Tools in analyzing linear approximation for Boolean functions related to FLIP

Subhamoy Maitra[1], Bimal Mandal[1], Thor Martinsen[2],
Dibyendu Roy[3] and Pantelimon Stănică[2]

[1] Indian Statistical Institute, Kolkata, India
[2] Naval Postgraduate School, Monterey, USA
[3] National Institute of Science Education and Research (HBNI), Bhubaneswar, India

Indocrypt, 2018

Outline

- Preliminaries
  - Boolean function
  - FLIP design
  - Existing results
- Our approach
- Biased Walsh–Hadamard transform
- Comparisons with the existing work
  - For a small Boolean function
  - For the actual nonlinear filter function of $\mathrm{FLIP}_{530}(42, 128, 360)$
- Conclusion

Preliminaries: Boolean functions

Preliminaries: Boolean functions

- $\mathbb{F}_2 = \{0, 1\}$ is a binary field.
- $\mathbb{F}_2^n = \{\mathbf{x} = (x_1, x_2, \ldots, x_n) : \ x_i \in \mathbb{F}_2, \ 1 \leq i \leq n\}$ is an n-dimensional vector space over $\mathbb{F}_2$.
- The weight of an element $\mathbf{x} \in \mathbb{F}_2^n$, $wt(\mathbf{x})$, is defined as $wt(\mathbf{x}) = \sum_{i=1}^{n} x_i$, the sum is over $\mathbb{Z}$.
- Any function $f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ is said to be a Boolean function involving n-variables.
- The set of all Boolean functions in n variables is denoted by $\mathcal{B}_n$.
- Algebraic normal form (ANF) of $f \in \mathcal{B}_n$ is defined as

$$f(x_1, x_2 \ldots, x_n) = \sum_{\mathbf{a} \in \mathbb{F}_2^n} \mu_{\mathbf{a}} \prod_{i=1}^{n} x_i^{a_i}.$$

Preliminaries: Boolean functions

- The algebraic degree of $f \in \mathcal{B}_n$ is defined as

$$\deg(f) = \max_{\mathbf{a} \in \mathbb{F}_2^n} \{wt(\mathbf{a}) : \mu_{\mathbf{a}} \neq 0\}.$$

- If algebraic degree of $f \in \mathcal{B}_n$ is at most 1 then the function is said to be affine function. The set of all affine functions involving n-variables is denoted by,

$$\mathcal{A}_n = \{l_{\mathbf{a},\varepsilon} : \ \mathbf{a} \in \mathbb{F}_2^n, \varepsilon \in \mathbb{F}_2\}, \text{ where } l_{\mathbf{a},\varepsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + \varepsilon.$$

- The correlation between two Boolean functions $f, g \in \mathcal{B}_n$ is defined by

$$corr(f, g) = \left| \frac{|\{\mathbf{x} : \ f(\mathbf{x}) = g(\mathbf{x})\}| - |\{\mathbf{x} : \ f(\mathbf{x}) \neq g(\mathbf{x})\}|}{2^n} \right|.$$

Preliminaries: Boolean functions

- The Walsh–Hadamard transform of $f \in \mathcal{B}_n$ at $\mathbf{a} \in \mathbb{F}_2^n$ is defined by

$$\mathcal{W}_f(\mathbf{a}) = \frac{1}{2^n} \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{f(\mathbf{x}) + \mathbf{a} \cdot \mathbf{x}} \ (= \pm corr(f, l_{\mathbf{a},0})).$$

- Parseval's identity:

$$\sum_{\mathbf{a} \in \mathbb{F}_2^n} \mathcal{W}_f^2(\mathbf{a}) = 1.$$

- $E_{n,i} = \{\mathbf{x} \in \mathbb{F}_2^n : \ wt(\mathbf{x}) = i\}$, for all $0 \leq i \leq n$.
- The correlation between two Boolean functions $f, l_{\mathbf{a},0} \in \mathcal{B}_n$ in a restricted domain $E_{n,k}, \ 0 \leq k \leq n$ is defined by

$$\begin{aligned} corr^{(k)}(f, l_{\mathbf{a},0}) &= \left| \frac{|\{\mathbf{x} : f(\mathbf{x}) = l_{\mathbf{a},0}(\mathbf{x})\}| - |\{\mathbf{x} : f(\mathbf{x}) \neq l_{\mathbf{a},0}(\mathbf{x})\}|}{|E_{n,k}|} \right| \\ &= |\mathcal{W}_f^{(k)}(\mathbf{a})| \end{aligned}$$

