langle b \,|\, S(X)\rangle\right)\\ &= 2\operatorname{prob}\left(\langle a \,|\, X\rangle = \langle b \,|\, S(X)\rangle\right) - 1\\ &= \frac{1}{2^k}\sum_{x \in \mathbb{F}_2^k} (-1)^{\langle a \,|\, x\rangle}(-1)^{\langle b \,|\, S(x)\rangle}\\ &\in [-1,1],\end{aligned}$$

which is the correlation between the chosen linear forms on input and output of S.

The value used in Definition 2 in Saarinen (2012) equals $\frac{1}{2}\left|\operatorname{bias}_S(\beta_i, \beta_o)\right|$. We prefer our definition — also found in Daemen & Rijmen (2002), for example — since it makes Matsui's piling-up lemma much nicer and also equals the correlation of the selected input and output bits.

2.3. Algebraic attacks. For algebraic attacks polynomial relations of input and output bits are essential. The smaller the degree is, the stronger the attack can be. Actually, at most quadratic and possibly cubic equations seem to be relevant in practice. For example, when looking for cubic equations we try to fulfill

$$\alpha_\square + \sum_i \alpha_i z_i + \sum_{i<j} \alpha_{ij} z_i z_j + \sum_{i<j<k} \alpha_{ijk} z_i z_j z_k = 0$$

for all $z = (x_0, \ldots, x_{k-1}, y_0, \ldots, y_{\ell-1}) \in \mathbb{F}_2^{k+\ell}$ with $y = S(x)$ when S is a k-to-ℓ-bit S-box. For $k = 6$, $\ell = 4$ and degree 3, this is a linear system with 64 equations for the 176 coefficients $[\alpha_*]$. Since we have $u^2 = u$ for $u \in \mathbb{F}_2$, we usually consider only *multilinear* polynomials, ie. polynomials of degree at most 1 with respect to each variable.

DEFINITION 2.3 (General algebraic relations). *An* algebraic relation *is a polynomial* $p \in \mathbb{F}_2[x, y]$ *such that* $p(x, S(x)) = 0$ *for all* $x \in \mathbb{F}_2^k$. *Given* $d \in \mathbb{N}$ *we define the number*

$$\operatorname{dimrel}(S)_d := \dim \left\{ p \in \mathbb{F}_2[x,y] \;\middle|\; \begin{array}{l} p \text{ multilinear } \wedge\ \deg p \le d \wedge \\ \forall x \in \mathbb{F}_2^k \colon p(x, S(x)) = 0 \end{array} \right\}$$

of independent (multilinear) relations for S *in degree* d. *The algebraic degree*

$$\operatorname{algdeg}(S) := \min \{ d \in \mathbb{N}_{\le k} \mid \operatorname{dimrel}(S)_d > 0 \}$$

of S *is the smallest degree that allows a non-trivial algebraic relation.*

In general, $\operatorname{dimrel}(S)_d \ge \sum_{j \le d} \binom{k+\ell}{j} - 2^k$ and $\operatorname{algdeg}(S) \le k$.

We wish to have the number of independent relations as small as possible to avoid algebraic attacks as far as possible. In particular, for DES conditions we would like to have $\operatorname{dimrel}(S)_2 = 0$ and $\operatorname{dimrel}(S)_d = \sum_{j \le d} \binom{10}{j} - 2^6$ for $i \ge 3$.

Saarinen (2012), Definition 3, considers more special relations:

