Linear and Differential Properties of S-boxes with Respect to Modular Addition

Matúš Jókay  Peter Špaček  Pavol Zajac

Institute of Computer Science and Mathematics
Slovak University of Technology

pavol.zajac@stuba.sk

Central European Conference on Cryptology 2019

1Supported by grant VEGA 1/0159/17.
## Outline

### Introduction

### Definitions
- D-spectrum
- L-spectrum

### Modular affine equivalence (MAE)

### Experimental results
- All MAE classes
- Optimal S-boxes

### Summary
Introduction

S-boxes are typically studied in the context of Boolean functions:

- \( S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m \),
- Linear profile: \( Pr(a^T \cdot x = b^T \cdot S(x)) \).
- Differential profile: \( Pr(S(x) \oplus S(x \oplus \delta_x) = \delta_y) \).
- Small S-boxes can be easily characterised using affine equivalence \(^2\) (302 classes):

\[
S_2(x) = A_1 \cdot S_1(A_2 \cdot x \oplus b_2) \oplus b_1
\]

---

Modular S-box properties

Research question
What properties have small bijective S-boxes with respect to modular addition?

Research question refinement
Do the modular properties depend on the quality of S-box w.r.t. standard S-box criteria?
Motivation

• Alternative cipher designs:
  • Rotor machines: clocking can be expressed as $S(x + t)$, with $+$ over some $\mathbb{Z}_n$
  • GOST, Kalyna (and others): Key addition or linear layer with $+$ over some $\mathbb{Z}_{2^n}$

• Theoretical generalizations of non-linearity properties

• Attacks based on alternative operations

---

4 Calderini M., Sala M.: Elementary abelian regular subgroups as hidden sums for cryptographic trapdoors. 2017
5 Civino R, Blondeau C, Sala M. Differential attacks: using alternative operations. 2019
Notation

- We work in ring $\mathbb{Z}_{2^n} = (\mathbb{Z}/2^n\mathbb{Z})$
- Addition/subtraction: $+/-$
- Multiplication: $ax$
- Division: $x/a = a^{-1}x$, for $a$ with $\gcd(a, 2) = 1$
- Affine permutations:

$$A(x) = ax + b, \text{ with } \gcd(a, 2) = 1$$
Differential properties

- Table of differences:
  \[ D_{(d_x, d_y)} = \left| \{ x, S(x + d_x) - S(x) = d_y \} \right| \]

- D-spectrum: multiset \( \{ D_{(d_x, d_y)} \} \)

- D-criterium: \( D(S) = \max \{ D_{(d_x, d_y)} \} \)

- Affine function: \( D(f) = 2^n \)
Linear properties

- Linear approximation:
  \[ L_{(a,b)} = \left| \{ x, S(x) = ax + b \} \right| \]

- L-spectrum: multiset \( \{ L_{(a,b)} \} \)

- L-criterium: \( L(S) = \max \{ L_{(a,b)} \} \)

- Affine function: \( L(f) = 2^n \)
Modular affine equivalence

To explore (modular) S-box properties, we can use (modular) affine equivalence (MAE):

\[ S_1 \equiv S_2 \text{ iff } A_1 \circ S_1 = S_2 \circ A_2 \]

Explicitly:

\[ \forall x : S_2(x) = a_1 \cdot S_1(a_2 \cdot x + b_2) + b_1 \]

S-box criteria \( L(S) \) and \( D(S) \) are invariant under MAE.
Modular affine equivalence

- **Class size:** at most $2^{4n-2}$
  - $n = 3$: 58 classes
  - $n = 4$: 1277100855 classes ($\approx 2^{30}$)

- **Representatives:**
  - can always normalize to $S(0) = 0$, $S(1) = 1$
  - representative is the first S-box in lex order
Modular S-box properties and affine equivalence

Research question reformulation
What is the statistical distribution of L- and D-criterium in MAE classes of small S-boxes?
All MAE classes

Statistics of class representatives based on exhaustive enumeration of 4-bit S-boxes:

<table>
<thead>
<tr>
<th>D \ L</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.78%</td>
<td>5.04%</td>
<td>2.10%</td>
<td>0.20%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>4.84%</td>
<td>30.44%</td>
<td>15.09%</td>
<td>2.77%</td>
<td>0.22%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>2.82%</td>
<td>15.92%</td>
<td>7.94%</td>
<td>1.89%</td>
<td>0.36%</td>
<td>0.03%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.70%</td>
<td>3.78%</td>
<td>2.44%</td>
<td>0.83%</td>
<td>0.24%</td>
<td>0.05%</td>
<td>0.00%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.12%</td>
<td>0.53%</td>
<td>0.28%</td>
<td>0.10%</td>
<td>0.04%</td>
<td>0.02%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.02%</td>
<td>0.10%</td>
<td>0.13%</td>
<td>0.07%</td>
<td>0.03%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
All MAE classes

