Linear and differential cryptanalysis

Tim Beyne

tim@cryptanalysis.info

KU Leuven

March 11, 2025

Differential cryptanalysis

Overview

- Differentials and differential characteristics
- Quasidifferential transition matrices
- Quasidifferential trails
- Cost analysis
- Key-recovery techniques
- Exercises

Differentials

- Probabilistic relation between an input difference a and an output difference b

$$\mathsf{F}(x + a) \approx \mathsf{F}(x) + b$$

- Pair (a, b) of differences $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2^m$ determines the differential

Differentials

- Probabilistic relation between an input difference a and an output difference b

$$\mathsf{F}(x + a) \approx \mathsf{F}(x) + b$$

- Pair (a, b) of differences $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2^m$ determines the differential

- If F is a uniform random function, then the number of inputs x such that $\mathsf{F}(x + a) = \mathsf{F}(x) + b$ is $2^n/2^m$ on average

- Probability of a differential:

$$p = \frac{\#\{x \in \mathbb{F}_2^n \mid \mathsf{F}(x + a) = \mathsf{F}(x) + b\}}{2^n} = \Pr_{\boldsymbol{x}}\left[\mathsf{F}(\boldsymbol{x} + a) = \mathsf{F}(\boldsymbol{x}) + b\right]$$

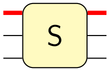
Differentials

Example

- 3-bit S-box $S\colon \mathbb{F}_2^3 \to \mathbb{F}_2^3$

x	000	001	010	011	100	101	110	111
$S(x)$	111	010	100	101	001	110	011	000

- Differential $(a, b) = (001, 001)$

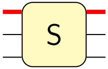


Differentials
Example

- 3-bit S-box $\mathsf{S}\colon \mathbb{F}_2^3 \to \mathbb{F}_2^3$

x	000	001	010	011	100	101	110	111
$\mathsf{S}(x)$	111	010	100	101	001	110	011	000

- Differential $(a, b) = (001, 001)$



- Probability $\Pr_{\boldsymbol{x}} \left[\mathsf{S}(\boldsymbol{x} + a) = \mathsf{S}(\boldsymbol{x}) + b\right] = \frac{2}{8} = \frac{1}{4}$

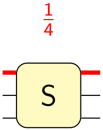
Differentials

Distinguishers

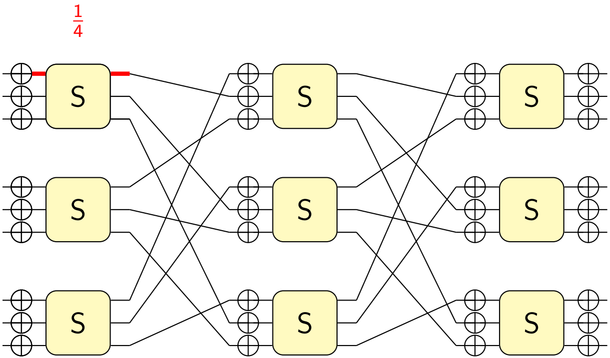
- Sample q input pairs $(x_1, x_1 + a), \ldots, (x_q, x_q + a)$ at random
- Average number of pairs with output difference b is pq
- $q \approx 1/p$ samples are enough for a distinguisher because right pairs are uncommon (assuming p is not too small or large)

Number of samples depends on true- and false-positive probabilities (see later)

Differentials



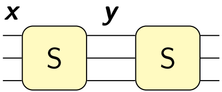
Differentials



Propagation through a sequence of operations?

Differentials

Example



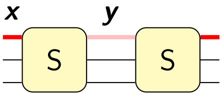
$$\Pr[S^2(\boldsymbol{x} + a) = S^2(\boldsymbol{x}) + b] \approx \Pr[S(\boldsymbol{x} + a) = S(\boldsymbol{x}) + c \text{ and } S(\boldsymbol{y} + c) = S(\boldsymbol{y}) + b]$$

Pretend that $\boldsymbol{x}$ and $\boldsymbol{y}$ are independent:

$$\Pr[S^2(\boldsymbol{x} + a) = S^2(\boldsymbol{x}) + b] \approx \Pr[S(\boldsymbol{x} + a) = S(\boldsymbol{x}) + c] \times \Pr[S(\boldsymbol{y} + c) = S(\boldsymbol{y}) + b]$$

Differentials

Example



$$\Pr[S^2(\boldsymbol{x} + a) = S^2(\boldsymbol{x}) + b] \approx \Pr[S(\boldsymbol{x} + a) = S(\boldsymbol{x}) + c \text{ and } S(\boldsymbol{y} + c) = S(\boldsymbol{y}) + b]$$

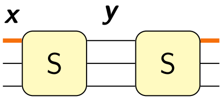
Pretend that $\boldsymbol{x}$ and $\boldsymbol{y}$ are independent:

$$\Pr[S^2(\boldsymbol{x} + a) = S^2(\boldsymbol{x}) + b] \overset{\text{☠}}{\approx} \Pr[S(\boldsymbol{x} + a) = S(\boldsymbol{x}) + c] \times \Pr[S(\boldsymbol{y} + c) = S(\boldsymbol{y}) + b]$$

