Commutative Cryptanalysis Made Practical

Jules Baudrin

jules.baudrin@inria.fr

Inria, Paris, France

Joint work with P. Felke, G. Leander, P. Neumann, L. Perrin & L. Stennes.

Séminaire Crypto, UVSQ, 2023

Symmetric cryptography : a bit of context

Plaintext $x \in X$, ciphertext $y \in Y$, key $k \in K$. $X = \mathbb{F}_2^{n_X}, Y = \mathbb{F}_2^{n_Y}, K = \mathbb{F}_2^{n_K}$.

Symmetric cryptography : a bit of context

Plaintext $x \in X$, ciphertext $y \in Y$, key $k \in K$. $X = \mathbb{F}_2^{n_X}, Y = \mathbb{F}_2^{n_Y}, K = \mathbb{F}_2^{n_K}$.

Block cipher

A family $(E_k)_{k \in K}$, where: $\forall\, k \in K, \quad E_k : X \to Y$ is bijective.

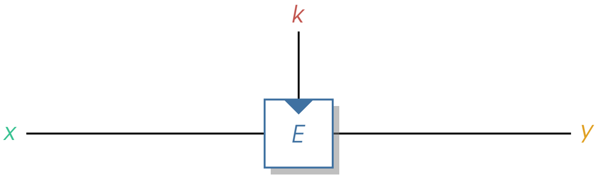
$(\implies n_X = n_Y)$

Plaintext $x \in X$, ciphertext $y \in Y$, key $k \in K$. $X = \mathbb{F}_2^{n_X}, Y = \mathbb{F}_2^{n_Y}, K = \mathbb{F}_2^{n_K}$.

Block cipher

A family $(E_k)_{k \in K}$, where: $\forall\, k \in K, \quad E_k : X \to Y$ is bijective.

$(\implies n_X = n_Y)$

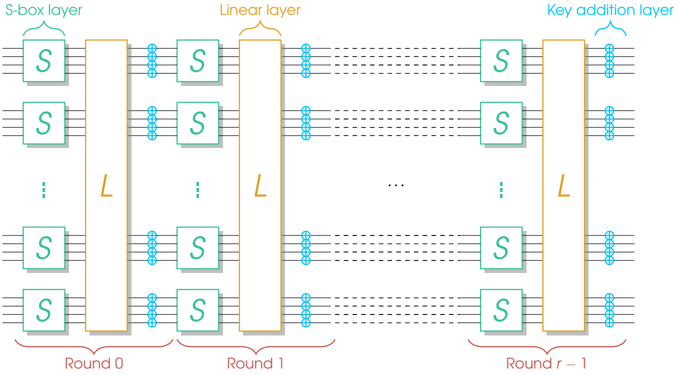


$$y = E_k(x) \iff x = (E_k)^{-1}(y)$$

$\implies$ Shared key for encryption & decryption.

Substitution Permutation Network (SPN)

- Subclass of block ciphers
- Round function, a 3-step process:
  - Local non-linear layer,
  - global linear layer,
  - and key/constant addition
- Repeat r times



Security goals

Security

- Modes + block cipher = confidentiality, integrity, authenticity.
- If the block cipher is secure.

Security

- Modes + block cipher = confidentiality, integrity, authenticity.
- If the block cipher is secure.

Indistinguishability

[$k \xleftarrow{\$} K, \quad E_k$] indistinguishable from [random $F \xleftarrow{\$} \mathrm{Bij}(\mathbb{F}_2^{n_X}, \mathbb{F}_2^{n_Y})$].

Security

- Modes + block cipher = confidentiality, integrity, authenticity.
- If the block cipher is secure.

Indistinguishability

[$k \xleftarrow{\$} K, \quad E_k$] indistinguishable from [random $F \xleftarrow{\$} \mathrm{Bij}(\mathbb{F}_2^{n_X}, \mathbb{F}_2^{n_Y})$].

Security goals

Security

- Modes + block cipher = confidentiality, integrity, authenticity.
- If the block cipher is secure.

Indistinguishability

$[\ k \xleftarrow{\$} K, \quad E_k\ ]$ indistinguishable from $[$ random $F \xleftarrow{\$} \mathrm{Bij}(\mathbb{F}_2^{n_X}, \mathbb{F}_2^{n_Y})\ ]$.

Differential distinguisher

Find α, β st. for many k, $\quad E_k(x + \alpha) = E_k(x) + \beta \quad$ has many solutions x.

Security goals

Security

- Modes + block cipher = confidentiality, integrity, authenticity.
- If the block cipher is secure.

Indistinguishability

[$k \xleftarrow{\$} K, \quad E_k$] indistinguishable from [random $F \xleftarrow{\$} \mathrm{Bij}(\mathbb{F}_2^{n_X}, \mathbb{F}_2^{n_Y})$].

Differential distinguisher

Find α, β st. for many k, $\quad E_k(x+\alpha) = E_k(x) + \beta \quad$ has many solutions x.

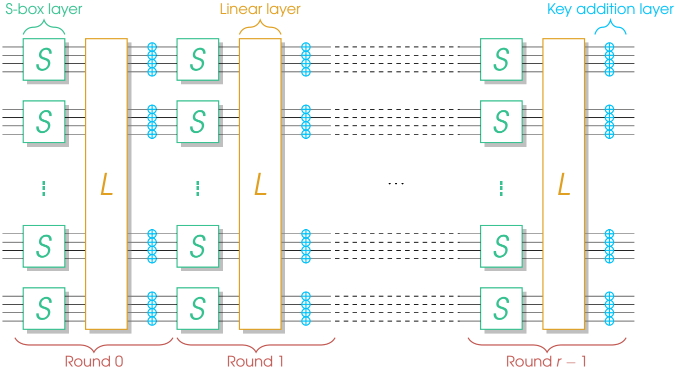
Random permutation F

$F(x+\alpha) + F(x) = \beta$ with proba 2^{-n}.

