Practical Decorrelation

Thomas Baignères
EPFL

ESC'08

Related Work

[Vau03] Vaudenay. *Decorrelation: A Theory for Block Cipher Security.* JOC 16(4) 2003

[BV05] Baignères, Vaudenay. *Proving the Security of AES Substitution-Permutation Network.* SAC 2005

[BF06a] Baignères, Finiasz. *Dial C for Cipher.* SAC 2006

[BF06b] Baignères, Finiasz. *KFC: the Krazy Feistel Cipher.* Asiacrypt 2006.

(Provable) Security For Block Ciphers

Today, most of the block ciphers that we use in practice (AES, FOX,...) are practically secure:

None of the smart cryptanalysts who attacked them was able to break them (yet).

(Provable) Security For Block Ciphers

Today, most of the block ciphers that we use in practice (AES, FOX,...) are practically secure:

None of the smart cryptanalysts who attacked them was able to break them (yet).

- Are there constructions that show something “stronger”?

(Provable) Security For Block Ciphers

Today, most of the block ciphers that we use in practice (AES, FOX,...) are practically secure:

None of the smart cryptanalysts who attacked them was able to break them (yet).

- Are there constructions that show something “stronger”?
- If there are, to what extend are they really “stronger”?

Outline

- Basic Security Notions
- The Decorrelation Theory
- Construction 1 : $\mathcal{C}$
- Construction 2 : KFC
- Critics

Outline

- The Decorrelation Theory
- Construction 1 : $\mathcal{C}$
- Construction 2 : KFC
- Critics

- Basic Security Notions
- The Luby-Rackoff Model
- The quantity to minimize: the advantage of an adversary $\mathcal{A}$

Outline

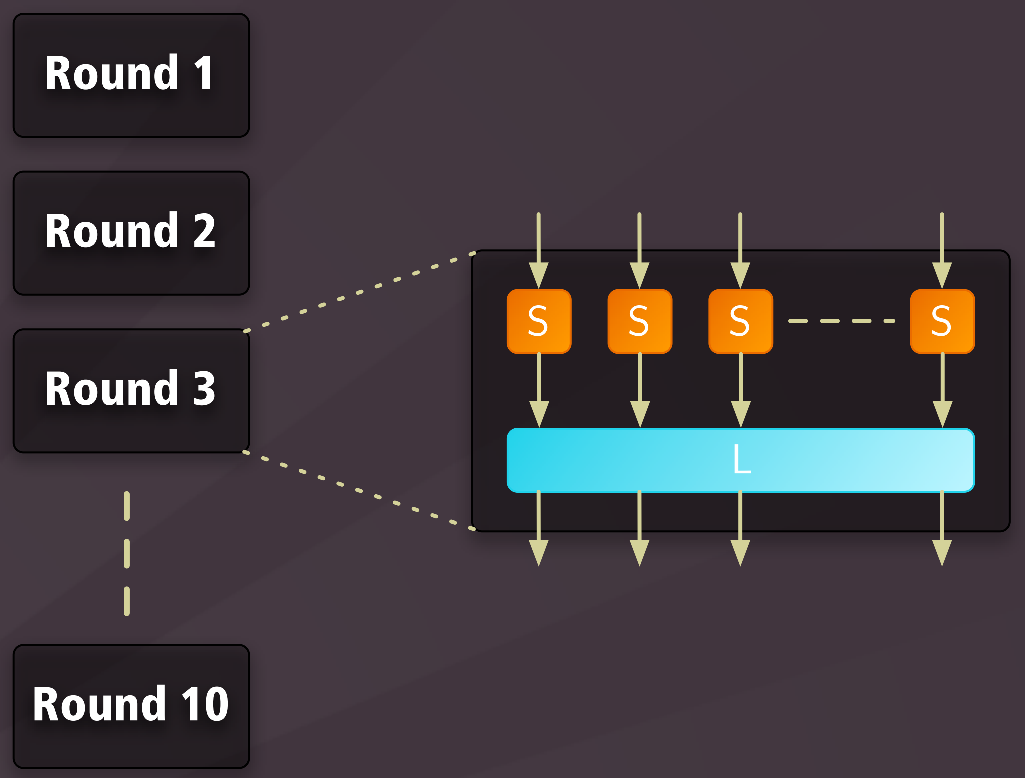
- Basic Security Notions
- Construction 1 : $\mathcal{C}$
- Construction 2 : KFC
- Critics

- The Decorrelation Theory
  - The distribution matrix of a block cipher
  - Link between the advantage of $\mathcal{A}$ and the distance between distribution matrices
  - Basic properties

Outline

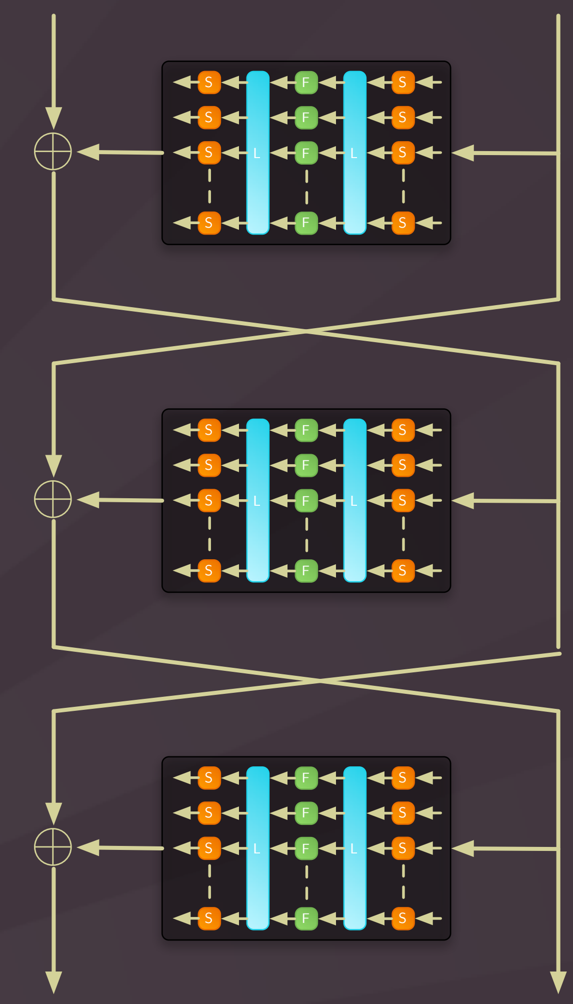
- Basic Security Notions
- The Decorrelation Theory
- Construction 2 : KFC
- Critics

- Construction 1 : C



Outline

- Basic Security Notions
- The Decorrelation Theory
- Construction 1 : $\mathcal{C}$
- Critics
- Construction 2 : KFC



Outline

- Basic Security Notions
- The Decorrelation Theory
- Construction 1 : $\mathcal{C}$
- Construction 2 : KFC

- Critics
- Independence of round keys
- Couldn't we use the one-time-pad instead?
- What about cash-timing attacks?