DEFINITION 2.4 (Output-linear algebraic relations). *An* output-linear algebraic relation *is a pair* (p, b) *with a multilinear polynomial* $p \in \mathbb{F}_2[x]$ *and a (non-trivial) linear form* $b \in \mathbb{F}_2^{\ell\vee}$ *so that* $\langle b \,|\, S(x) \rangle = p(x)$ *for all* $x \in \mathbb{F}_2^k$. *The smallest possible degree for* p *is called the* degree $\deg \langle b \,|\, S(\cdot) \rangle$ *of* $\langle b \,|\, S(\cdot) \rangle$. *We define the* output-linear degree

$$\operatorname{outlindeg}(S) := \min \left\{ \deg(\langle b \,|\, S(\cdot) \rangle) \;\middle|\; b \in \mathbb{F}_2^{\ell\vee} \setminus \{0\} \right\}$$

as the least degree needed for p *when varying over all* b. *More details are revealed by the number*

$$\operatorname{dimoutlinrel}(S)_d := \dim \left\{ (p,b) \in \mathbb{F}_2[x] \times \mathbb{F}_2^\ell \;\middle|\; \begin{array}{l} p \text{ multilinear } \wedge\ \deg p \le d \wedge \\ \forall x \in \mathbb{F}_2^k \colon \langle b \,|\, S(x) \rangle = p(x) \end{array} \right\}$$

of independent output-linear relations for S *in degree* d. *Now,* $\operatorname{outlindeg}(S) = \min \{ d \in \mathbb{N}_{\le k} \mid \operatorname{dimoutlinrel}(S)_d > 0 \}$.

In general, $\operatorname{dimoutlinrel}(S)_d \ge \sum_{j \le d} \binom{k}{j} - 2^k$ and $\operatorname{outlindeg}(S) \le k$. Since each output-linear algebraic relation is an algebraic relation $\langle b \,|\, y \rangle - p(x) = 0$ of same degree, we have that $\operatorname{dimoutlindeg}(S)_d > 0$ implies $\operatorname{dimrel}(S)_d > 0$. However, the converse is wrong. In other words, $\operatorname{algdeg}(S) \le \operatorname{outlindeg}(S)$ but equality may not hold. The DES S-box S1 actually has $\operatorname{algdeg}(\mathrm{S1}) = 2$ and $\operatorname{outlindeg}(\mathrm{S1}) = 3$.

2.4. Avalanche effect. Finally, we define the differential and linear branch numbers.

Definition 2.5 (Branch numbers). *The* differential branch number *is defined as*

$$\operatorname{diffbranch}(S) = \min\{\operatorname{wt}(\Delta x) + \operatorname{wt}(\Delta y) \mid \operatorname{diff}_S(\Delta x \to \Delta y) \neq 0\}$$

where $\Delta x \in \mathbb{F}_2^k$, $\Delta y \in \mathbb{F}_2^\ell$, $(\Delta x, \Delta y) \neq (0,0)$. *The* linear branch number *is defined as*

$$\operatorname{linbranch}(S) = \min\left\{\operatorname{wt}(a) + \operatorname{wt}(b) \;\middle|\; \operatorname{bias}_S(a,b) \neq 0\right\}$$

where $a \in \mathbb{F}_2^{k\vee}$, $b \in \mathbb{F}_2^{\ell\vee}$, $(a,b) \neq (0,0)$.

The larger the branch numbers are, the stronger the avalanche effect should be. Saarinen (2012), Definition 4, only considers differential branch numbers.

3. Equivalence classes

Given one S-box S we can derive others by affine transformation:

$$T\colon \begin{array}{ccc} \mathbb{F}_2^k & \longrightarrow & \mathbb{F}_2^\ell, \\ x & \longmapsto & JS(Ix+s)+t, \end{array}$$

where $I\colon \mathbb{F}_2^k \to \mathbb{F}_2^k$, $J\colon \mathbb{F}_2^\ell \to \mathbb{F}_2^\ell$ are invertible linear maps and $s \in \mathbb{F}_2^k$, $t \in \mathbb{F}_2^\ell$.