Some security arguments for an SPN

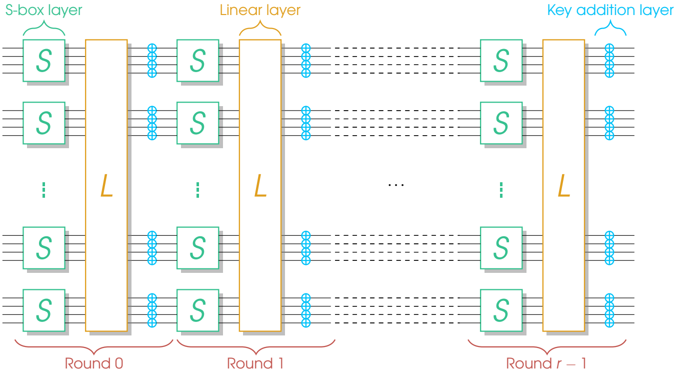
Substitution Permutation Network (SPN)

- Sbox layer, $\rightsquigarrow$ $S(x+\alpha) = S(x) + \beta$ must have few solutions for all α, β
- Linear layer, $\rightsquigarrow$ must diffuse a lot
- Key addition $\rightsquigarrow$ hard to handle...



Substitution Permutation Network (SPN)

- Sbox layer, $\rightsquigarrow S(x+\alpha) = S(x) + \beta$ must have few solutions for all α, β
- Linear layer, $\rightsquigarrow$ must diffuse a lot
- Key addition $\rightsquigarrow$ hard to handle...

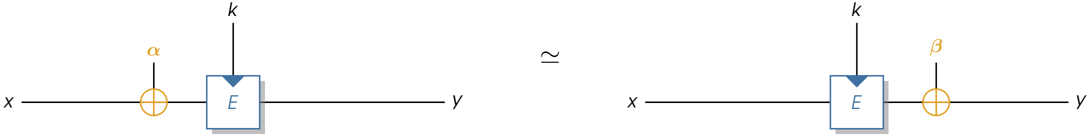


As a designer

Estimate $\mathbb{E}_{k \xleftarrow{\$} K}(\#\{x \text{ st. } E_k(x+\alpha) = E_k(x) + \beta\})$ and assume representativeness.

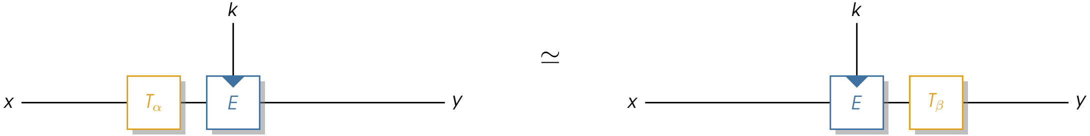
Other cryptanalysis techniques

$$E(x + \alpha) = E(x) + \beta$$



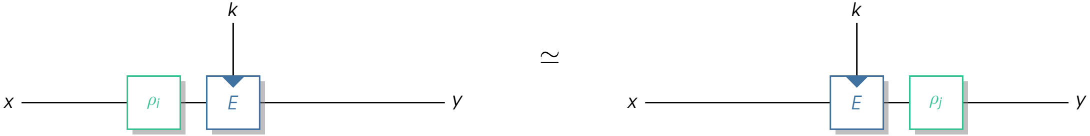
Other cryptanalysis techniques

$$E \circ T_\alpha(x) = T_\beta \circ E(x)$$

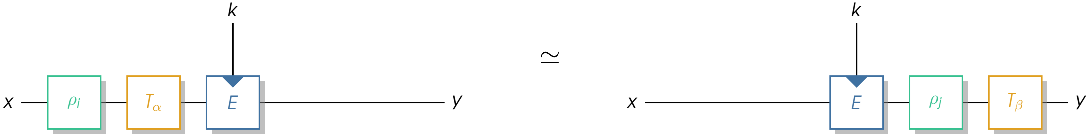


Other cryptanalysis techniques

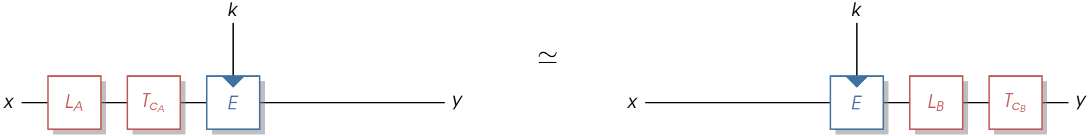
$$E \circ \rho_i(x) = \rho_j \circ E(x)$$



$$E \circ T_\alpha \circ \rho_i(x) = T_\beta \circ \rho_j \circ E(x)$$



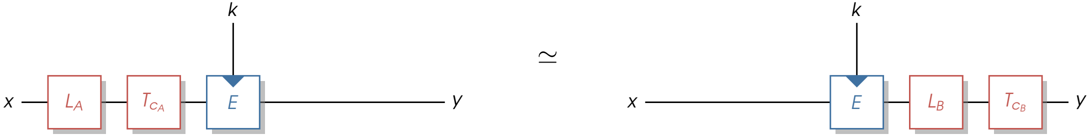
$$E \circ T_{c_A} \circ L_A(x) = T_{c_b} \circ L_B \circ E(x).$$



where $A(x) = L_A(x) + c_A, B(x) = L_B(x) + c_B$

Other cryptanalysis techniques

$$E \circ T_{c_A} \circ L_A(x) = T_{c_b} \circ L_B \circ E(x).$$



where $A(x) = L_A(x) + c_A, B(x) = L_B(x) + c_B$

A tempting desire of unification

Mathematically elegant, better understanding & new attacks

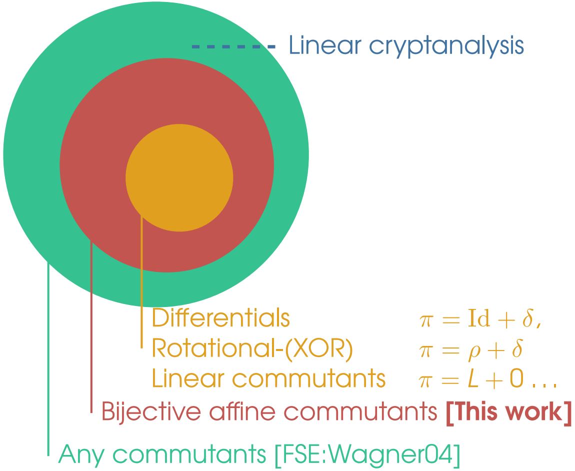
A 20-year-old idea [Wagner, FSE 2004]

Commutative diagram cryptanalysis: not so fruitful[1] since.

