Linear Cryptanalysis of Non Binary Ciphers

(with an Application to SAFER)

Thomas Baignères
EPFL

Jacques Stern
ENS

Serge Vaudenay
EPFL

Selected Areas in Cryptography - SAC 07
Ottawa, Canada

Motivations

- A Block Cipher is commonly described as“ a set of permutations $C_k : \{0,1\}^\ell \rightarrow \{0,1\}^\ell$ indexed by a key k ”
- The data is not always *binary*, e.g. credit card numbers, social security numbers, string of alphabetical characters, etc.
- *We don't want to restrict to block ciphers defined over binary strings.*

	efficiency	simplicity (security analysis)
encode prior encryption		
dedicated cipher		

Motivations

- A Block Cipher is commonly described as“ a set of permutations $\mathsf{C}_k : \{0,1\}^\ell \rightarrow \{0,1\}^\ell$ indexed by a key k ”
- The data is not always *binary*, e.g. credit card numbers, social security numbers, string of alphabetical characters, etc.
- *We don’t want to restrict to block ciphers defined over binary strings.*

	efficiency	simplicity (security analysis)
encode prior encryption	✗	
dedicated cipher	✓	

Motivations

- A Block Cipher is commonly described as“ a set of permutations $C_k : \{0,1\}^\ell \to \{0,1\}^\ell$ indexed by a key k ”
- The data is not always *binary*, e.g. credit card numbers, social security numbers, string of alphabetical characters, etc.
- *We don’t want to restrict to block ciphers defined over binary strings.*

	efficiency	simplicity (security analysis)
encode prior encryption	✗	≈
dedicated cipher	✓	

Motivations

- A Block Cipher is commonly described as" a set of permutations $C_k : \{0,1\}^\ell \rightarrow \{0,1\}^\ell$ indexed by a key k "
- The data is not always *binary*, e.g. credit card numbers, social security numbers, string of alphabetical characters, etc.

We don't want to restrict to block ciphers defined over binary strings.

	efficiency	simplicity (security analysis)
encode prior encryption	✗	≈
dedicated cipher	✓	?

Motivations

- Arbitrary cardinality the only structure we assume is that of *Abelian Group*.
- Do the typical binary security notions easily generalize to this more general assumption?

Motivations

- Arbitrary cardinality the only structure we assume is that of *Abelian Group*.
- Do the typical binary security notions easily generalize to this more general assumption?
- Not always! Linear cryptanalysis is based on a metric called the linear probability, that sticks to the $\oplus$ operation over binary strings.
- Granboulan et al. [FSE'07] provide a generalization of the LP which is not completely sound (no duality with DC, no means to compute the exact attack complexity, no linear hull effect).

Outline (New Tools...)

Distinguishing a random source over an Abelian group:

- Optimal distinguisher
- Linear distinguisher
- Links between the two
- Distinguishing in practice using compression

Outline (New Tools...)

Distinguishing a random source over an Abelian group:

- Optimal distinguisher
- Linear distinguisher
- Links between the two
- Distinguishing in practice using compression

Distinguishing a random permutation over an Abelian group:

- From random sources to random permutations
- A Toolbox for Linear Cryptanalysis of block ciphers defined over an Abelian group

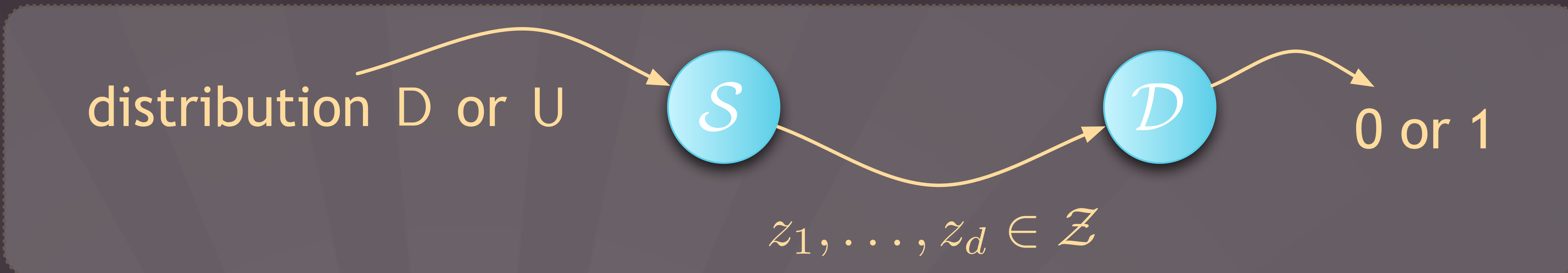
Outline (...in Practice)

- A $\boxplus$-linear cryptanalysis of SAFER K/SK (our generalization also improves results in the binary case!)
- An attack on TOY100 (toy cipher proposed by Granboulan et al. at FSE'07)
- "New" toy block cipher proposal: DEAN18 (Digital Encryption Algorithm for Numbers)

Distinguishing a Random Source over an Abelian Group

The Game

- D is an arbitrary distribution over some set $\mathcal{Z}$.
- U is the uniform distribution over $\mathcal{Z}$.



- $\mathcal{S}$ is a source that generates d samples $z_1, \ldots, z_d \in \mathcal{Z}$ according to distribution D (prob. 1/2) or U (prob. 1/2).
- $\mathcal{D}$ is a distinguisher that outputs 1 if it guesses that the correct distribution is D and 0 otherwise.

The Game

- D is an arbitrary distribution over some set $\mathcal{Z}$.
- U is the uniform distribution over $\mathcal{Z}$.

- $\mathcal{S}$ is a source that generates d samples $z_1, \ldots, z_d \in \mathcal{Z}$ according to distribution D (prob. 1/2) or U (prob. 1/2).
- $\mathcal{D}$ is a distinguisher that outputs 1 if it guesses that the correct distribution is D and 0 otherwise.
- The ability of $\mathcal{D}$ to distinguish D from U is its *advantage*:

$$\mathrm{Adv}_{\mathcal{D}}^{d} = \left|\Pr_{\mathsf{U}^d}[\mathcal{D} \to 1] - \Pr_{\mathsf{D}^d}[\mathcal{D} \to 1]\right|$$

Best Distinguisher

- Using maximum-likelihood techniques, one can describe an "optimal" distinguisher (i.e., maximizing the advantage).
- Defining the Squared Euclidean Imbalance of D as:

$$\Delta(\mathsf{D}) = |\mathcal{Z}| \sum_{z \in \mathcal{Z}} \left(P_{\mathsf{D}}(z) - \frac{1}{|\mathcal{Z}|} \right)^2$$

