Analysis of *split-and-lookup*-based ZK-friendly primitives

Antoine Bak[1, 2] Léo Perrin[1]

[1]INRIA Paris

[2]DGA, Paris

March 28, 2025

Plan

Linear cryptanalysis over $\mathbb{F}_p$

Linear cryptanalysis over $\mathbb{F}_2$

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^m, a \in \mathbb{F}_2^m, b \in \mathbb{F}_2^n, c \in \mathbb{F}_2$:

We can study the number of solutions to a linear equation on f:

$$L^f(a, b, c) = \#\{x \in \mathbb{F}_2^n \mid a \cdot f(x) \oplus b \cdot x = c\} .$$

When this quantity is far apart from 2^{n-1}, the function f is **highly biased**.

Linear cryptanalysis over $\mathbb{F}_2$

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^m, a \in \mathbb{F}_2^m, b \in \mathbb{F}_2^n, c \in \mathbb{F}_2$:

We can study the number of solutions to a linear equation on f:

$$L^f(a, b, c) = \#\{x \in \mathbb{F}_2^n \mid a \cdot f(x) \oplus b \cdot x = c\} \ .$$

When this quantity is far apart from 2^{n-1}, the function f is **highly biased**.

Alternatively, one can study the *Walsh* coefficients of f:

$$C^f(a, b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot f(x) \oplus b \cdot x} \ .$$

Linear cryptanalysis over $\mathbb{F}_2$

Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^m, a \in \mathbb{F}_2^m, b \in \mathbb{F}_2^n, c \in \mathbb{F}_2$:

We can study the number of solutions to a linear equation on f:

$$L^f(a, b, c) = \#\{x \in \mathbb{F}_2^n \mid a \cdot f(x) \oplus b \cdot x = c\} \ .$$

When this quantity is far apart from 2^{n-1}, the function f is **highly biased**.

Alternatively, one can study the *Walsh* coefficients of f:

$$C^f(a, b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot f(x) \oplus b \cdot x} \ .$$

These quantities are closely related through:

$$C^f(a, b) = 2 \cdot (L^f(a, b, 0) - 2^{n-1}) \ .$$

Generalizations over $\mathbb{F}_p$

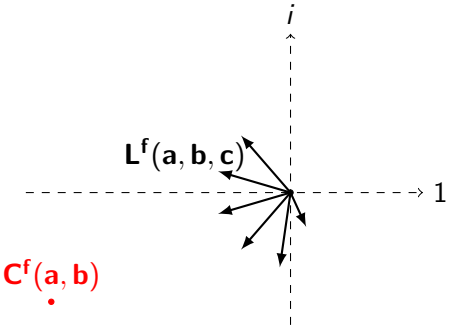
Let ζ_p a primitive p-th root of unity, over $\mathbb{F}_p$, we have the following generalizations [BSV07]:

$$L^f(a,b,c) = \#\{x \in \mathbb{F}_p^n \mid a \cdot f(x) = b \cdot x + c\} \text{ and } C^f(a,b) = \sum_{x \in \mathbb{F}_p^n} \zeta_p^{a \cdot f(x) - b \cdot x} .$$

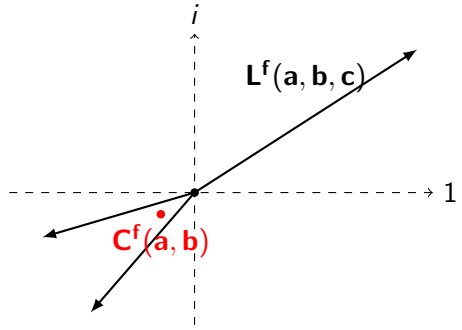
Generalizations over $\mathbb{F}_p$

Let ζ_p a primitive p-th root of unity, over $\mathbb{F}_p$, we have the following generalizations [BSV07]:

$$L^f(a,b,c) = \#\{x \in \mathbb{F}_p^n \mid a \cdot f(x) = b \cdot x + c\} \text{ and } C^f(a,b) = \sum_{x \in \mathbb{F}_p^n} \zeta_p^{a \cdot f(x) - b \cdot x} .$$



(c) Hich correlation, low probabilities.