---

[1] to the best of our knowledge...

$$
\begin{tikzcd}
X \arrow[r, "E"] \arrow[d, "\pi_i"] \arrow[dr, phantom, "\circlearrowleft"] & Y \arrow[d, "\pi_o"] \\
X' \arrow[r, "E'"] & Y'
\end{tikzcd}
$$



In this talk

Affine commutation with **probability 1**: theory + practice

A **surprising differential** interpretation

A few words about the **probabilistic case**

Goal

Find **bijective affine** A, B st. for many k: $\boxed{E_k \circ A = B \circ E_k}$ (all x are solutions)

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. for many k: $\boxed{E_k \circ A = B \circ E_k}$ (all x are solutions)

$$E = R_{r-1} \circ \cdots \circ R_1 \circ R_0$$

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. for many k: $\boxed{E_k \circ A = B \circ E_k}$ (all x are solutions)

$$E = R_{r-1} \circ \cdots \circ R_1 \circ R_0$$

Sufficient condition for **iterated** constructions

There exist $A_0, \cdots, A_r$ st. for all i $\boxed{A_{i+1} \circ R_i = R_i \circ A_i}$.

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. for many k: $\boxed{E_k \circ A = B \circ E_k}$ (all x are solutions)

$$E = R_{r-1} \circ \cdots \circ R_1 \circ R_0$$

Sufficient condition for **iterated** constructions

There exist $A_0, \cdots, A_r$ st. for all i $\boxed{A_{i+1} \circ R_i = R_i \circ A_i}$.

$$\begin{aligned}
E \circ A_0 &= R_{r-1} \circ \cdots \circ (R_0 \circ A_0) \\
&= R_{r-1} \circ \cdots \circ R_1 \circ (A_1 \circ R_0) \\
&= \cdots \\
&= A_r \circ R_{r-1} \circ \cdots \circ R_0 \\
&= A_r \circ E
\end{aligned}$$

Goal

Find **bijective affine** A, B st. for many k: $\boxed{E_k \circ A = B \circ E_k}$ (all x are solutions)

$$E = R_{r-1} \circ \cdots \circ R_1 \circ R_0$$

Sufficient condition for **iterated** constructions

There exist $A_0, \cdots, A_r$ st. for all i $\boxed{A_{i+1} \circ R_i = R_i \circ A_i}$.

$$\begin{aligned} E \circ A_0 &= R_{r-1} \circ \cdots \circ (R_0 \circ A_0) \\ &= R_{r-1} \circ \cdots \circ R_1 \circ (A_1 \circ R_0) \\ &= \cdots \\ &= A_r \circ R_{r-1} \circ \cdots \circ R_0 \\ &= A_r \circ E \end{aligned}$$

$$\begin{tikzcd} x_0 \arrow[r, "R_0"] \arrow[d, "A_0"] & x_1 \arrow[r, dashed] \arrow[d, "A_1"] \arrow[r, phantom, shift right=6ex, "\circlearrowleft"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "A_{r-1}"] & E(x_0) \arrow[d, "A_r"] \\ z_0 \arrow[r, "R_0"'] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, "R_{r-1}"'] & E(z_0) \end{tikzcd}$$

$\Longrightarrow$ **round-by-round** and **layer-by-layer** studies.

Layer-by-layer probability-1 trail

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A \colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

Layer-by-layer probability-1 trail

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A\colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

S-box layer

$\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A} \iff A \circ S = S \circ A \implies$ $\boxed{\text{self-affine equivalent S-box.}}$

Effective search for small m (4, 8 bits).

Layer-by-layer probability-1 trail

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A \colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

S-box layer

$\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A} \iff A \circ S = S \circ A \implies \boxed{\text{self-affine equivalent S-box.}}$

Effective search for small m (4, 8 bits).

Constant addition

$T_c(x) := x + c, \quad A(x) := L_A(x) + c_A.$

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A \colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

S-box layer

$\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A} \iff A \circ S = S \circ A \implies$ $\boxed{\text{self-affine equivalent S-box.}}$

Effective search for small m (4, 8 bits).

Constant addition

$T_c(x) := x + c, \quad A(x) := L_A(x) + c_A.$

$$A \circ T_c(x) = L_A(x) + L_A(c) + c_A \quad \text{and} \quad T_c \circ A(x) = L_A(x) + c + c_A$$

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A \colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

S-box layer

$\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A} \iff A \circ S = S \circ A \implies$ $\boxed{\text{self-affine equivalent S-box.}}$

Effective search for small m (4, 8 bits).

Constant addition

$T_c(x) := x + c, \quad A(x) := L_A(x) + c_A.$

$$A \circ T_c(x) = L_A(x) + L_A(c) + c_A \quad \text{and} \quad T_c \circ A(x) = L_A(x) + c + c_A$$

$A \circ T_c = T_c \circ A \iff \boxed{c \in \mathrm{Fix}(L_A).}$

Simplified setting for this presentation

- Commutation only: $E \circ \mathcal{A} = \mathcal{A} \circ E$ (case $\mathcal{A} = \mathcal{B}$)
- Parallel mappings: $\mathcal{A} := A \times A \times \cdots \times A$, where $A \colon \mathbb{F}_2^m \to \mathbb{F}_2^m$.

S-box layer

$\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A} \iff A \circ S = S \circ A \implies$ $\boxed{\text{self-affine equivalent S-box.}}$

Effective search for small m (4, 8 bits).