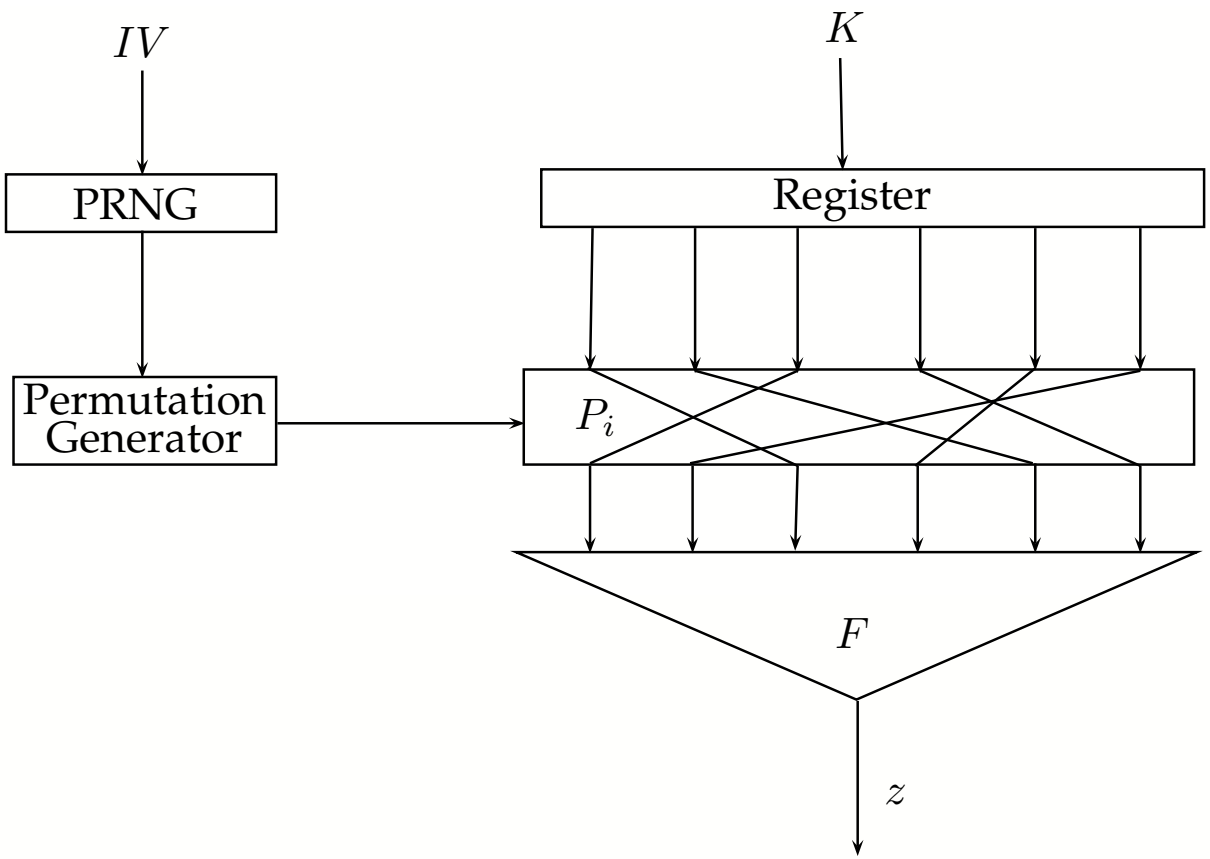
Preliminaries: FLIP design

Preliminaries: FLIP design

- FLIP is based on three components:
  1. One register of length n.
  2. One pseudorandom number generator (PRNG).
  3. One nonlinear filter function $F = f_1 + f_2 + f_3$ involving n-variables.