Definition 3.1 (Equivalence). *We call* S *and* T linear-affine equivalent *if there is an affine transformation.*

We call S *and* T permutation-affine equivalent *if there is an affine transformation where the matrices are even permutation matrices.*

Note that permutation matrices are exactly those invertible linear maps that respect the weight. One can check that two permutation equivalent S-boxes S and T have strongly related differential probabilities and biases. Namely, considered as matrices $\operatorname{diff}_S(\cdot \to \cdot)$ and $\operatorname{diff}_T(\cdot \to \cdot)$ are obtained from each other by permuting rows and columns according to the permutations I and J^{-1}; for $\operatorname{bias}_S(\cdot,\cdot)$ and $\operatorname{bias}_T(\cdot,\cdot)$ additionally some signs change depending on the shifts s and t. The algebraic quantities, namely $\operatorname{dimrel}(S)$, $\operatorname{algdeg}(S)$, $\operatorname{dimoutlinrel}(S)$ and $\operatorname{outlindeg}(S)$, are even invariant under linear-affine transformation. This still holds for the finer multiset $\left\{\deg\langle b \mid S(\cdot)\rangle \;\middle|\; b \in \mathbb{F}_2^4 \setminus \{0\}\right\}$. The branch numbers do not change under permutation equivalence. Summarizing: all quantities introduced in Section 2 are essentially invariant under permutation-affine transformations.

4. Properties of the proposed S-box

The differential probabilities are found in Figure 4.2. The different colors mark values affected by different properties from Figure 1.1. Figure 4.2 shows that diffbranch$(U) = 2$. Actually, the orange area, reflecting **S4**, and its surrounding shows that it only fails marginally to reach 3: There are only five differentials $\Delta x \to 0000$ with $\mathrm{wt}(\Delta x) = 2$ that have non-zero probability, there is no nontrivial differential $\Delta x \to \Delta y$ with $\Delta y \neq 0$ and $\mathrm{wt}(\Delta x) + \mathrm{wt}(\Delta y) = 2$. Most of this behavior is part of the design properties for the DES S-boxes.

The biases are found in Figure 4.3. Note that for all $a \in \mathbb{F}_2^{6\vee}$ and $b \in \mathbb{F}_2^{4\vee}$ with a, b not both zero, we have $|\mathrm{bias}_U(a,b)| \leq \frac{24}{64}$, improving the previous best bound $\frac{28}{64}$ again. The yellow area is all zero due to $\mathbf{Q3}^+$ and that implies that U has

$$\mathrm{linbranch}(U) = 3.$$

All earlier 6-to-4-bit S-boxes examined only have linear branch number 2.

Next, we consider the number of independent algebraic relations:

$$\mathrm{dimrel}(U) = [0, 0, 0, 112, 322, \dots].$$

These are the minimal possible numbers: $\mathrm{dimrel}(U)_d = \sum_{j \leq d} \binom{10}{j} - 2^6$ for $d \geq 3$. Consequently, we have the optimal value

$$\mathrm{algdeg}(U) = 3.$$

When restricting to special relations we find

$$\mathrm{outlindeg}(U) = 4.$$

Optimal would be 5, but this is only achieved by some DES S-boxes which are worse

Property	Optimal	U	DESL	DES1	DES2	DES3	DES4	DES5	DES6	DES7	DES8
diffbranch	2?	2	2	2	2	2	2	2	2	2	2
linbranch	3?	3	2	2	2	2	2	2	2	2	2
algdeg	3	3	2	2	3	3	2	2	3	3	3
dimrel_2	0	0	1	1	0	0	5	1	0	0	0
outlindeg	5	4	4	4	4	4	3	4	5	5	4
$\mathrm{dimoutlinrel}_3$	0	0	0	0	0	0	1	0	0	0	0
$\mathrm{dimoutlinrel}_4$	0	4	2	1	2	1	3	1	0	0	1
$\mathrm{dimoutlinrel}_5$	3	4	4	4	4	4	4	4	4	4	4

Figure 4.1: Comparison of 6-to-4-bit S-boxes

in other aspects. Actually, there are four independent (ie. $2^4 - 1$ in total) degree 4 relations. Equivalently, for $b \in \mathbb{F}_2^4 \setminus \{0\}$ each $\langle b \,|\, U(\cdot) \rangle$ has degree 4. However, degree 4 is already a very high single round degree for algebraic attacks, anyways.