Constant addition

$T_c(x) := x + c, \quad A(x) := L_A(x) + c_A.$

$$A \circ T_c(x) = L_A(x) + L_A(c) + c_A \quad \text{and} \quad T_c \circ A(x) = L_A(x) + c + c_A$$

$A \circ T_c = T_c \circ A \iff \boxed{c \in \mathrm{Fix}(L_A).}$

Linear layer

Let $\mathcal{L} = (\mathcal{L}_{ij})$ be an invertible block matrix with m-size blocks $\mathcal{L}_{ij}$.

$\mathcal{L} \circ \mathcal{A} = \mathcal{A} \circ \mathcal{L} \iff \boxed{\mathcal{L}_{ij} \circ L_A = L_A \circ \mathcal{L}_{ij} \text{ for all } i, j \text{ and } c_{\mathcal{A}} \in \mathrm{Fix}(\mathcal{L}).}$

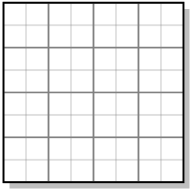
Midori [AC:BBISHA15] in a nutshell

A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- ...yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$



A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- ...yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$

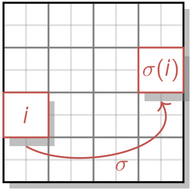
S	S	S	S
S	S	S	S
S	S	S	S
S	S	S	S

A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- ...yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$

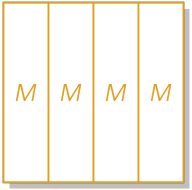


A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- . . . yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$

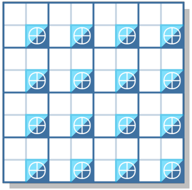


A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- ...yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$



A (not so) standard SPN

- AES-like,
- Standard wide-trail analysis,
- ...yet weak-key probability-1 (non)-linear approximations [TLS19, Bey18]
- due to (excessive) lightweightness and sparsity.

The round function

$$p = AK \circ AC \circ MC \circ PC \circ S$$

The Midori case

$$p = AK \circ AC \circ MC \circ PC \circ S$$

$$p = AK \circ AC \circ MC \circ PC \circ S$$

Sbox layer

There exists a single non-trivial $A^\star$ st. $A^\star \circ S = S \circ A^\star$.

S	S	S	S
S	S	S	S
S	S	S	S
S	S	S	S

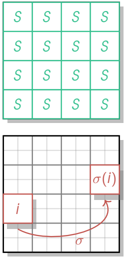
$$p = AK \circ AC \circ MC \circ PC \circ S$$

Sbox layer

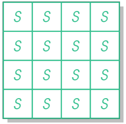
There exists a single non-trivial $A^\star$ st. $A^\star \circ S = S \circ A^\star$.

Cells permutation

Parallel mapping $\mathcal{A}$: free commutation.

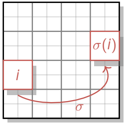


$$p = AK \circ AC \circ MC \circ PC \circ S$$



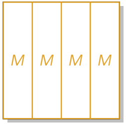
Sbox layer

There exists a single non-trivial $A^\star$ st. $A^\star \circ S = S \circ A^\star$.



Cells permutation

Parallel mapping $\mathcal{A}$: free commutation.



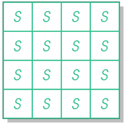
Linear layer

- $M_{ij} \circ L_A = L_A \circ M_{ij}\ \forall\, i, j$. But $M_{ij} \in \{0_4, \mathrm{Id}_4\}$.
- $c_{\mathcal{A}} \in \mathrm{Fix}(\mathcal{L})$. But $M(c,c,c,c) = (c,c,c,c)$.

$\implies$ Any $\mathcal{A}$ would work.

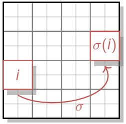
The Midori case

$$p = AK \circ AC \circ MC \circ PC \circ S$$



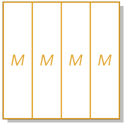
Sbox layer

There exists a single non-trivial $A^\star$ st. $A^\star \circ S = S \circ A^\star$.



Cells permutation

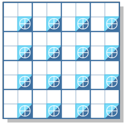
Parallel mapping $\mathcal{A}$: free commutation.



Linear layer

- $M_{ij} \circ L_A = L_A \circ M_{ij}\ \forall\, i, j.$ But $M_{ij} \in \{0_4, \mathrm{Id}_4\}$.
- $c_{\mathcal{A}} \in \mathrm{Fix}(\mathcal{L})$. But $M(c,c,c,c) = (c,c,c,c)$.

$\implies$ Any $\mathcal{A}$ would work.



Constants

$\mathrm{Fix}(L_{A^\star}) = \langle \mathtt{0x2}, \mathtt{0x5}, \mathtt{0x8} \rangle$. $\rightsquigarrow$ Consider **variants** with modified constants.

Weak-keys: 1-bit condition per nibble $\rightsquigarrow 2^{96}$ out of 2^{128}.

The Midori case, part 2

Recap

$\mathcal{A}^\star \circ P = P \circ \mathcal{A}^\star$ for every layer P (given weak constants/keys).

$\mathcal{A}^\star \circ E_k = E_k \circ \mathcal{A}^\star$ for 1 out of 2^{32} keys k.

Recap

$\mathcal{A}^\star \circ P = P \circ \mathcal{A}^\star$ for every layer P (given weak constants/keys).

$\mathcal{A}^\star \circ E_k = E_k \circ \mathcal{A}^\star$ for 1 out of 2^{32} keys k.