Preliminaries: FLIP design

- Design specification of FLIP cipher:



Preliminaries: FLIP design

- In FLIP cipher:
  1. $n = n_1 + n_2 + n_3$.
  2. Register takes input with weight $\frac{n}{2}$ only.
  3. $z_t = F(s^t) = f_1(s_1^t) + f_2(s_2^t) + f_3(s_3^t)$.
     (a) f_1: Linear function involving n_1 number of variables.
     (b) f_2: Quadratic function involving n_2 number of variables.
     (c) f_3: Sum of r Triangular functions each involves b number of variables.

- The ANFs of the component functions of F are described below:
  1. **L-type function:** $L_n(x_0, x_1, \ldots, x_{n-1}) = \sum_{i=0}^{n-1} x_i$.
  2. **Q-type function:** $Q_{2n}(x_0, x_1, \ldots, x_{2n-1}) = \sum_{i=0}^{n-1} x_{2i}x_{2i+1}$.
  3. **T-type function:**
     $T_n(x_0, x_1, \ldots, x_{\frac{n(n+1)}{2}-1}) = \sum_{i=1}^{n} \prod_{j=0}^{i-1} x_{j+\sum_{\ell=0}^{i-1} \ell}$.

Existing results

Existing results

- In Journées Codage et Cryptographie - JC2 2015, Méaux first proposed the stream cipher FLIP.
- At Crypto 2016, Duval et al. proposed an attack on the old version of the FLIP stream cipher as introduced by Méaux in In: Journées Codage et Cryptographie - JC2 2015.
- At EUROCRYPT 2016, Méaux et al. proposed a modified design of FLIP.
- Recently, Carlet et al. (IACR ToSC 2018) and Mesnager et al. (CCDS 2018) derived several properties of Boolean functions in restricted domain.

Our work

Our work

- In this paper, we analyze the cipher $\mathrm{FLIP}_{530}(42, 128, {}^{8}\Delta^{9})$ proposed by Méaux et al. at EUROCRYPT 2016. Here $n_1 = 42, n_2 = 2 \cdot 64 = 128, n_3 = 8 \cdot (1 + 2 + \cdots + 9) = 360$.
  1. First 42 variables are involved in the linear function.
  2. Next 128 variables (43 to 170) are involved in the quadratic bent function.
  3. Last 360 variables (171 to 530) are involved in 8 triangular functions of degree 9.