Basic Security Notions

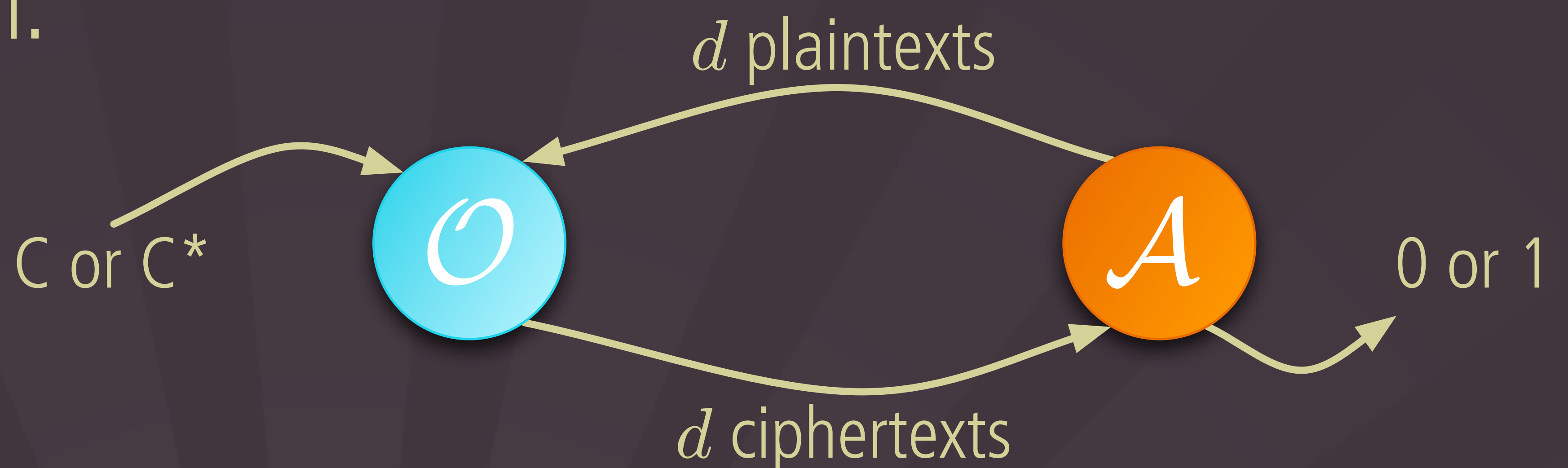
The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:

- computationally unbounded
- limited to d queries to an oracle $\mathcal{O}$ implementing either
  - a random instance C of the block cipher
  - a random instance C* of the perfect cipher
- the objective of $\mathcal{A}$ being to guess which is the case.

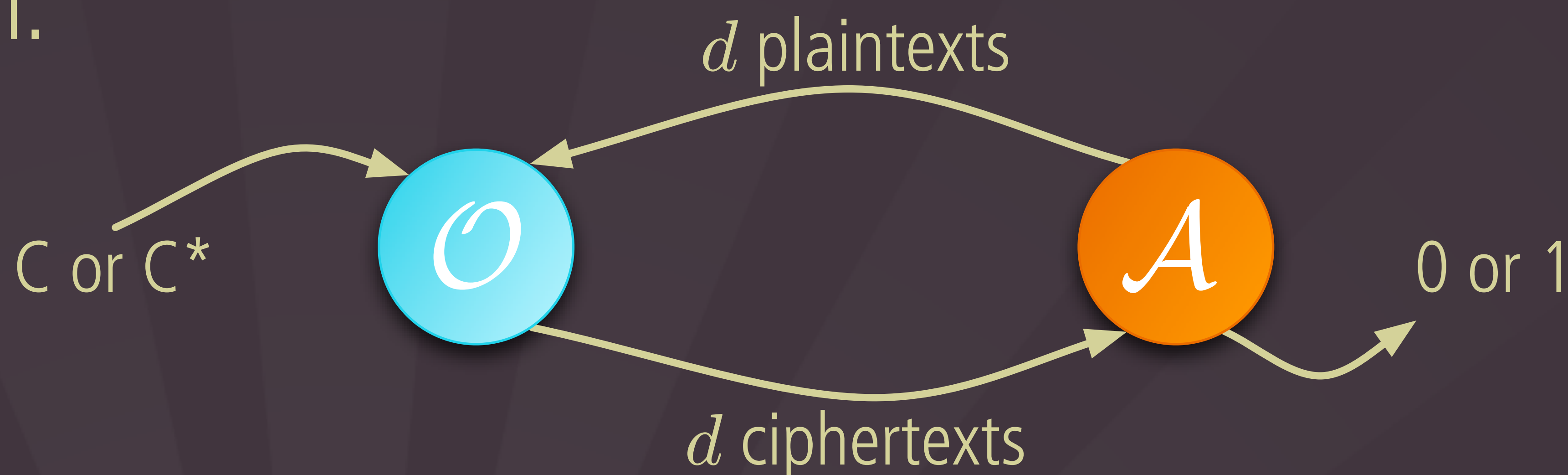
The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:



The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:



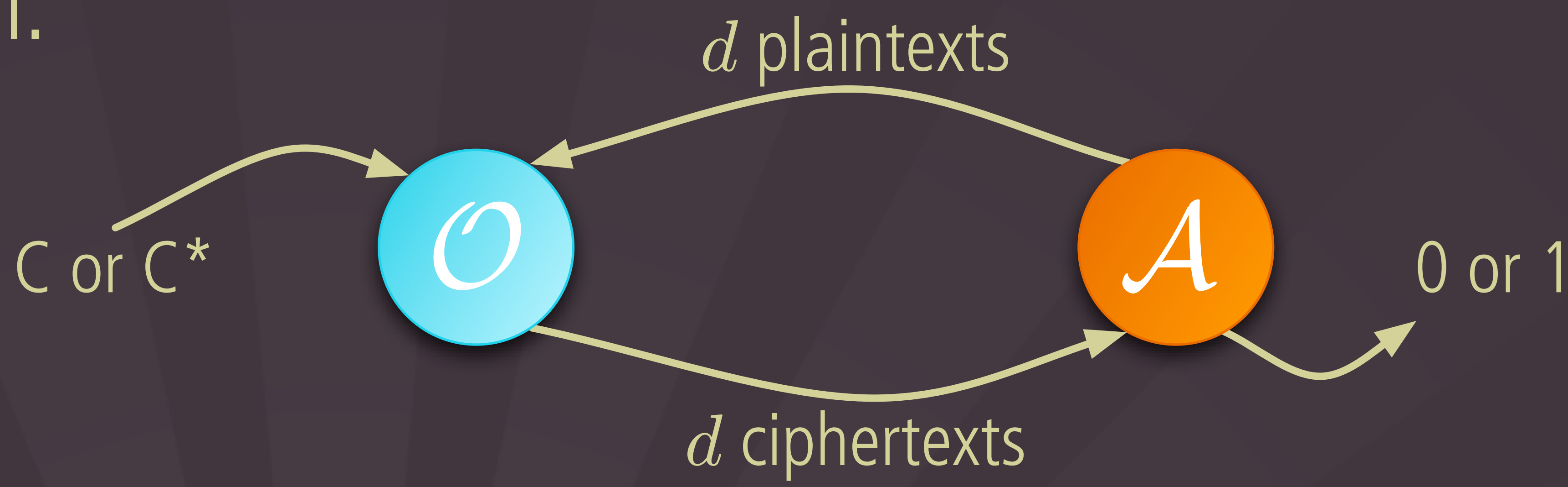
$$\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = |\Pr[\mathcal{A}(\mathsf{C}) = 1] - \Pr[\mathcal{A}(\mathsf{C}^*) = 1]|$$

Advantage of the d-limited adversary $\mathcal{A}$ between C and C*

The block cipher C is secure if the advantage of $\mathcal{A}$ is negligible for all $\mathcal{A}$'s.

