Commutative Cryptanalysis Made Practical

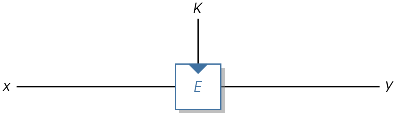
Jules Baudrin

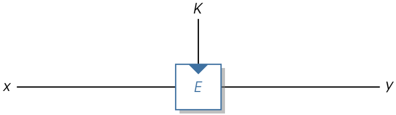
jules.baudrin@inria.fr

Inria, Paris, France

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

Journées Codage et Cryptographie 2023

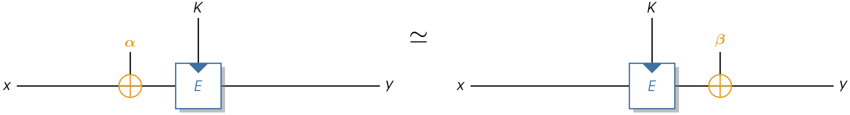

K
x
E
y



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

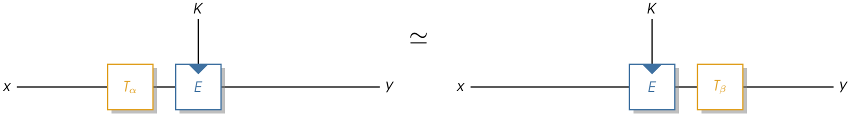
Overview of symmetric cryptanalysis

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

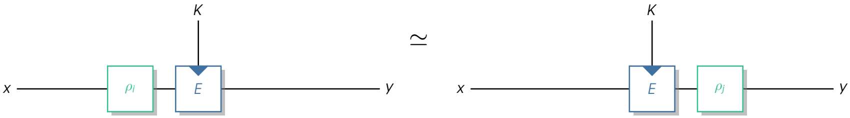


Overview of symmetric cryptanalysis

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

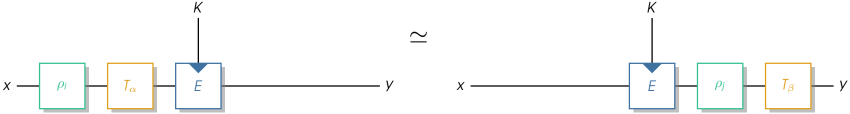


$$E \circ \rho_i(x) = \rho_j \circ E(x)$$

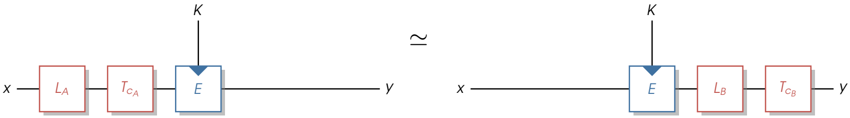


Overview of symmetric cryptanalysis

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



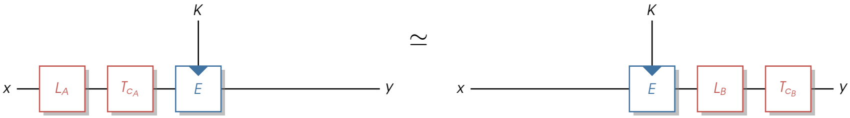
$$E \circ T_{c_A} \circ L_A(x) = T_{c_b} \circ L_B \circ E(x).$$



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

Overview of symmetric cryptanalysis

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



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

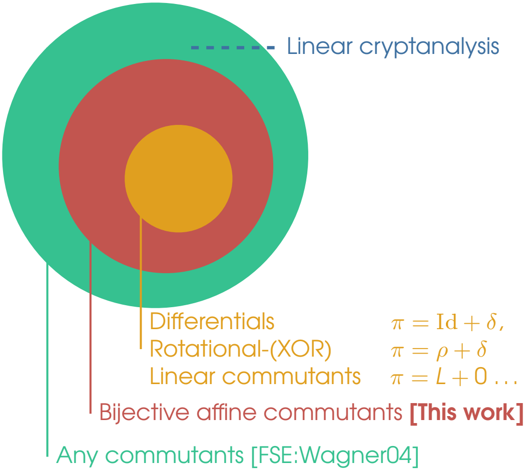
A 20-year-old idea [Wagner, FSE 2004]

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

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

Commutative (diagram) cryptanalysis

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



In this talk

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

A **surprising differential** interpretation

A few words about the **probabilistic case**

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. : $\boxed{E \circ A = B \circ E}$ (for many k, if $E = (E_k)_k$).

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. : $\boxed{E \circ A = B \circ E}$ (for many k, if $E = (E_k)_k$).