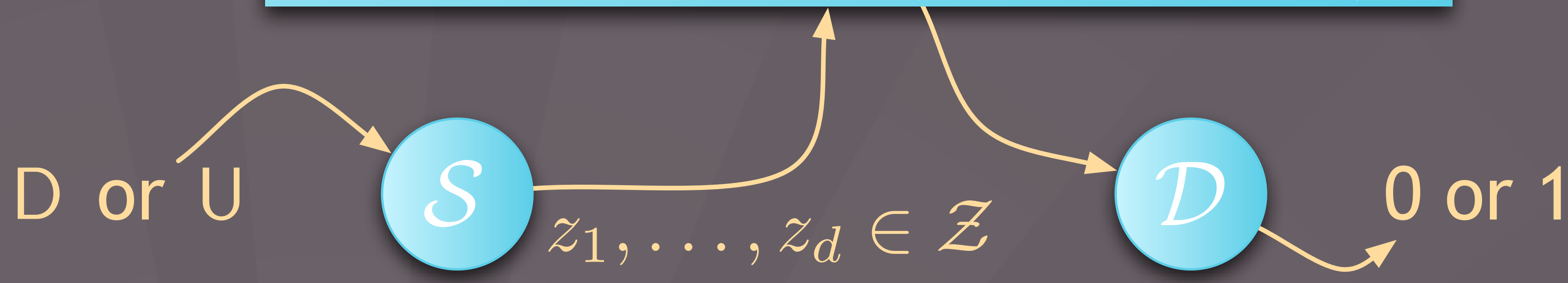
- ... the best distinguisher has an advantage equal to

$$\mathrm{Adv}^d_{\mathcal{D}} \approx 1 - 2 \cdot \Phi \left(-\sqrt{d \cdot \Delta(\mathsf{D})}/2 \right)$$

- Using $d \approx 1/\Delta(\mathsf{D})$ samples is sufficient to reach a significant advantage. [BJV04]

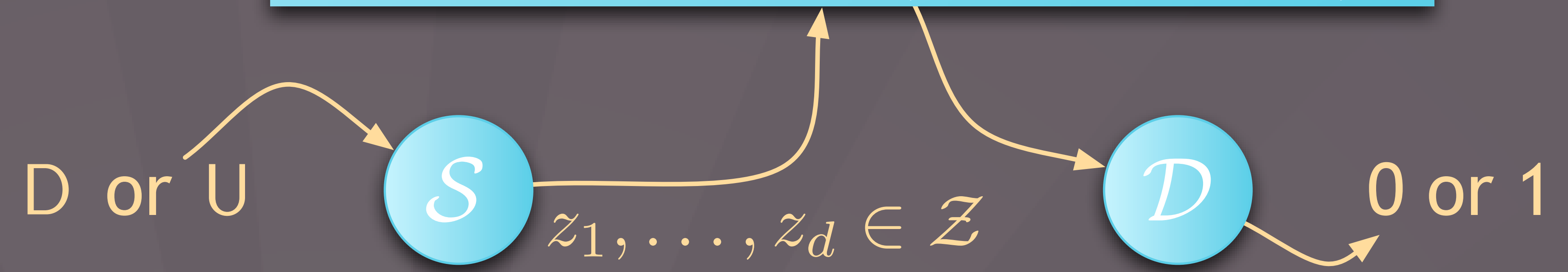
Linear Distinguisher

$$\mathrm{lp}(z_1,\ldots,z_d;u) = \left|\frac{1}{d}\sum_{i=1}^{d}(-1)^{u\cdot z_i}\right|^2$$



Linear Distinguisher

$$\mathrm{lp}(z_1,\ldots,z_d;u) = \left|\frac{1}{d}\sum_{i=1}^{d}(-1)^{u\cdot z_i}\right|^2$$



- When the distribution is U :

$$\mathrm{lp}(z_1,\ldots,z_d;u) \xrightarrow[d\to\infty]{} \left|\mathrm{E}_{\mathsf{U}}((-1)^{u\cdot X})\right|^2 = 0$$

- When the distribution is D :

$$\mathrm{lp}(z_1,\ldots,z_d;u) \xrightarrow[d\to\infty]{} \left|\mathrm{E}_{\mathsf{D}}((-1)^{u\cdot X})\right|^2 > 0$$

- Linear Distinguisher based on $\mathrm{LP}(u) = (2\Pr[u \cdot X] - 1)^2$.

Linear Distinguisher

- This description makes no sense when $\mathcal{Z}$ is not $\{0,1\}^n$!!

Thomas Baignères, Jacques Stern, Serge Vaudenay

Linear Distinguisher

- This description makes no sense when $\mathcal{Z}$ is not $\{0,1\}^n$!!

Definition:

The Linear Probability of D over the group $\mathcal{Z}$ with respect to the character χ is defined by

$$\mathrm{LP}_{\mathsf{D}}(\chi) = \left|\mathrm{E}_{\mathsf{D}}(\chi(X))\right|^2$$

Linear Distinguisher

- This description makes no sense when $\mathcal{Z}$ is not $\{0,1\}^n$!!

Definition:

The Linear Probability of D over the group $\mathcal{Z}$ with respect to the character χ is defined by

$$\mathrm{LP}_{\mathsf{D}}(\chi) = |\mathrm{E}_{\mathsf{D}}(\chi(X))|^2$$

- A character of $\mathcal{Z}$ is a homomorphism $\chi : \mathcal{Z} \to \mathbf{C}^\times$.

Linear Distinguisher

This description makes no sense when $\mathcal{Z}$ is not $\{0,1\}^n$!!

Definition:

The Linear Probability of D over the group $\mathcal{Z}$ with respect to the character χ is defined by

$$\mathrm{LP}_{\mathsf{D}}(\chi) = |\mathrm{E}_{\mathsf{D}}(\chi(X))|^2$$

- A character of $\mathcal{Z}$ is a homomorphism $\chi : \mathcal{Z} \to \mathbf{C}^{\times}$.
- Example: when $\mathcal{Z} = \{0,1\}^n$ we have $\chi(a) = (-1)^{u \cdot a}$.

Linear Distinguisher

This description makes no sense when $\mathcal{Z}$ is not $\{0,1\}^n$!!

