Linear and differential cryptanalysis

Tim Beyne

tim@cryptanalysis.info

KU Leuven

August 15, 2023

Symmetric-key cryptography

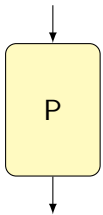
- Symmetric-key cryptography provides
  - Encryption
  - Authentication
  - Cryptographic hashing
  - ...
- Obtained by (carefully) combining building blocks ('primitives")

Symmetric-key cryptography

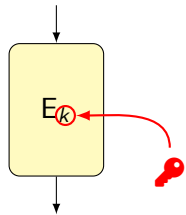
- Symmetric-key cryptography provides
  - Encryption
  - Authentication
  - Cryptographic hashing
  - ...
- Obtained by (carefully) combining building blocks ('primitives")
- Symmetric-key primitives are not based on reductions to other problems
- Cryptanalysis is necessary to understand their design and security

Primitives

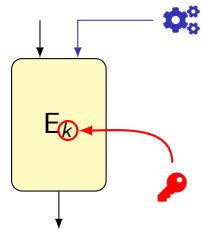
Permutations



Block ciphers

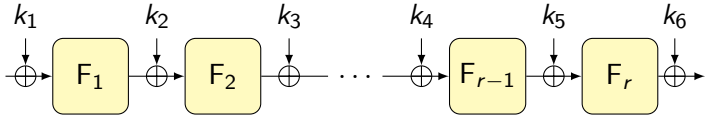


Tweakable block ciphers



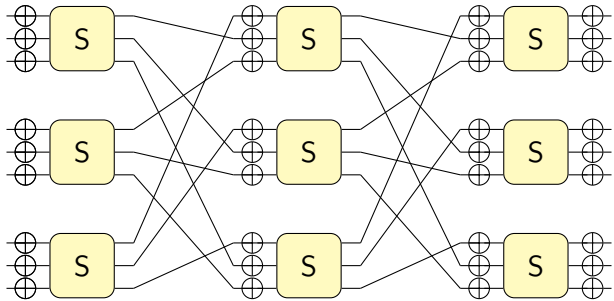
Primitives

- Iterated constructions
- Multiple *rounds* $\mathsf{F}_1, \ldots, \mathsf{F}_r$
- Many block ciphers are *key-alternating*



Primitives

Example



Cryptanalysis

- Often starts with an usual combinatorial property of (part of) the primitive
- Property should be useful to attack applications of the primitive, this depends on
  - Access model
  - Cost of evaluation (queries, time, memory, failure probability, false-positive rate)
- Most important examples:
  - Linear cryptanalysis
  - Differential cryptanalysis
  - Integral cryptanalysis

Overview

- Linear cryptanalysis
  - Correlation matrices and linear trails
  - Cost analysis
  - Key-recovery techniques
- Differential cryptanalysis
  - Quasidifferential transition matrices and trails
  - Cost analysis
  - Key-recovery techniques
- We will follow a semi-classical approach

Linear cryptanalysis

Linear approximations

- Function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, e.g. a cryptographic primitive
- Probabilistic linear relation between x and $y = F(x)$ (Tardy-Corfdir & Gilbert)

$$\underbrace{\sum_{i=1}^{m} v_i\, y_i}_{v^T y} \approx \underbrace{\sum_{i=1}^{n} u_i\, x_i}_{u^T x}$$

- Short notation $v^T y \approx u^T x$
- Pair (u, v) of masks $u \in \mathbb{F}_2^n$ and $v \in \mathbb{F}_2^m$ determines the linear approximation

Linear approximations

Correlation

- If x and $\mathsf{F}(x)$ are 'unrelated', the number of x such that $v^{\mathsf{T}}\mathsf{F}(x) = u^{\mathsf{T}}x$ is $2^n/2$
- Correlation

$$
\begin{aligned}
c &= 2 \times \left(\frac{\#\{x \in \mathbb{F}_2^n \mid v^{\mathsf{T}}\mathsf{F}(x) = u^{\mathsf{T}}x\}}{2^n} - \frac{1}{2} \right) \\
&= 2 \Pr_{\boldsymbol{x}} \left[v^{\mathsf{T}}\mathsf{F}(\boldsymbol{x}) = u^{\mathsf{T}}\boldsymbol{x} \right] - 1
\end{aligned}
$$

Linear approximations

Correlation

- Technical result: if $\boldsymbol{r}$ is a random variable on $\mathbb{F}_2$, then

$$2 \Pr_{\boldsymbol{r}}[\boldsymbol{r} = 0] - 1 = \Pr_{\boldsymbol{r}}[\boldsymbol{r} = 0] - \Pr_{\boldsymbol{r}}[\boldsymbol{r} = 1] = \mathrm{E}_{\boldsymbol{r}}\left[(-1)^{\boldsymbol{r}}\right]$$

- Applied to $\boldsymbol{r} = v^{\mathsf{T}}\mathrm{F}(\boldsymbol{x}) + u^{\mathsf{T}}\boldsymbol{x}$, this gives

$$2 \Pr_{\boldsymbol{x}}\left[v^{\mathsf{T}}\mathrm{F}(\boldsymbol{x}) = u^{\mathsf{T}}\boldsymbol{x}\right] - 1 = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{v^{\mathsf{T}}\mathrm{F}(x) + u^{\mathsf{T}}x}$$

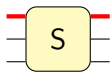
Linear approximations

Example

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

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

- Linear approximation $(u, v) = (001, 001)$



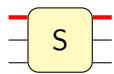
Linear approximations

Example

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

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

- Linear approximation $(u, v) = (001, 001)$



- Correlation $2\Pr_{\boldsymbol{x}}\left[v^{\mathsf{T}}S(\boldsymbol{x}) = u^{\mathsf{T}}\boldsymbol{x}\right] - 1 = 2 \cdot \frac{2}{8} - 1 = -\frac{1}{2}$

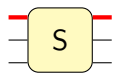
Linear approximations

Example

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

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

- Linear approximation $(u, v) = (001, 001)$

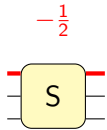


- Correlation $2\Pr_{\boldsymbol{x}}\left[v^{\mathsf{T}}S(\boldsymbol{x}) = u^{\mathsf{T}}\boldsymbol{x}\right] - 1 = 2 \cdot \frac{2}{8} - 1 = -\frac{1}{2}$
$= (-1 - 1 + 1 + 1 - 1 - 1 - 1 - 1)/8$

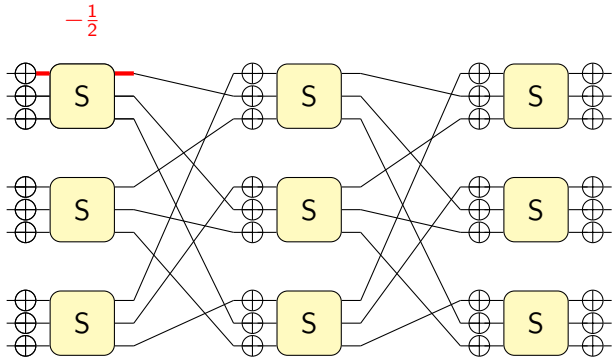
Linear approximations
Distinguishers

- Statistical attack: sample q inputs at random and estimate correlation
- Estimation error will be about $1/\sqrt{q}$
- $q \approx 1/c^2$ samples are enough for a distinguisher
  (assuming c is not too small/large)

Linear approximations



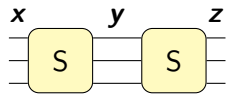
Linear approximations



Propagation through a sequence of operations?

Linear approximations

Piling up approximations



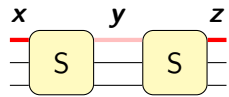
$$\frac{\begin{aligned} r_1 &= u^\mathsf{T} x + w^\mathsf{T} y \\ r_2 &= w^\mathsf{T} y + v^\mathsf{T} z \end{aligned}}{r_1 + r_2 = u^\mathsf{T} x + v^\mathsf{T} z}$$

Pretend that r_1 and r_2 are independent:

$$\underbrace{\mathrm{E}\left[(-1)^{r_1 + r_2}\right]}_{2\Pr[v^\mathsf{T} z = u^\mathsf{T} x] - 1} \stackrel{☠}{=} \underbrace{\mathrm{E}\left[(-1)^{r_1}\right]}_{2\Pr[w^\mathsf{T} y = u^\mathsf{T} x] - 1} \times \underbrace{\mathrm{E}\left[(-1)^{r_2}\right]}_{2\Pr[w^\mathsf{T} y = v^\mathsf{T} z] - 1}$$

Linear approximations

Piling up approximations



$$\frac{\begin{aligned} r_1 &= u^\mathsf{T} x + w^\mathsf{T} y \\ r_2 &= w^\mathsf{T} y + v^\mathsf{T} z \end{aligned}}{r_1 + r_2 = u^\mathsf{T} x + v^\mathsf{T} z}$$