- For example: $a = b = c = 001$ gives $1/4 \times 1/4 = 1/16$
- Unfortunately, this is wrong (the correct result is $1/4$)

Differentials

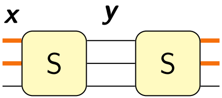
Example



$$
\begin{aligned}
&\Pr[\mathsf{S}^2(\boldsymbol{x} + 001) = \mathsf{S}^2(\boldsymbol{x}) + 001] \\
= \; &\Pr[\mathsf{S}(\boldsymbol{x} + 001) = \mathsf{S}(\boldsymbol{x}) + 001 \text{ and } \mathsf{S}(\boldsymbol{y} + 001) = \mathsf{S}(\boldsymbol{y}) + 001] + \\
&\Pr[\mathsf{S}(\boldsymbol{x} + 001) = \mathsf{S}(\boldsymbol{x}) + 011 \text{ and } \mathsf{S}(\boldsymbol{y} + 011) = \mathsf{S}(\boldsymbol{y}) + 001] + \\
&\Pr[\mathsf{S}(\boldsymbol{x} + 001) = \mathsf{S}(\boldsymbol{x}) + 101 \text{ and } \mathsf{S}(\boldsymbol{y} + 101) = \mathsf{S}(\boldsymbol{y}) + 001] + \\
&\Pr[\mathsf{S}(\boldsymbol{x} + 001) = \mathsf{S}(\boldsymbol{x}) + 111 \text{ and } \mathsf{S}(\boldsymbol{y} + 111) = \mathsf{S}(\boldsymbol{y}) + 001] \\
\approx \; &\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}
\end{aligned}
$$

Differentials

Example



$$
\begin{aligned}
& \Pr[S^2(\boldsymbol{x} + 011) = S^2(\boldsymbol{x}) + 011] \\
= \; & \Pr[S(\boldsymbol{x} + 011) = S(\boldsymbol{x}) + 001 \text{ and } S(\boldsymbol{y} + 001) = S(\boldsymbol{y}) + 011] + \\
& \Pr[S(\boldsymbol{x} + 011) = S(\boldsymbol{x}) + 010 \text{ and } S(\boldsymbol{y} + 010) = S(\boldsymbol{y}) + 011] + \\
& \Pr[S(\boldsymbol{x} + 011) = S(\boldsymbol{x}) + 101 \text{ and } S(\boldsymbol{y} + 101) = S(\boldsymbol{y}) + 011] + \\
& \Pr[S(\boldsymbol{x} + 001) = S(\boldsymbol{x}) + 110 \text{ and } S(\boldsymbol{y} + 110) = S(\boldsymbol{y}) + 001] \\
\approx \; & \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}
\end{aligned}
$$

- Unfortunately, this is still wrong (the correct result is 0)
- It is not reasonable to assume independence

Differential characteristics

- Suppose $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$ and let $\boldsymbol{x}_i = \mathsf{F}_i(\boldsymbol{x}_{i-1})$ with $\boldsymbol{x}_0 = \boldsymbol{x}$
- Law of total probability:

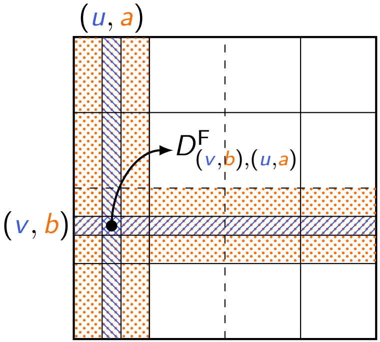
$$\Pr[\mathsf{F}(\boldsymbol{x} + a_1) = \mathsf{F}(\boldsymbol{x}) + a_{r+1}] = \sum_{a_2,\ldots,a_r} \Pr\left[\bigwedge_{i=1}^{r} \mathsf{F}_i(\boldsymbol{x}_i + a_i) = \mathsf{F}(\boldsymbol{x}_i) + a_{i+1}\right]$$

- A sequence $(a_1, a_2, \ldots, a_{r+1})$ is called a differential characteristic
- How to calculate the probability of a characteristic?

Quasidifferential transition matrices

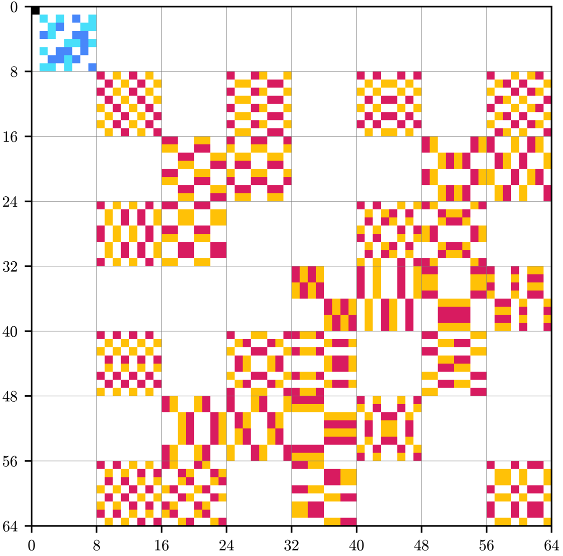
- $2^{2m} \times 2^{2n}$ matrix corresponding to $\mathsf{F}\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$

$$D^{\mathsf{F}}_{(v,b),(u,a)} = \left(2 \Pr_{\boldsymbol{x}}\left[v^{\mathsf{T}}\mathsf{F}(\boldsymbol{x}) = u^{\mathsf{T}}\boldsymbol{x} \mid \mathsf{F}(\boldsymbol{x}+a) = \mathsf{F}(\boldsymbol{x}) + b\right] - 1\right) \times \Pr_{\boldsymbol{x}}\left[\mathsf{F}(\boldsymbol{x}+a) = \mathsf{F}(\boldsymbol{x}) + b\right]$$