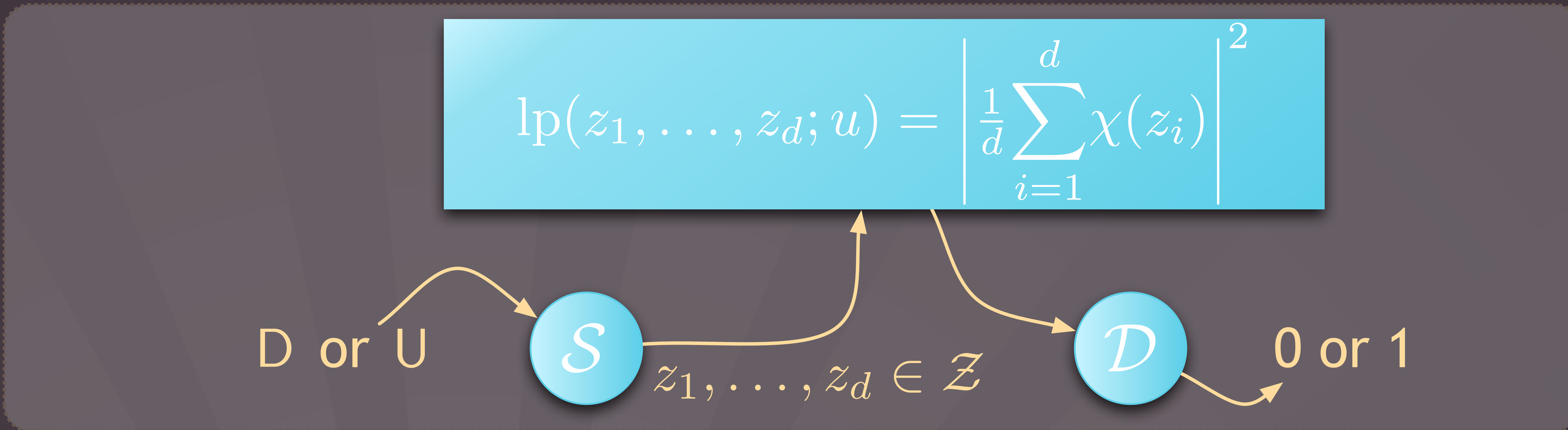
Definition:

The Linear Probability of D over the group $\mathcal{Z}$ with respect to the character χ is defined by

$$\mathrm{LP}_{\mathsf{D}}(\chi) = \left|\mathrm{E}_{\mathsf{D}}(\chi(X))\right|^2$$

- A character of $\mathcal{Z}$ is a homomorphism $\chi : \mathcal{Z} \to \mathbf{C}^\times$.
- Example: when $\mathcal{Z} = \{0,1\}^n$ we have $\chi(a) = (-1)^{u \cdot a}$.
- Consequence: when $\mathcal{Z} = \{0,1\}^n$ this new definition corresponds to the old one!

Linear Distinguisher



- When the distribution is U :

$$\mathrm{lp}(z_1,\ldots,z_d;u) \xrightarrow[d\to\infty]{} \mathrm{LP}_{\mathsf{U}}(\chi) = 0$$

- When the distribution is D :

$$\mathrm{lp}(z_1,\ldots,z_d;u) \xrightarrow[d\to\infty]{} \mathrm{LP}_{\mathsf{D}}(\chi) > 0$$

Linear Distinguisher

Theorem 7. *Let* G *be a finite Abelian group and let* $\chi \in \widehat{\mathsf{G}}$. *Using heuristic approximations, the advantage* $\mathrm{Adv}^d_{\mathcal{D}}$ *of a d-limited linear distinguisher* $\mathcal{D}$ *trying to distinguish the uniform distribution* U *from* D *is such that* $\mathrm{Adv}^d_{\mathcal{D}}(\chi) \succeq 1 - 2 \cdot e^{-\frac{d}{4}\mathrm{LP}_{\mathsf{D}}(\chi)}$ *(resp.* $\mathrm{Adv}^d_{\mathcal{D}}(\chi) \succeq 1 - 4 \cdot \Phi\left(-\frac{1}{2}\sqrt{d \cdot \mathrm{LP}_{\mathsf{D}}(\chi)}\right)$*) for* χ *of order at least 3 (resp. of order 2), when d is large enough and under the heuristic assumption that the covariance matrix of* $\mathrm{lp}(\boldsymbol{Z}^d;\chi)$ *is the same for both distributions.*[3]

Linear Distinguisher

A linear distinguisher needs $d \approx \frac{1}{\mathrm{LP}_{\mathsf{D}}(\chi)}$ samples to distinguish D from U.

Nice Properties

- Link Between Optimal and Linear Distinguishers:

Theorem: $\Delta(\mathsf{D}) = \sum_{\chi \neq \mathrm{id}} \mathrm{LP}_{\mathsf{D}}(\chi)$

- Link Between Linear and Differential Distinguishers:

Property: For all $u \in \mathcal{Z}$: $\widehat{\mathrm{LP}}_{\mathsf{D}}(u) = |\mathcal{Z}|\mathrm{DP}_{\mathsf{D}}(u)$

(where $\widehat{\mathrm{LP}}$ is the Fourier transform of the LP and where $\mathrm{DP}_{\mathsf{D}}(u) = \Pr[A \cdot u = B]$)

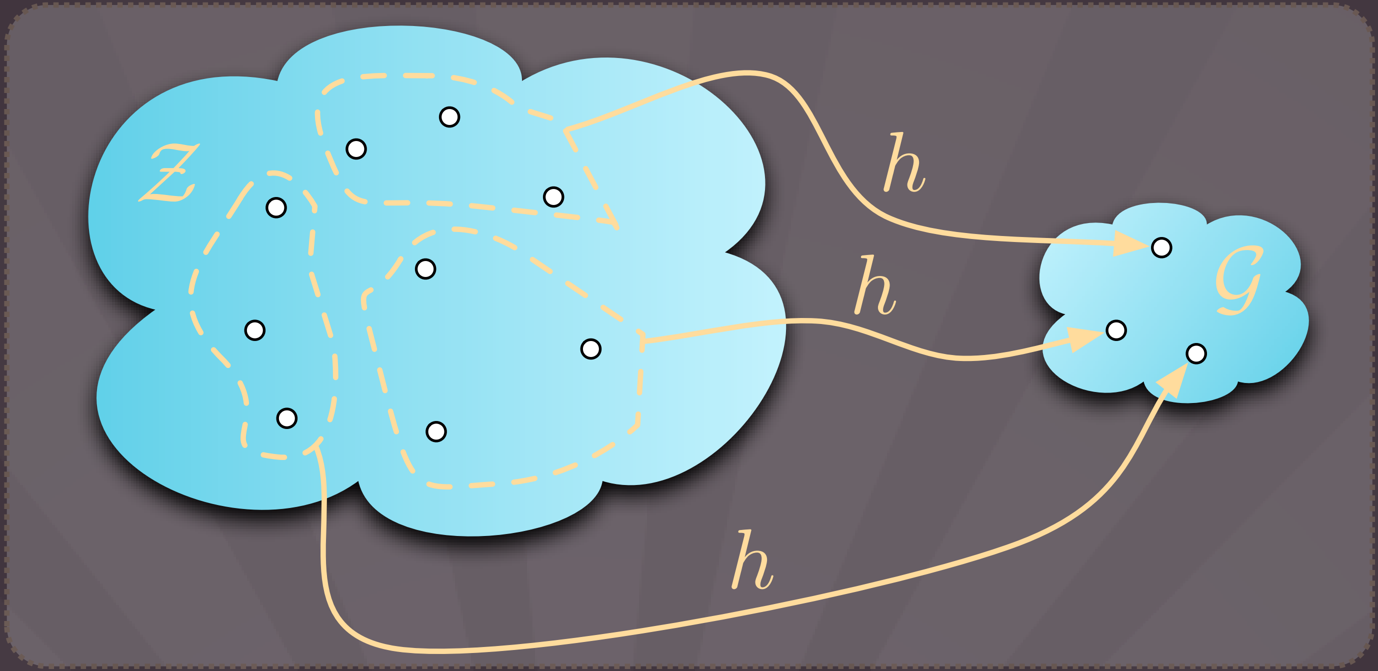
- Other links with the best distinguishers (see the paper).