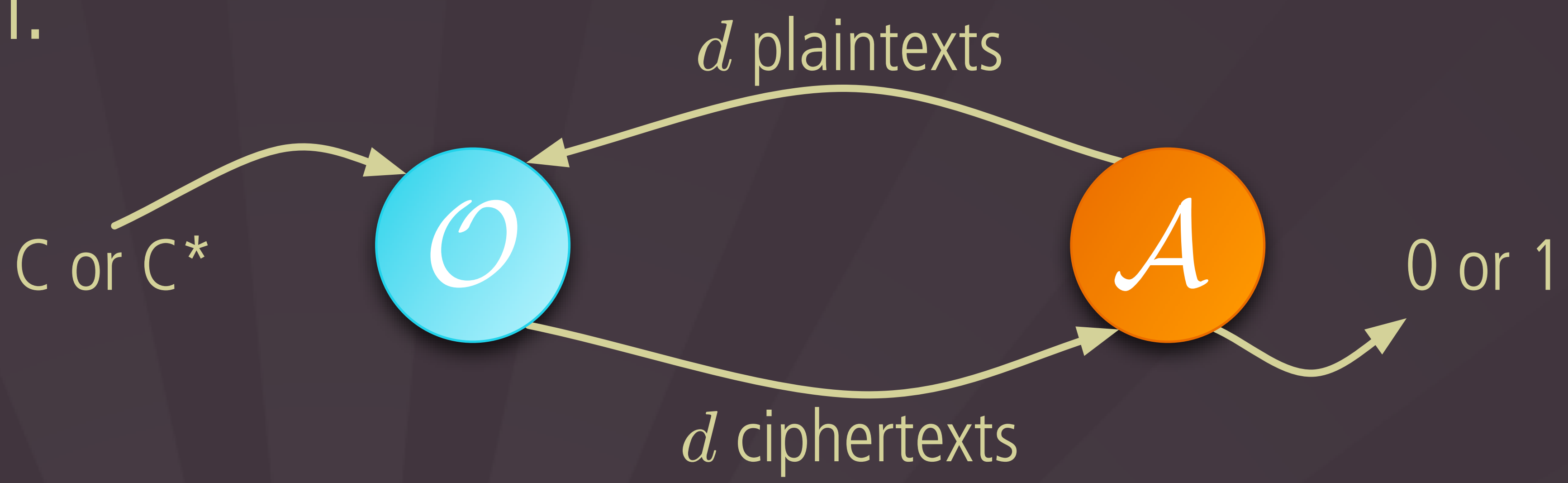
The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:



The Luby-Rackoff Model

We consider a d-limited adversary $\mathcal{A}$ in the Luby-Rackoff model:



$\mathcal{A}$ is non-adaptive if the d plaintexts are chosen "at once".

$\mathcal{A}$ is adaptive if plaintext i depends on ciphertexts 1,...,i-1.

The Decorrelation Theory

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$

- Computing the advantage is not a trivial task in general.
- Possible solution: use Vaudenay's Decorrelation Theory.

$$\max_{\mathcal{A}} \mathsf{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = \frac{1}{2} \left\| [\mathsf{C}]^d - [\mathsf{C}^*]^d \right\|$$

[Vau03]

Computing $\mathsf{Adv}_{\mathcal{A}}(\mathsf{C},\mathsf{C}^*)$

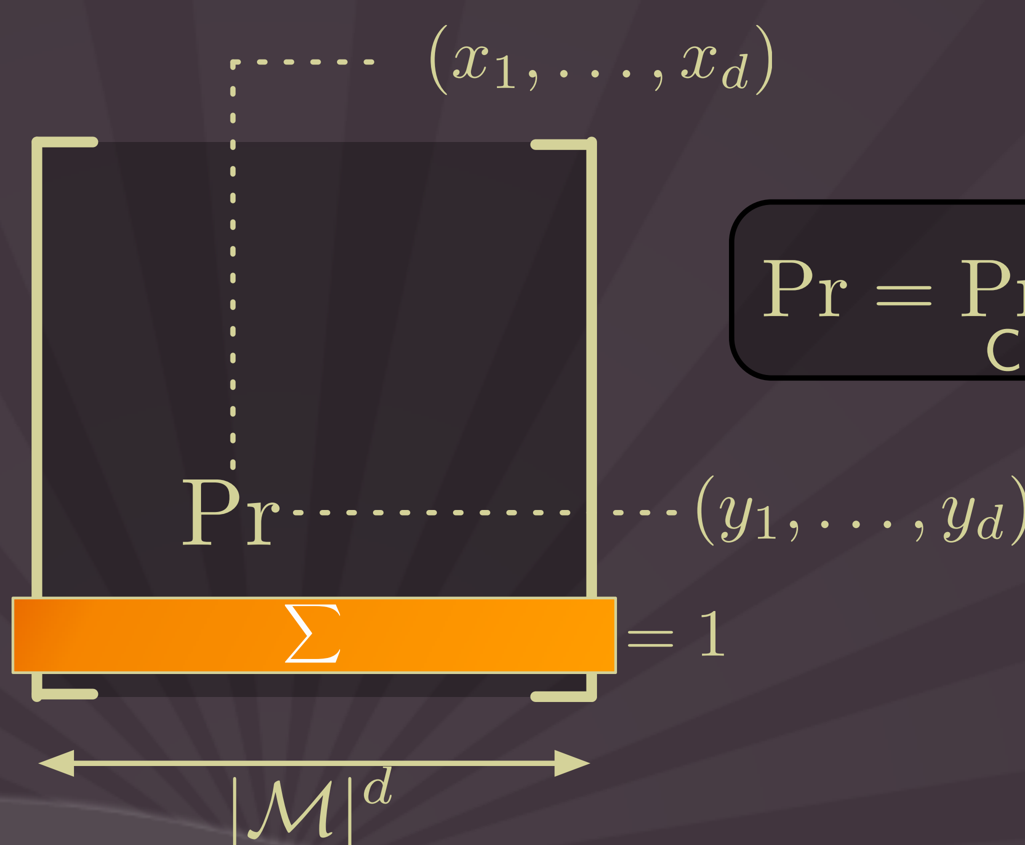
- Computing the advantage is not a trivial task in general.
- Possible solution: use Vaudenay's Decorrelation Theory.

$$\max_{\mathcal{A}} \mathsf{Adv}_{\mathcal{A}}(\mathsf{C},\mathsf{C}^*) = \frac{1}{2}\left|\left|[\mathsf{C}]^d - [\mathsf{C}^*]^d\right|\right|$$

[Vau03]

$$\Pr = \Pr_{\mathsf{C}}[\mathsf{C}(x_1)=y_1,\ldots,\mathsf{C}(x_d)=y_d]$$

$[\mathsf{C}]^d =$



Example !

On the set {1,2,3}, the distribution matrices of the perfect cipher look that this (at orders 1 and 2):

$$[\mathsf{C}^\star]^1 = \begin{array}{cc} \begin{array}{ccc} (1) & (2) & (3) \end{array} & \\ \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} & \begin{array}{c} (1) \\ (2) \\ (3) \end{array} \end{array}$$

$$[\mathsf{C}^\star]^2 = \begin{array}{cc} \begin{array}{ccccccccc} (1,1) & (1,2) & (1,3) & (2,1) & (2,2) & (2,3) & (3,1) & (3,2) & (3,3) \end{array} & \\ \begin{bmatrix} 1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \end{bmatrix} & \begin{array}{c} (1,1) \\ (1,2) \\ (1,3) \\ (2,1) \\ (2,2) \\ (2,3) \\ (3,1) \\ (3,2) \\ (3,3) \end{array} \end{array}$$

Adaptive vs. non-Adaptive Adversaries

- The norm used to compute the distance between two distribution matrices depends on the kind of adversary we consider.

Adaptive vs. non-Adaptive Adversaries

- The norm used to compute the distance between two distribution matrices depends on the kind of adversary we consider.
- If $\mathcal{A}$ is adaptive:

$$\max_{\mathcal{A}_\mathrm{a}} \mathrm{Adv}_{\mathcal{A}_\mathrm{a}}(\mathsf{C}, \mathsf{C}^*) = \frac{1}{2} \| [\mathsf{C}]^d - [\mathsf{C}^\star]^d \|_\mathrm{a}$$

$$\|M\|_\mathrm{a} = \max_{x_1} \sum_{y_1} \cdots \max_{x_d} \sum_{y_d} |M_{x,y}|$$

Adaptive vs. non-Adaptive Adversaries

- The norm used to compute the distance between two distribution matrices depends on the kind of adversary we consider.
- If $\mathcal{A}$ is adaptive:

$$\boxed{\max_{\mathcal{A}_\mathrm{a}} \mathrm{Adv}_{\mathcal{A}_\mathrm{a}}(\mathsf{C}, \mathsf{C}^*) = \tfrac{1}{2} \| [\mathsf{C}]^d - [\mathsf{C}^\star]^d \|_\mathrm{a}}$$

$$\|M\|_\mathrm{a} = \max_{x_1} \sum_{y_1} \cdots \max_{x_d} \sum_{y_d} |M_{x,y}|$$