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

Commutative cryptanalysis principle

Goal

Find **bijective affine** A, B st. : $\boxed{E \circ A = B \circ E}$ (for many k, if $E = (E_k)_k$).

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

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "A_0", "(\star)"'] & x_1 \arrow[r, dashed] \arrow[d, "A_1"] \arrow[dr, phantom, "\circlearrowleft"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "A_{r-1}"] & E(x_0) \arrow[d, "A_r", "(\star)"'] \\
y_0 \arrow[r, "R_0"'] & y_1 \arrow[r, dashed] & y_{r-1} \arrow[r, "R_{r-1}"'] & E(y_0)
\end{tikzcd}$$

$$\begin{aligned}
(\star) \quad & y_0 = A_0(x_0) \\
(\star) \quad & E(y_0) = A_r \circ E(x_0) \\
\Longrightarrow \quad & \boxed{E \circ A_0(x_0) = A_r \circ E(x_0)}
\end{aligned}$$

Sufficient condition for iterated constructions

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

$\Longrightarrow$ **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

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

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

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

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

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

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

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

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

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

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

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

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

Linear layer

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

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

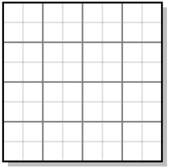
Midori [AC:BBISHA15] in a nutshell

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

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

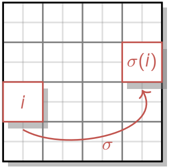
S	S	S	S
S	S	S	S
S	S	S	S
S	S	S	S

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

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



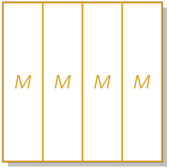
Midori [AC:BBISHA15] in a nutshell

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

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

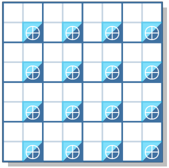


A (not so) standard SPN

- 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

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



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

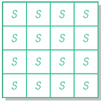
$$p = AK \circ AC \circ MC \circ PC \circ S$$

The Midori case

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

The Midori case

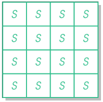
$$p = AK \circ AC \circ MC \circ PC \circ S$$



Sbox layer

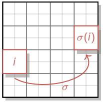
There exists a single non-trivial $A^\star$ st. $A^\star \circ S = S \circ A^\star$.

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



Sbox layer

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

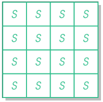


Cells permutation

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

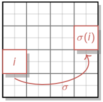
The Midori case

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



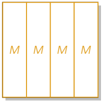
Sbox layer

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

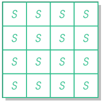
- $M_{ij} \circ L_A = L_A \circ M_{ij}\ \forall\, i, j$. But $M_{ij} \in \{0_4, \mathrm{Id}_4\}$.
- $c_{\mathcal{A}} \in \mathrm{Fix}(\mathcal{L})$. But $M(c, c, c, c) = (c, c, c, c)$.

Any $\mathcal{A}$ would work.

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

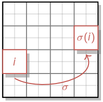
Sbox layer

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



Cells permutation

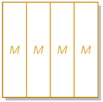
Parallel mapping $\mathcal{A}$: free commutation.



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 $M(c, c, c, c) = (c, c, c, c)$.

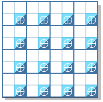
Any $\mathcal{A}$ would work.



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/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).

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

$\mathcal{A}^\star \circ E_k = E_k \circ \mathcal{A}^\star$ 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).

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

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

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\mathcal{A}^\star"] \\
y_0 \arrow[r, "R_0"'] & y_1 \arrow[r, dashed] & y_{r-1} \arrow[r, "R_{r-1}"'] & E(y_0)
\end{tikzcd}$$

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

The Midori case, part 2

Recap

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

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

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

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\mathcal{A}^\star"] \\
y_0 \arrow[r, "R_0"'] & y_1 \arrow[r, dashed] & y_{r-1} \arrow[r, "R_{r-1}"'] & E(y_0)
\end{tikzcd}$$

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

Surprising differential interpretation

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

The Midori case, part 2

Recap

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

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

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

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\mathcal{A}^\star"] \\
y_0 \arrow[r, "R_0"'] & y_1 \arrow[r, dashed] & y_{r-1} \arrow[r, "R_{r-1}"'] & E(y_0)
\end{tikzcd}$$

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

Surprising differential interpretation

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

- $\mathcal{A}^\star$: $\quad \mathbb{P}_{\boldsymbol{x}}\left(\mathcal{A}^\star(x) = x + \texttt{0xf}\right) = \frac{1}{2} \quad \mathbb{P}_{\boldsymbol{x}}\left(\mathcal{A}^\star(x) = x + \texttt{0xa}\right) = \frac{1}{2}.$
- $\mathcal{A}^\star$: $\quad \forall\, x, \quad x + \mathcal{A}^\star(x) \in \{\delta, \delta'\}^{16}$