$\xi \setminus \eta$	0000	0001	0010	0100	1000	0011	0101	0110	1001	1010	1100	0111	1011	1101	1110	1111
000000	64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000001	.	.	.	.	.	6	6	2	8	4	6	2	6	4	4	16
000010	.	.	.	.	.	4	16	4	8	4	8	8	.	8	.	4
000100	.	.	.	.	.	8	8	4	8	.	4	12	4	4	4	8
001000	.	.	.	.	.	6	4	4	6	8	4	10	8	2	12	.
010000	.	.	.	.	.	8	8	8	4	4	4	4	8	4	8	4
100000	.	.	.	.	.	2	4	6	6	4	2	4	12	12	4	8
000011	8	4	6	.	4	8	4	10	4	10	4	.	.	.	2	.
000101	8	.	8	6	6	6	.	4	2	4	6	.	2	4	6	2
000110	.	.	12	8	4	8	.	4	8	.	4	.	4	.	12	.
001001	2	6	6	6	4	2	8	2	6	6	4	.	4	4	2	2
001010	.	.	10	4	2	8	.	6	8	4	6	4	4	4	4	.
001100	.	.	.	.	.	2	4	.	10	16	8	2	4	2	8	8
010001	2	6	4	6	6	2	6	4	4	4	8	6	.	2	2	2
010010	.	12	.	12	.	.	.	4	.	8	4	.	12	.	4	8
010100	.	4	8	4	16	.	8	4	.	.	4	8	.	8	.	.
011000	.	4	4	4	12	2	.	4	2	.	4	10	4	10	.	4
100001	10	2	6	8	10	2	4	6	4	2	.	2	.	2	2	4
100010	.	.	2	4	6	.	.	4	4	16	8	6	4	2	8	.
100100	.	2	8	10	8	4	4	4	4	8	4	.	2	.	2	4
101000	.	10	6	6	6	4	4	2	4	4	6	.	2	4	6	.
110000	.	8	6	.	2	4	8	8	.	8	4	2	4	6	4	.
000111	6	10	.	4	.	2	8	2	4	.	6	8	2	6	2	4
001011	2	2	.	2	8	4	2	6	6	4	2	6	6	4	4	6
001101	6	2	2	4	2	10	10	8	4	6	4	6	.	.	.	.
001110	.	.	2	4	2	4	16	2	8	8	6	4	.	4	.	4
010011	2	2	4	10	6	6	8	2	4	4	2	2	4	2	2	4
010101	6	10	.	.	8	2	.	2	2	4	.	8	8	6	8	.
010110	.	12	12	12	12	.	.	.	.	.	.	.	8	8	.	.
011001	.	4	10	8	6	2	.	4	2	2	2	8	6	2	4	4
011010	.	12	10	4	10	.	.	2	.	4	2	.	.	8	4	8
011100	.	8	4	8	4	6	.	8	2	4	4	2	4	2	.	8
100011	2	4	.	2	6	6	10	4	2	.	.	4	4	10	2	8
100101	.	2	.	4	4	6	2	6	8	6	.	4	14	4	.	4
100110	4	2	2	2	2	6	8	2	6	.	2	6	6	6	6	4
101001	.	10	4	6	2	8	4	6	4	4	2	2	.	4	4	4
101010	12	2	2	2	2	4	4	6	4	.	2	8	2	4	2	8
101100	.	12	10	8	2	6	4	4	2	.	4	4	.	.	4	4
110001	2	4	6	.	.	4	4	4	6	4	6	4	6	2	10	2
110010	8	4	.	4	.	6	.	2	2	4	6	12	8	4	.	4
110100	.	6	6	2	6	6	4	2	6	4	2	2	2	2	6	8
111000	.	6	8	10	4	2	4	4	6	4	4	2	6	2	2	.
001111	8	.	2	2	.	2	2	6	6	6	8	2	4	2	4	10
010111	.	.	6	6	6	4	.	2	.	2	4	10	10	4	6	4
011011	.	8	10	8	2	2	6	2	2	2	.	.	2	10	8	2
011101	8	4	2	2	.	2	2	6	4	2	2	2	6	10	6	6
011110	.	12	2	4	2	8	.	10	.	4	2	.	4	.	8	8
100111	4	4	6	6	.	2	.	8	2	8	8	2	2	4	8	.
101011	2	6	.	6	8	10	.	6	.	.	4	2	2	2	6	10
101101	2	6	2	2	8	4	2	2	.	12	6	4	2	6	2	4
101110	16	4	2	.	6	6	4	4	2	.	4	4	4	4	.	4
110011	6	2	6	2	4	6	2	4	4	.	6	.	8	2	8	4
110101	8	4	4	.	2	.	10	4	2	.	6	6	.	8	8	2
110110	8	2	4	2	4	4	.	4	4	.	4	4	2	4	10	8
111001	4	.	4	6	4	2	4	.	2	2	8	8	2	4	4	10
111010	8	2	4	2	.	2	8	4	6	4	4	6	2	6	2	4
111100	.	4	4	12	4	4	.	2	4	.	6	2	4	6	4	8
011111	6	6	4	.	6	4	2	2	6	4	6	4	4	4	4	2
101111	4	6	6	6	2	2	2	2	4	8	4	4	.	8	.	6
110111	8	2	8	2	2	2	8	.	8	4	2	6	6	4	2	.
111011	6	4	6	2	10	2	4	2	6	8	2	2	6	2	.	2
111101	6	4	6	6	2	6	2	4	6	2	4	2	8	2	2	2
111110	8	.	.	.	12	4	8	6	4	8	2	2	4	2	4	.
111111	.	4	.	6	.	2	6	6	6	4	6	12	4	.	6	2