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\mathcal{A}^\star"] \\
z_0 \arrow[r, "R_0"'] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, "R_{r-1}"'] & E(z_0)
\end{tikzcd}$$

$$\mathbb{P}_{x \xleftarrow{\$} X}(\underbrace{\mathcal{A}^\star \to \mathcal{A}^\star \to \cdots \to \mathcal{A}^\star}_{r \text{ times}}) = 1, \quad \text{for any } r!$$

The Midori case, part 3

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\Delta_0"', "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\Delta_1"', "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\Delta_{r-1}"', "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\Delta_r"', "\mathcal{A}^\star"] \\
z_0 \arrow[r, "R_0"'] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, "R_{r-1}"'] & E(z_0)
\end{tikzcd}$$

$$\Delta_i := x_i \oplus z_i = x_i \oplus \mathcal{A}^\star(x_i)$$

$$\begin{tikzcd} x_0 \arrow[r, "R_0"] \arrow[d, "\Delta_0"', "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\Delta_1"', "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\Delta_{r-1}"', "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\Delta_r"', "\mathcal{A}^\star"] \\ z_0 \arrow[r, "R_0"'] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, "R_{r-1}"'] & E(z_0) \end{tikzcd}$$

$$\Delta_i := x_i \oplus z_i = x_i \oplus \mathcal{A}^\star(x_i)$$

Surprising differential interpretation

$\delta = \texttt{0xf}, \quad \Delta = \delta^{\otimes 16}, \quad \delta' = \texttt{0xa}, \quad \Delta' = \delta'^{\otimes 16}.$

$$
\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, swap, "\Delta_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, swap, "\Delta_1"] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, swap, "\Delta_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, swap, "\Delta_r"] \arrow[d, "\mathcal{A}^\star"] \\
z_0 \arrow[r, swap, "R_0"] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, swap, "R_{r-1}"] & E(z_0)
\end{tikzcd}
$$

$$\Delta_i := x_i \oplus z_i = x_i \oplus \mathcal{A}^\star(x_i)$$

Surprising differential interpretation

$\delta = \texttt{0xf}, \quad \Delta = \delta^{\otimes 16}, \quad \delta' = \texttt{0xa}, \quad \Delta' = \delta'^{\otimes 16}.$

- $\mathbb{P}_{x \xleftarrow{\$} X} (A^\star(x) = x + \delta) = \frac{1}{2} \quad \mathbb{P}_{x \xleftarrow{\$} X} (A^\star(x) = x + \delta') = \frac{1}{2}.$
- $\forall\, x, \quad x + \mathcal{A}^\star(x) \in \{\delta, \delta'\}^{16}.$

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\Delta_0"', "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\Delta_1"', "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\Delta_{r-1}"', "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\Delta_r"', "\mathcal{A}^\star"] \\
z_0 \arrow[r, "R_0"'] & z_1 \arrow[r, dashed] & z_{r-1} \arrow[r, "R_{r-1}"'] & E(z_0)
\end{tikzcd}$$

$$\Delta_i := x_i \oplus z_i = x_i \oplus \mathcal{A}^\star(x_i)$$

Surprising differential interpretation

$\delta = \texttt{0xf}$, $\quad \Delta = \delta^{\otimes 16}$, $\quad \delta' = \texttt{0xa}$, $\quad \Delta' = \delta'^{\otimes 16}$.

- $\mathbb{P}_{x \xleftarrow{\$} X}(A^\star(x) = x + \delta) = \frac{1}{2} \quad \mathbb{P}_{x \xleftarrow{\$} X}(A^\star(x) = x + \delta') = \frac{1}{2}$.
- $\forall\, x, \quad x + \mathcal{A}^\star(x) \in \{\delta, \delta'\}^{16}$.

$$\Delta \xrightarrow{2^{-16}} \mathcal{A}^\star \xrightarrow{1} \cdots \xrightarrow{1} \mathcal{A}^\star \xrightarrow{2^{-16}} \Delta$$

Weak-key Differential interpretation

Recap

If k is **weak**:

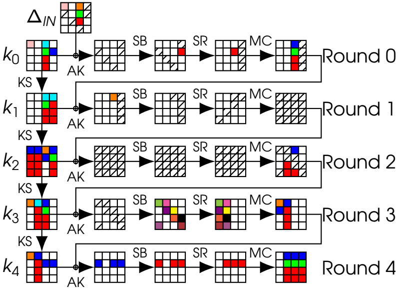
- $\mathbb{P}_{x \xleftarrow{\$} X}(\Delta \to \Delta') = 2^{-32}$ for any $\Delta, \Delta' \in \{\delta, \delta'\}^{16}$.
- $\mathbb{P}_{x \xleftarrow{\$} X}\left(\Delta \to \{\delta, \delta'\}^{16}\right) = 2^{-16}$ for any $\Delta \in \{\delta, \delta'\}^{16}$.
- For any number of rounds, activate all S-boxes.

Recap

If k is **weak**:

- $\mathbb{P}_{x \xleftarrow{\$} X}(\Delta \to \Delta') = 2^{-32}$ for any $\Delta, \Delta' \in \{\delta, \delta'\}^{16}$.
- $\mathbb{P}_{x \xleftarrow{\$} X}(\Delta \to \{\delta, \delta'\}^{16}) = 2^{-16}$ for any $\Delta \in \{\delta, \delta'\}^{16}$.
- For any number of rounds, activate all S-boxes.

Standard case : quite low $\mathbb{P}_{k,x}$



Part of 9-round chosen-key distinguisher for AES-128.
Figure by J. Jean, extracted from Tikz for Cryptographers [Jean16].

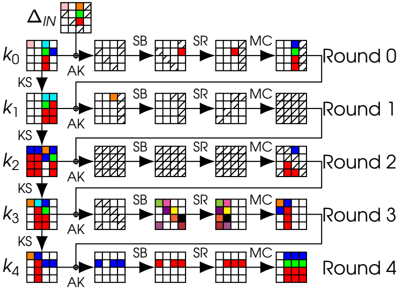
Weak-key Differential interpretation

Recap

If k is **weak**:

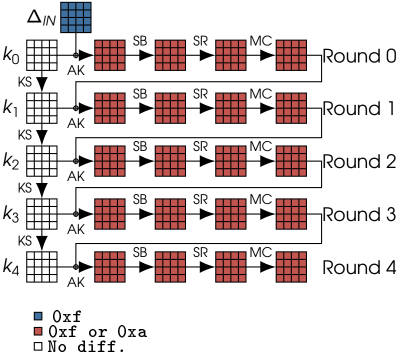
- $\mathbb{P}_{x \xleftarrow{\$} X}(\Delta \rightarrow \Delta') = 2^{-32}$ for any $\Delta, \Delta' \in \{\delta, \delta'\}^{16}$.
- $\mathbb{P}_{x \xleftarrow{\$} X}(\Delta \rightarrow \{\delta, \delta'\}^{16}) = 2^{-16}$ for any $\Delta \in \{\delta, \delta'\}^{16}$.
- For any number of rounds, activate all S-boxes.