- If $\mathcal{A}$ is non-adaptive:

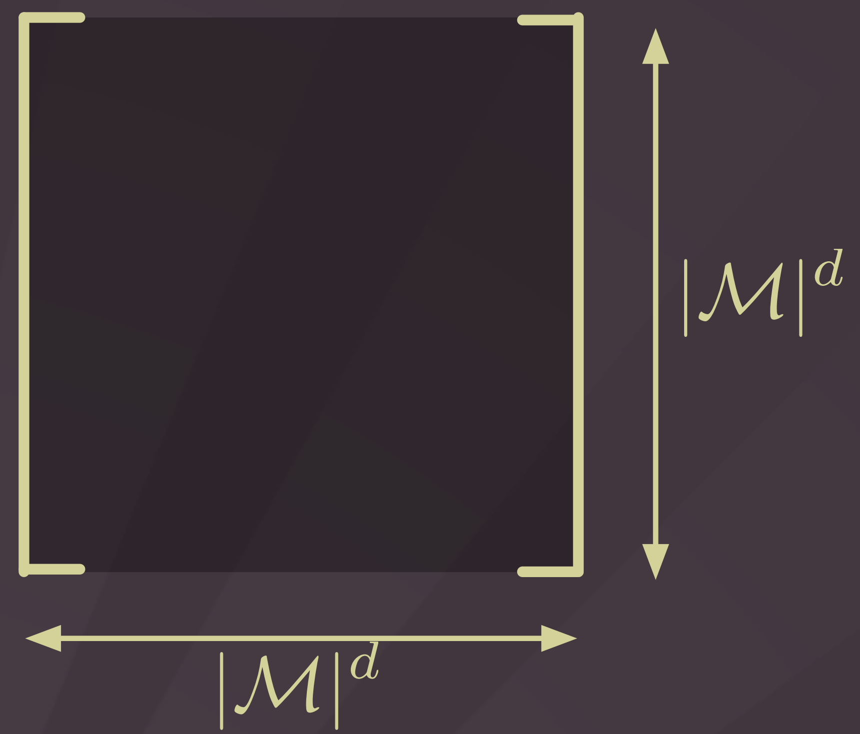
$$\boxed{\max_{\mathcal{A}} \mathrm{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*) = \tfrac{1}{2} \| [\mathsf{C}]^d - [\mathsf{C}^\star]^d \|_\infty}$$

$$\|M\|_\infty = \max_{x_1,\ldots,x_d} \sum_{y_1,\ldots,y_d} |M_{x,y}|$$

[Vau03]

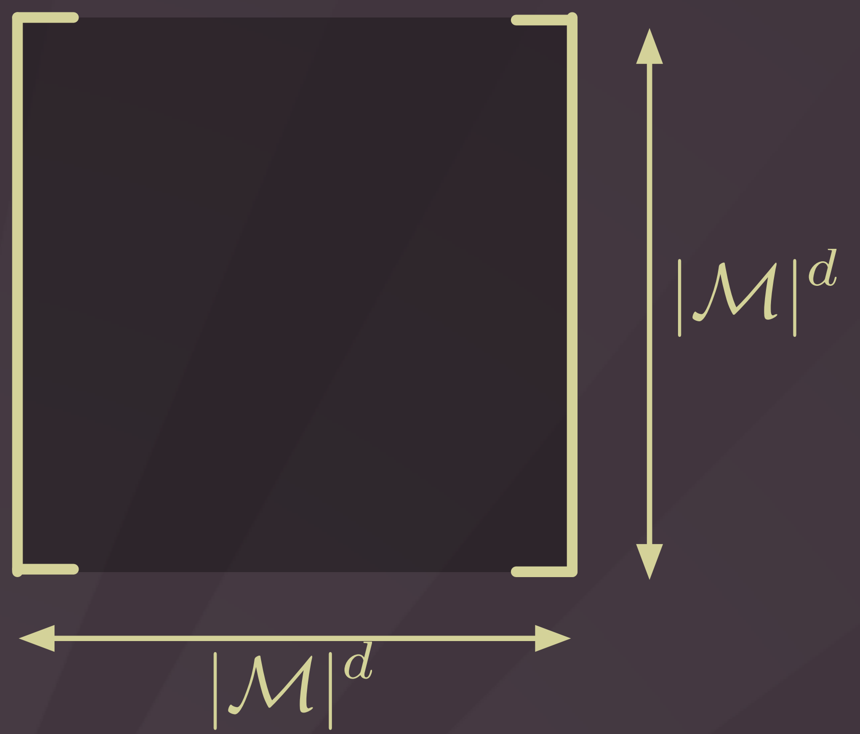
Are we done then? Not Quite :-<

$[\mathsf{C}]^d =$



Are we done then? Not Quite :-<

$[\mathsf{C}]^d =$



$|\mathcal{M}^d| = 2^{128 \cdot d}$ for a 128-bits block cipher

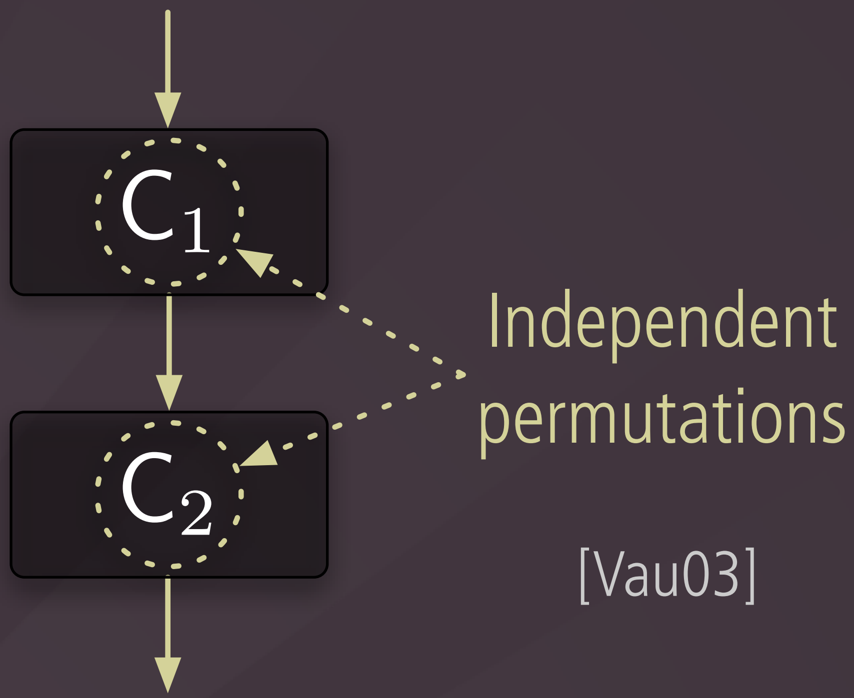
Tricks for Computing $\mathrm{Adv}_{\mathcal{A}}(\mathsf{C}, \mathsf{C}^*)$

To deal with the size of the distribution matrices:

Tricks for Computing $\mathrm{Adv}_{\mathcal{A}}(C, C^*)$

To deal with the size of the distribution matrices:

- ✔ $[C_2 \circ C_1]^d = [C_1]^d \times [C_2]^d$



[Vau03]