Quasidifferential transition matrices

Example



Quasidifferential transition matrices

Multiplication property

- If $\mathsf{F} = \mathsf{F}_2 \circ \mathsf{F}_1$, then
$$D^{\mathsf{F}} = D^{\mathsf{F}_2} D^{\mathsf{F}_1}$$

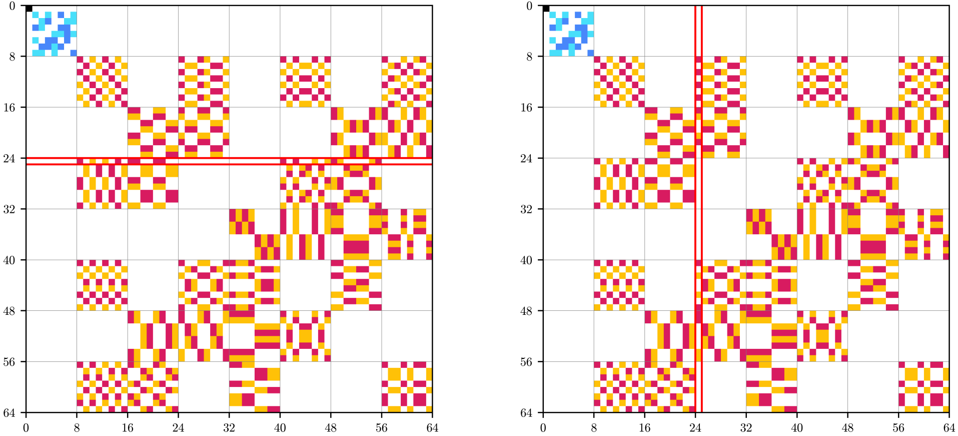
Proof by calculation

- This is the most important property of quasidifferential transition matrices

- There are more conceptual (but more abstract) proofs without calculation

Quasidifferential transition matrices

Multiplication property: example



Quasidifferential trails

- If $\mathsf{F} = \mathsf{F}_r \circ \cdots \mathsf{F}_2 \circ \mathsf{F}_1$, then $D^{\mathsf{F}} = D^{\mathsf{F}_r} \cdots D^{\mathsf{F}_2} D^{\mathsf{F}_1}$, so

$$D^{\mathsf{F}}_{\varpi_{r+1},\varpi_1} = \sum_{\varpi_2,\ldots,\varpi_r} D^{\mathsf{F}_r}_{\varpi_{r+1},\varpi_r} \cdots D^{\mathsf{F}_2}_{\varpi_3,\varpi_2} D^{\mathsf{F}_1}_{\varpi_2,\varpi_1}$$

  with $\varpi_i = (u_i, a_i)$ for $i \in \{1, \ldots, r\}$

- A quasidifferential trail is a sequence $(\varpi_1, \ldots, \varpi_{r+1})$ with correlation $\prod_{i=1}^{r} D^{\mathsf{F}_i}_{\varpi_{i+1},\varpi_i}$

- Analysis relies on the assumption that there exists a set Λ of 'dominant trails':

$$D^{\mathsf{F}}_{\varpi_{r+1},\varpi_1} = \sum_{\varpi\in\Lambda} \prod_{i=1}^{r} D^{\mathsf{F}_i}_{\varpi_{i+1},\varpi_i} + \varepsilon$$

Quasidifferential trails

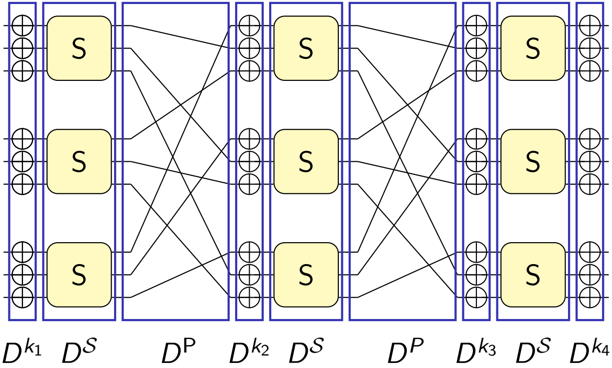
- $D^{\mathsf{F}}_{(0,a_{r+1}),(0,a_1)}$ is the probability of the differential (a_1, a_{r+1})
- Quasidifferential trails can be used to compute the probability of a differential
- Quasidifferential trails can be used to compute the probability of a characteristic:

$$\sum_{u_2,\ldots,u_r} \prod_{i=1}^{r} D^{\mathsf{F}_i}_{(u_{i+1},a_{i+1}),(u_i,a_i)}$$