Standard case : quite low $\mathbb{P}_{k,x}$

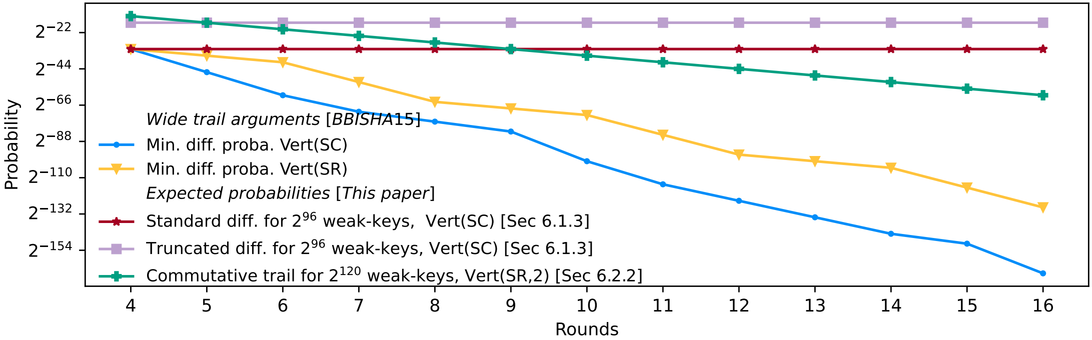


Part of 9-round chosen-key distinguisher for AES-128.
Figure by J. Jean, extracted from Tikz for Cryptographers [Jean16].

This work: high $\mathbb{P}_x$ for some k



Weak-key Differential interpretation, part 2



The designers' work

Estimate $\mathbb{E}_{k \xleftarrow{\$} K} \left(\# \{x \text{ st. } E_k(x+\alpha) = E_k(x) + \beta\}\right)$
and assume representativeness.
Blue curve.

This work

Find non-average keys with easily-distinguishable property.
Purple and red curves.

Is it that easy to detect this behavior ?

Yes !
Small demo here.

A bigger weak-key space ?

Constants

$\mathrm{Fix}(L_{A^\star}) = \langle \mathtt{0x2}, \mathtt{0x5}, \mathtt{0x8} \rangle$.

Weak-keys: 1-bit condition per nibble $\rightsquigarrow 2^{96}$ out of 2^{128}.

Constants

$\mathrm{Fix}(L_{A^\star}) = \langle \mathtt{0x2}, \mathtt{0x5}, \mathtt{0x8} \rangle$.

Weak-keys: 1-bit condition per nibble $\rightsquigarrow 2^{96}$ out of 2^{128}.

$\implies$ "active" S-boxes reduce the key-space.

A bigger weak-key space ?

Constants

$\mathrm{Fix}(L_{A^\star}) = \langle \mathtt{0x2}, \mathtt{0x5}, \mathtt{0x8} \rangle$.

Weak-keys: 1-bit condition per nibble $\rightsquigarrow 2^{96}$ out of 2^{128}.

$\Longrightarrow$ "active" S-boxes reduce the key-space.

$$\begin{pmatrix} A & A & A & A \\ A & A & A & A \\ A & A & A & A \\ A & A & A & A \end{pmatrix} \rightsquigarrow \widetilde{\mathcal{A}}_i = \begin{pmatrix} A^{i_0} & A^{i_4} & A^{i_8} & A^{i_c} \\ A^{i_1} & A^{i_5} & A^{i_9} & A^{i_d} \\ A^{i_2} & A^{i_6} & A^{i_a} & A^{i_e} \\ A^{i_3} & A^{i_7} & A^{i_b} & A^{i_f} \end{pmatrix}, \text{ where } A^0 = \mathrm{Id}, A^1 = A$$

Constants

$\mathrm{Fix}(L_{A^\star}) = \langle \mathtt{0x2}, \mathtt{0x5}, \mathtt{0x8} \rangle$.

Weak-keys: 1-bit condition per nibble $\rightsquigarrow 2^{96}$ out of 2^{128}.

$\Longrightarrow$ "active" S-boxes reduce the key-space.

$$\begin{pmatrix} A & A & A & A \\ A & A & A & A \\ A & A & A & A \\ A & A & A & A \end{pmatrix} \rightsquigarrow \widetilde{A}_i = \begin{pmatrix} A^{i_0} & A^{i_4} & A^{i_8} & A^{i_c} \\ A^{i_1} & A^{i_5} & A^{i_9} & A^{i_d} \\ A^{i_2} & A^{i_6} & A^{i_a} & A^{i_e} \\ A^{i_3} & A^{i_7} & A^{i_b} & A^{i_f} \end{pmatrix}, \text{ where } A^0 = \mathrm{Id}, A^1 = A$$

New study

- Constants : 1-bit condition if $i_j = 1$ else, 0.
- S-box: $S \circ A = A \circ S, S \circ \mathrm{Id} = \mathrm{Id} \circ S$
- Cell permutation:

$$\text{ShiftRows} \begin{pmatrix} X_0 & X_4 & X_8 & X_c \\ X_1 & X_5 & X_9 & X_d \\ X_2 & X_6 & X_a & X_e \\ X_3 & X_7 & X_b & X_f \end{pmatrix} = \begin{pmatrix} X_0 & X_4 & X_8 & X_c \\ X_5 & X_9 & X_d & X_1 \\ X_a & X_e & X_2 & X_6 \\ X_f & X_3 & X_7 & X_b \end{pmatrix}$$

A bigger weak-key space ? part 2

New pattern

$$\begin{pmatrix} A & \mathrm{Id} & A & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ A & \mathrm{Id} & A & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \end{pmatrix}$$

Commutes with S-box layer, cells perm. and weak-constants/weak-key addition.