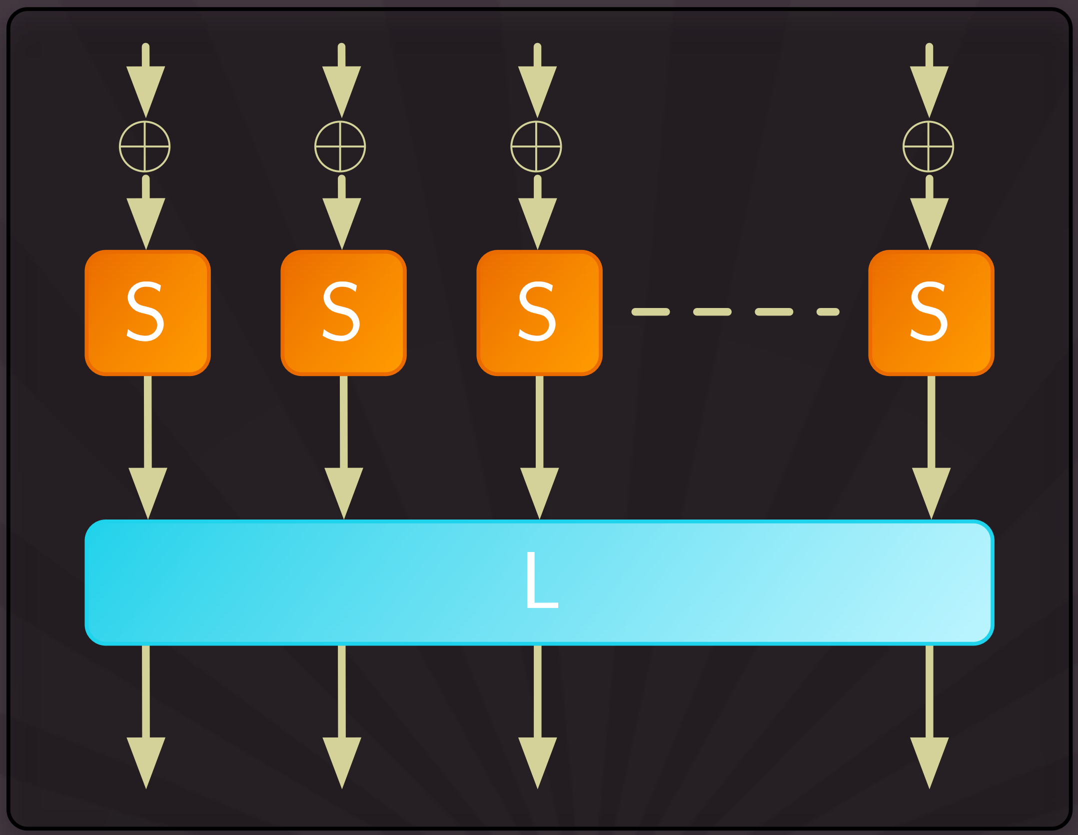
- ✔ Take advantage of the symmetries of the block cipher in order to compute the distribution matrix of each round.

Dial $\mathcal{C}$ for Cipher

Description of $\mathcal{C}$

$\mathcal{C}$ corresponds to the AES where " addRoundKeys → SubBytes" is replaced by mutually independent random permutations.

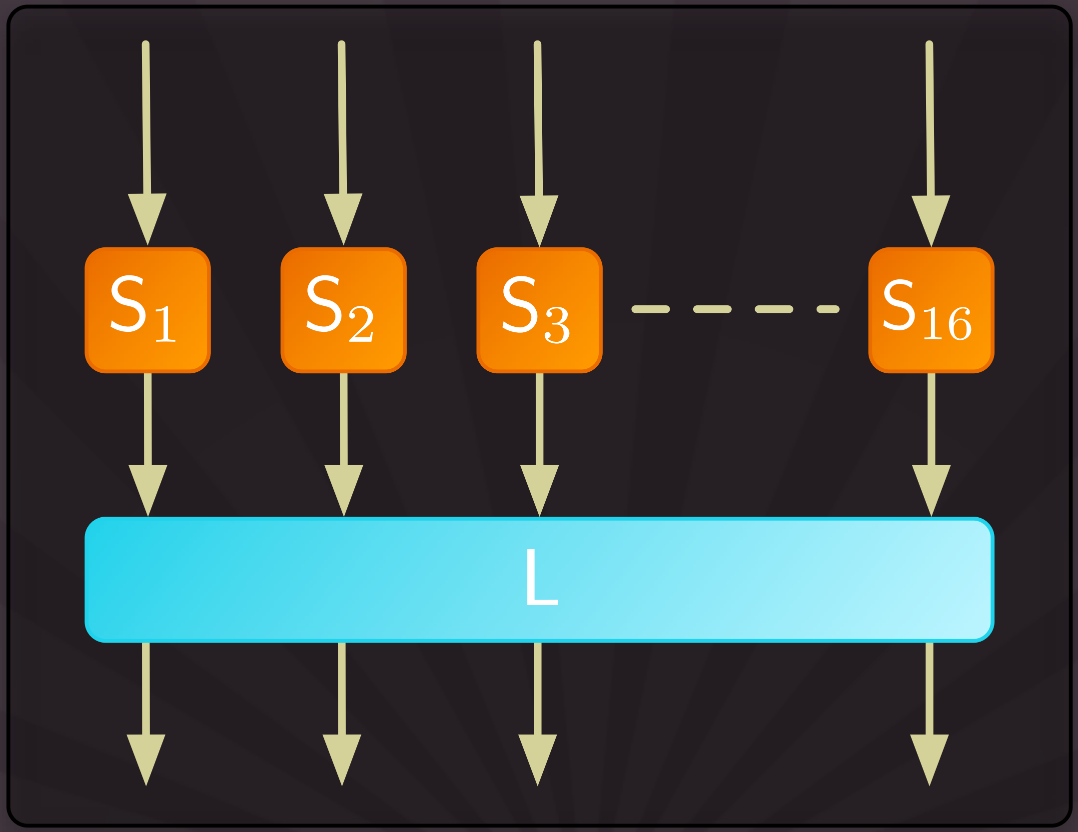
AES



Description of C

C corresponds to the AES where " addRoundKeys → SubBytes" is replaced by mutually independent random permutations.

AES ⤳ C



- C is made of 9 identical rounds, followed by a layer of substitution boxes.
- C uses $16 \cdot 10 = 160$ mutually independent random 8-bits substitution boxes.

Notations...

- A plaintext (or ciphertext) of $\mathcal{C}$ is a 4x4 array of elements of GF(256).
- The support of a plaintext is the 4x4 array with 0's where the plaintext has 0's and 1's anywhere else.

Notations...

- A plaintext (or ciphertext) of C is a 4x4 array of elements of GF(256).
- The support of a plaintext is the 4x4 array with 0's where the plaintext has 0's and 1's anywhere else.

plaintext

0x2f	0x00	0xaa	0x90
0xc2	0x43	0x12	0x01
0x01	0x26	0x00	0x2f
0xf1	0x00	0x55	0x7b

corresponding support

1	0	1	1
1	1	1	1
1	1	0	1
1	0	1	1

weight pattern

4	2	3	4

Shape of $[S]^2$

Denoting S one layer of substitution boxes of C:

$$[S]^2_{(x,x'),(y,y')} = \frac{\mathbf{1}_{\mathrm{supp}(x \oplus x') = \mathrm{supp}(y \oplus y')}}{q^{16} q^{\mathrm{w}(x \oplus x')}}$$

where $q = 2^8$.