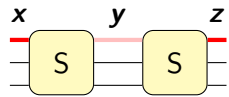
Pretend that r_1 and r_2 are independent:

$$\underbrace{\mathrm{E}\left[(-1)^{r_1 + r_2}\right]}_{2\Pr[v^\mathsf{T} z = u^\mathsf{T} x] - 1} \overset{☠}{=} \underbrace{\mathrm{E}\left[(-1)^{r_1}\right]}_{2\Pr[w^\mathsf{T} y = u^\mathsf{T} x] - 1} \times \underbrace{\mathrm{E}\left[(-1)^{r_2}\right]}_{2\Pr[w^\mathsf{T} y = v^\mathsf{T} z] - 1}$$

- For example: $u = w = v = 001$ gives $-1/2 \times -1/2 = 1/4$

Linear approximations

Piling up approximations



$$\frac{\begin{aligned} r_1 &= u^\mathsf{T} x + w^\mathsf{T} y \\ r_2 &= w^\mathsf{T} y + v^\mathsf{T} z \end{aligned}}{r_1 + r_2 = u^\mathsf{T} x + v^\mathsf{T} z}$$

Pretend that r_1 and r_2 are independent:

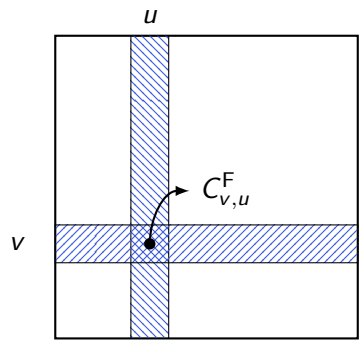
$$\underbrace{\mathrm{E}\left[(-1)^{r_1+r_2}\right]}_{2\Pr[v^\mathsf{T} z = u^\mathsf{T} x]-1} \overset{☠}{=} \underbrace{\mathrm{E}\left[(-1)^{r_1}\right]}_{2\Pr[w^\mathsf{T} y = u^\mathsf{T} x]-1} \times \underbrace{\mathrm{E}\left[(-1)^{r_2}\right]}_{2\Pr[w^\mathsf{T} y = v^\mathsf{T} z]-1}$$

- For example: $u = w = v = 001$ gives $-1/2 \times -1/2 = 1/4$
- Unfortunately, this is wrong (the correct result is zero)

Correlation matrices

- $2^m \times 2^n$ matrix containing correlations of linear approximations of $\mathrm{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$

$$C^{\mathrm{F}}_{v,u} = 2 \Pr_{\boldsymbol{x}} \left[v^{\mathsf{T}} \mathrm{F}(\boldsymbol{x}) = u^{\mathsf{T}} \boldsymbol{x} \right] - 1$$



ℹ 'Matrix' rather than table because C^{F} really does represent a linear map

Correlation matrices

Example

$$C^{\mathsf{S}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

Correlation matrices

Example

$$C^{\mathsf{S}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

$$C^{\mathsf{F}}_{0,u} = 2\Pr[u^{\mathsf{T}}\boldsymbol{x} \quad = 0] - 1 = \begin{cases} 1 & \text{if } u = 0 \\ 0 & \text{else} \end{cases}$$

$$C^{\mathsf{F}}_{v,0} = 2\Pr[v^{\mathsf{T}}\mathsf{F}(\boldsymbol{x}) = 0] - 1 = \begin{cases} 1 & \text{if } v = 0 \\ 0 & \text{else} \end{cases}$$

(second property if F is invertible)

Correlation matrices

Multiplication property

- If $\mathsf{F} = \mathsf{F}_2 \circ \mathsf{F}_1$, then $C^{\mathsf{F}} = C^{\mathsf{F}_2} C^{\mathsf{F}_1}$
- Proof:

$$
\begin{aligned}
\left(C^{\mathsf{F}_2} C^{\mathsf{F}_1}\right)_{v,u} &= \sum_w C^{\mathsf{F}_2}_{v,w} C^{\mathsf{F}_1}_{w,u} \\
&= \sum_w \left(\frac{1}{2^m} \sum_y (-1)^{v^{\mathsf{T}} \mathsf{F}_2(y) + w^{\mathsf{T}} y} \right) \left(\frac{1}{2^n} \sum_x (-1)^{w^{\mathsf{T}} \mathsf{F}_1(x) + u^{\mathsf{T}} x} \right) \\
&= \frac{1}{2^n} \sum_{x,y} (-1)^{v^{\mathsf{T}} \mathsf{F}_2(y) + u^{\mathsf{T}} x} \frac{1}{2^m} \sum_w (-1)^{w^{\mathsf{T}} y + w^{\mathsf{T}} \mathsf{F}_1(x)} \\
&= \frac{1}{2^n} \sum_x (-1)^{v^{\mathsf{T}} \mathsf{F}_2(\mathsf{F}_1(x)) + u^{\mathsf{T}} x}
\end{aligned}
$$

- There is an easier proof without calculation

Correlation matrices
Multiplication property

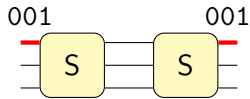
- If F is invertible, then $C^{\mathsf{F}^{-1}} = \left(C^{\mathsf{F}}\right)^{-1}$
- If F is invertible, then C^{F} is orthogonal
  Proof: $C^{\mathsf{F}^{-1}} = \left(C^{\mathsf{F}}\right)^{\mathsf{T}}$ because

$$C^{\mathsf{F}^{-1}}_{v,u} = 2 \Pr_{\boldsymbol{y}}\left[v^{\mathsf{T}}\mathsf{F}^{-1}(\boldsymbol{y}) = u^{\mathsf{T}}\boldsymbol{y}\right] - 1 = 2 \Pr_{\boldsymbol{x}}\left[v^{\mathsf{T}}\boldsymbol{x} = u^{\mathsf{T}}\mathsf{F}(\boldsymbol{x})\right] - 1 = C^{\mathsf{F}}_{u,v}$$

- $\boldsymbol{x} = \mathsf{F}^{-1}(\boldsymbol{y})$ is still uniform random because F is invertible

Correlation matrices

Multiplication property: example



Correlation matrices

Multiplication property: example

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

- Correlation of $(001, 001)$ over $S \circ S$:

$$\frac{1}{4} - \frac{1}{4} - \frac{1}{4} + \frac{1}{4} = 0$$

- Correct result, but doesn't scale

Linear trails

- If $\mathsf{F} = \mathsf{F}_r \circ \cdots \circ \mathsf{F}_2 \circ \mathsf{F}_1$, then $C^{\mathsf{F}} = C^{\mathsf{F}_r} \cdots C^{\mathsf{F}_2} C^{\mathsf{F}_1}$
- Writing out this product of matrices gives

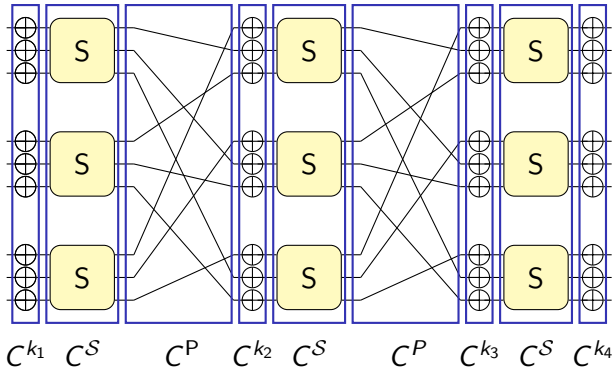
$$C^{\mathsf{F}}_{u_{r+1},u_1} = \sum_{u_2,\ldots,u_r} C^{\mathsf{F}_r}_{u_{r+1},u_r} \cdots C^{\mathsf{F}_2}_{u_3,u_2} C^{\mathsf{F}_1}_{u_2,u_1}$$

- A linear trail is a sequence $(u_1, u_2, \ldots, u_{r+1})$ and has correlation $\prod_{i=1}^{r} C^{\mathsf{F}_i}_{u_{i+1},u_i}$
- Most analysis relies on the assumption that there exist a set Λ of 'dominant trails':

$$C^{\mathsf{F}}_{u_{r+1},u_1} = \sum_{u \in \Lambda} \prod_{i=1}^{r} C^{\mathsf{F}_i}_{u_{i+1},u_i} + \varepsilon$$