Figure 4.2: $2^6 \cdot \mathrm{diff}_U(\xi \to \eta)$

$a \setminus b$	0000	0001	0010	0100	1000	0011	0101	0110	1001	1010	1100	0111	1011	1101	1110	1111
000000	64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000001	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000100	.	.	.	.	.	.	−8	.	8	8	.	−8	.	.	8	−8
001000	.	.	.	.	.	−8	8	.	8	−8	.	16	8	−16	−8	.
010000	.	.	.	.	.	.	8	.	.	.	.	−8	−16	8	.	−24
100000	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000011	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
000101	.	.	.	.	.	.	−8	.	−8	8	.	−8	16	16	−24	8
000110	.	−8	−8	.	8	.	.	−8	8	−8	8	8	−8	16	−8	.
001001	.	.	.	.	.	8	8	.	−8	−8	.	.	8	.	−8	.
001010	.	.	8	−8	.	.	−16	16	−8	.	−8	16	.	8	8	−16
001100	.	.	.	.	.	−8	.	.	.	16	.	8	8	.	16	24
010001	.	.	.	.	.	.	−8	.	.	.	.	8	16	24	.	−8
010010	.	8	.	16	.	8	.	.	−8	16	.	.	−8	.	.	.
010100	.	.	−8	8	.	−8	−8	16	8	.	8	16	8	.	−8	.
011000	.	8	8	.	.	8	−8	8	16	.	.	−8	8	.	.	−8
100001	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
100010	.	.	.	.	.	16	16	−16	.	.	.	16	16	16	16	−16
100100	.	−8	−8	8	8	.	−8	.	8	8	.	.	8	−8	.	−8
101000	.	−8	−16	8	−16	16	8	8	.	−8	8	.	.	.	.	.
110000	.	.	.	.	.	.	−8	.	.	.	.	8	−16	24	.	24
000111	.	8	8	16	8	.	.	−8	8	8	−8	24	8	.	−8	.
001011	.	.	−8	−8	16	.	16	.	−8	.	8	16	.	8	8	16
001101	.	.	.	.	.	8	.	.	.	16	.	−8	−8	.	−16	8
001110	.	−8	.	8	8	.	.	−8	−16	8	16	8	−8	−8	.	.
010011	.	8	.	.	.	8	.	16	−8	−16	16	.	−8	.	16	.
010101	.	16	−8	−8	.	8	8	.	8	.	−8	.	8	.	8	.
010110	.	.	−16	8	−8	.	8	−8	16	16	16	−8	−8	8	8	.
011001	.	8	8	.	.	−8	8	8	.	.	.	−8	8	.	.	8
011010	.	.	.	8	.	24	8	8	−24	8	−8	.	8	.	.	.
011100	.	−8	16	−8	.	.	−24	−8	−8	.	−8	.	.	8	8	.