Proof: similar as for the multiplication property (exercise)
visual proof (💻)

Quasidifferential trails

Example



- To analyze trails we need to determine D^{k_i}, $D^{\mathcal{S}}$ and D^{P}

Quasidifferential trails
Bricklayer functions

- If $\mathsf{F}(x_1 \| x_2) = \mathsf{F}_1(x_1) \| \mathsf{F}_2(x_2)$, then

$$D^{\mathsf{F}}_{(v_1\|v_2, b_1\|b_2),(u_1\|u_2, a_1\|a_2)} = D^{\mathsf{F}_1}_{(v_1, b_1),(u_1, a_1)} D^{\mathsf{F}_2}_{(v_2, b_2),(u_2, a_2)}$$

Proof by calculation

- Equivalently, $D^{\mathsf{F}} = D^{\mathsf{F}_1} \otimes D^{\mathsf{F}_2}$
- For the S-box layer: $D^{\mathcal{S}} = D^{S} \otimes D^{S} \otimes D^{S}$

Quasidifferential trails

Translations and linear functions

- If $\mathsf{F}(x) = x + k$, then

$$D^{\mathsf{F}}_{(v,b),(u,a)} = \begin{cases} (-1)^{v^\top k} & \text{if } u = v \text{ and } a = b \\ 0 & \text{else.} \end{cases}$$

Proof

Quasidifferential trails

Translations and linear functions

- If $\mathsf{F}(x) = x + k$, then

$$D^{\mathsf{F}}_{(v,b),(u,a)} = \begin{cases} (-1)^{v^\top k} & \text{if } u = v \text{ and } a = b \\ 0 & \text{else.} \end{cases}$$

Proof

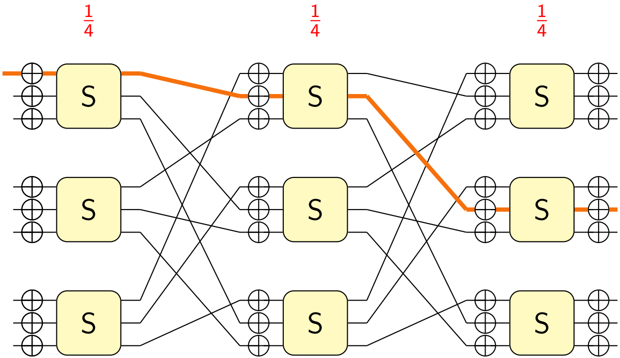
- If $\mathsf{F}(x) = Mx$, then

$$D^{\mathsf{F}}_{(v,b),(u,a)} = \begin{cases} 1 & \text{if } u = M^\top v \text{ and } b = Ma \\ 0 & \text{else.} \end{cases}$$

Proof

Quasidifferential trails

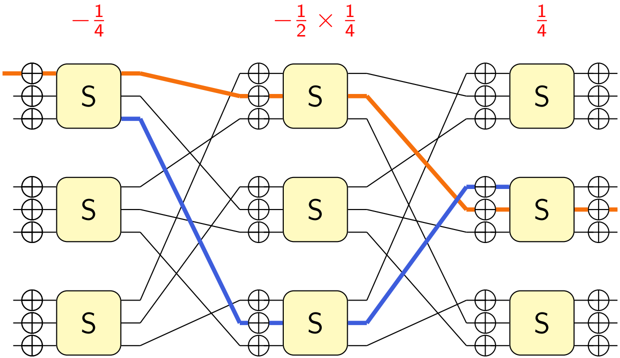
Example: 3-round differential (characteristic 1)



$$\frac{1}{2^6}$$

Quasidifferential trails

Example: 3-round differential (characteristic 1)

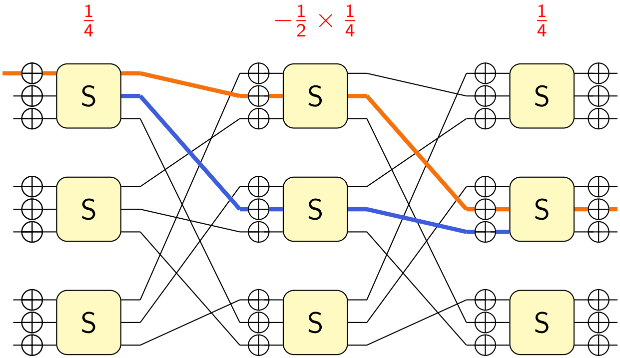


$$\frac{1}{2^6} + (-1)^{\kappa_1} \frac{1}{2^7}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$

Quasidifferential trails

Example: 3-round differential (characteristic 1)