Linear trails

Example



- To analyze trails we need to determine C^{k_1}, $C^{\mathcal{S}}$ and C^{P}

Correlation matrices

Bricklayer functions

- If $\mathsf{F}(x_1 \| x_2) = \mathsf{F}_1(x_1) \| \mathsf{F}_2(x_2)$, then $C^{\mathsf{F}}_{v_1\|v_2, u_1\|u_2} = C^{\mathsf{F}_1}_{v_1,u_1} C^{\mathsf{F}_2}_{v_2,u_2}$
- Proof: if $u = u_1 \| u_2$ and $v = v_1 \| v_2$, then

$$\begin{aligned} C^{\mathsf{F}}_{v,u} &= \mathop{\mathrm{E}}_{\boldsymbol{x}}\left[(-1)^{v^{\mathsf{T}}\mathsf{F}(\boldsymbol{x})+u^{\mathsf{T}}\boldsymbol{x}}\right] \\ &= \mathop{\mathrm{E}}_{\boldsymbol{x}_1,\boldsymbol{x}_2}\left[(-1)^{v_1^{\mathsf{T}}\mathsf{F}_1(\boldsymbol{x}_1)+v_2^{\mathsf{T}}\mathsf{F}_2(\boldsymbol{x}_2)+u_1^{\mathsf{T}}\boldsymbol{x}_1+u_2^{\mathsf{T}}\boldsymbol{x}_2}\right] \\ &= \mathop{\mathrm{E}}_{\boldsymbol{x}_1}\left[(-1)^{v_1^{\mathsf{T}}\mathsf{F}_1(\boldsymbol{x}_1)+u_1^{\mathsf{T}}\boldsymbol{x}_1}\right] \mathop{\mathrm{E}}_{\boldsymbol{x}_2}\left[(-1)^{v_2^{\mathsf{T}}\mathsf{F}_2(\boldsymbol{x}_2)+u_2^{\mathsf{T}}\boldsymbol{x}_2}\right] \\ &= C^{\mathsf{F}_1}_{v_1,u_1} C^{\mathsf{F}_2}_{v_2,u_2} \end{aligned}$$

Correlation matrices

Bricklayer functions

- If $\mathsf{F}(x_1 \| x_2) = \mathsf{F}_1(x_1) \| \mathsf{F}_2(x_2)$, then $C^{\mathsf{F}}_{v_1 \| v_2, u_1 \| u_2} = C^{\mathsf{F}_1}_{v_1, u_1} C^{\mathsf{F}_2}_{v_2, u_2}$
- Proof: if $u = u_1 \| u_2$ and $v = v_1 \| v_2$, then

$$\begin{aligned} C^{\mathsf{F}}_{v,u} &= \mathop{\mathrm{E}}_{\boldsymbol{x}} \left[(-1)^{v^{\mathsf{T}} \mathsf{F}(\boldsymbol{x}) + u^{\mathsf{T}} \boldsymbol{x}} \right] \\ &= \mathop{\mathrm{E}}_{\boldsymbol{x}_1, \boldsymbol{x}_2} \left[(-1)^{v_1^{\mathsf{T}} \mathsf{F}_1(\boldsymbol{x}_1) + v_2^{\mathsf{T}} \mathsf{F}_2(\boldsymbol{x}_2) + u_1^{\mathsf{T}} \boldsymbol{x}_1 + u_2^{\mathsf{T}} \boldsymbol{x}_2} \right] \\ &= \mathop{\mathrm{E}}_{\boldsymbol{x}_1} \left[(-1)^{v_1^{\mathsf{T}} \mathsf{F}_1(\boldsymbol{x}_1) + u_1^{\mathsf{T}} \boldsymbol{x}_1} \right] \mathop{\mathrm{E}}_{\boldsymbol{x}_2} \left[(-1)^{v_2^{\mathsf{T}} \mathsf{F}_2(\boldsymbol{x}_2) + u_2^{\mathsf{T}} \boldsymbol{x}_2} \right] \\ &= C^{\mathsf{F}_1}_{v_1, u_1} C^{\mathsf{F}_2}_{v_2, u_2} \end{aligned}$$

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

Correlation matrices

Translations and linear functions

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

$$C^{\mathsf{F}}_{v,u} = \begin{cases} (-1)^{v^{\mathsf{T}}k} & \text{if } u = v \\ 0 & \text{else}. \end{cases}$$

Proof: $C^{\mathsf{F}}_{v,u} = \mathsf{E}_{\boldsymbol{x}}\left[(-1)^{v^{\mathsf{T}}(\boldsymbol{x}+k)+u^{\mathsf{T}}\boldsymbol{x}}\right] = (-1)^{v^{\mathsf{T}}k}\,\mathsf{E}\left[(-1)^{(u+v)^{\mathsf{T}}\boldsymbol{x}}\right]$

Correlation matrices

Translations and linear functions

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

$$C^{\mathsf{F}}_{v,u} = \begin{cases} (-1)^{v^{\mathsf{T}}k} & \text{if } u = v \\ 0 & \text{else}. \end{cases}$$

  Proof: $C^{\mathsf{F}}_{v,u} = \mathsf{E}_{\boldsymbol{x}}\left[(-1)^{v^{\mathsf{T}}(\boldsymbol{x}+k)+u^{\mathsf{T}}\boldsymbol{x}}\right] = (-1)^{v^{\mathsf{T}}k}\,\mathsf{E}\left[(-1)^{(u+v)^{\mathsf{T}}\boldsymbol{x}}\right]$

- If $\mathsf{F}(x) = Mx$ with $M \in \mathbb{F}_2^{m\times n}$ then

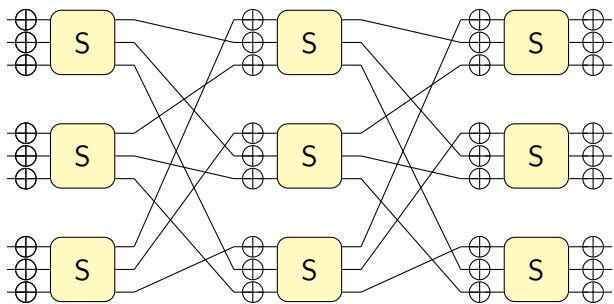
$$C^{\mathsf{F}}_{v,u} = \begin{cases} 1 & \text{if } u = M^{\mathsf{T}}v \\ 0 & \text{else}. \end{cases}$$

  Proof: $C^{\mathsf{F}}_{v,u} = \mathsf{E}_{\boldsymbol{x}}\left[(-1)^{v^{\mathsf{T}}M\boldsymbol{x}+u^{\mathsf{T}}\boldsymbol{x}}\right] = \mathsf{E}\left[(-1)^{(u+M^{\mathsf{T}}v)^{\mathsf{T}}\boldsymbol{x}}\right]$

- Bit permutation P satisfies $\mathsf{P}^{\mathsf{T}} = \mathsf{P}^{-1}$

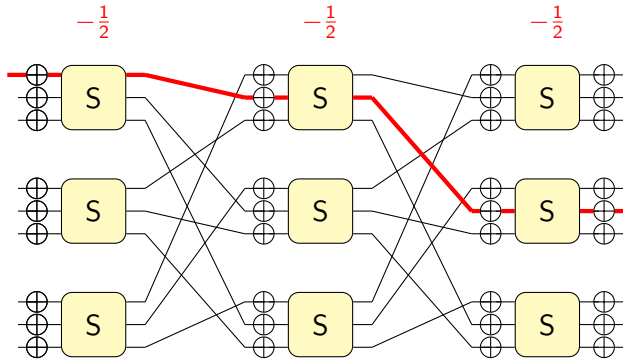
Linear trails

Example: 3-round approximation



Linear trails

Example: 3-round approximation

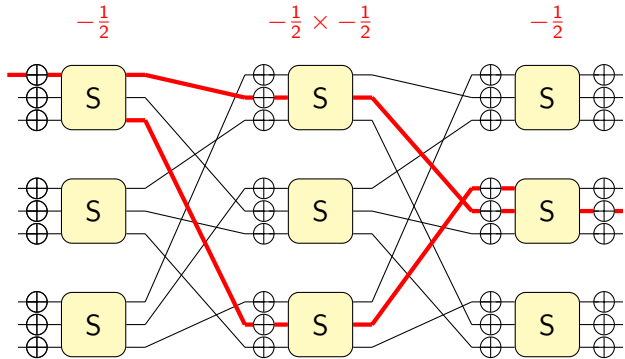


$$(-1)^{\kappa_1}/8$$

with $\kappa_1 = k_{1,1} + k_{2,2} + k_{2,5} + 1$

Linear trails

Example: 3-round approximation

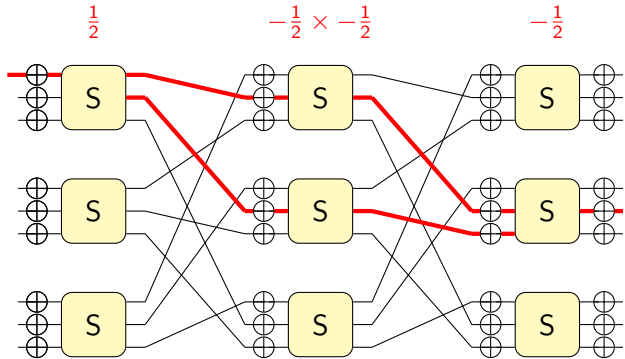


$$(-1)^{\kappa_1}/8 + (-1)^{\kappa_1+\kappa_2}/16$$

with $\kappa_1 = k_{1,1} + k_{2,2} + k_{2,5} + 1$, $\kappa_2 = k_{2,8} + k_{3,4}$