Statistical Dist.

- If $|\mathcal{Z}|$ is too large, the best dist. cannot be implemented.

Statistical Dist.

- If $|\mathcal{Z}|$ is too large, the best dist. cannot be implemented.
- Possible solution: reduce the sample size using a *projection*:



- Distinguish in $\mathcal{G}$ instead of $\mathcal{Z}$
- This reduces the power of the distinguisher.

(Informal) Theorem:

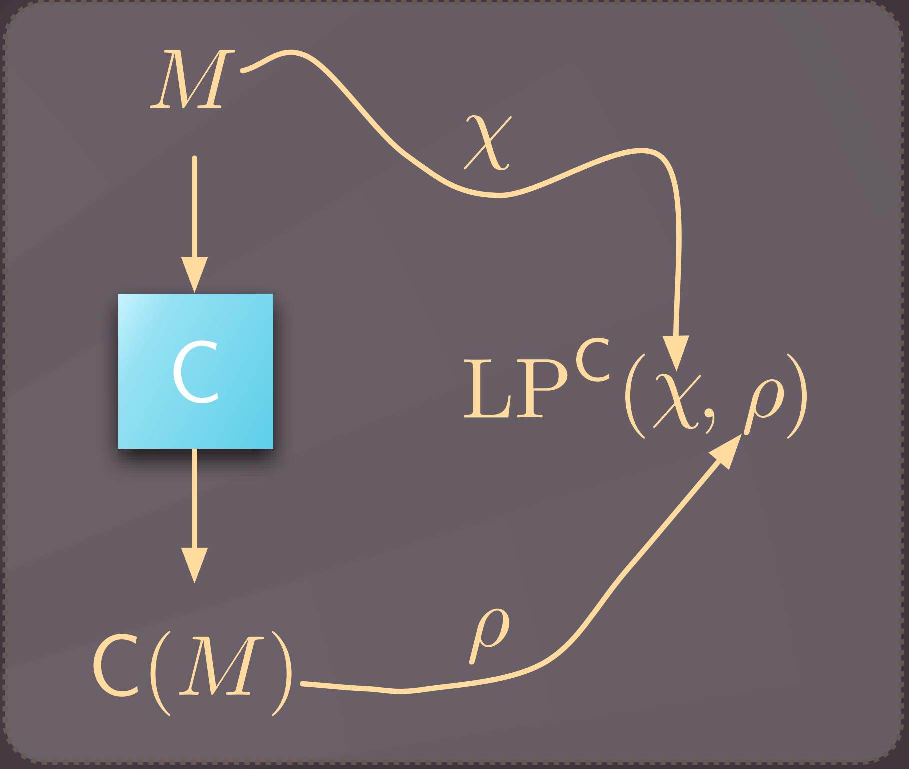
If we can efficiently distinguish using some projection, we can also do it *linearly*.

Linear Cryptanalysis of Block Ciphers

Dist. Permutations

- A simple trick allows to turn distinguishers of random sources into distinguisher of random permutations (block ciphers) [BJV04].
- All the results on random sources apply to random permutations
- In the case of linear cryptanalysis:

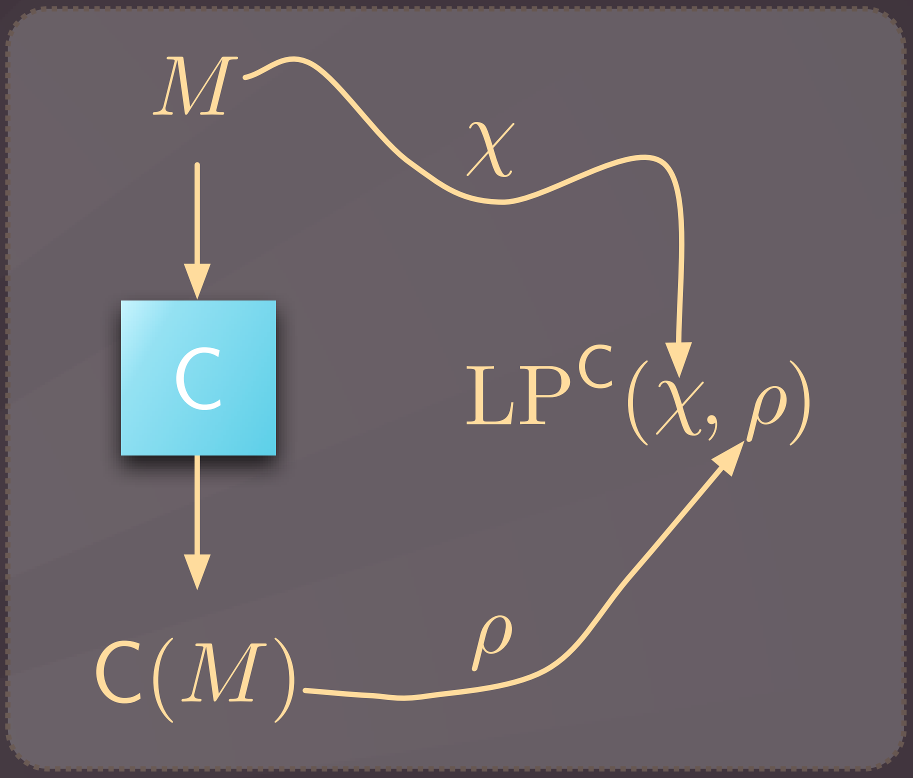
$$\mathrm{LP}^{\mathsf{C}}(\chi,\rho) = \left|\mathrm{E}_{M\in_{\mathsf{U}}\mathcal{M}}\left(\overline{\chi}(M)\rho(\mathsf{C}(M))\right)\right|^2$$