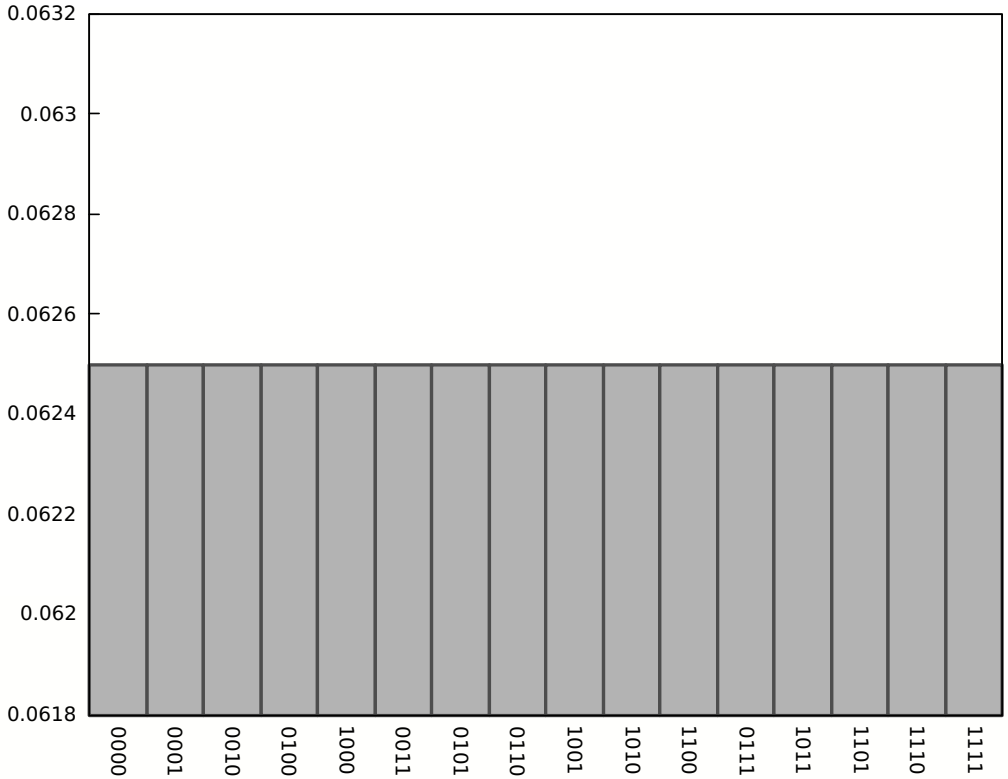
Our approach

Our approach

- The properties of a Boolean function change significantly when the inputs are from restricted domain.
- The nonlinear filter function of FLIP is a special type of function.
- Changes in the distributions.

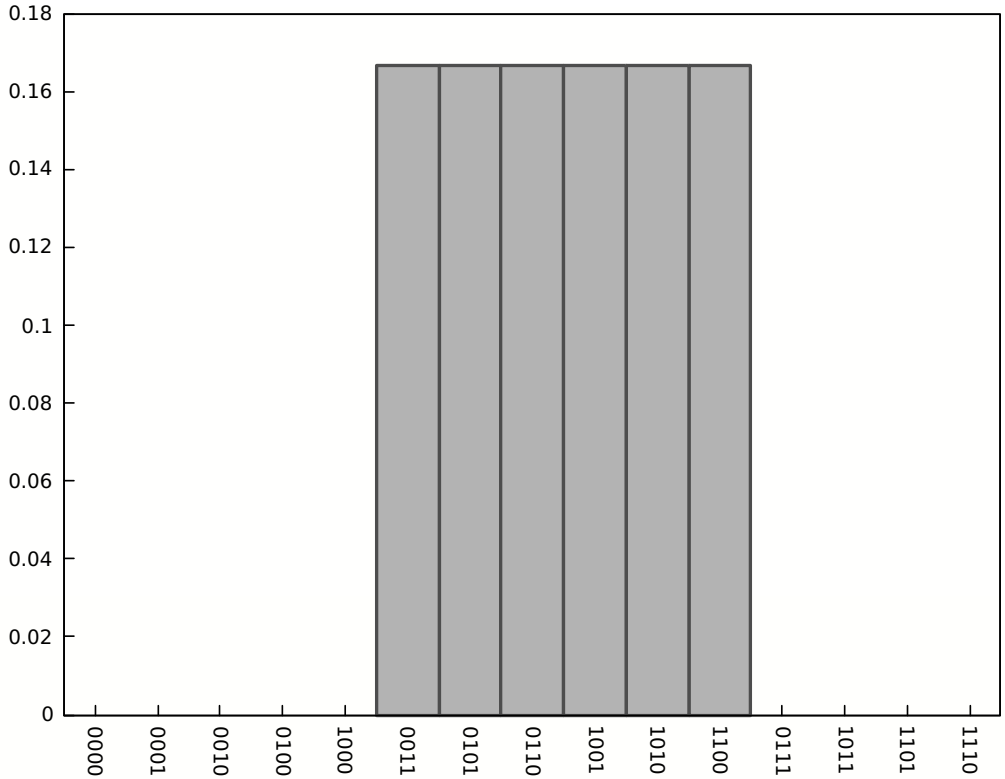
Our approach

- All weight uniform case $n = 4$.