Shape of $[S]^2$

Denoting S one layer of substitution boxes of C:

$$[S]^2_{(x,x'),(y,y')} = \frac{\mathbf{1}_{\mathrm{supp}(x\oplus x')=\mathrm{supp}(y\oplus y')}}{q^{16} q^{\mathrm{w}(x\oplus x')}}$$

where $q = 2^8$.

$$[S]^2 = \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix}$$

$$PS_{(x,x'),\gamma} = 1_{\gamma=\mathrm{supp}(x\oplus x')}$$

$$SP_{\gamma',(y,y')} = 1_{\gamma'=\mathrm{supp}(y\oplus y')} q^{-16} q^{-w(\gamma')}$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = [\mathsf{S}]^2 \times [\mathsf{L}]^2 \times [\mathsf{S}]^2 \times [\mathsf{L}]^2 \times [\mathsf{S}]^2$$

Shape of $[\mathrm{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathrm{C}]^2 = \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix} \times \begin{bmatrix} [\mathrm{L}]^2 \end{bmatrix} \times \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix} \times \begin{bmatrix} [\mathrm{L}]^2 \end{bmatrix} \times \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix}$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix} \times \begin{bmatrix} [\mathsf{L}]^2 \end{bmatrix} \times \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix} \times \begin{bmatrix} [\mathsf{L}]^2 \end{bmatrix} \times \begin{bmatrix} PS \end{bmatrix} \times \begin{bmatrix} SP \end{bmatrix}$$

$$\begin{bmatrix} SW \end{bmatrix} \times \begin{bmatrix} \bar{\mathsf{L}} \end{bmatrix} \times \begin{bmatrix} WS \end{bmatrix}$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = \left[PS\right] \times \left[SP\right] \times \left[[\mathsf{L}]^2\right] \times \left[PS\right] \times \left[SP\right] \times \left[[\mathsf{L}]^2\right] \times \left[PS\right] \times \left[SP\right]$$

$$\left[SW\right] \times \left[\bar{\mathsf{L}}\right] \times \left[WS\right] \qquad \left[SW\right] \times \left[\bar{\mathsf{L}}\right] \times \left[WS\right]$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP}$$

$$= \boxed{PS} \times \boxed{SW} \times \boxed{\bar{\mathsf{L}}} \times \boxed{WS} \times \boxed{SW} \times \boxed{\bar{\mathsf{L}}} \times \boxed{WS} \times \boxed{SP}$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP}$$

$$= \boxed{PS} \times \boxed{SW} \times \boxed{\bar{\mathsf{L}}} \times \boxed{WS} \times \boxed{SW} \times \boxed{\bar{\mathsf{L}}} \times \boxed{WS} \times \boxed{SP}$$

$$\boxed{WS} \times \boxed{SW} = \boxed{\overline{\mathsf{W}}}$$

Shape of $[\mathsf{C}]^2$

Considering C reduced to 3 rounds:

$$[\mathsf{C}]^2 = \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP} \times \boxed{[\mathsf{L}]^2} \times \boxed{PS} \times \boxed{SP}$$

$$= \boxed{PS} \times \boxed{SW} \times \boxed{\overline{\mathsf{L}}} \times \boxed{WS} \times \boxed{SW} \times \boxed{\overline{\mathsf{L}}} \times \boxed{WS} \times \boxed{SP}$$

$$= \boxed{PS} \times \boxed{SW} \times \boxed{\overline{\mathsf{L}}} \times \boxed{\overline{\mathsf{W}}} \times \boxed{\overline{\mathsf{L}}} \times \boxed{WS} \times \boxed{SP}$$

$\overline{\mathsf{W}}$ and $\overline{\mathsf{L}}$ are 625×625 matrices.

Advantage against 2-limited Adversaries

Using the previous expression of $[C]^2$, we manage to compute the exact values of

$$\frac{1}{2}\|[C]^2 - [C^\star]^2\|_a \quad \text{and} \quad \frac{1}{2}|||[C]^2 - [C^\star]^2|||_\infty$$

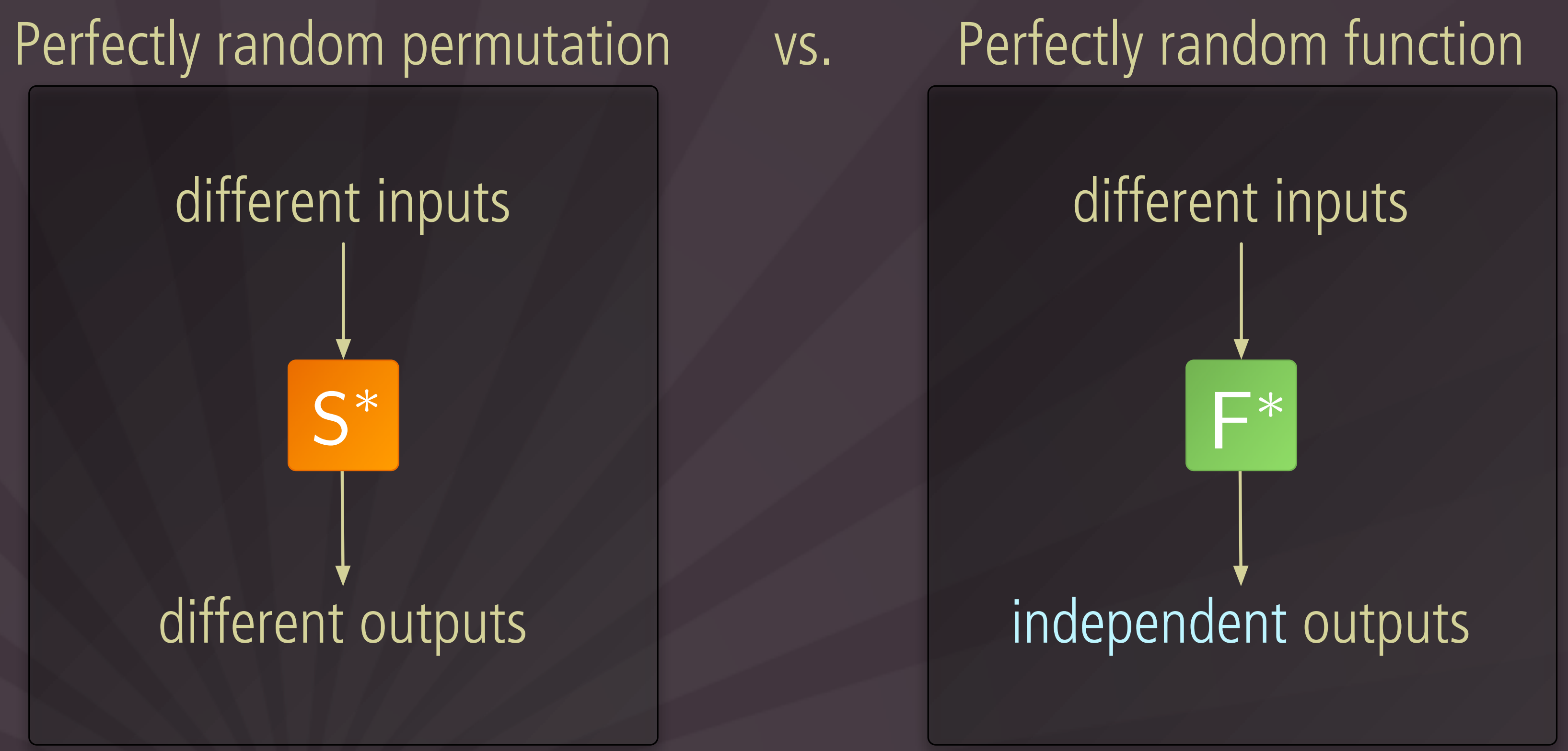
which appear to be the same.

r	2	3	4	5	6	7	8	9	10	11	12
BestAdv	1	2^{-4}	2^{-23}	2^{-45}	2^{-71}	2^{-126}	2^{-141}	2^{-163}	2^{-185}	2^{-210}	2^{-238}