The Midori case, part 2

Recap

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

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

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

$$\begin{tikzcd}
x_0 \arrow[r, "R_0"] \arrow[d, "\mathcal{A}^\star"] & x_1 \arrow[r, dashed] \arrow[d, "\mathcal{A}^\star"] & x_{r-1} \arrow[r, "R_{r-1}"] \arrow[d, "\mathcal{A}^\star"] & E(x_0) \arrow[d, "\mathcal{A}^\star"] \\
y_0 \arrow[r, "R_0"'] & y_1 \arrow[r, dashed] & y_{r-1} \arrow[r, "R_{r-1}"'] & E(y_0)
\end{tikzcd}$$

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

Surprising differential interpretation

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

- $\mathcal{A}^\star$: $\quad \mathbb{P}_{\boldsymbol{x}}\left(\mathcal{A}^\star(x) = x + \mathtt{0xf}\right) = \frac{1}{2} \quad \mathbb{P}_{\boldsymbol{x}}\left(\mathcal{A}^\star(x) = x + \mathtt{0xa}\right) = \frac{1}{2}.$
- $\mathcal{A}^\star$: $\quad \forall\, x, \quad x + \mathcal{A}^\star(x) \in \{\delta, \delta'\}^{16}$

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

Fixed-key Differential interpretation

Recap

If k is weak (**fixed-key** setting):

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

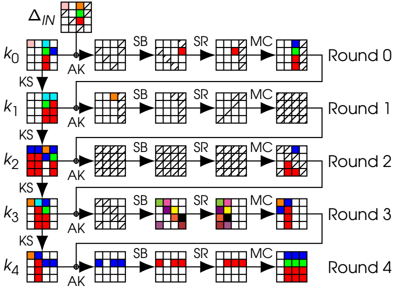
Fixed-key Differential interpretation

Recap

If k is weak (**fixed-key** setting):

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

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



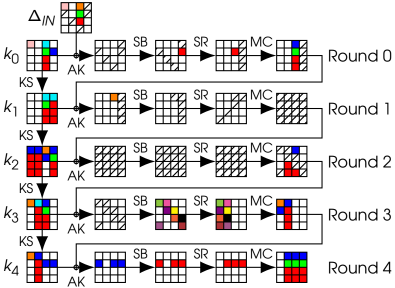
Part of 9-round chosen-key distinguisher for AES-128.
Figure by J. Jean, extracted from Tikz for Cryptographers [Jean16].

Recap

If k is weak (**fixed-key** setting):

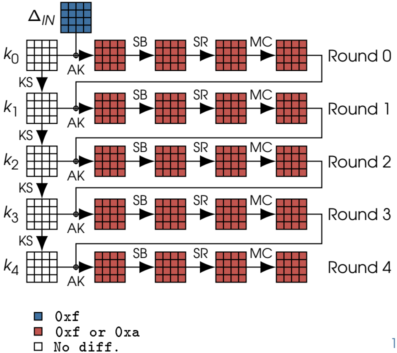
- $\mathbb{P}_{\boldsymbol{x}}(\Delta \to \Delta') = 2^{-32}$ for any $\Delta, \Delta' \in \{\delta, \delta'\}^{16}$.
- $\mathbb{P}_{\boldsymbol{x}}\left(\Delta \to \{\delta, \delta'\}^{16}\right) = 2^{-16}$ for any $\Delta \in \{\delta, \delta'\}^{16}$.
- For any number of rounds, activate all S-boxes.

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



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

This work: high $\mathbb{P}_{\boldsymbol{x}}$ for some k



What about probabilistic commutative trails?

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.

What about probabilistic commutative trails?

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.

In practice

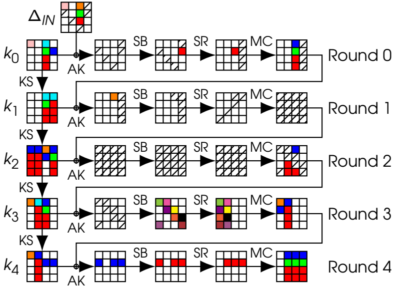
- Trade-offs: number-of-weak-keys VS probability-of-success.
- Independence of rounds must be supposed . . .
- . . . but often too optimistic.

Conclusion

Further studies

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

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



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

This work: high $\mathbb{P}_{\boldsymbol{x}}$ for some k