(d) High probability, low correlation.

Plan

Case of ZK-friendly primitives

Allowed operations

Proof systems can prove the following relations:

- linear relations (essentially for free),
- low degree maps (eg. x^α),
- arbitrary lookups ($\sigma : X \to Y$ where X, Y small subsets of $\mathbb{F}_p$).

Allowed operations

Proof systems can prove the following relations:

- linear relations (essentially for free),
- low degree maps (eg. x^α),
- arbitrary lookups ($\sigma : X \to Y$ where X, Y small subsets of $\mathbb{F}_p$).

ZK-friendly primitive: ($p \sim 2^{64}$ prime number)

$$f : \mathbb{F}_p^n \to \mathbb{F}_p^m \ ,$$

such that

$$f(x) = y \Leftrightarrow \exists z, P_1(x, y, z) = \cdots = P_k(x, y, z) = 0 \ .$$

where the P_i can be proven efficiently.

The *split-and-lookup* (S&L) S-boxes

Introduced in [GKL+21] to yield both efficient evaluation and proving time:

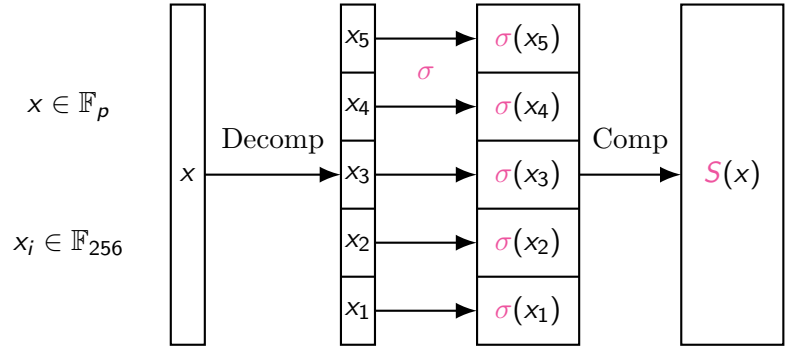


Figure: *Split-and-lookup* function.

Role: guaranteeing security against algebraic attacks.

The *split-and-lookup* (S&L) S-boxes

Introduced in [GKL+21] to yield both efficient evaluation and proving time:

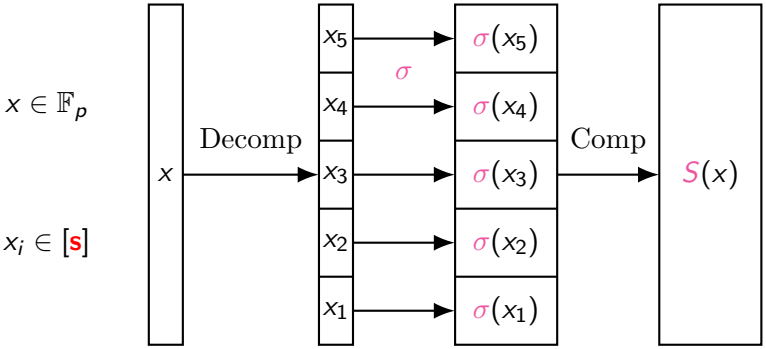


Figure: *Split-and-lookup* function.

Role: guaranteeing security against algebraic attacks.

Linear properties of ZK-friendly nonlinear maps

Low degree maps: x^α

- $L^{x^\alpha}(a, b, c) \leq \alpha$.
- $|C^{x^\alpha}(a, b)| \leq (\alpha - 1)\sqrt{p}$ (Weil bound).

Split-and-lookups: S

- $L^S(a, b, c)$: **?**
- $|C^S(a, b)|$: **?**

Linear properties of ZK-friendly nonlinear maps

Low degree maps: x^α

- $L^{x^\alpha}(a, b, c) \leq \alpha$.
- $|C^{x^\alpha}(a, b)| \leq (\alpha - 1)\sqrt{p}$ (Weil bound).