100011	.	.	.	.	.	16	−16	−16	.	.	.	−16	16	−16	16	16
100101	.	−8	8	8	−8	16	−8	16	8	8	16	16	−8	−8	.	8
100110	.	.	.	8	.	16	.	−8	8	−8	−8	.	.	−8	.	.
101001	.	−8	16	8	16	.	8	8	16	−8	8	−16	.	16	.	.
101010	.	−8	8	.	.	−8	.	−8	.	.	16	.	8	−8	.	.
101100	.	.	8	.	−8	.	.	−8	8	−16	−8	.	−8	8	.	24
110001	.	.	.	.	.	.	8	.	.	.	.	−8	16	8	.	8
110010	.	−8	.	.	.	8	.	.	8	−16	−16	.	−8	.	.	.
110100	.	8	.	16	−8	−8	−8	.	8	.	−8	8	.	8	16	.
111000	.	.	−8	8	16	.	8	16	8	.	−24	−8	.	.	8	8
001111	.	8	.	−8	−8	.	.	8	−16	−8	16	−8	8	8	.	.
010111	.	.	.	−8	−8	−16	8	−8	.	.	.	8	−8	−8	8	.
011011	.	.	−16	−8	−16	−8	8	8	8	8	−8	.	8	.	−16	.
011101	.	8	16	8	.	.	−8	8	8	.	8	.	.	−8	−8	.
011110	.	8	.	.	−8	−16	.	16	.	8	8	−8	8	8	8	.
100111	.	16	.	−8	16	.	.	8	−8	8	−8	.	−16	−8	.	.
101011	.	−8	−8	.	16	−8	.	8	.	.	.	.	8	−8	.	.
101101	.	.	−8	.	8	.	.	8	−8	−16	8	.	8	−8	.	8
101110	.	8	8	8	.	8	.	.	8	8	8	.	8	.	.	.
110011	.	−8	.	−16	.	8	.	16	8	16	.	.	−8	.	16	.
110101	.	−8	16	.	−24	−8	8	.	−8	.	−8	8	.	−8	.	.
110110	.	8	8	.	.	.	8	−8	.	16	.	.	.	.	.	.
111001	.	.	−8	8	16	−16	−8	−16	−8	.	8	−8	.	.	8	−8
111010	.	8	.	.	.	.	8	.	.	8	.	.	.	.	8	.
111100	.	8	8	−8	8	8	8	.	.	.	.	8	.	.	−8	.
011111	.	8	.	−16	8	.	.	.	16	−8	8	8	8	−8	8	.
101111	.	24	−8	24	.	−8	.	.	−8	−8	−8	.	−8	.	.	.
110111	.	−24	8	16	16	.	8	8	.	.	.	.	.	.	.	.
111011	.	8	16	−16	16	.	8	.	.	8	.	.	.	.	−8	.
111101	.	−8	24	8	−8	−8	24	.	.	.	.	−8	.	.	8	.
111110	.	−8	−8	−16	.	8	.	−8	8	8	.	.	−8	.	−8	.
111111	.	24	8	.	.	8	.	−8	8	−8	16	.	−8	.	−8	.

Figure 4.3: $2^6 \cdot \mathrm{bias}_U(a, b)$

5. Bijective 4-to-4-bit sub-boxes

We briefly analyze the bijective 4-to-4-bit sub-boxes of the S-box U.