KFC
the Krazy Feistel Cipher

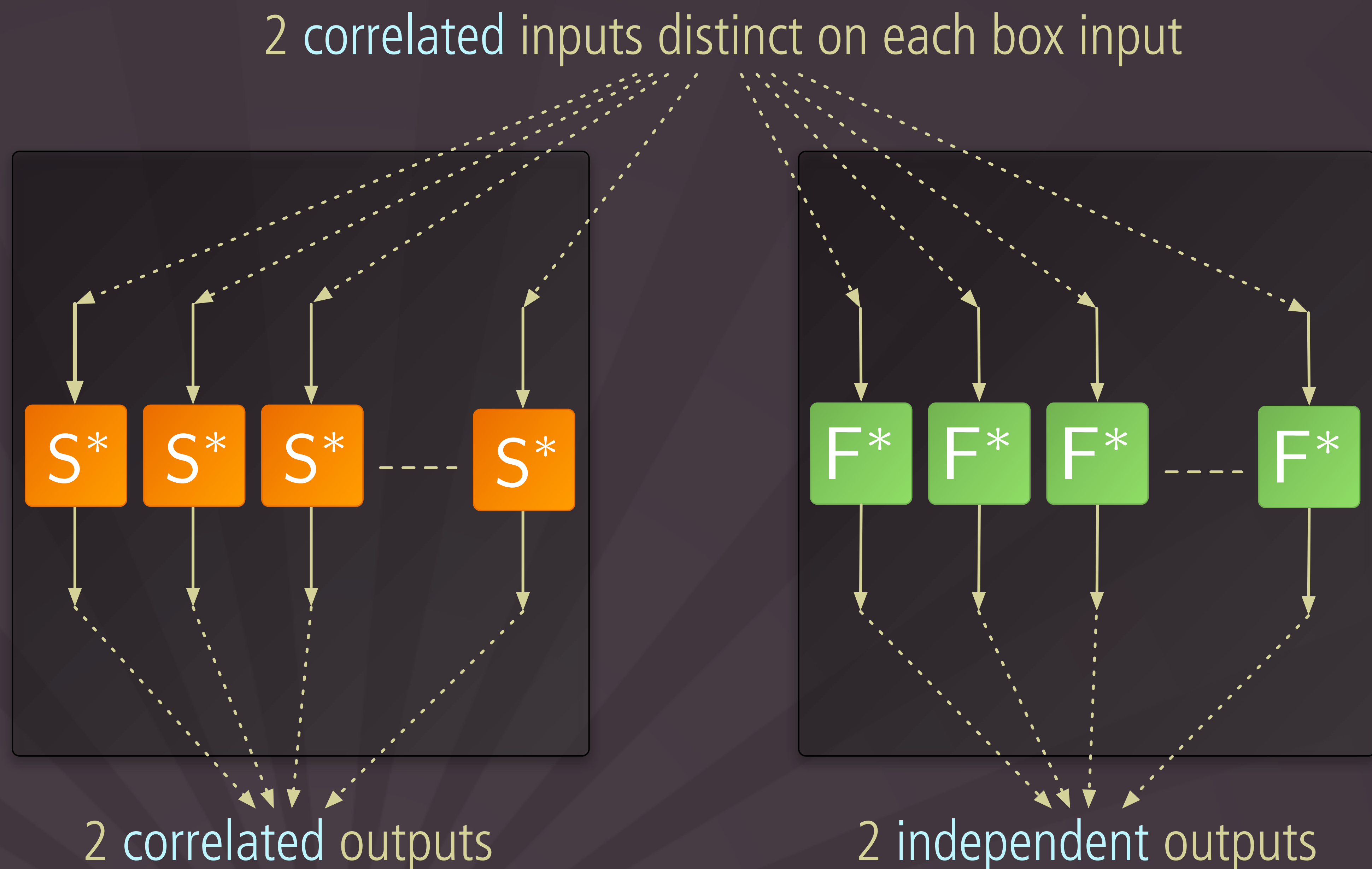
What about Higher Orders?

- We did not manage to prove the security of C against higher d-limited adversaries for $d > 2$.
- Idea: try to bound the advantage of the best d-limited adversary by that of the best (d-1)-limited adversary.