Split-and-lookups: S

- $L^S(a, b, c)$: **?**
- $|C^S(a, b)|$: **?**

In the following, we study the linear properties of S&L:

- Efficient algorithms for computing those quantities.
- Study of the properties of the S&L used in the literature.

Plan

Linear correlation of S&L

A simple formula for the linear correlation of S&L

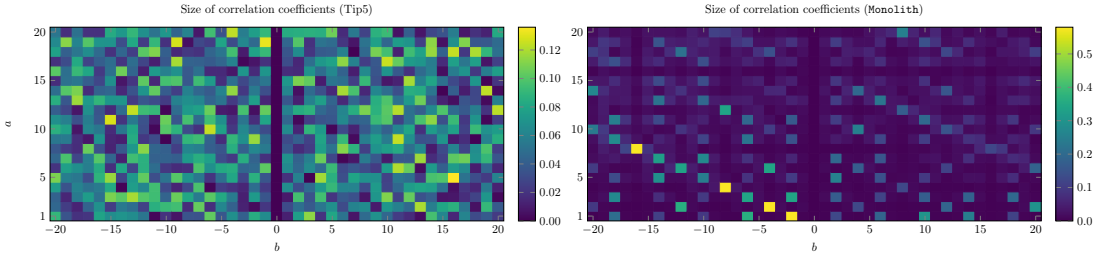
Problem: given a, b, compute $C^S(a,b) = \sum_{x \in \mathbb{F}_p} \zeta_p^{a \cdot S(x) - b \cdot x}$.

We found the following formula:

$$C^S(a,b) = \sum_{i=1}^{n} \prod_{j=1}^{i-1} C_{[s]}^{\sigma,j}(a,b) \cdot C_{[p_i]}^{\sigma,i}(a,b) \cdot \prod_{k=i+1}^{n} C_{\{p_k\}}^{\sigma,k}(a,b) \ ,$$

where $C_X^{\sigma,i}(a,b) = \sum_{x \in X} \zeta_p^{(a \cdot \sigma(x) - b \cdot x) \cdot s^{i-1}}$, and $(p_i)_i$ is the base-s decomposition of p.

Application to some primitives



We observed that for **small** a, b, the $|C^S(a, b)|$ are **high** (the order of p).

- We provided a partial explanation using the established **formula**.
- Not a problem if other components provide security against correlation attacks (eg. x^α).

Plan

Linear probability of S&L

Trivial estimates

Evaluate

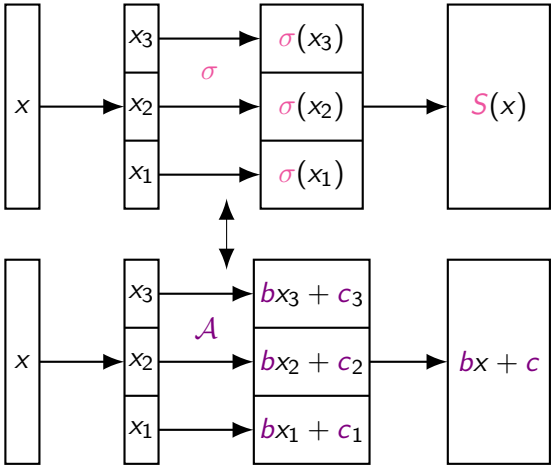
$$L^S(a,b,c) = \#\{x \in \mathbb{F}_p \mid a \cdot S(x) = b \cdot x + c\} \ .$$

Method 1: Iterate through all $x \in \mathbb{F}_p$?

- Cost: $\mathcal{O}(p)$.
- In the case of $\mathrm{Tip5}$, $p \sim 2^{64}$: too costly.