Like any DES-suitable S-box, the S-box U is composed of four bijective 4-to-4-bit S-boxes, namely, $U(a****f)$ for $a, f \in \mathbb{F}_2$. Their canonical representatives in Saarinen's language are:

af	$U(a****f)$	permutation-affine canonical form	linear-affine canonical form
00	0972BEC53FD841A6	035F78E1BD24C69A	012345768A9BCEFD
01	B68F215CDAE37409	035674ED9F28CAB1	012345896ACEFDB7
10	E48D271B5A639CF0	03596AFCB42ED187	012345768A9BCEFD
11	1D42F8A76095CB3E	0358749EF6AD2BC1	012345768ACE9BFD

It turns out that the first and the third are linear-affine equivalent.

Due to property **Q1'** there is a second way to compose the S-box U from four 4-to-4-bit S-boxes, namely by taking $U(****ef)$ for $e, f \in \mathbb{F}_2$. These are:

ef	$U(****ef)$	permutation-affine canonical form	linear-affine canonical form
00	07BC3D4AE821569F	0358A46FE9B7D21C	012345896ABCE7DF
01	B825DE7014FA69C3	0356789FDABCE142	012345786ABCE9FD
10	92E5F8164D7BA3C0	0358A46FE9B7D21C	012345896ABCE7DF
11	6F1CA349D28705BE	0356789FADCB1E24	012345786ABCE9FD

Here, the first and third are even permutation-affine equivalent and the second and fourth are linear-affine equivalent.

These eight 4-to-4-bit S-boxes are the only bijective 4-to-4-bit sub-boxes of U. Even, when allowing linear-affine sub-boxes there is no further bijective 4-to-4-bit sub-box of U. None of the them is linear-affine equivalent to a golden S-box in terms of Saarinen (2012).

6. How to find good boxes

In order to find good S-boxes, we have written a C program that searches in a depth first manner the tree of all partially defined 6-to-4-bit S-boxes, where the leaves are all totally defined boxes. Thus the tree has depth k and each node has 2^4 child nodes, representing one new value for a so far undefined position. Since the complete tree has $(2^4)^{2^6} = 2^{256}$ leaves it is obviously infeasible to traverse the whole tree. Also, many S-boxes within this tree are trivially unsuitable or do not fulfill the necessary properties like the ones stated in Coppersmith (1994) or Leander *et al.* (2007).

Our search algorithm computes the table of differential probabilities and the bias tables incrementally along the path from the root to a leaf. During this progress the values for the differential probabilities increase monotonically by $0/2^6$, $1/2^6$ or

$2/2^6$ in each step and table position. Once a differential probability has bypassed a bound from a condition, no completion can fulfill that condition any more and thus that entire branch can be purged from the tree. In contrast, the values for the biases change by $\pm 1/2^6$ in each step. Thus a similar technique for the bias values is more tricky. Finally, conditions like **Q5**, involving products of bias values, can only be checked at the leaves. These techniques allow us to reduce the size of the tree to estimated 2^{48} S-boxes, which now seems feasible. After 14.66 CPU-years on 12 to 16 Intel Xeon 3.00GHz processors, it has scanned an estimated fraction of 2–7% of the purged search tree. (The large interval is due to the inherent difficulty of telling the size of the subtrees below a given node. We thus use two different heuristics to estimate the processed fraction of the tree.) We estimate that the algorithm would finish with the whole tree in roughly 12 to 39 years on our tiny cluster. However, we do not expect to find even better S-boxes nor S-boxes inequivalent to U.

7. Conclusion

We propose a family of 6-to-4-bit S-boxes with linear branch number 3. Besides that extraordinary feature, its biases are bounded by $\frac{24}{64}$, which is very small, and it also fulfills most other design criteria of Coppersmith (1994) and Leander *et al.* (2007). It may thus serve as a building block for DES-like ciphers.

We briefly discussed the structure of bijective 4-to-4-bit sub-boxes and observed that these sub-boxes are not golden but in a surprising way closely related to each other.

Acknowledgements

This work was funded by the B-IT foundation and the state of North Rhine-Westphalia.