Linear trails

Example: 3-round approximation

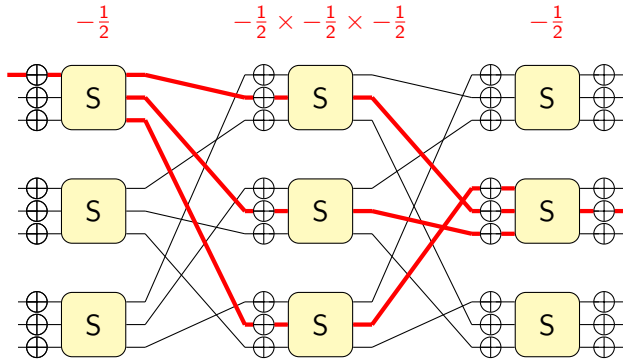


$$(-1)^{\kappa_1}/8 + (-1)^{\kappa_1+\kappa_2}/16 + (-1)^{\kappa_1+\kappa_3}/16$$

with $\kappa_1 = k_{1,1} + k_{2,2} + k_{2,5} + 1$, $\kappa_2 = k_{2,8} + k_{3,4}$ and $\kappa_3 = k_{2,5} + k_{3,6}$

Linear trails

Example: 3-round approximation

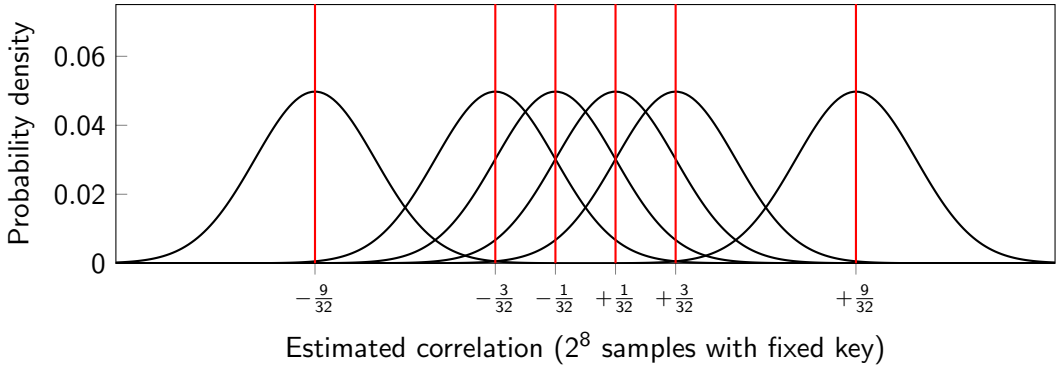


$$(-1)^{\kappa_1}/8 + (-1)^{\kappa_1+\kappa_2}/16 + (-1)^{\kappa_1+\kappa_3}/16 + (-1)^{\kappa_1+\kappa_2+\kappa_3}/32$$

with $\kappa_1 = k_{1,1} + k_{2,2} + k_{2,5} + 1$, $\kappa_2 = k_{2,8} + k_{3,4}$ and $\kappa_3 = k_{2,5} + k_{3,6}$

Linear trails

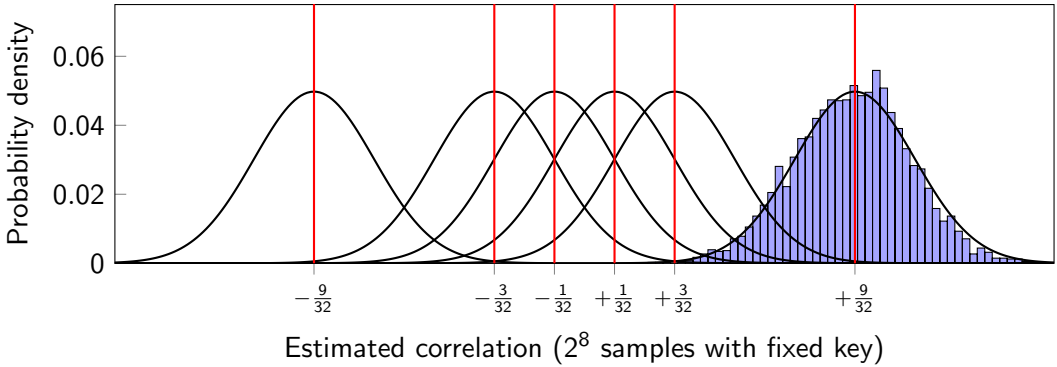
Example: 3-round approximation



- $C^{\mathsf{F}}_{001,001} = (-1)^{\kappa_1}/8\,\left(1+(-1)^{\kappa_1+\kappa_2}/2\right)\left(1+(-1)^{\kappa_1+\kappa_3}/2\right) \in \left\{\pm\frac{1}{32}, \pm\frac{3}{32}, \pm\frac{9}{32}\right\}$
- Correlation reveals something about the key (but we will see better methods later)

Linear trails

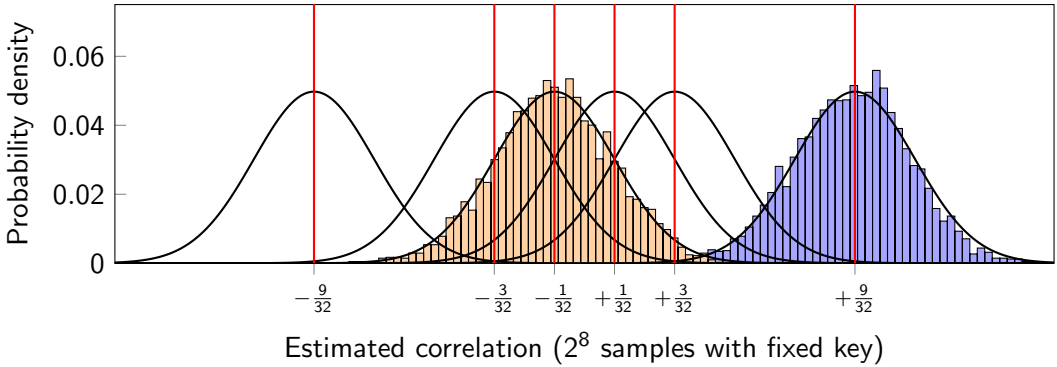
Example: 3-round approximation



- $C^{\mathsf{F}}_{001,001} = (-1)^{\kappa_1}/8\,\left(1+(-1)^{\kappa_1+\kappa_2}/2\right)\left(1+(-1)^{\kappa_1+\kappa_3}/2\right) \in \left\{\pm\frac{1}{32}, \pm\frac{3}{32}, \pm\frac{9}{32}\right\}$
- Correlation reveals something about the key (but we will see better methods later)

Linear trails

Example: 3-round approximation



- $C^{\mathsf{F}}_{001,001} = (-1)^{\kappa_1}/8\,\left(1+(-1)^{\kappa_1+\kappa_2}/2\right)\left(1+(-1)^{\kappa_1+\kappa_3}/2\right) \in \left\{\pm\frac{1}{32}, \pm\frac{3}{32}, \pm\frac{9}{32}\right\}$
- Correlation reveals something about the key (but we will see better methods later)

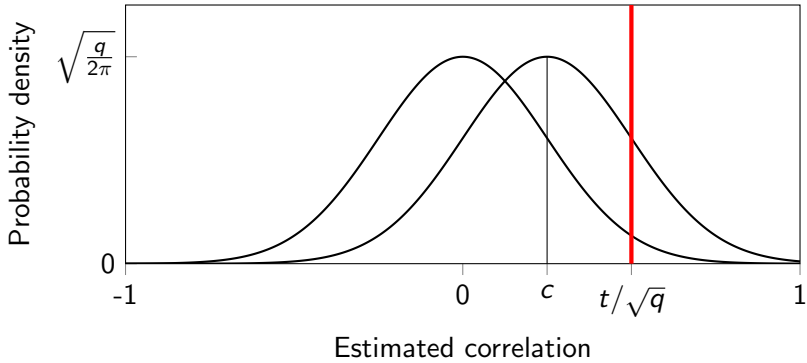
Cost analysis

- Suppose the correlation is c and we use q independent samples:

$$\widehat{\boldsymbol{c}} = \frac{1}{q}\sum_{i=1}^{q}(-1)^{u^{\mathsf{T}}\boldsymbol{x}_i + v^{\mathsf{T}}\boldsymbol{y}_i}$$

- Simplifications:
  - Correlation is zero in the ideal case
  - q is not too small and c is not too large
- Distribution of $\widehat{\boldsymbol{c}}$ is close to normal with mean c and variance $(1 - c^2)/q \approx 1/q$
- Hypothesis test: $|\widehat{\boldsymbol{c}}| \geq t/\sqrt{q}$?