Method 2: Wordwise linear approximation.

$$L^S(a,b,c) \geq \prod_{i=1}^{n} L^\sigma(a,b,c_i) \ .$$



Carry propagation phenomenon

In general, those quantities are not equal, eg.

$$a \cdot \sigma(x_2) = b \cdot x_2 + c_2 - 1$$

and

$$a \cdot \sigma(x_1) = b \cdot x_1 + c_1 + s \ ,$$

gives

$$a \cdot S(x) = b \cdot x + c \ .$$

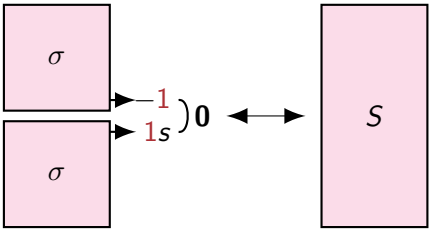


Figure: Two carries cancelling out.

Carry matrices

Idea:

$$m_{ij}^{\sigma,X,a,b,c} = \#\{x \in X \mid a \cdot \sigma(x) = b \cdot x + c + i \cdot s - j\} .$$

We have that:

$$m_{ij}^{S,X,a,b,c} = \sum_{\mathbf{k}} m_{i\mathbf{k}}^{\sigma,X_2,a,b,c_2} m_{\mathbf{k}j}^{\sigma,X_1,a,b,c_1} .$$

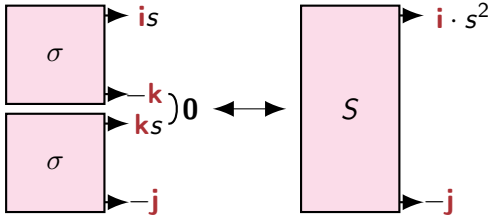


Figure: Two carries cancelling out.

Carry matrices

Idea:

$$m_{ij}^{\sigma,X,a,b,c} = \#\{x \in X \mid a \cdot \sigma(x) = b \cdot x + c + i \cdot s - j\} .$$

We have that:

$$m_{ij}^{S,X,a,b,c} = \sum_{\mathbf{k}} m_{i\mathbf{k}}^{\sigma,X_2,a,b,c_2} m_{\mathbf{k}j}^{\sigma,X_1,a,b,c_1} .$$

This is a matrix multiplication!

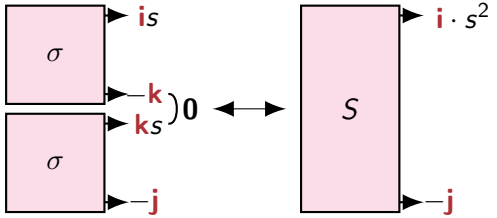


Figure: Two carries cancelling out.

Computing the number of solutions

Fact: The values of i, j, k can be restricted to an interval $I_{a,b}$ of size $|a| + |b|$.

Let

$$M^{\sigma,X}_{a,b,c} = \left(m^{\sigma,X,a,b,c}_{ij}\right)_{i,j \in I_{a,b}} .$$

If $\mathrm{Decomp}(c) = (c_1, \ldots, c_n)$, $X = X_1 \times \ldots \times X_n$:

$$M^{S,X}_{a,b,c} = M^{\sigma,X_n}_{a,b,c_n} \cdot \ldots \cdot M^{\sigma,X_1}_{a,b,c_1} .$$

We can compute $\#\{x \in X \mid a \cdot S(x) = b \cdot x + c\}$ in time

$$n(|a| + |b|)^3 .$$

Applications to Tip5 and Monolith

Primitive	p (log)	a	b	c	$L^S_{a,b,c}$ (log)	source
Tip5	p_{goldi} (64)	fixed points			7.9	[SLS+23]
		1	1	c_{Tip5}†	12.68	trivial bound
					15.81	carry matrices
Monolith	p_{goldi} (64)	1	-2	0	40.7	trivial bound/ [BGK+25]
					41.76	carry matrices
	p_{mers} (31)	fixed points			4	[GKL+23]
		1	-2	2^{24}	17.58	trivial bound
					18.35	carry matrices

†$c_{\text{Tip5}} = \texttt{0x086a0896f76a0894}$.