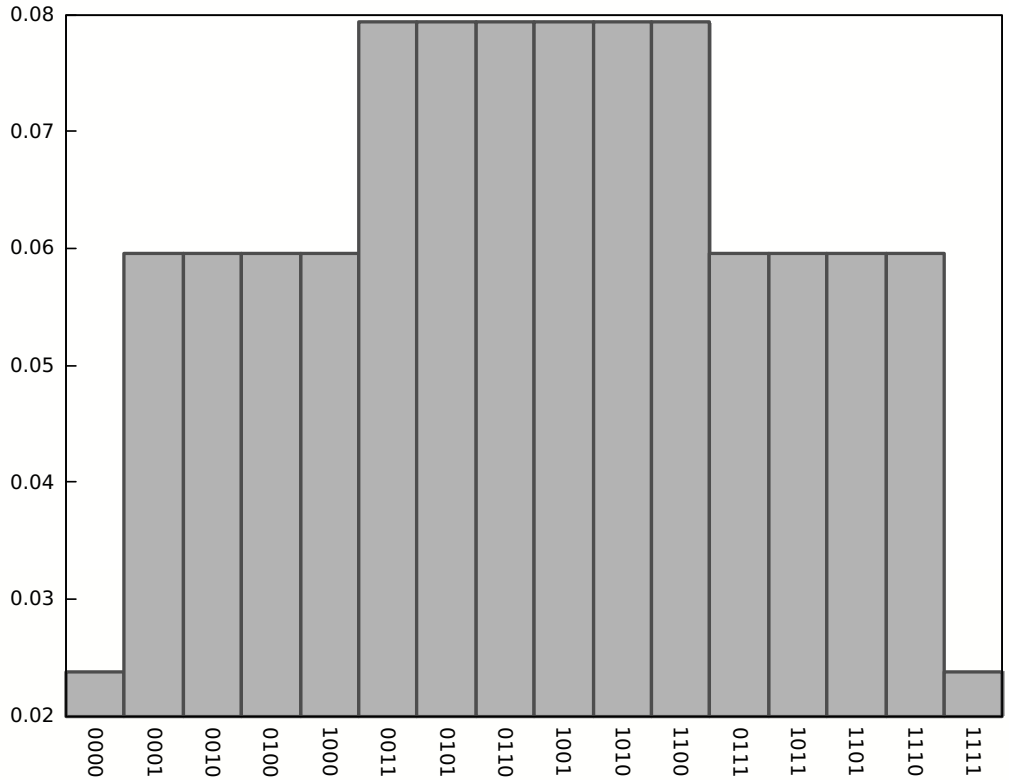
Our approach

- Weight 2 uniform case.



Our approach

- For restricted case $n = 10$, $n_1 = 4$, $k = 5$.



Our approach

- Let $n = 6$. The $|E_{6,3}| = 20$ and the following tables provide the frequency distributions of $\mathbb{F}_2^{n_1}$, $n_1 = 2, 3$ and 4.

x_2x_1	Frequency
00	4
01	6
10	6
11	4

$x_3x_2x_1$	Frequency
000	1
001	3
010	3
011	3
100	3
101	3
110	3
111	1

$x_4x_3x_2x_1$	Frequency
0000	0
0001	1
0010	1
0011	2
0100	1
0101	2
0110	2
0111	1
1000	1
1001	2
1010	2
1011	1
1100	2
1101	1
1110	1
1111	0

Our results

Biased Walsh–Hadamard transform

- Let $p(\mathbf{a})$ be the probability of $\mathbf{a} \in \mathbb{F}_2^n$, not necessarily uniform.
- The *biased Hamming distance* between $f, g \in \mathcal{B}_n$ is

$$d_H^B(f,g) = \frac{1}{2} - \frac{1}{2} \sum_{\mathbf{x} \in \mathbb{F}_2^n} p(\mathbf{x})(-1)^{f(\mathbf{x})+g(\mathbf{x})}.$$

- In particular, $d_H^B(f, l_{\mathbf{a},\varepsilon}) = \frac{1}{2} - \frac{(-1)^\varepsilon}{2} \mathcal{W}_f^B(\mathbf{a})$, where $\mathcal{W}_f^B(\mathbf{a}) = \sum_{\mathbf{x} \in \mathbb{F}_2^n} p(\mathbf{x})(-1)^{f(\mathbf{x})+\mathbf{a}\cdot\mathbf{x}}$ is the *biased Walsh–Hadamard transform* of $f \in \mathcal{B}_n$ at $\mathbf{a} \in \mathbb{F}_2^n$.
- Further, $\mathrm{corr}^B(f, l_{\mathbf{a},0}) = \left|\mathcal{W}_f^B(\mathbf{a})\right|$.