Cost analysis



- True-positive rate:

$$\begin{aligned} P_{\mathsf{S}} &= \Pr\left[|\widehat{\boldsymbol{c}}| \geq t/\sqrt{q}\right] \\ &= \Pr\left[\widehat{\boldsymbol{c}} \geq t/\sqrt{q}\right] + \Pr\left[\widehat{\boldsymbol{c}} \leq -t/\sqrt{q}\right] \\ &= \Phi(c\sqrt{q} - t) + \Phi(-c\sqrt{q} - t) \end{aligned}$$

- False-positive rate: $P_{\mathsf{F}} = 2\Phi(-t)$

Cost analysis

- Eliminating t gives

$$P_S = \Phi\left(\Phi^{-1}(P_F/2) + c\sqrt{q}\right) + \Phi\left(\Phi^{-1}(P_F/2) - c\sqrt{q}\right)$$

- If $|c|\sqrt{q}$ is large enough, one of both terms is dominant so

$$q = \left(\frac{\Phi^{-1}(P_S) - \Phi^{-1}(P_F/2)}{c}\right)^2$$

- One can show that this is essentially optimal

⚠ under important assumptions

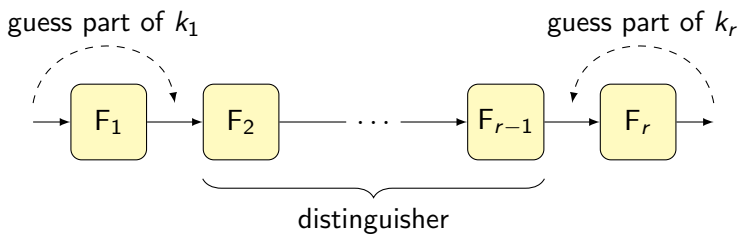
- If c depends on the key, need to average the formulas above

Key recovery

- Correlation depends on the key, and this can be used for key-recovery
  Extreme case with one dominant trail

$$C^{\mathsf{F}}_{v,u} \approx (-1)^{w^{\mathsf{T}}k} c$$

- Guessing key material from the first and/or last round is usually more powerful



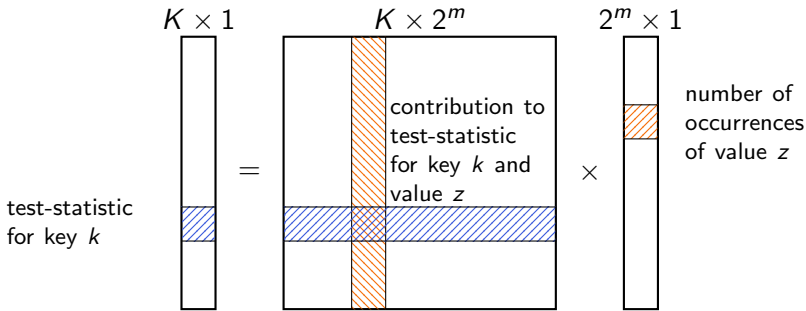
- Naive cost: $\mathcal{O}(qK)$ for K candidate keys if the distinguisher uses q data
  on average $P_{\mathsf{F}}K$ incorrect candidates remain

Key recovery

Matsui's method

- Samples $(x_1, y_1), \ldots, (x_q, y_q) \to$ reduced values $z_1, \ldots, z_q \in \mathbb{F}_2^m$
- For candidate key k, the estimated correlation is of the form

$$\widehat{c}_k = \sum_{i=1}^{q} f_k(z_i) = \sum_{z \in \mathbb{F}_2^m} f_k(z) \, \#\{1 \le i \le q \mid z_i = z\}$$



- Cost: $\mathcal{O}(q + K2^m)$ time and $\mathcal{O}(q + K + 2^m)$ memory

Key recovery

Fast Fourier transformation method

- Matrix $\left[f_k(z)\right]_{k,z}$ often has exploitable structure
- In particular: $f_k(z) = f(z \boxplus k)$, i.e. a circulant matrix (equivalently: convolution)
- Multiplication with a circulant matrix can be done using the FFT:

$$\mathscr{F}(\widehat{c}_k) = \mathscr{F}(f_0) \odot \mathscr{F}(w)$$

  with $w(z) = \#\{1 \leq i \leq q \mid z_i = z\}$ and $\odot$ the coordinate-wise product

- Cost: $\mathcal{O}(q + K \log K)$ time and $\mathcal{O}(q + K)$ memory

Differential cryptanalysis

Differentials

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

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

- 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

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

- 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 $\mathrm{F}(x + a) = \mathrm{F}(x) + b$ is $2^n/2^m$ on average

- Probability of a differential:

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

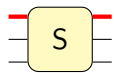
Differentials

Example

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

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

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



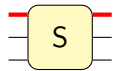
Differentials

Example

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

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

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

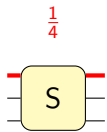


- Probability $\Pr_{\boldsymbol{x}}\left[S(\boldsymbol{x} + a) = S(\boldsymbol{x}) + b\right] = \frac{2}{8} = \frac{1}{4}$

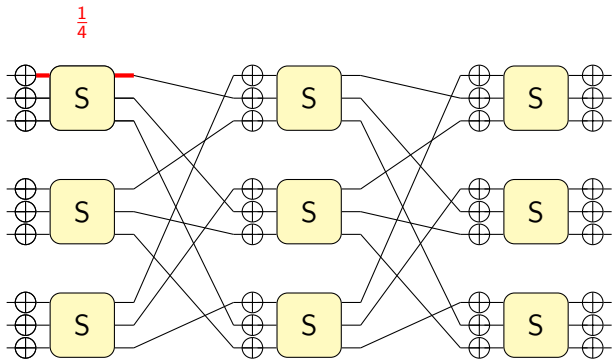
Differentials
Distinguishers

- 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)

Differentials



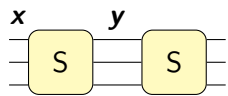
Differentials



Propagation through a sequence of operations?

Differentials

Example



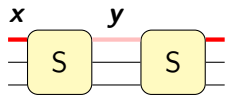
$$\Pr[S^2(x+a) = S^2(x) + b] \geq \Pr[S(x+a) = S(x) + c \text{ and } S(y+c) = S(y) + b]$$

Pretend that x and y are independent:

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

Differentials

Example



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

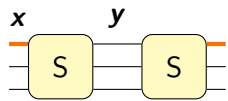
Pretend that x and y are independent:

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

- 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

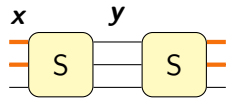
Example



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

Differentials

Example



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

- Unfortunately, this is still wrong in general (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:

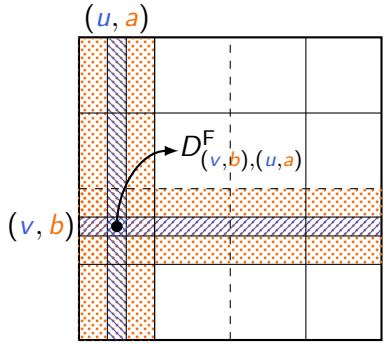
$$\Pr[\mathsf{F}(\boldsymbol{x} + a_1) = \mathsf{F}(\boldsymbol{x}) + a_{r+1}] = \sum_{a_2,\ldots,a_r} \Pr\left[\bigwedge_{i=1}^{r} \mathsf{F}_i(\boldsymbol{x}_i + a_i) = \mathsf{F}(\boldsymbol{x}_i) + a_{i+1}\right]$$

- 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

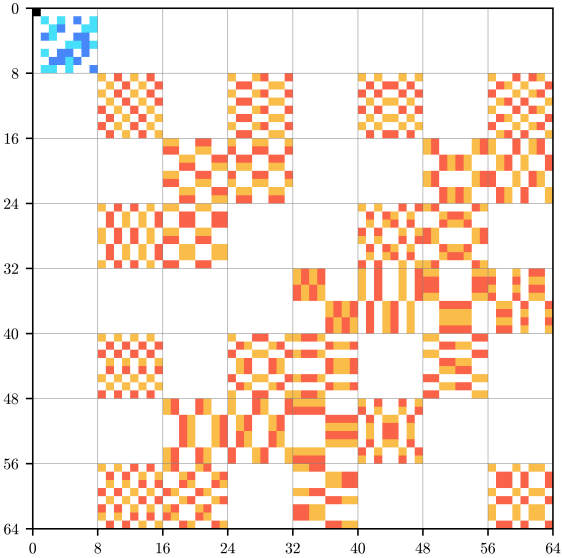
- $2^{2m} \times 2^{2n}$ matrix corresponding to $\mathrm{F} : \mathbb{F}_2^n \to \mathbb{F}_2^m$