A bigger weak-key space ? part 2

New pattern

$$\begin{pmatrix} A & \mathrm{Id} & A & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ A & \mathrm{Id} & A & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \end{pmatrix}$$

Commutes with S-box layer, cells perm. and weak-constants/weak-key addition.

What about M ?

$$M := \begin{pmatrix} 0 & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & 0 & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & 0 & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & 0 \end{pmatrix}$$

We indeed have $M_{ij} \in \{0, \mathrm{Id}\}$ and $M(c, c, c, c) = (c, c, c, c)$.

A bigger weak-key space ? part 3

$$M := \begin{pmatrix} 0 & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & 0 & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & 0 & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & 0 \end{pmatrix}$$

A bigger weak-key space ? part 3

$$M := \begin{pmatrix} 0 & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & 0 & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & 0 & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & 0 \end{pmatrix}$$

$$M \begin{pmatrix} Ax_0 \\ x_1 \\ Ax_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + Ax_2 + x_3 \\ Ax_0 + Ax_2 + x_3 \\ Ax_0 + x_1 + x_3 \\ Ax_0 + x_1 + Ax_2 \end{pmatrix} = \begin{pmatrix} x_1 + L_A x_2 + x_3 + c_A \\ L_A x_0 + L_A x_2 + x_3 \\ L_A x_0 + x_1 + x_3 + c_A \\ L_A x_0 + x_1 + L_A x_2 \end{pmatrix} \tag{1}$$

A bigger weak-key space ? part 3

$$M := \begin{pmatrix} 0 & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & 0 & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & 0 & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & 0 \end{pmatrix}$$

$$M \begin{pmatrix} Ax_0 \\ x_1 \\ Ax_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + Ax_2 + x_3 \\ Ax_0 + Ax_2 + x_3 \\ Ax_0 + x_1 + x_3 \\ Ax_0 + x_1 + Ax_2 \end{pmatrix} = \begin{pmatrix} x_1 + L_A x_2 + x_3 + c_A \\ L_A x_0 + L_A x_2 + x_3 \\ L_A x_0 + x_1 + x_3 + c_A \\ L_A x_0 + x_1 + L_A x_2 \end{pmatrix} \tag{1}$$

$$A \times \mathrm{Id} \times A \times \mathrm{Id} \circ M(x_0, x_1, x_2, x_3) = \begin{pmatrix} A(x_1 + x_2 + x_3) \\ x_0 + x_2 + x_3 \\ A(x_0 + x_1 + x_3) \\ x_0 + x_1 + x_2 \end{pmatrix} = \begin{pmatrix} L_A x_1 + L_A x_2 + L_A x_3 + c_A \\ x_0 + x_2 + x_3 \\ L_A x_0 + L_A x_1 + L_A x_3 + c_A \\ x_0 + x_1 + x_2 \end{pmatrix} \tag{2}$$

A bigger weak-key space ? part 3

$$M := \begin{pmatrix} 0 & \mathrm{Id} & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & 0 & \mathrm{Id} & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & 0 & \mathrm{Id} \\ \mathrm{Id} & \mathrm{Id} & \mathrm{Id} & 0 \end{pmatrix}$$

$$M \begin{pmatrix} Ax_0 \\ x_1 \\ Ax_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + Ax_2 + x_3 \\ Ax_0 + Ax_2 + x_3 \\ Ax_0 + x_1 + x_3 \\ Ax_0 + x_1 + Ax_2 \end{pmatrix} = \begin{pmatrix} x_1 + L_A x_2 + x_3 + c_A \\ L_A x_0 + L_A x_2 + x_3 \\ L_A x_0 + x_1 + x_3 + c_A \\ L_A x_0 + x_1 + L_A x_2 \end{pmatrix} \tag{1}$$

$$A \times \mathrm{Id} \times A \times \mathrm{Id} \circ M(x_0, x_1, x_2, x_3) = \begin{pmatrix} A(x_1 + x_2 + x_3) \\ x_0 + x_2 + x_3 \\ A(x_0 + x_1 + x_3) \\ x_0 + x_1 + x_2 \end{pmatrix} = \begin{pmatrix} L_A x_1 + L_A x_2 + L_A x_3 + c_A \\ x_0 + x_2 + x_3 \\ L_A x_0 + L_A x_1 + L_A x_3 + c_A \\ x_0 + x_1 + x_2 \end{pmatrix} \tag{2}$$

$$(1) = (2) \iff \begin{pmatrix} L_A x_1 + x_1 + L_A x_3 + x_3 \\ L_A x_0 + x_0 + L_A x_2 + x_2 \\ L_A x_1 + x_1 + L_A x_3 + x_3 \\ L_A x_0 + x_0 + L_A x_2 + x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{matrix} (L_A + \mathrm{Id})(x_0 + x_2) = 0 \\ (L_A + \mathrm{Id})(x_1 + x_3) = 0 \end{matrix}$$

A bigger weak-key space ? final

Recap

$A \times \mathrm{Id} \times A \times \mathrm{Id} \circ M(x_0, x_1, x_2, x_3) = M\,(Ax_0, x_1, Ax_2, x_3)$ if and only if $x_0 + x_2 \in \ker(L_A + \mathrm{Id})$ and $x_1 + x_3 \in \ker(L_A + \mathrm{Id})$

Recap

$A \times \mathrm{Id} \times A \times \mathrm{Id} \circ M(x_0, x_1, x_2, x_3) = M\left(Ax_0, x_1, Ax_2, x_3\right)$ if and only if $x_0 + x_2 \in \ker(L_A + \mathrm{Id})$ and $x_1 + x_3 \in \ker(L_A + \mathrm{Id})$

Fact

$\dim(\ker(L_A + \mathrm{Id})) = 2.$

First probabilistic commutation, first trade-off

$$\mathbb{P}_{x \xleftarrow{\$} X}\left(\mathcal{A} \circ \mathcal{M}(x) = \mathcal{M} \circ \mathcal{A}(x)\right) = \frac{2^2}{2^4} \times \frac{2^2}{2^4} = 2^{-4}.$$

For $2^{128-2\times4} = 2^{120}$ weak keys, $\mathbb{P}_{x \xleftarrow{\$} X}\left(R \circ \mathcal{M}(x) = \mathcal{M} \circ R(x)\right) = 2^{-4}$.