Biased Walsh–Hadamard transform

- Let $f(\mathbf{x}) = f_1(\mathbf{x}') + f_2(\mathbf{x}'')$, where $\mathbf{x} = \mathbf{x}'||\mathbf{x}''$. Then $\mathcal{W}_f^B(\mathbf{a})$ is equal to

$$\sum_{\mathbf{x}'' \in \mathbb{F}_2^{n_2}} p(\mathbf{x}'')(-1)^{f_2(\mathbf{x}'')+\mathbf{a}''\cdot\mathbf{x}''} \sum_{\mathbf{x}' \in \mathbb{F}_2^{n_1}} p(\mathbf{x}'/\mathbf{x}'')(-1)^{f_1(\mathbf{x}')+\mathbf{a}'\cdot\mathbf{x}'},$$

  where $p(\mathbf{x}'/\mathbf{x}'') = Pr[\mathbf{x}'/\mathbf{x}'']$.

- The problem is that in general $Pr[\mathbf{x}'/\mathbf{x}''] \neq Pr[\mathbf{x}']$. So we are unable to directly calculate the biased Walsh–Hadamard transform of $f = f_1 + f_2$ knowing the biased Walsh–Hadamard transform of two component functions f_1 and f_2.

Biased Walsh–Hadamard transform

Theorem 1 (Restricted Domain Convolution)

Let $n = n_1 + n_2$ *and* $f = f_1 + f_2$, *where* $f_i \in \mathcal{B}_{n_i}$, $i \in \{1, 2\}$. *Then, for any* $\mathbf{a} = \mathbf{a}'||\mathbf{a}'' \in \mathbb{F}_2^n$ *and* $0 \leq k \leq n$,

$$\mathcal{W}_f^{B(k)}(\mathbf{a}) = p_{n,k} \sum_{i=0}^{k} \binom{n_1}{i} \binom{n_2}{k-i} \mathcal{W}_{f_1}^{(i)}(\mathbf{a}') \mathcal{W}_{f_2}^{(k-i)}(\mathbf{a}'')$$

$$= \sum_{i=0}^{k} \frac{p_{n,k}}{q_{n_1,i} q_{n_2,k-i}} \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}') \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}''),$$

where $q_{n_1,i} = \frac{\binom{n_2}{k-i}}{\binom{n}{k}}, q_{n_2,k-i} = \frac{\binom{n_1}{i}}{\binom{n}{k}}$.

Biased Walsh–Hadamard transform

Corollary 1

Let $n = n_1 + n_2$ and $f = f_1 + f_2$, where $f_i \in \mathcal{B}_{n_i}$, $i \in \{1, 2\}$. For any $\mathbf{a} = \mathbf{a}'||\mathbf{a}'' \in \mathbb{F}_2^n$, $\mathcal{W}_f^B(\mathbf{a})$ is equal to

$$\sum_{k=0}^{n} \mathcal{W}_f^{B(k)}(\mathbf{a}) = \sum_{k=0}^{n} p_{n,k} \sum_{i=0}^{k} \binom{n_1}{i} \binom{n_2}{k-i} \mathcal{W}_{f_1}^{(i)}(\mathbf{a}') \mathcal{W}_{f_2}^{(k-i)}(\mathbf{a}'')$$

$$= \sum_{k=0}^{n} \sum_{i=0}^{k} \frac{p_{n,k}}{q_{n_1,i} q_{n_2,k-i}} \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}') \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}'').$$

Biased Walsh–Hadamard transform

Lemma 1 (Carlet et al. (IACR ToSC 2018))

Let $n = n_1 + n_2$ *and* $f = f_1 + f_2$, *where* $f_i \in \mathcal{B}_{n_i}$, $i \in \{1, 2\}$.

$$\max_{\mathbf{a}} \left|\mathcal{W}_f^{(k)}(\mathbf{a})\right| \leq \sum_{i=0}^{k} p_{n,k} \Bigg\{ \max_{\mathbf{a}_1} \left| \sum_{\mathbf{x}_1 \in E_{n_1,i}} (-1)^{f_1(\mathbf{x}_1)+\mathbf{a}_1\cdot\mathbf{x}_1} \right| \max_{\mathbf{a}_2} \left| \sum_{\mathbf{x}_2 \in E_{n_2,k-i}} (-1)^{f_1(\mathbf{x}_2)+\mathbf{a}_2\cdot\mathbf{x}_2} \right| \Bigg\}$$