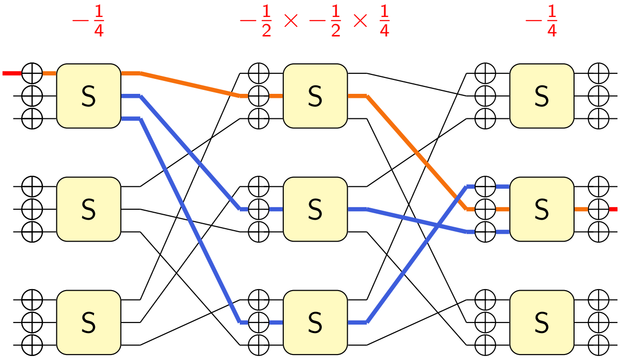


$$\frac{1}{2^6} + (-1)^{\kappa_1} \frac{1}{2^7} + (-1)^{\kappa_2} \frac{1}{2^7}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$

Quasidifferential trails

Example: 3-round differential (characteristic 1)

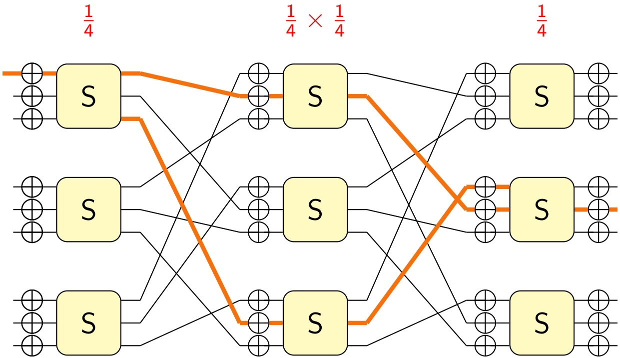


$$\frac{1}{2^6} + (-1)^{\kappa_1}\frac{1}{2^7} + (-1)^{\kappa_2}\frac{1}{2^7} + (-1)^{\kappa_1+\kappa_2}\frac{1}{2^8}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$

Quasidifferential trails

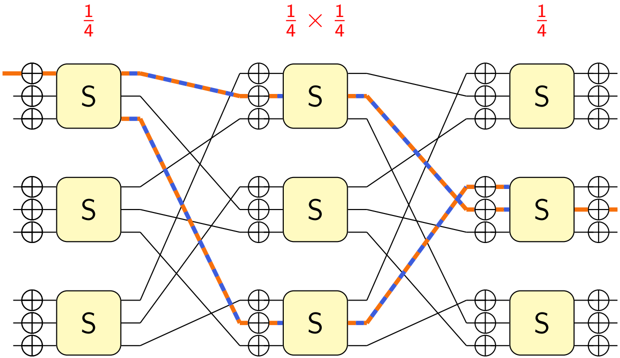
Example: 3-round differential (characteristic 2)



$$\frac{1}{2^8}$$

Quasidifferential trails

Example: 3-round differential (characteristic 2)

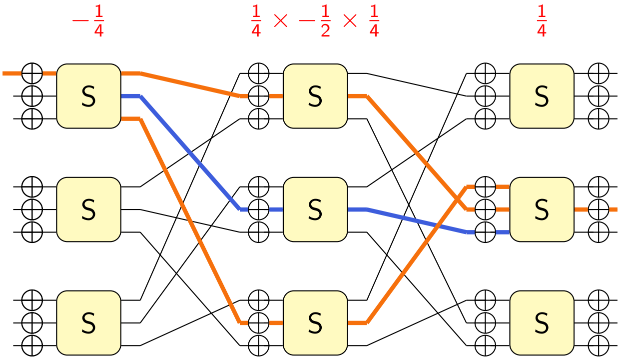


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^8}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 2)

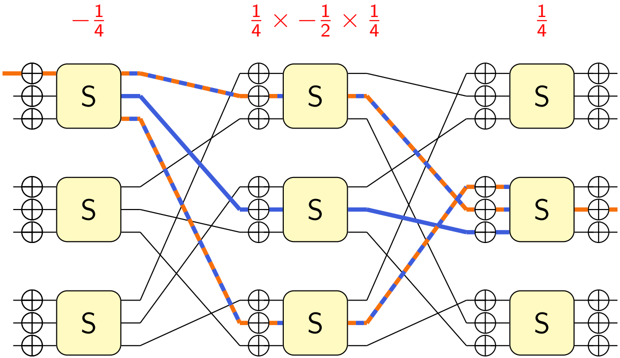


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_2}\frac{1}{2^9}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 2)

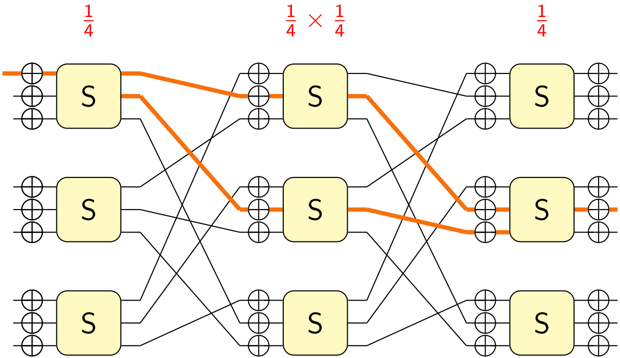


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_2}\frac{1}{2^9} + (-1)^{\kappa_1+\kappa_2+\kappa_3}\frac{1}{2^9}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