Dist. Permutations

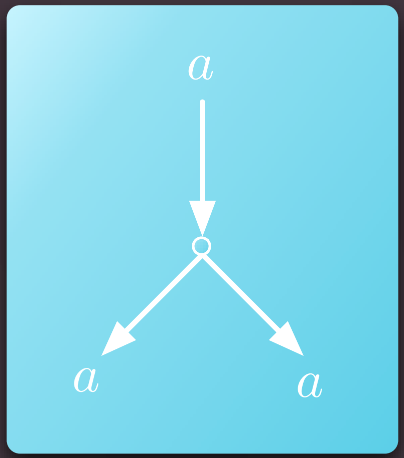
- A simple trick allows to turn distinguishers of random sources into distinguisher of random permutations (block ciphers) [BJV04].
- All the results on random sources apply to random permutations
- In the case of linear cryptanalysis:

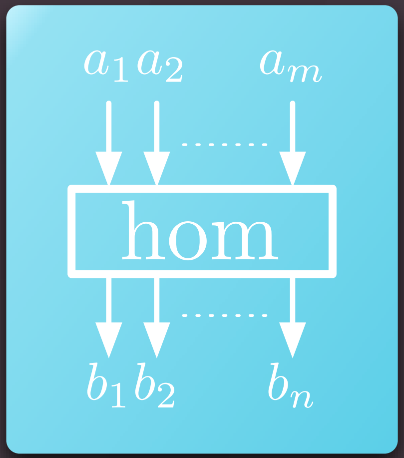
$$\mathrm{LP}^{\mathsf{C}}(\chi, \rho) = \left| \mathrm{E}_{M \in_{\mathsf{U}} \mathcal{M}} \left(\overline{\chi}(M) \rho(\mathsf{C}(M)) \right) \right|^2$$