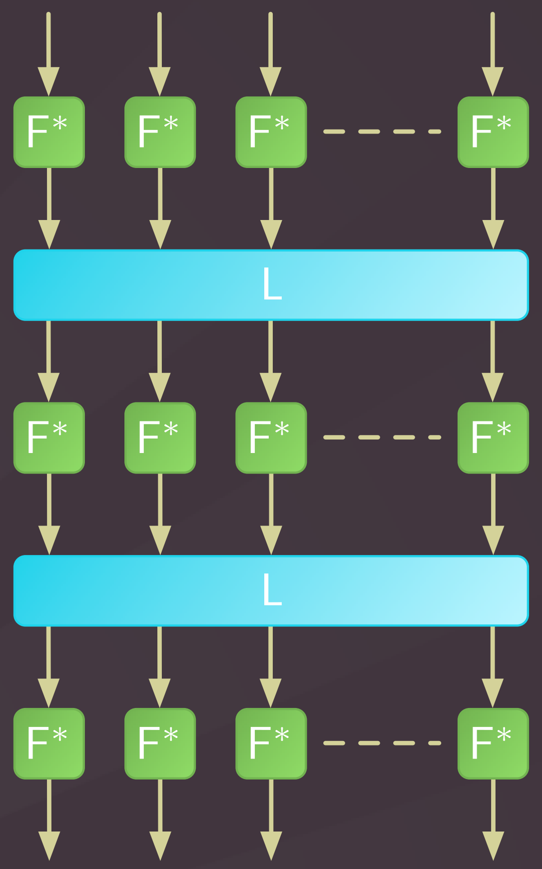


Rand. Permutations vs. Rand. Functions

Rand. Permutations vs. Rand. Functions

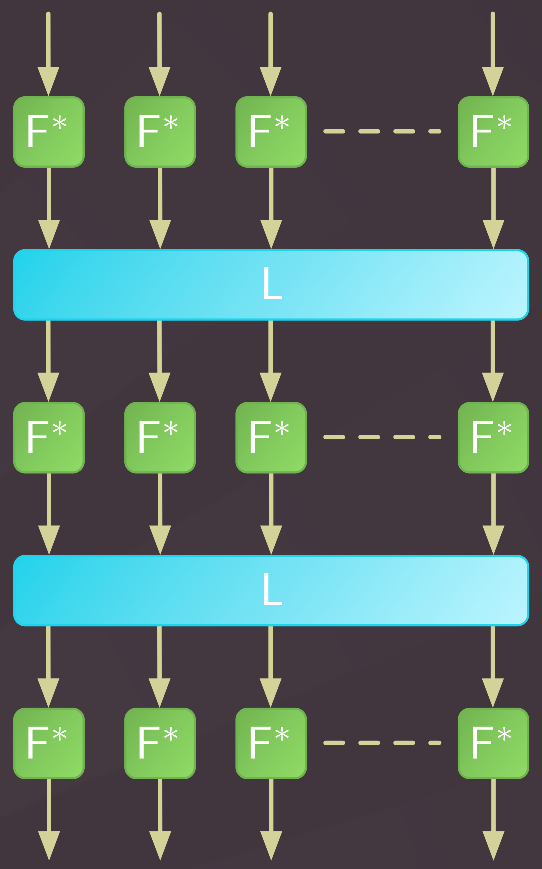


Towards a new Construction



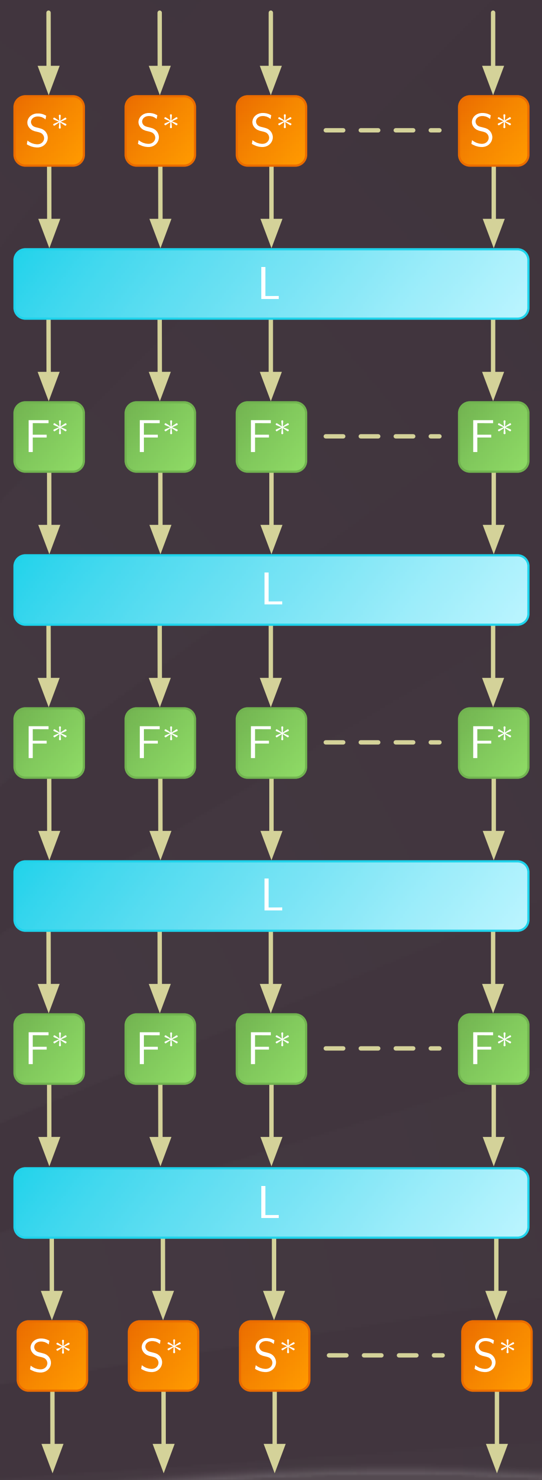
Towards a new Construction

- Non negligible risk of collision after a F box.



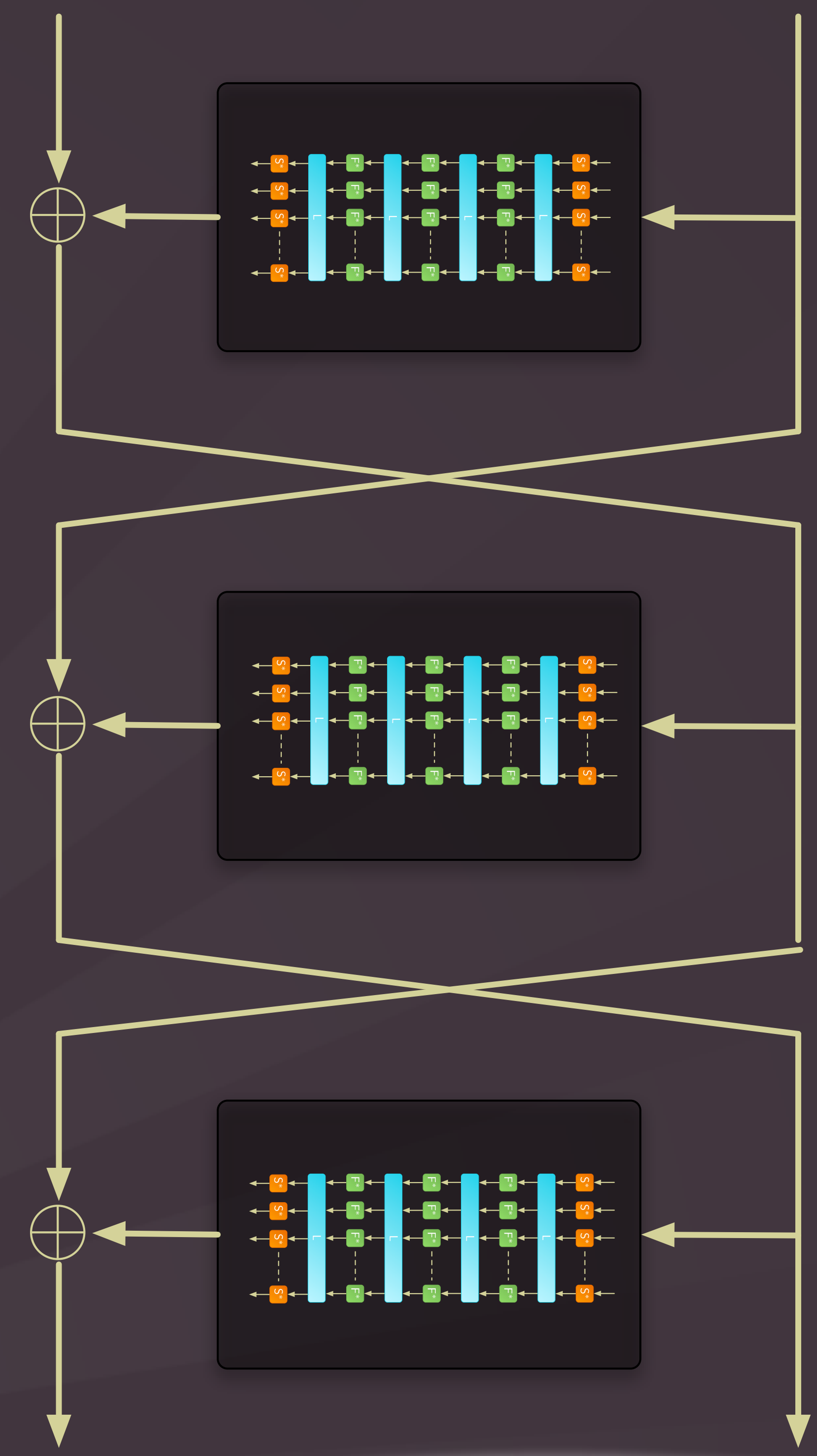
Towards a new Construction

- Non negligible risk of collision after a F box.
- Use the "sandwich technique" to obtain (almost) pairwise independent inputs before the layer of random functions.



Towards a new Construction

- Non negligible risk of collision after a F box.
- Use the "sandwich technique" to obtain (almost) pairwise independent inputs before the layer of random functions.
- The construction is not invertible. We plug it in a Feistel scheme.



Results obtained on KFC

- With this approach, we manage to prove the security against adversaries up to order 70 (for an unreasonable set of parameters).
- The bound is not tight at all   it is certainly possible to improve our results.

Results obtained on KFC

- With this approach, we manage to prove the security against adversaries up to order 70 (for an unreasonable set of parameters).
- The bound is not tight at all ⇝ it is certainly possible to improve our results.

Critics !

Requirements & Uncovered Attacks

- C might never fit, say, RFID tags (in the best case we need 160kB of memory to store the tables).
- We proposed so-called “provably secure” block ciphers...
- ... which are not provably secure against all know attacks.
- e.g., C is not provably secure against cache attacks.

Requirements & Uncovered Attacks

- C might never fit, say, RFID tags (in the best case we need 160kB of memory to store the tables).
- We proposed so-called "provably secure" block ciphers...
- ... which are not provably secure against all know attacks.
- e.g., C is not provably secure against cache attacks.
- "You should worry about things that are not in the security proofs." (Preneel, ESC08)

On the Decorrelation Theory

- The Decorrelation Theory tells more than what we used:
  - Resistance against 2-limited adversaries is sufficient to resist basic LC and DC.
  - Resistance against $2d$-limited adversaries is sufficient to resist iterated attacks of order d.
- The constructions that we proposed are not based on decorrelation modules (perfect constructions up to a given order, possibly weak beyond). We rely on the symmetries within the constructions themselves.

On the Independence of the Round Keys

- Our proofs assume that the rounds are mutually independent.
- This is not true in practice: thousands of bit of randomness are derived from a 128 bit key.
- Using a cryptographically secure PRNG, we can show that if an attack applies on the block cipher with the key schedule, but not on the block cipher with mutually independent rounds, then PRNG’s sequence can be distinguished from pure random.

Pessimistic View (not my favorite one)

- Should we use BBS or QUAD in practice?
- Well... since we need more bits of randomness to generate the boxes than the number of bits we are allowed to encrypt, why not use the bits as a one-time-pad... and throw away all the constructions? ☹

Optimistic View

- The assumption about the independence of the round keys has nothing to do with the block cipher itself, but with the key schedule.
- If a "provably secure" block cipher is broken by an attack against which it should resist, it should be sufficient to make its key schedule stronger.
- Making sure that the distribution matrix of the block cipher considered is close to that of the perfect cipher appears to be very natural. Independently of the key schedule, it seems to be a strong security argument.

Thank you for your Attention

Thank you for your Attention

☺