References

Alex Biryukov, Christophe De Cannière, An Braeken & Bart Preneel (2003). A Toolbox for Cryptanalysis: Linear and Affine Equivalence Algorithms. In *Advances in Cryptology: Proceedings of EUROCRYPT 2003,* Warsaw, Poland, Eli Biham, editor, volume 2656 of *Lecture Notes in Computer Science*, 33–50. Springer-Verlag. ISBN 978-3-540-14039-9. URL `http://dx.doi.org/10.1007/3-540-39200-9_3`. 0302-9743.

Don Coppersmith (1994). The Data Encryption Standard (DES) and its strength against attacks. *IBM Journal of Research and Development* **38**(3), 243–250. URL `http://dx.doi.org/10.1147/rd.383.0243`.

Nicolas T. Courtois & Grgory V. Bard (2007). Algebraic Cryptanalysis of the Data Encryption Standard. In *Cryptography and Coding 11th IMA International Conference, Cirencester, UK, December 18-20, 2007.*, Steven D. Galbraith, editor, volume 4887 of

Lecture Notes in Computer Science, 152–169. Springer-Verlag, Berlin, Heidelberg. ISBN 978-3-540-77272-9. ISSN 0302-9743. URL http://dx.doi.org/10.1007/978-3-540-77272-9_10.

Joan Daemen & Vincent Rijmen (2002). *The Design of Rijndael. AES - The Advanced Encryption Standard*. Springer, Berlin, Heidelberg, New York. ISBN 3-540-42580-2.

Kwang-jo Kim, Sang-jin Lee, Sang-jun Park & Dai-ki Lee (1994). DES can be Immune to Linear Cryptanalysis. In *Proceedings of the Workshop on Selected Areas in Cryptography SAC'94*, 70–81. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.37.711.

Kwangjo Kim, Sangjin Lee, Sangjun Park & Daiki Lee (1995). Securing DES S-boxes against Three Robust Cryptanalysis. In *Proceedings of the Workshop on Selected Areas in Cryptography SAC '95*, 145–157.

Gregor Leander, Christof Paar, Axel Poschmann & Kai Schramm (2007). New Lightweight DES Variants. In *Fast Software Encryption 2007, 14th International Workshop, FSE 2007,* Luxembourg, Luxembourg, Alex Biryukov, editor, volume 4593 of *Lecture Notes in Computer Science*, 192–210. Springer-Verlag, Berlin, Heidelberg. ISBN 3-540-74617-X. ISSN 0302-9743. URL http://dx.doi.org/10.1007/978-3-540-74619-5_13.

Mitsuru Matsui (1994). Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology: Proceedings of EUROCRYPT 1993,* Lofthus, Norway, Tor Helleseth, editor, volume 765 of *Lecture Notes in Computer Science*, 386–397. Springer-Verlag, Heidelberg. ISBN 3-540-57600-2. ISSN 0302-9743. URL http://dx.doi.org/10.1007/3-540-48285-7_33.

Markku-Juhani O. Saarinen (2012). Cryptographic Analysis of All 4×4-Bit S-Boxes. In *Selected Areas in Cryptography 2011*, Ali Miri & Serge Vaudenay, editors, number 7118 in Lecture Notes in Computer Science, 118–133. Springer-Verlag, Heidelberg, Dordrecht, London, New York. ISBN 978-3-642-28495-3, e-ISBN 978-3-642-28496-0. ISSN 0302-9743, e-ISSN 1611-3349. URL http://dx.doi.org/10.1007/978-3-642-28496-0_7. Preprint available at http://eprint.iacr.org/2011/218.

Daniel Loebenberger
b-it
Universität Bonn
Dahlmannstr. 2
D53113 Bonn
daniel@bit.uni-bonn.de
http://cosec.bit.uni-bonn.de/~loebenberger/

Michael Nüsken
b-it
Universität Bonn
Dahlmannstr. 2
D53113 Bonn
nuesken@bit.uni-bonn.de
http://cosec.bit.uni-bonn.de/~nuesken/