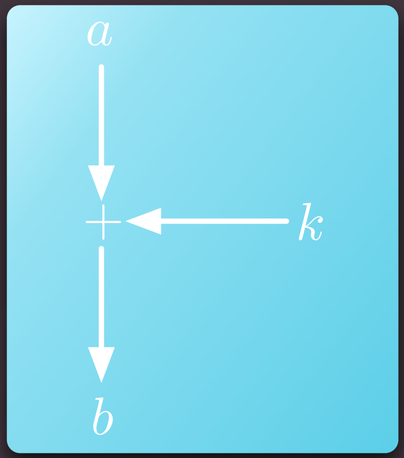


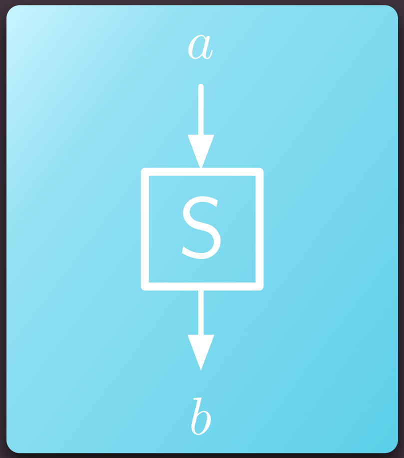
Characters ≈ Masks

Toolbox

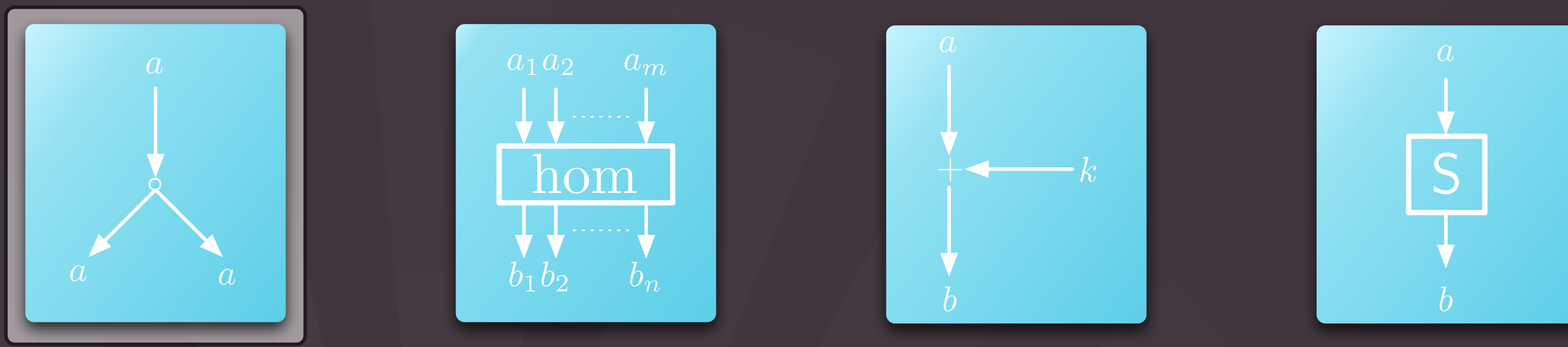






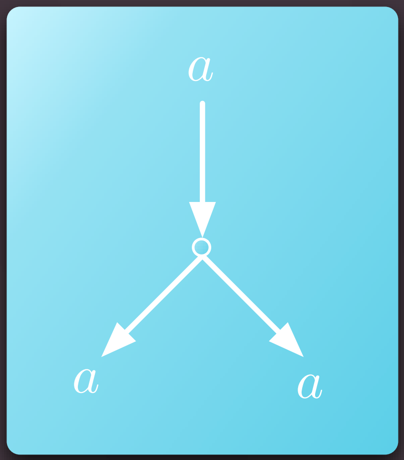


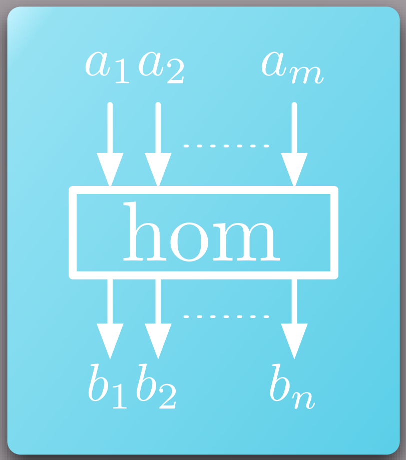
Toolbox

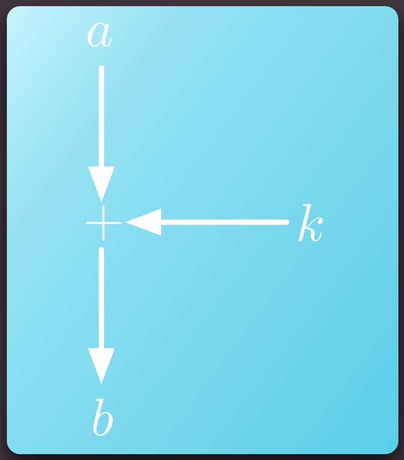


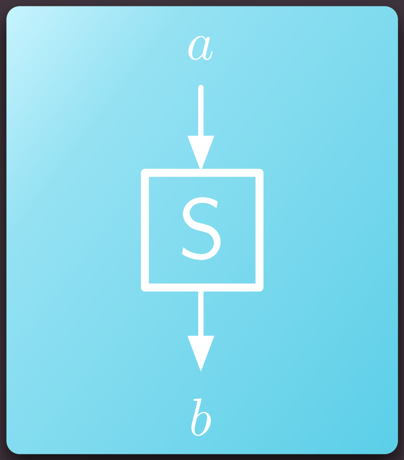
$$\mathrm{LP}(\chi_1\chi_2, \chi_1\|\chi_2) = 1$$

Toolbox





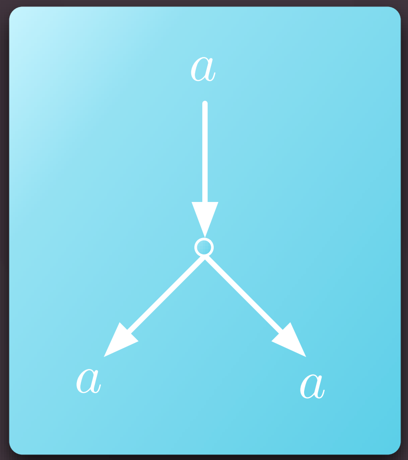


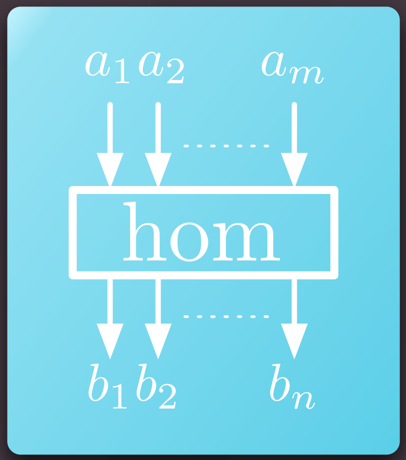


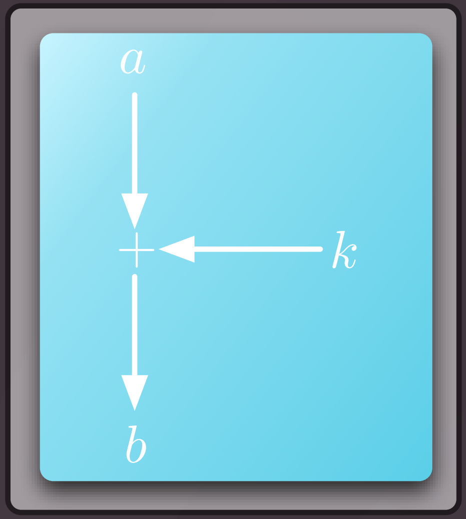
With $\chi = \chi_1 \| \cdots \| \chi_n$

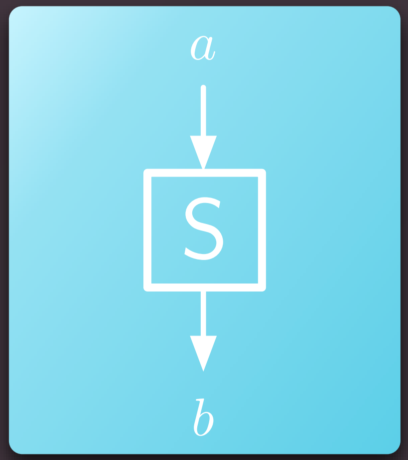
$$\mathrm{LP}(\chi \circ \mathrm{hom}, \chi) = 1$$