Table: Best known linear approximations of *split-and-lookup* S-boxes.

Plan

Conclusion

Conclusion

Linear properties:

- New tools to study the properties of *split-and-lookups*.
- Comparison of the S-boxes from the literature.

Conclusion

Linear properties:

- New tools to study the properties of *split-and-lookups*.
- Comparison of the S-boxes from the literature.

Other works: Cryptanalysis of S&L-based hash functions

- Collision attack on reduced $\mathrm{Tip5}$.
- Practical distinguisher on full Skyscraper.

Conclusion

Linear properties:

- New tools to study the properties of *split-and-lookups*.
- Comparison of the S-boxes from the literature.

Other works: Cryptanalysis of S&L-based hash functions

- Collision attack on reduced $\mathrm{Tip5}$.
- Practical distinguisher on full Skyscraper.

Thank you!

References

- Clémence Bouvier, Lorenzo Grassi, Dmitry Khovratovich, Katharina Koschatko, Christian Rechberger, Fabian Schmid, and Markus Schofnegger, *Skyscraper: Fast hashing on big primes*, Cryptology ePrint Archive (2025).
- Thomas Baignères, Jacques Stern, and Serge Vaudenay, *Linear cryptanalysis of non binary ciphers: (with an application to safer)*, International Workshop on Selected Areas in Cryptography, Springer, 2007, pp. 184–211.
- Lorenzo Grassi, Dmitry Khovratovich, Reinhard Lüftenegger, Christian Rechberger, Markus Schofnegger, and Roman Walch, *Reinforced concrete: A fast hash function for verifiable computation*, Cryptology ePrint Archive, Paper 2021/1038, 2021, `https://eprint.iacr.org/2021/1038`.
- _____, *Monolith: Circuit-friendly hash functions with new nonlinear layers for fast and constant-time implementations*, Cryptology ePrint Archive, Paper 2023/1025, 2023, `https://eprint.iacr.org/2023/1025`.
- Christian Rechberger, *On the history of fhempczk-friendly symmetric crypto*, `https://who.paris.inria.fr/Leo.Perrin/rescale/slides/Christian-STAP-2023.pdf`, 2023.
- Alan Szepieniec, Alexander Lemmens, Jan Ferdinand Sauer, Bobbin Threadbare, et al., *The tip5 hash function for recursive starks*, Cryptology ePrint Archive (2023).

Plan

Types of AO designs

Strategies of design

Three design strategies for AO [Rec23]:

Type I

- Low degree

$$y = x^{\alpha} \ .$$

- Fast in Plain.
- Many rounds.
- More constraints.
- Poseidon, MiMC...

Type II

- CCZ-equivalence

$$y = x^{1/\alpha} \ .$$

- Slow in Plain.
- Fewer rounds.
- Fewer constraints.
- `Anemoi`, *Rescue*...

Type III

- Lookup tables

$$y = \sigma(x) \ .$$

- Even faster in Plain.
- Even fewer rounds.
- Constraints depend on proof system.
- Reinforced Concrete, `Monolith`...

Plan

Algorithms for linear probability

Computing the carry matrix

Algorithm Computing the matrix $M^{S}_{a,b,c}$

Require: $|a| + |b| + 2 \leq \min s_i$

$M^{[1.-1],s}, M^{[1.0],s}, M^{[n.n],v}, M^{[n+1.n],v} \leftarrow Id$

for $k = 1$ to $n-1$ **do**

$M^{[1.k],s} \leftarrow M^{\sigma^k,[\![s_k]\!]}_{a,b,c_k} M^{[1.k-1],s}$

$M^{[n-k.n],v} \leftarrow M^{[n-k+1.n],v} M^{\sigma^{n-k},\{v_{n-k}\}}_{a,b,c_{n-k}}$

end for

$M \leftarrow 0$

for $k = 0$ to n **do**

$M \leftarrow M + M^{[k+1.n],v} M^{\sigma^k,[\![v_k]\!]}_{a,b,c_k} M^{[1.k-1],s}$

end for

return M

Complexity: $\mathcal{O}\left(n\left(|a| + |b|\right)^{\omega} + \sum_{i=1}^{n} s_i\right)$.