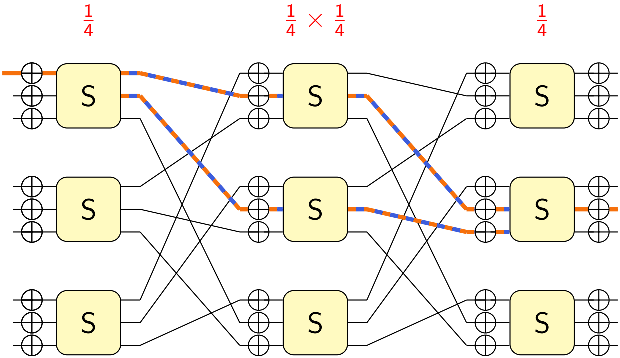
Example: 3-round differential (characteristic 3)



$$\frac{1}{2^8}$$

Quasidifferential trails

Example: 3-round differential (characteristic 3)

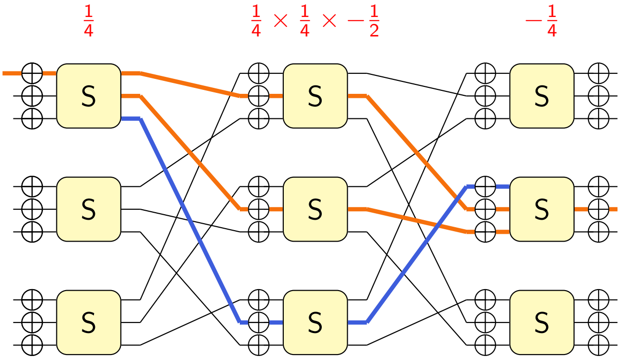


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^8}$$

with $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 3)

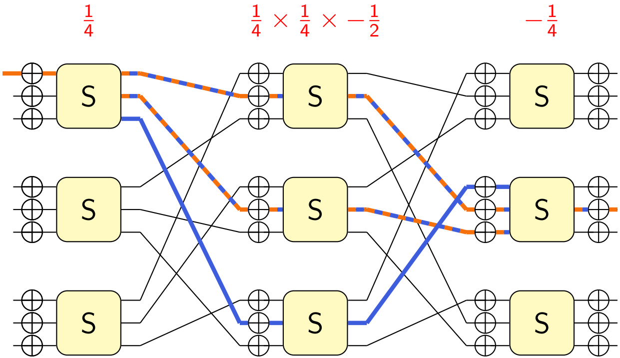


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_1}\frac{1}{2^9}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 3)

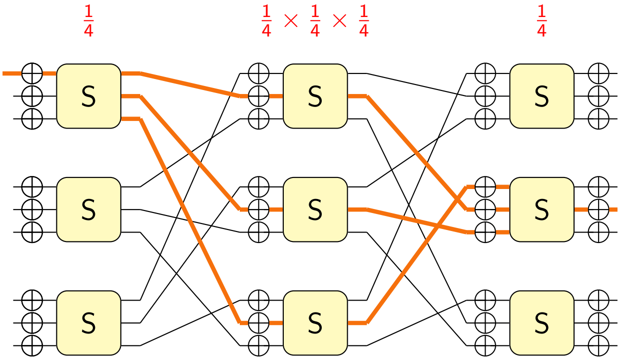


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_1}\frac{1}{2^9} + (-1)^{\kappa_1+\kappa_2+\kappa_3}\frac{1}{2^9}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

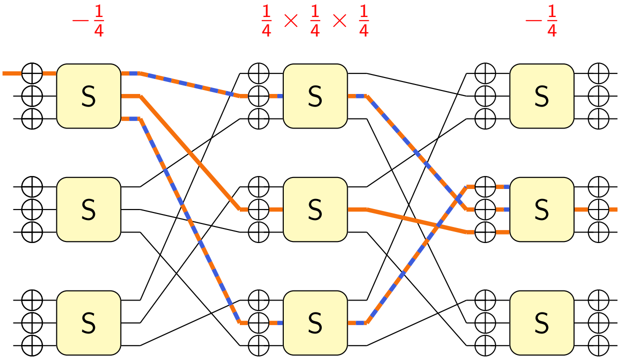
Example: 3-round differential (characteristic 4)



$$\frac{1}{2^{10}}$$

Quasidifferential trails

Example: 3-round differential (characteristic 4)

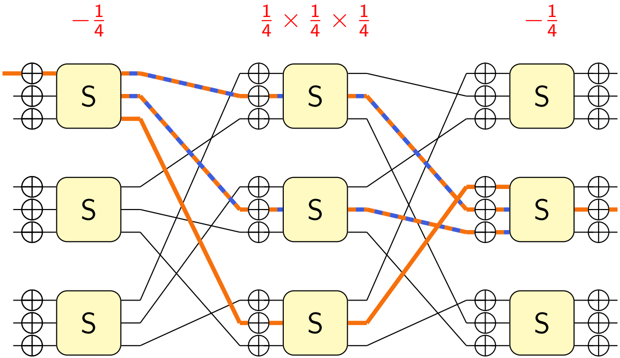


$$\frac{1}{2^{10}} + (-1)^{\kappa_1 + \kappa_3} \frac{1}{2^{10}}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 4)