Biased Walsh–Hadamard transform

Theorem 2

For all $0 \leq k \leq n$, the following inequality holds

$$\sum_{i=0}^{k} p_{n,k} \max_{\mathbf{a}_1} \left| \sum_{\mathbf{x}_1 \in E_{n_1,i}} (-1)^{f_1(\mathbf{x}_1)+\mathbf{a}_1 \cdot \mathbf{x}_1} \right| \max_{\mathbf{a}_2} \left| \sum_{\mathbf{x}_2 \in E_{n_2,k-i}} (-1)^{f_1(\mathbf{x}_2)+\mathbf{a}_2 \cdot \mathbf{x}_2} \right|$$
$$\geq \sum_{i=0}^{k} \max_{\mathbf{a}_1 \in \mathbb{F}_2^{n_1}} \left| \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}_1) \right| \max_{\mathbf{a}_2 \in \mathbb{F}_2^{n_2}} \left| \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}_2) \right|.$$

- From the above inequality we can theoretically claim that the bound provided by Carlet et al. is much higher than our bound

$$G = \sum_{i=0}^{k} \max_{\mathbf{a}_1 \in \mathbb{F}_2^{n_1}} |\mathcal{W}_{f_1}^{B(i)}(\mathbf{a}_1)| \max_{\mathbf{a}_2 \in \mathbb{F}_2^{n_2}} |\mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}_2)|.$$

Biased Walsh–Hadamard transform

Lemma 2

Let a_i be positive numbers and b_i be any integers (positive or negative), where $i = 0, 1, \ldots, k$. If

$$\left|\left|\sum_{i=0}^{k} a_i b_i\right| - \left|\sum_{i,j=0; i\neq j=0}^{k} a_i b_j\right|\right| \leq \left|\sum_{i=0}^{k} a_i b_i\right|,$$

and the sums $\sum_{i=0}^{k} a_i b_i$, $\sum_{i,j=0;\ i\neq j}^{k} a_i b_j$ have opposite signs, then

$$\left|\sum_{i=0}^{k} a_i b_i\right| \geq \left(\sum_{i=0}^{k} a_i\right) \left|\sum_{j=0}^{k} b_j\right|.$$

Biased Walsh–Hadamard transform

Theorem 3

Let $f = f_1 + f_2 \in \mathcal{B}_n$, $f_i \in \mathcal{B}_{n_i}$, $i = 1, 2$, $A_i := q_{n_1,i} q_{n_2,k-i}$ *and* $B_i := \sum_{\mathbf{x}_1 \in E_{n_1,i}} (-1)^{f_1(\mathbf{x}_1)+\mathbf{a}_1 \cdot \mathbf{x}_1} \sum_{\mathbf{x}_2 \in E_{n_2,k-i}} (-1)^{f_2(\mathbf{x}_2)+\mathbf{a}_2 \cdot \mathbf{x}_2}$, *for all* $0 \le i \le k$ *(here* $q_{n_1,i} = \frac{\binom{n_2}{k-i}}{\binom{n}{k}}$, $q_{n_2,k-i} = \frac{\binom{n_1}{i}}{\binom{n}{k}}$*). Then*

$$\max_{\mathbf{a} \in \mathbb{F}_2^n} \left| \mathcal{W}_f^{(k)}(\mathbf{a}) \right| \le \sum_{i=0}^{k} \max_{\mathbf{a}_1 \in \mathbb{F}_2^{n_1}} \left| \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}_1) \right| \max_{\mathbf{a}_2 \in \mathbb{F}_2^{n_2}} \left| \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}_2) \right|,$$

if $\left| \left| \sum_{i=0}^{k} A_i B_i \right| - \left| \sum_{i=0}^{k} A_i B_i - p_{n,k} \sum_{j=0}^{k} B_j \right| \right| \le \left| \sum_{i=0}^{k} A_i B_i \right|$, *where* $p_{n,k} = \frac{1}{\binom{n}{k}}$, *and, the expressions* $\sum_{i=0}^{k} A_i B_i$, $p_{n,k} \sum_{j=0}^{k} B_j - \sum_{i=0}^{k} A_i B_i$ *have opposite signs.*