Toolbox



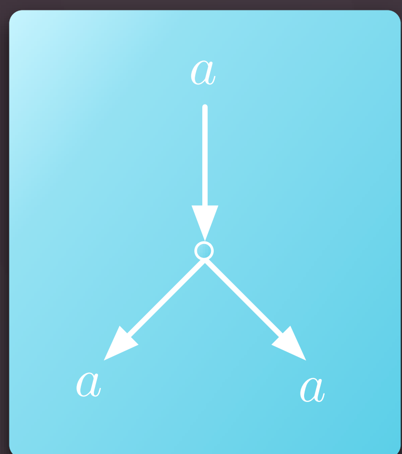


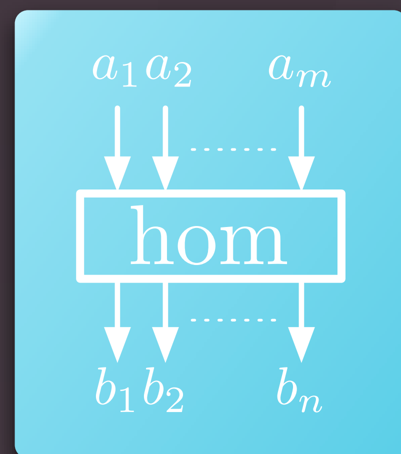


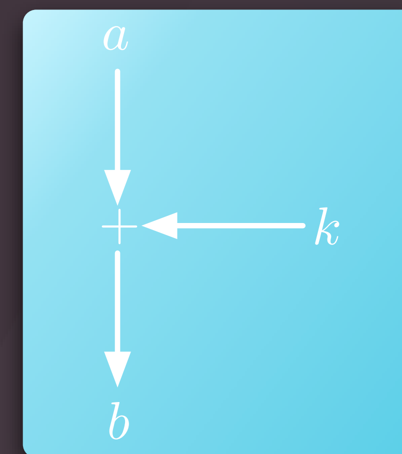


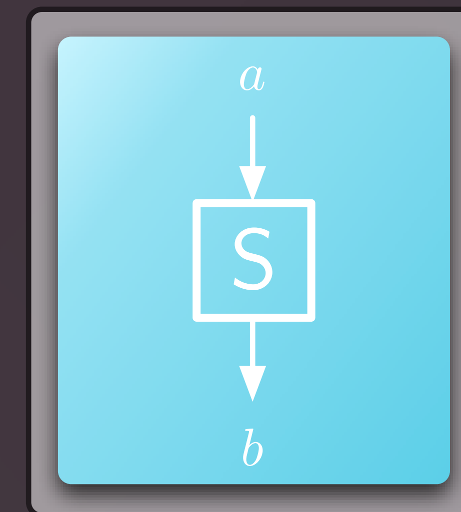
$$\mathrm{LP}(\chi, \chi) = 1$$

Toolbox









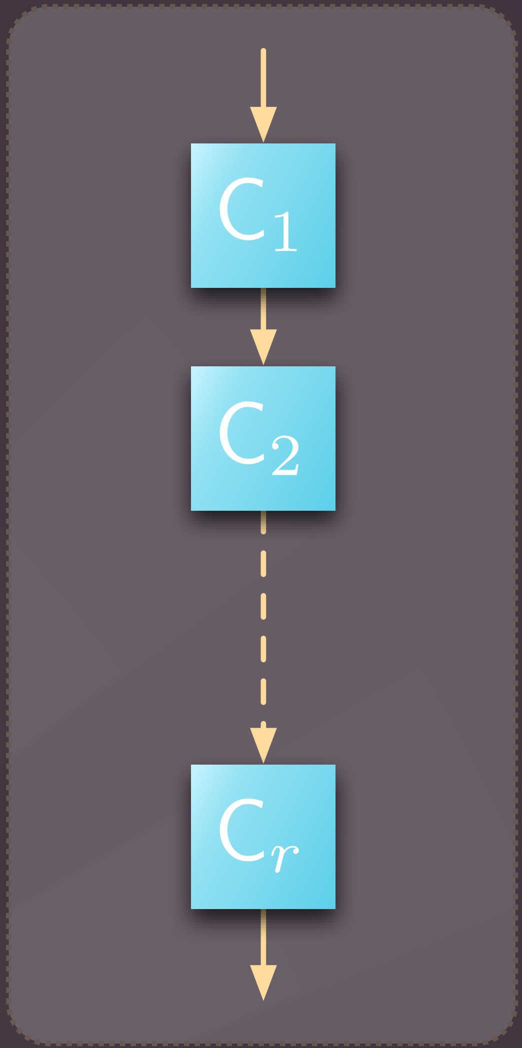
Compute $\mathrm{LP}(\chi, \rho)$ “by hand”

Toolbox

- Consider the product cipher:

$$\mathsf{C} = \mathsf{C}_r \circ \cdots \circ \mathsf{C}_1$$

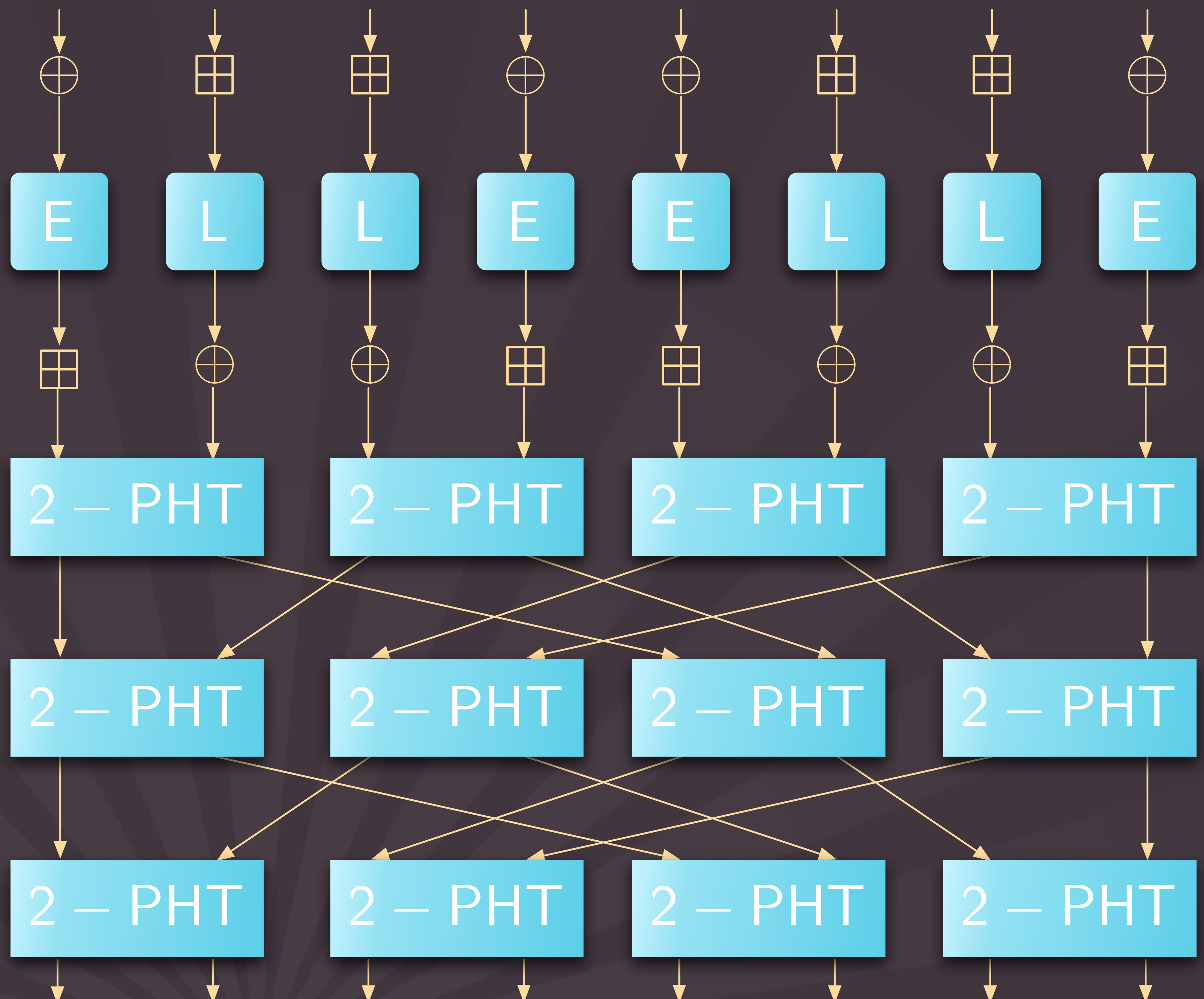
- made of independent Markov ciphers...
- we obtain Nyberg's Linear hull effect