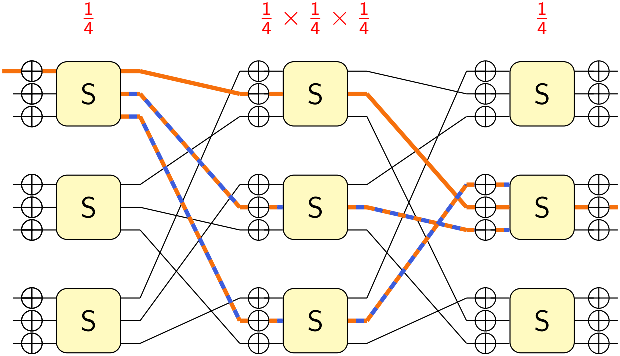


$$\frac{1}{2^{10}} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^{10}} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^{10}}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential (characteristic 4)



$$\frac{1}{2^{10}} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^{10}} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^{10}} + (-1)^{\kappa_1+\kappa_2}\frac{1}{2^{10}}$$

with $\kappa_1 = k_{2,8} + k_{3,4}$, $\kappa_2 = k_{2,5} + k_{3,6}$, $\kappa_3 = k_{2,2} + k_{3,5}$

Quasidifferential trails

Example: 3-round differential

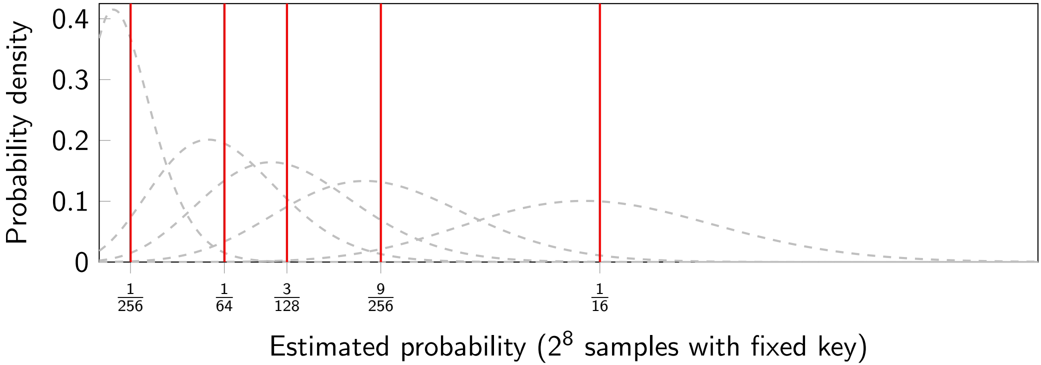
- Overall probability depends on three key bits

$$
\begin{aligned}
&\frac{1}{2^6}\ (1+(-1)^{\kappa_1}/2)(1+(-1)^{\kappa_2}/2) \\
+&\frac{1}{2^8}\ (1+(-1)^{\kappa_1+\kappa_3})(1+(-1)^{\kappa_1}/2) \\
+&\frac{1}{2^8}\ (1+(-1)^{\kappa_2+\kappa_3})(1+(-1)^{\kappa_2}/2) \\
+&\frac{1}{2^{10}}(1+(-1)^{\kappa_1+\kappa_3})(1+(-1)^{\kappa_2+\kappa_3}) \\
\in&\left\{\frac{1}{256},\frac{1}{64},\frac{3}{128},\frac{9}{256},\frac{1}{16}\right\}
\end{aligned}
$$

⚠ Characteristics with ≥ 4 active S-boxes can contribute significantly

Quasidifferential trails

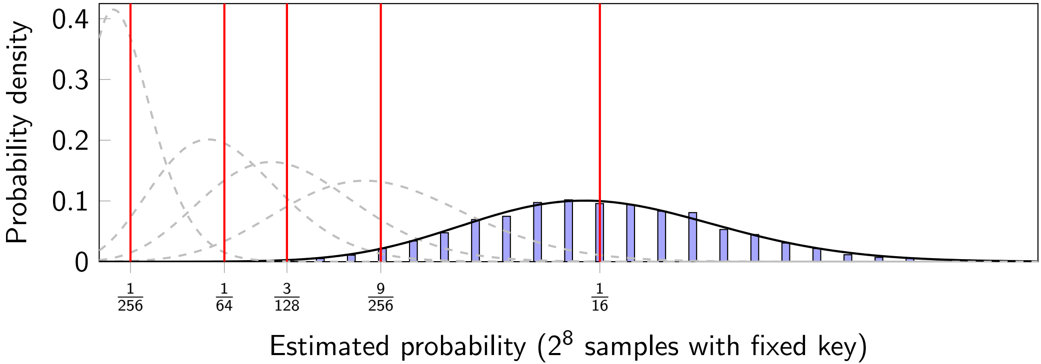
Example: 3-round differential



- Probability reveals something about the key (but we will see better methods later)