Computing the number of approximations

Algorithm Computing the number $L^S_{a,b,c}$

$L \leftarrow 0$
for k in $[\![-(a^- + b^+ + 1), (a^+ + b^-)]\!]$ **do**
  $c' + k' \cdot \prod_{i=1}^n s_i \leftarrow c + k \cdot p$ {Euclidean division}
  $L \leftarrow L + M^S_{a,b,c'}[k', 0]$ {Call to the previous algorithm}
end for
return L

Complexity: $\mathcal{O}\left(n\left(|a| + |b|\right)^{\omega+1} + \sum_{i=1}^n s_i\right)$.

Plan

S-boxes definitions

Definition of the Tip5 S-box

- Decompose $x \in \mathbb{F}_p, p_{\text{goldi}} = 2^{64} - 2^{32} + 1$ into 8 chunks of 8 bits.
- Apply the following to each chunk:

$$\begin{array}{rccc} \sigma & : & \mathbb{F}_{2^8+1} \setminus \{2^8\} & \to & \mathbb{F}_{2^8+1} \setminus \{2^8\} \\ & & x & \mapsto & (x+1)^3 - 1 \ . \end{array}$$

- σ has fixed points 0 and 255.

Definition of the MONOLITH S-boxes

Case $p_{\text{goldi}} = 2^{64} - 2^{32} + 1$:

- Decompose $x \in \mathbb{F}_p$ into 8 chunks of 8 bits.
- Apply the following to each chunk:

$$\begin{array}{rccc} \sigma & : & \mathbb{F}_2^8 & \to & \mathbb{F}_2^8 \\ & & x & \mapsto & (x \oplus ((\overline{x} \lll 1) \odot (x \lll 2) \odot (x \lll 3))) \lll 1 \ . \end{array}$$

- σ has fixed points 0 and 255.

Definition of the MONOLITH S-boxes

Case $p_{\text{mers}} = 2^{31} - 1$:

- Decompose $x \in \mathbb{F}_p$ into 3 chunks of 8 bits and 1 chunk of 7 bits.
- Apply the following to each 8-bits chunk:

$$\begin{array}{rccc} \sigma_8 & : & \mathbb{F}_2^8 & \to & \mathbb{F}_2^8 \\ & & x & \mapsto & (x \oplus ((\overline{x} \lll 1) \odot (x \lll 2) \odot (x \lll 3))) \lll 1 \ . \end{array}$$

- Apply the following to each 7-bits chunk:

$$\begin{array}{rccc} \sigma_7 & : & \mathbb{F}_2^7 & \to & \mathbb{F}_2^7 \\ & & x & \mapsto & (x \oplus ((\overline{x} \lll 1) \odot (x \lll 2))) \lll 1 \ . \end{array}$$

- σ_8 has fixed points 0 and 255 and σ_7 has fixed points 0 and 127.

Plan

Best known linear approximations

The best known approximation for Tip5

S-box in Tip5 (p_{goldi}):

The best approximation found for S is:

$$S(X) = -X + c_{\text{Tip5}} \ ,$$

where $c_{\text{Tip5}} = 0\text{x}086\text{a}0896\text{f}76\text{a}0894$.
It works for $57360 \sim 2^{15.81}$ values of X.

The best known approximation for Monolith

S-box in Monolith (p_{goldi}):

The best approximation found for S is:

$$S(X) = 2 \cdot X \ .$$

It works for $3726693261568 \sim 2^{41.76}$ values of X.

S-box in Monolith (p_{mers}):

The best approximation found for S is:

$$S(X) = 2 \cdot X + c \ ,$$

where $c = 2^{24}$.
It works for $335928 \sim 2^{18.35}$ values of X.