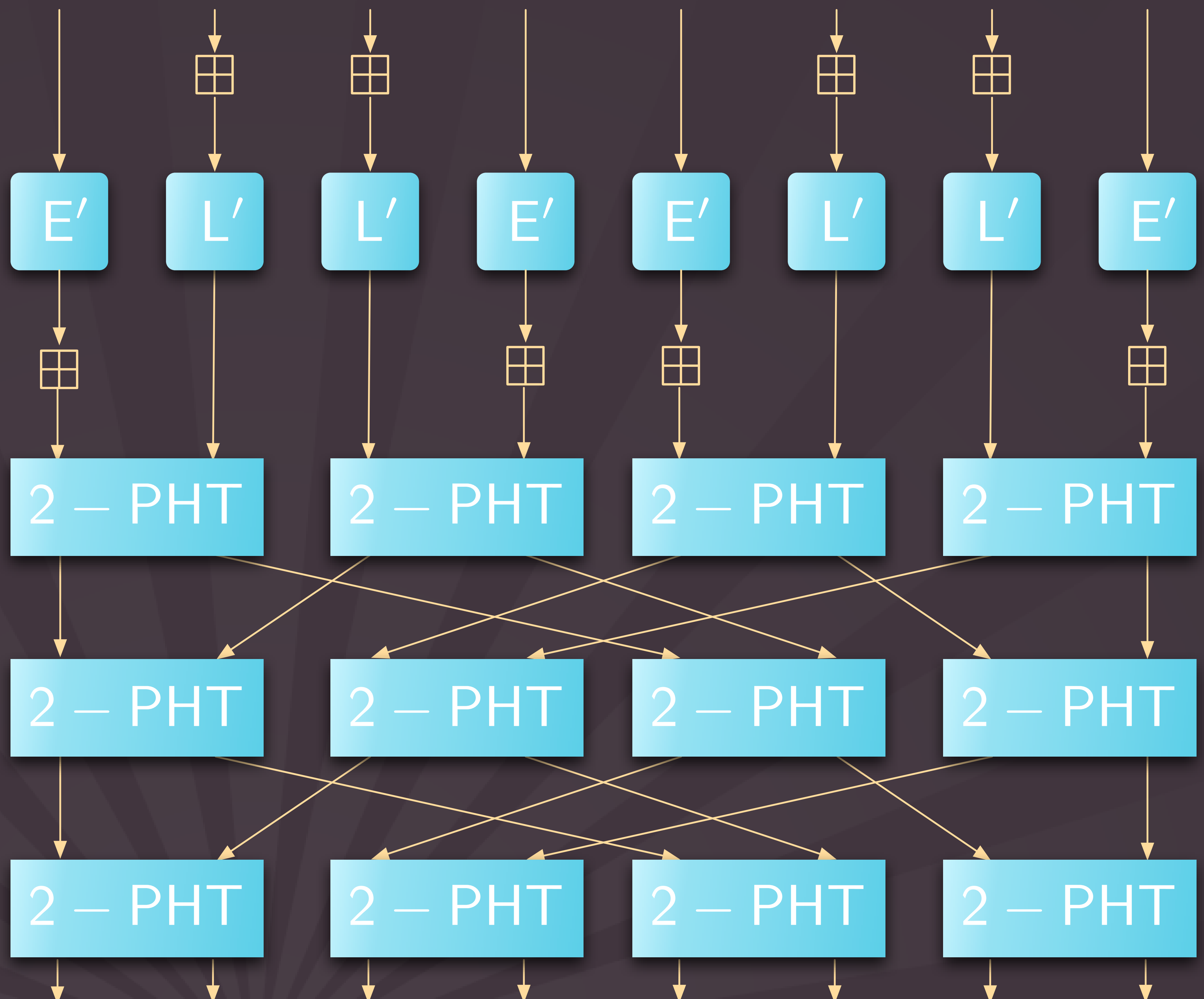
$$\mathrm{ELP}^{\mathsf{C}}(\chi_0, \chi_r) = \sum_{\chi_1} \sum_{\chi_2} \cdots \sum_{\chi_{r-1}} \prod_{i=1}^{r} \mathrm{ELP}^{\mathsf{C}_i}(\chi_{i-1}, \chi_i)$$

Applications !

An Attack on SAFER



An Attack on SAFER



An Attack on SAFER

- We attack SAFER with a $\boxplus$-linear cryptanalysis.
- We have to consider characters in $\mathbf{Z}_{2^8}^8$:

$$\begin{array}{rccc}\chi_{\mathbf{a}}: & \mathbf{Z}_{2^8}^8 & \longrightarrow & \mathbf{C}^\times \\ & \mathbf{x} = (x_1, \ldots, x_8) & \longmapsto & e^{\frac{2\pi i}{256} \sum_{\ell=1}^8 a_\ell x_\ell}\end{array}$$

An Attack on SAFER

- We attack SAFER with a $\boxplus$-linear cryptanalysis.
- We have to consider characters in $\mathbf{Z}_{2^8}^8$:

$$\begin{array}{rccc} \chi_{\mathbf{a}} : & \mathbf{Z}_{2^8}^8 & \longrightarrow & \mathbf{C}^\times \\ & \mathbf{x} = (x_1, \ldots, x_8) & \longmapsto & e^{\frac{2\pi i}{256} \sum_{\ell=1}^8 a_\ell x_\ell} \end{array}$$

- Use the toolbox to find characteristics with SAFER K/SK
- To compute the complexities of our attacks we consider several characteristics among the hull (i.e., all characteristics share the same input/output characters).
- To turn distinguishing attacks into key recovery attacks, we also take advantage of the linearity of the key schedule.

An attack on SAFER

Attack Complexities:

Nbr Rounds	Complexity
2	$2^{29}/2^{37}$
3	2^{36}
4	2^{47}
5	2^{59}

An attack on SAFER

Attack Complexities:

Nbr Rounds	Complexity
2	$2^{29}/2^{37}$
3	2^{36}
4	2^{47}
5	2^{59}

- These are *not* the best attacks on SAFER.
- Yet, SAFER was though to be secure against linear cryptanalysis.
- These attacks only work on “old” SAFER K/SK. Do they apply to SAFER+, SAFER++ ?

Other Applications

- Cryptanalysis of 9 rounds (out of 12) of TOY100 (Granboulan et al. at FSE'07).
- We use the fact that the diffusion of TOY100 is not based on a multipermutation (MDS matrix).
- Replacement toy block cipher: DEAN18
  - Structure close to that of the AES.
  - Security analysis against Linear Cryptanalysis.

Conclusion

- Generalization of Linear Cryptanalysis
- Seems to be sound:
  - Equivalent to the classical LC in the binary case
  - Link with DC
  - Linear hull effect
  - Complexity analysis
  - Link with other distinguishers (perfect, statistical,...)
- Applications, even in the binary case (SAFER)
- Open doors to new applications:
  - Specific designs
  - New attacks (think of IDEA, RCx,....)

Thank you for your attention!