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



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:

$$\left(D^{\mathsf{F}_2} D^{\mathsf{F}_1}\right)_{(v,b),(u,a)}$$

$$= \sum_{w,c} \left(\frac{1}{2^m} \sum_{\substack{y \\ \mathsf{F}_2(y+c)=\mathsf{F}_2(y)+b}} (-1)^{v^{\mathsf{T}}\mathsf{F}_2(y)+w^{\mathsf{T}}y} \right) \left(\frac{1}{2^n} \sum_{\substack{x \\ \mathsf{F}_1(x+a)=\mathsf{F}_1(x)+c}} (-1)^{w^{\mathsf{T}}\mathsf{F}_1(x)+u^{\mathsf{T}}x} \right)$$

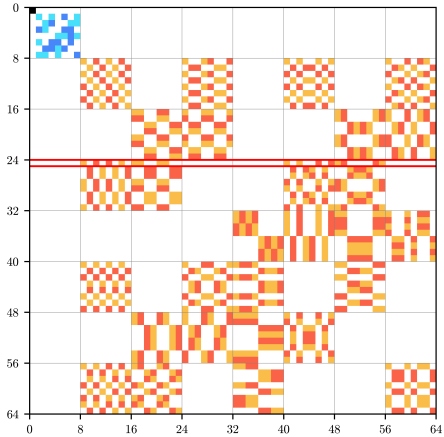
$$= \frac{1}{2^n} \sum_{\substack{x,y \\ \mathsf{F}_2(y+c)=\mathsf{F}_2(y)+b \\ c=\mathsf{F}_1(x+a)-\mathsf{F}_1(x)}} (-1)^{v^{\mathsf{T}}\mathsf{F}_2(y)+u^{\mathsf{T}}x} \frac{1}{2^m} \sum_{w} (-1)^{w^{\mathsf{T}}y+w^{\mathsf{T}}\mathsf{F}_1(x)}$$

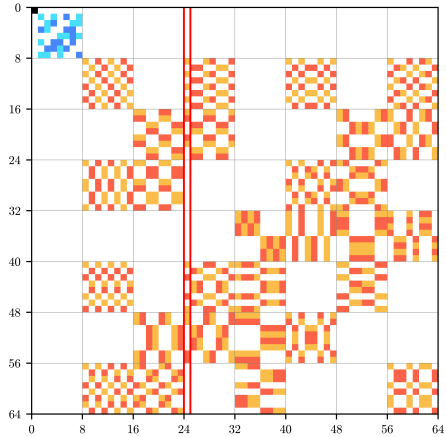
$$= \frac{1}{2^n} \sum_{\substack{x \\ \mathsf{F}_2(\mathsf{F}_1(x+a))=\mathsf{F}_2(\mathsf{F}_1(x))+b}} (-1)^{v^{\mathsf{T}}\mathsf{F}_2(\mathsf{F}_1(x))+u^{\mathsf{T}}x}$$

- There is an easier proof 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 $D^{\mathsf{F}} = D^{\mathsf{F}_r} \cdots D^{\mathsf{F}_2} D^{\mathsf{F}_1}$

- Writing out this product of matrices gives

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

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

- A quasidifferential trail is a sequence $(\varpi_1, \ldots, \varpi_{r+1})$ and has 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

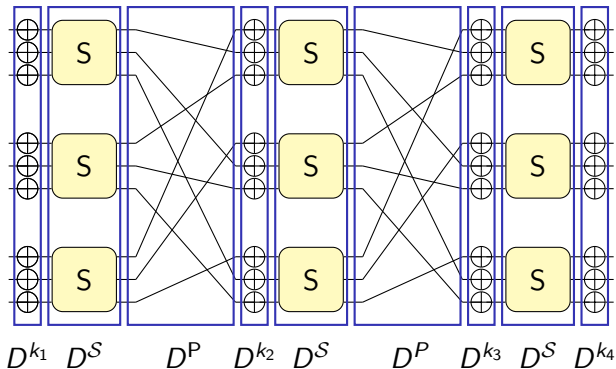
- $D^{\mathsf{F}}_{(0,a_{r+1}),(0,a_1)}$ is the probability of the differential (a_1, a_{r+1})
- Quasidifferential trails can be used to compute the probability of a differential
- Quasidifferential trails can be used to compute the probability of a characteristic:

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

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

Quasidifferential trails

Example



- To analyze trails we need to determine D^{k_1}, $D^{\mathcal{S}}$ 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: exercise.

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

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

Proof: exercise.

Quasidifferential trails

Translations and linear functions

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

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

  Proof: exercise.

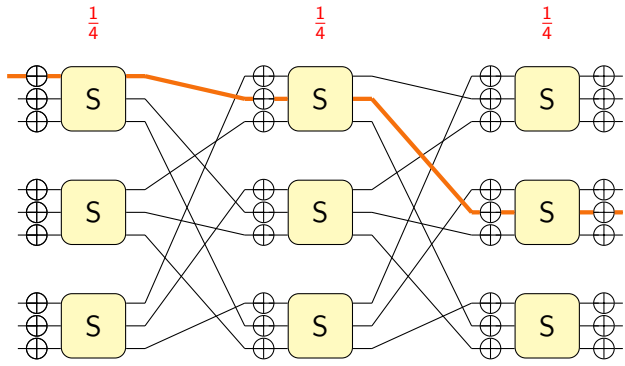
- If $\mathrm{F}(x) = Mx$, then

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

  Proof: exercise.

Quasidifferential trails

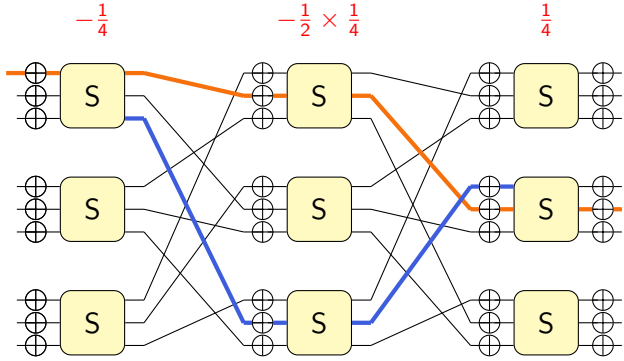
Example: 3-round differential (characteristic 1)



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

Quasidifferential trails

Example: 3-round differential (characteristic 1)

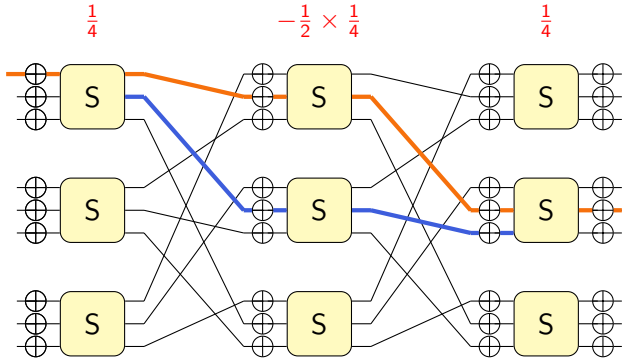


$$\frac{1}{2^6} + (-1)^{\kappa_1} \frac{1}{2^7}$$

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

Quasidifferential trails

Example: 3-round differential (characteristic 1)

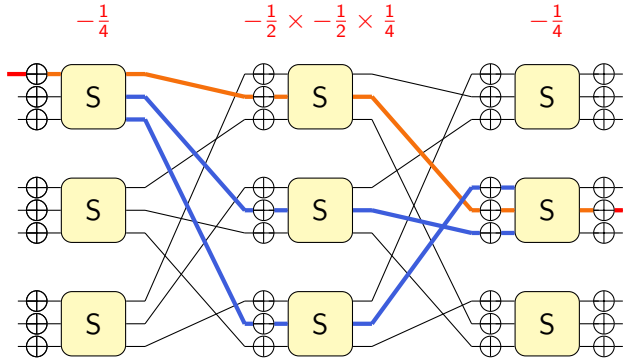


$$\frac{1}{2^6} + (-1)^{\kappa_1} \frac{1}{2^7} + (-1)^{\kappa_2} \frac{1}{2^7}$$

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)

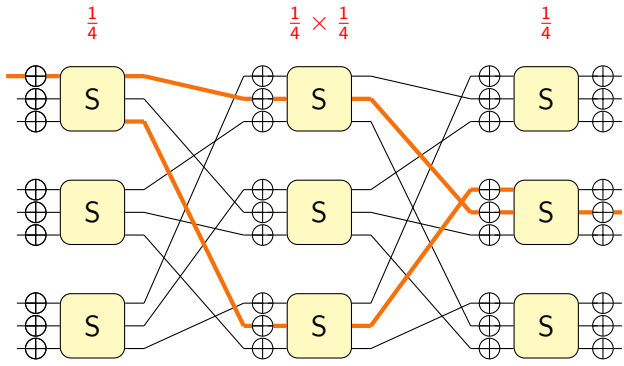


$$\frac{1}{2^6} + (-1)^{\kappa_1}\frac{1}{2^7} + (-1)^{\kappa_2}\frac{1}{2^7} + (-1)^{\kappa_1+\kappa_2}\frac{1}{2^8}$$