Biased Walsh–Hadamard transform

Theorem 4

Let $0 \leq i \leq k$, $\mathbf{c}_i \in \mathbb{F}_2^{n_1}$, $\mathbf{d}_i \in \mathbb{F}_2^{n_2}$, $q_{n_1,i} = \frac{\binom{n_2}{k-i}}{\binom{n}{k}}$, $q_{n_2,k-i} = \frac{\binom{n_1}{i}}{\binom{n}{k}}$, *and*

$$\max_{\mathbf{a}_1 \in \mathbb{F}_2^{n_1}} \left| \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}_1) \right| = q_{n_1,i} \left| \sum_{\mathbf{x}_1 \in E_{n_1,i}} (-1)^{f_1(\mathbf{x}_1)+\mathbf{c}_i \cdot \mathbf{x}_1} \right|,$$

$$\max_{\mathbf{a}_2 \in \mathbb{F}_2^{n_2}} \left| \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}_2) \right| = q_{n_2,k-i} \left| \sum_{\mathbf{x}_2 \in E_{n_2,k-i}} (-1)^{f_2(\mathbf{x}_2)+\mathbf{d}_i \cdot \mathbf{x}_2} \right|.$$

If $\sum_{\mathbf{x}_1 \in E_{n_1,i}} (-1)^{f_1(\mathbf{x}_1)+\mathbf{c}_i \cdot \mathbf{x}_1} \sum_{\mathbf{x}_2 \in E_{n_2,k-i}} (-1)^{f_2(\mathbf{x}_2)+\mathbf{d}_i \cdot \mathbf{x}_2}$ *has constant sign, for all* $0 \leq i \leq k$, *then,*

$$\sum_{i=0}^{k} \max_{\mathbf{a}_1 \in \mathbb{F}_2^{n_1}} \left| \mathcal{W}_{f_1}^{B(i)}(\mathbf{a}_1) \right| \max_{\mathbf{a}_2 \in \mathbb{F}_2^{n_2}} \left| \mathcal{W}_{f_2}^{B(k-i)}(\mathbf{a}_2) \right| \leq \max_{\mathbf{a} \in \mathbb{F}_2^{n}} \left| \mathcal{W}_{f}^{(k)}(\mathbf{a}) \right|.$$

Comparison with the existing result

Comparison with the existing result

- **For a small Boolean function** $f : E_{12,6} \to \{0,1\}$**:** Let $n = 12 = 2 + 4 + 6$ and the component functions be

$$f_1(x_0, x_1) = x_0 + x_1,$$
$$f_2(x_2, x_3, x_4, x_5) = x_2x_3 + x_4x_5,$$
$$f_1(x_6, x_7, \ldots, x_{11}) = x_6 + x_7x_8 + x_9x_{10}x_{11}.$$

Then

Original bias	≈ 0.264069
Carlet et al. (IACR ToSC 2018)	≤ 0.772727
This paper	≥ 0.20857

Comparison with the existing result

- **For the filter function of FLIP$_{530}(42, 128, 360)$:**
  Computationally, we found that all these functions

$$f_1 = x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9,$$
$$f_2 = x_0x_1 + x_2x_3 + x_4x_5 + x_6x_7,\ f_3 = x_0x_1x_2,$$
$$f_4 = x_0x_1x_2x_3,\ f_5 = x_0x_1x_2x_3x_4,\ f_6 = x_0x_1x_2x_3x_4x_5,$$
$$f_7 = x_0x_1x_2x_3x_4x_5x_6,\ f_8 = x_0x_1x_2x_3x_4x_5x_6x_7,$$
$$f_9 = x_0x_1x_2x_3x_4x_5x_6x_7x_8$$

  satisfy the required condition of Theorem 4.
  The comparison of the biases is given in the following table.

Carlet et al. (IACR ToSC 2018)	$\leq \frac{1}{2^{13.59}}$
This paper	$\geq \frac{1}{2^{18.49}}$

Conclusion

Conclusion

- We have studied the cryptographic properties of a Boolean function in biased domain.
- We have obtained a lower bound for the bias of the nonlinear filter function of the FLIP stream cipher.
- Our results provide a more accurate calculation of biases related to Boolean functions.

Thank you...!