Modular S-box criteria in numbers:

- 30% of S-boxes: $D = 4, L = 4$
- 95% of S-boxes: $D, L \in \{4, 5\}$
- 0.5% of S-boxes: $D \geq 8$ or $L \geq 8$
- $L = 2, D = 3$: 170 classes
- $L = 3, D = 2$: 411 classes
Selected S-boxes

Selected 4-bit S-boxes from (Saarinen, 2011)\(^6\):

- \(D, L \in \{3, 4, 5, 6, 7\}\), most of them: \(L = 4, D = 4\)
- DES S5-1: \(D = 7, L = 4\) (0.53\%):
  \[
  Pr(S(x + 3) - S(x) = 8) = 7/16
  \]
- GOST K8: \(D = 5, L = 7\) (0.36\%):
  \[
  Pr(S(x) = 5x + 1) = 7/16
  \]
- HAMSI, Serpent S2 (G1): \(D = 7, L = 3\) (0.12\%)

---

\(^6\)Saarinen MJ. Cryptographic analysis of all 4 x 4-bit S-boxes. SAC 2011.
Modular S-box properties and optimal S-box classes

Research question reformulation
What is the statistical distribution of L- and D-criterium in MAE classes of small S-boxes?

Additional question
What is the statistical distribution of L- and D-criterium in case of optimal S-boxes (in 16 optimal LA classes)?
Technical note

• To explore all S-boxes in 16 optimal classes would take $1303 \times$ more time than to explore all class representatives.

• Our computation:
  
  • Let $Aff = \{\mathcal{A}; \mathcal{A}(x) = A \cdot x \oplus c\}$,
  
  • $Aff_L$ contains reps. of $a\mathcal{A}(x) + b$ — 20160 permutations
  
  • $Aff_R$ contains reps. of $\mathcal{A}(ax + b)$ — 20160 permutations
  
  • compute $Aff_L \circ S \circ Aff_R$
All optimal classes

- Best S-boxes have always \((D, L) = (2, 3)\), or \((D, L) = (3, 2)\)
- Maximum \(L\) is 11 (G7, G9, G10, G13), or 12
- Maximum \(D\) is 12 (G1, G3, G7, G9, G10, G11, G15), or 13
Class G3 (finite field inverse)

<table>
<thead>
<tr>
<th>D</th>
<th>L</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.83%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>4.69%</td>
<td>11.31%</td>
<td>16.08%</td>
<td>2.96%</td>
<td>0.23%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>2.50%</td>
<td>15.19%</td>
<td>7.85%</td>
<td>1.86%</td>
<td>0.35%</td>
<td>0.03%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.55%</td>
<td>3.22%</td>
<td>1.99%</td>
<td>0.55%</td>
<td>0.19%</td>
<td>0.04%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.09%</td>
<td>0.44%</td>
<td>0.23%</td>
<td>0.08%</td>
<td>0.03%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.01%</td>
<td>0.07%</td>
<td>0.08%</td>
<td>0.05%</td>
<td>0.02%</td>
<td>0.01%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- \(D\) and \(L\) in range 2 to 12
- 95.75% of S-boxes with \(D, L \in \{4, 5\}\)
- 0.34% of S-boxes with \(D \geq 8\) or \(L \geq 8\)
S-box distribution within classes

![Graph showing S-box distribution within classes](image-url)
Experimental results summary:

1. Optimal S-boxes w.r.t. standard linear and differential cryptanalysis have similar properties w.r.t. modular addition (with all classes and between them).

2. A small fraction of S-boxes optimal w.r.t. standard linear and differential cryptanalysis have very bad properties w.r.t. modular addition.
Open questions

- General theoretical analysis and good algebraic constructions?
- What about other operations, are there S-boxes good against every approximation?
- Can we break standard SL designs with bad modular S-boxes?
- Can weak modular S-boxes be used to backdoor\(^7\) cipher designs?

\(^7\)A Biryukov, L Perrin, A Udovenko: Reverse-Engineering the S-Box of Streebog, Kuznyechik and STRIBOBr1, 2016