Quasidifferential trails

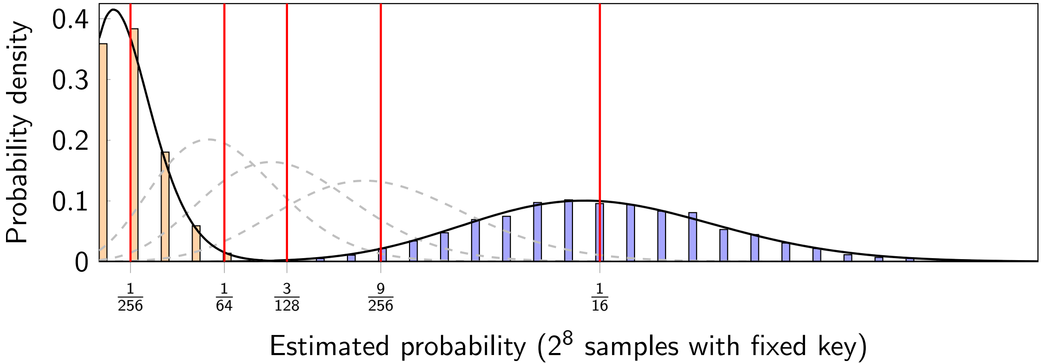
Example: 3-round differential



- Probability reveals something about the key (but we will see better methods later)

Quasidifferential trails

Example: 3-round differential



- Probability reveals something about the key (but we will see better methods later)

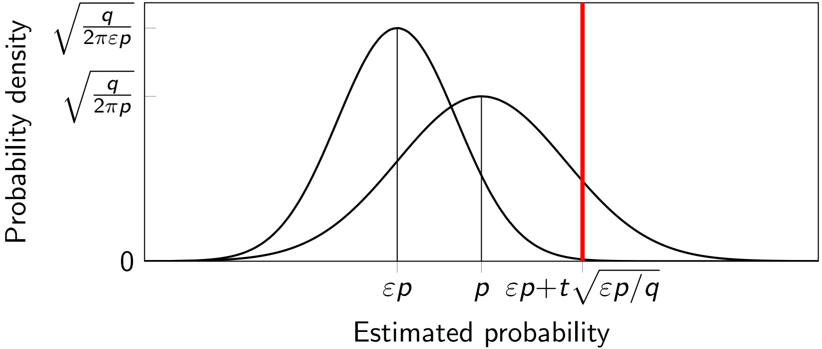
Cost analysis

- Using q independent samples $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q$ (so $2q$ chosen plaintexts):

$$\widehat{\boldsymbol{p}} = \frac{1}{q}\#\{1 \leq i \leq q \mid \mathsf{F}(\boldsymbol{x}_i + a) = \mathsf{F}(\boldsymbol{x}_i) + b\}$$

- Simplifications:
  - pq is not too small and probability p is not too large
  - Probability is εp for wrong keys
- Distribution of $\widehat{\boldsymbol{p}}$ is close to normal with mean p and variance $p(1-p)/q \approx p/q$
- Hypothesis test: $\widehat{\boldsymbol{p}} \geq \varepsilon p + t\sqrt{\varepsilon p/q}$

Cost analysis



- True-positive probability: $P_S = \Phi((1-\varepsilon)\sqrt{pq} - t\sqrt{\varepsilon})$
- False-positive probability: $P_F = \Phi(-t)$

Cost analysis

- Eliminating t gives

$$P_\mathsf{S} = \Phi\big(\Phi^{-1}(P_\mathsf{F})\sqrt{\varepsilon} + (1-\varepsilon)\sqrt{pq}\big)$$

- Inverting this gives

$$q = \frac{1}{p}\left(\frac{\Phi^{-1}\big(P_\mathsf{S}\big) - \Phi^{-1}(P_\mathsf{F})\sqrt{\varepsilon}}{1-\varepsilon}\right)^2$$

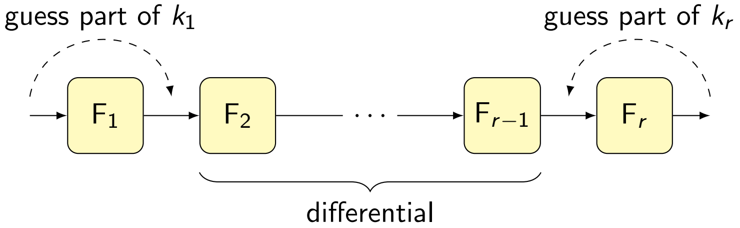
- If p depends on the key, need to average the formulas above

- $1/\varepsilon$ is sometimes called the 'signal-to-noise ratio'

This is essentially optimal *but important assumptions are made*

Key recovery

- If one characteristic is dominant:
  - (a) Differential probability depends on the key
  - (b) Part of the key can be deduced from the output difference
- Guessing key material from the first or last round is often more powerful



Key recovery

- Basic procedure
  - Count the number of right pairs per candidate key
  - Filter out invalid candidate keys using the hypothesis test
- For K candidate keys, $P_\mathsf{F} K$ incorrect candidates remain
- Optimizations of the counting phase

Further topics

- Optimization of differential characteristics and quasidifferential trails
- Key-recovery techniques
- Multiple differentials
- Impossible differentials
- Truncated differentials
- Hash function cryptanalysis
- Geometric approach