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

Quasidifferential trails

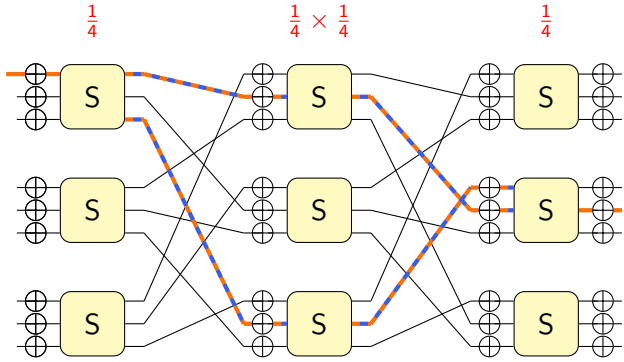
Example: 3-round differential (characteristic 2)



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

Quasidifferential trails

Example: 3-round differential (characteristic 2)

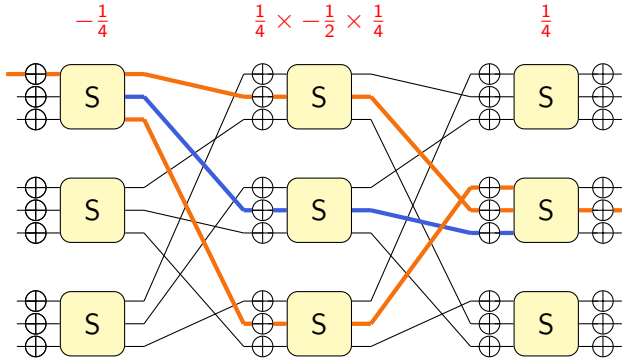


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3} \frac{1}{2^8}$$

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)

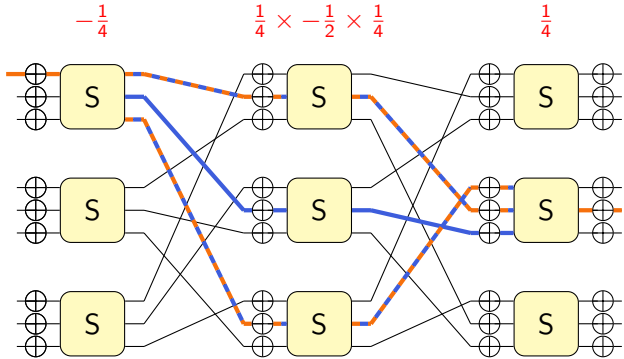


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_2}\frac{1}{2^9}$$

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)

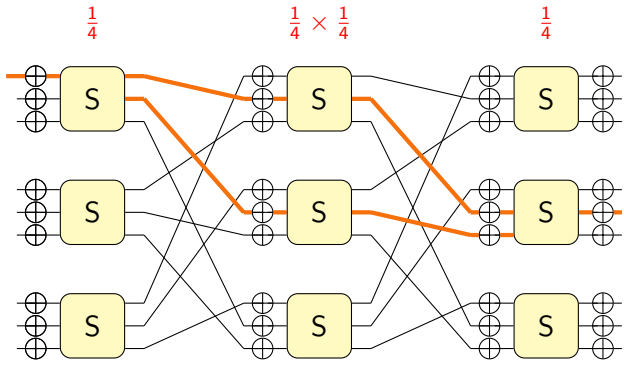


$$\frac{1}{2^8} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_2}\frac{1}{2^9} + (-1)^{\kappa_1+\kappa_2+\kappa_3}\frac{1}{2^9}$$

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

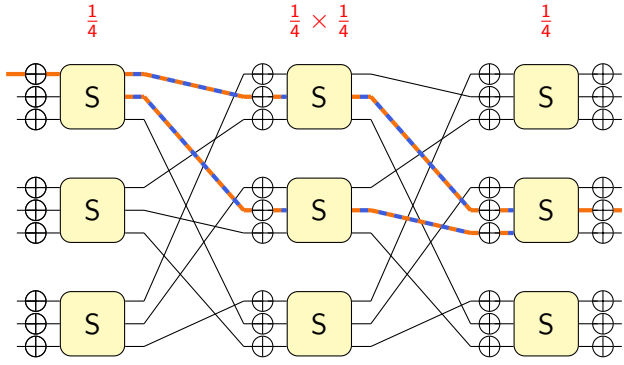
Example: 3-round differential (characteristic 3)



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

Quasidifferential trails

Example: 3-round differential (characteristic 3)

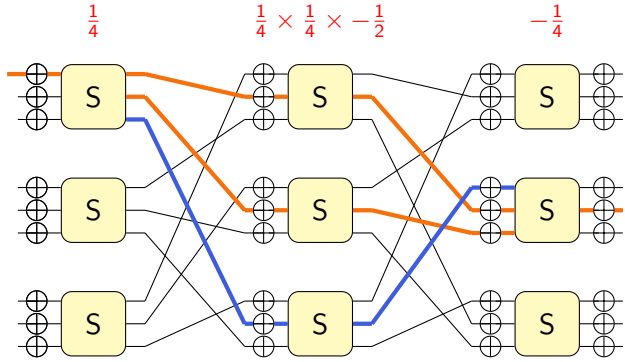


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3} \frac{1}{2^8}$$

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)

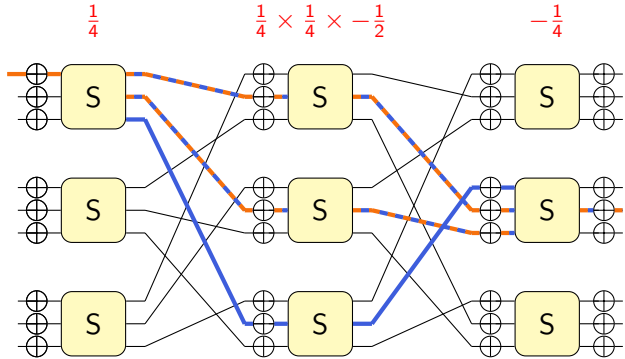


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_1}\frac{1}{2^9}$$

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)

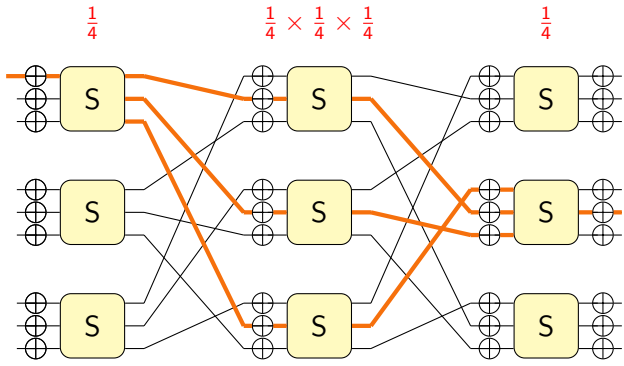


$$\frac{1}{2^8} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^8} + (-1)^{\kappa_1}\frac{1}{2^9} + (-1)^{\kappa_1+\kappa_2+\kappa_3}\frac{1}{2^9}$$

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

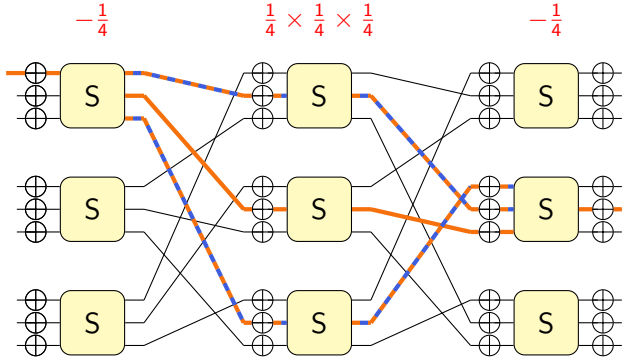
Example: 3-round differential (characteristic 4)



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

Quasidifferential trails

Example: 3-round differential (characteristic 4)

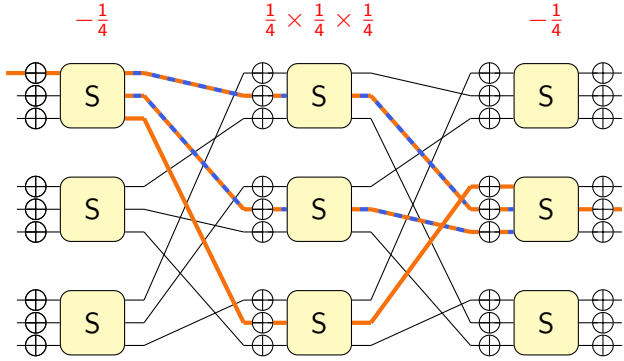


$$\frac{1}{2^{10}} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^{10}}$$

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)