A bigger weak-key space ? final

Recap

$A \times \mathrm{Id} \times A \times \mathrm{Id} \circ M(x_0, x_1, x_2, x_3) = M\,(Ax_0, x_1, Ax_2, x_3)$ if and only if $x_0 + x_2 \in \ker(L_A + \mathrm{Id})$ and $x_1 + x_3 \in \ker(L_A + \mathrm{Id})$

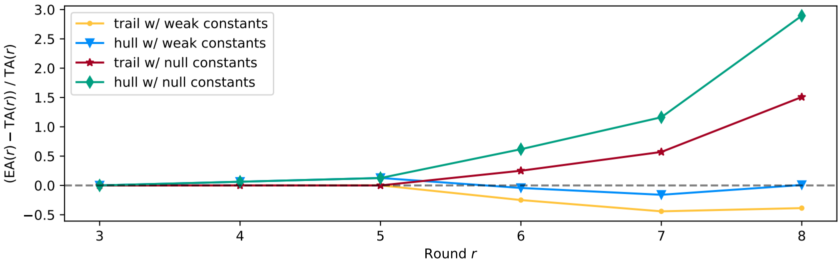
Fact

$\dim(\ker(L_A + \mathrm{Id})) = 2$.

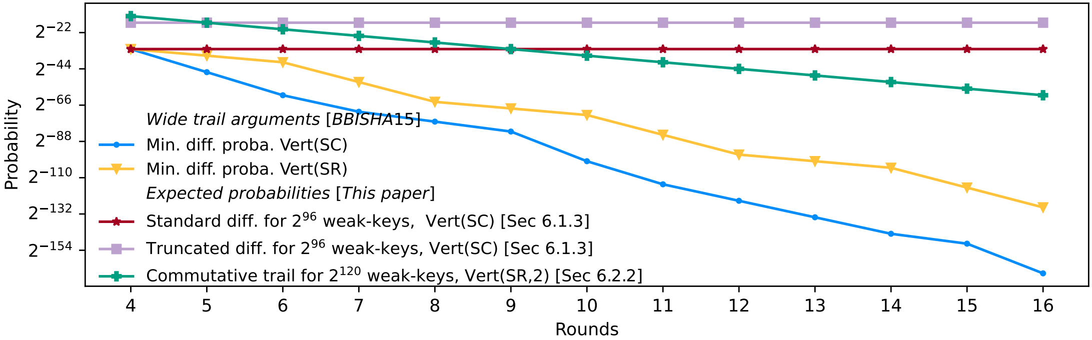
First probabilistic commutation, first trade-off

$$\mathbb{P}_{x \xleftarrow{\$} X}\left(\mathcal{A} \circ \mathcal{M}(x) = \mathcal{M} \circ \mathcal{A}(x)\right) = \frac{2^2}{2^4} \times \frac{2^2}{2^4} = 2^{-4}.$$

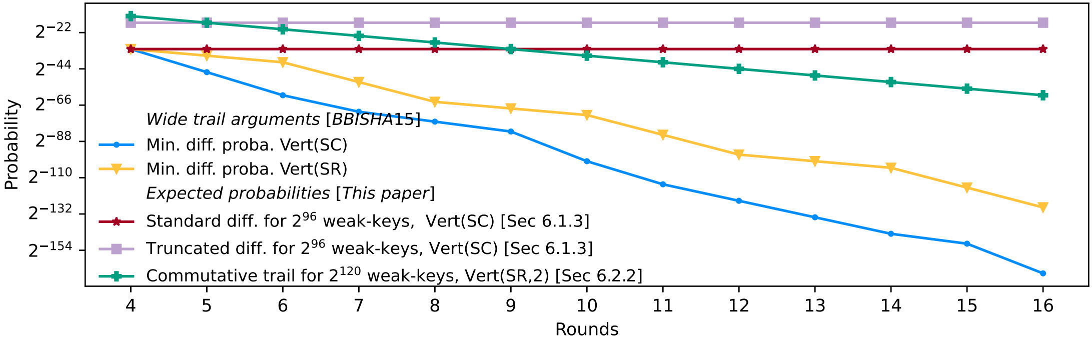
For $2^{128-2\times 4} = 2^{120}$ weak keys, $\mathbb{P}_{x \xleftarrow{\$} X}\left(R \circ \mathcal{M}(x) = \mathcal{M} \circ R(x)\right) = 2^{-4}$.



A few words about probabilistic commutation



A few words about probabilistic commutation



Probabilistic commutation with different layers

Let $p \in [0, 1]$.

- $A \circ T_k \overset{p}{=} T_k \circ B$: well-understood.
- $A \circ L \overset{p}{=} L \circ B$: manageable for parallel mappings.
- $A \circ S \overset{p}{=} S \circ B$: 4-bit mappings can be listed exhaustively.

Conclusion

In practice

- Trade-offs: number-of-weak-keys VS probability-of-success.
- Independence of rounds must be supposed ...but often too optimistic.

Further studies

- Algorithm for probabilistic affine-equivalence.
- Study the dependencies.
- Hybridization: *e.g.* commutative-differential ?

Conclusion

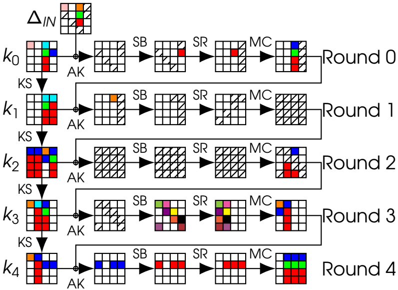
In practice

- Trade-offs: number-of-weak-keys VS probability-of-success.
- Independence of rounds must be supposed . . . but often too optimistic.

Further studies

- Algorithm for probabilistic affine-equivalence.
- Study the dependencies.
- Hybridization: *e.g.* commutative-differential ?

Standard case : quite low $\mathbb{P}_{k,x}$



This work: high $\mathbb{P}_x$ for some weak k