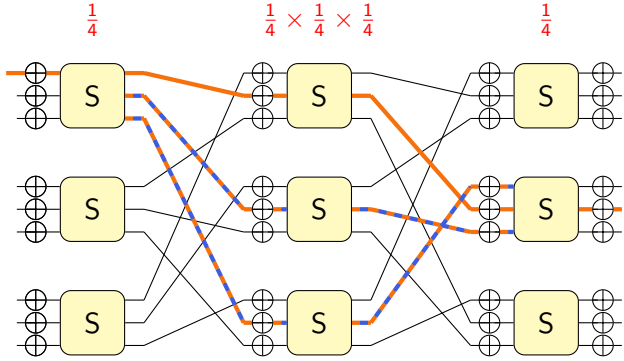


$$\frac{1}{2^{10}} + (-1)^{\kappa_1+\kappa_3}\frac{1}{2^{10}} + (-1)^{\kappa_2+\kappa_3}\frac{1}{2^{10}}$$

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

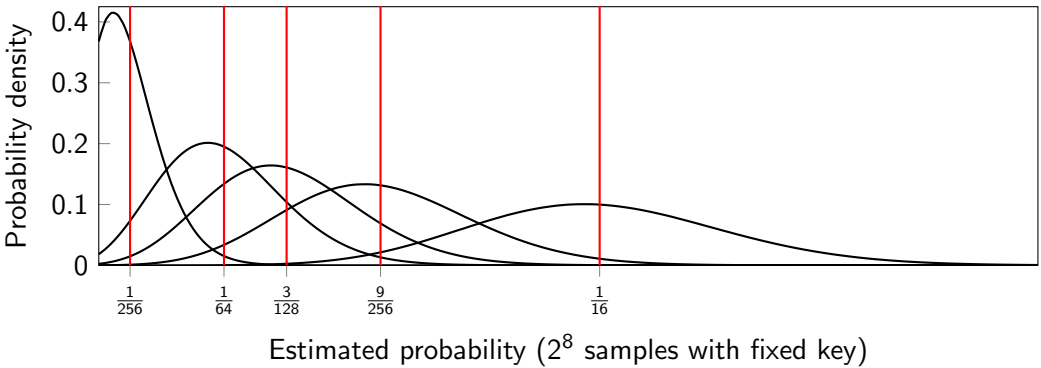
- Overall probability depends on three key bits

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

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

Quasidifferential trails

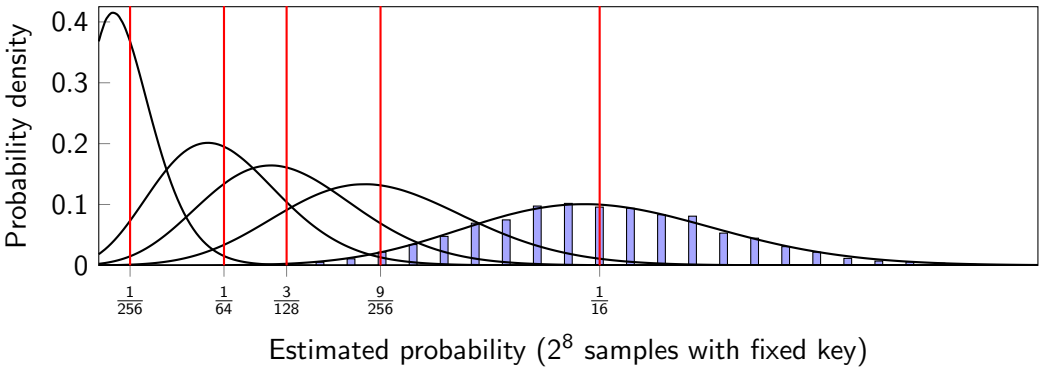
Example: 3-round differential



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

Quasidifferential trails

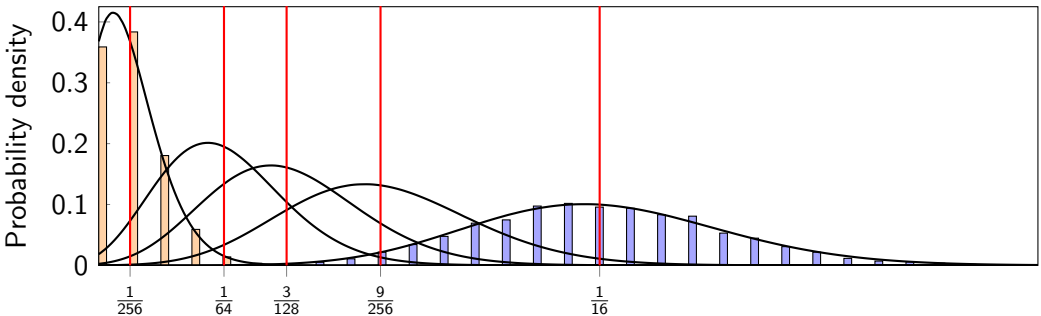
Example: 3-round differential



- 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)

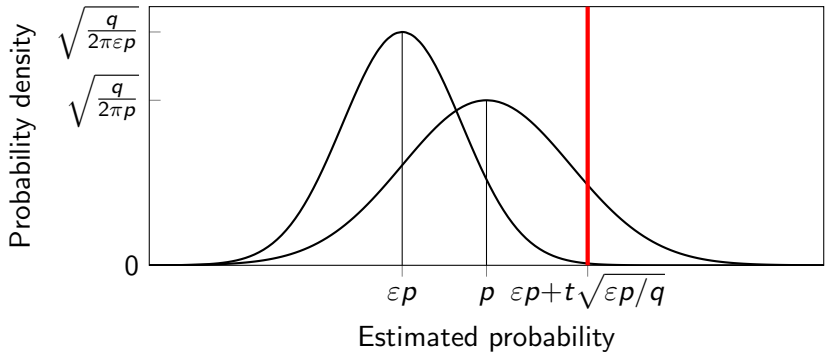
Cost analysis

- Suppose the probability is p and we use q independent samples:

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

- Simplifications:
  - Probability is εp in the ideal case
  - q is not too small and p is not too large
- Distribution of $\widehat{\boldsymbol{p}}$ is close to normal with mean p and variance $p(1-p)/q \approx p/q$
- Hypothesis test: $\widehat{\boldsymbol{p}} \ge \varepsilon p + t\sqrt{\varepsilon p/q}$

Cost analysis



- True-positive rate:

$$\begin{aligned} P_{\mathsf{S}} &= \Pr\left[\widehat{\boldsymbol{p}} \geq \varepsilon p + t\sqrt{\varepsilon p/q}\right] \\ &= \Pr\left[\widehat{\boldsymbol{p}} - p \geq t\sqrt{\varepsilon p/q} - (1-\varepsilon)p\right] \\ &= \Phi((1-\varepsilon)\sqrt{pq} - t\sqrt{\varepsilon}) \end{aligned}$$

- False-positive rate: $P_{\mathsf{F}} = \Phi(-t)$

Cost analysis

- Eliminating t gives

$$P_S = \Phi\big(\Phi^{-1}(P_F)\sqrt{\varepsilon} + (1-\varepsilon)\sqrt{pq}\big)$$

- Inverting this gives

$$q = \frac{1}{p}\left(\frac{\Phi^{-1}(P_S) - \Phi^{-1}(P_F)\sqrt{\varepsilon}}{1-\varepsilon}\right)^2$$

- One can show that this is essentially optimal

⚠ under important assumptions

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

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

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

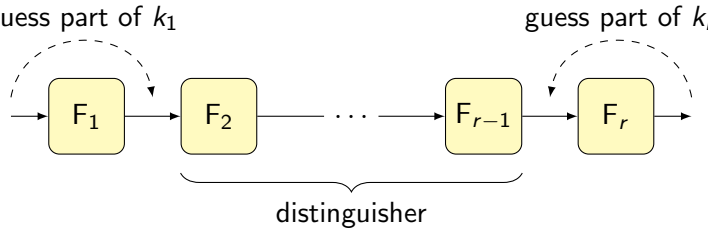
Key recovery

- 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
  - Count the number of right pairs per candidate key
  - Filter out invalid candidate keys based on the difference

- For K candidate keys, $P_F K$ incorrect candidates remain

ℹ Required amount of data depends on ε



Further topics

Outlook

- Further topics in linear cryptanalysis (not exhaustive)
  - Multiple linear cryptanalysis
  - Zero-correlation linear cryptanalysis
  - Nonlinear approximations
- Further topics in differential cryptanalysis (not exhaustive)
  - Truncated differentials
  - Impossible differentials
- It is worth learning the 'geometric approach' to understand how these fit together (see my talk at the SAC conference on Friday)

tim@cryptanalysis.info