Quantitative Security of Block Ciphers: Designs and Cryptanalysis Tools

Thomas Baignères

PhD Defense
November 14, 2008
Prologue
Cryptography: the Basics

Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Cryptography: the Basics

Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Cryptography: the Basics

Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Originally, cryptography aims at ensuring *confidentiality* through an insecure channel.

She’s got a dream today!

I have a *dream* today!

Bob

Alice
Originally, cryptography aims at ensuring **confidentiality** through an insecure channel.

Alice: I have a *dream* today!

Bob: She’s got a dream today!

Eve: Ha ha!! She’s got a dream today!
Originally, cryptography aims at ensuring **confidentiality** through an insecure channel.

She’s got a dream today!

I have a *dream* today!

_Eve_
Cryptography: the Basics

Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Cryptography: the Basics

Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Originally, cryptography aims at ensuring confidentiality through an insecure channel.
Originally, cryptography aims at ensuring confidentiality through an insecure channel.

I have a *dream* today!
Originally, cryptography aims at ensuring confidentiality through an insecure channel.

Alice

Bob

Cipher

Cipher

She’s got a dream today!

I have a dream today!
Originally, cryptography aims at ensuring *confidentiality* through an insecure channel.

She’s got a dream today!

I have a *dream* today!

%}@n4 ##/Wy<$ $$= ... ?????
Originally, cryptography aims at ensuring confidentiality through an insecure channel.

She’s got a dream today!

I have a dream today!

Bob

Cipher

Alice

Cipher

Eve

%]@n4 ##/Wy<$ $$=

Thomas Baignères

PhD Defense
What should we expect from the cipher?
What should we expect from the cipher?

Intuitively, turning %j@n4 ##/Wy<$ $$$= into *I have a dream today!* should be hard, except for Alice and Bob.
What should we expect from the cipher?

Intuitively, turning %j@n4 ##/Wy<$ $$= into *I have a dream today!* should be hard, except for Alice and Bob.

Fact: cryptographers are parnoïac ➔ they sometimes require more!
What should we expect from the cipher?

Intuitively, turning `%J@n4 ##/Wy<$ $$=` into *I have a dream today!* should be hard, except for Alice and Bob.

Fact: cryptographers are paranoiac they sometimes require more!

*I have a dream today!* \(\rightarrow\) Cipher \(\rightarrow\) `%J@n4 ##/Wy<$ $$=`
What should we expect from the cipher?

Intuitively, turning $^%@n4$ into $^{I \ have \ a \ dream \ today!}$ should be hard, except for Alice and Bob.

Fact: cryptographers are paranoïac they sometimes require more!

$I \ have \ a \ dream \ today! \rightarrow \ Cipher \rightarrow %j@n4$ $^{I \ have \ a \ dream \ today!}$

$2Hå$ $^\text{zę@+° ££ !`v65}$
What should we expect from the cipher?

Intuitively, turning $%j@n4 ##/Wy<$ $$=$$ into *I have a dream today!* should be hard, except for Alice and Bob.

Fact: cryptographers are paranoiac → they sometimes require more.

---

It should be hard for Eve to guess whether she’s looking at an encrypted message (ciphertext) or to pure rubbish (random string).
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

Thomas Baignères
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

Cipher
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

Cipher

Atom
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

1!$£_&\& ç%”1i87 : ;-)

Thomas Baignères
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

Cipher

| Cipher | or | ?? |

1!$£_& & ç%"1î87 : ;-)
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

- Eve wins if she guesses correctly.
The security requirements in terms of a game...

... or “Cryptographers will never grow up”.

- Eve wins if she guesses correctly.
- **Objective for the cryptographer:** make sure that Eve cannot do better than guessing correctly 50% of the time.
Outline

Distinguishers between two sources

Projection-based distinguishers between two sources

Practical Implications for block ciphers
Outline

- Distinguishers between two sources
  - Projection-based distinguishers between two sources
- Practical Implications for block ciphers
  - The game: distinguishing between two sources of randomness
  - The optimal solution
  - Complexity analysis: How many samples do we need to distinguish with a given efficiency?
Outline

Distinguishers between two sources

Projection-based distinguishers between two sources

Practical Implications for block ciphers

• What if the optimal solution cannot be implemented?

• Distinguishing in practice using compression

• Example: Generalized linear distinguisher
Outline

Distinguishers between two sources

Projection-based distinguishers between two sources

Practical Implications for block ciphers

- Cryptanalysis of SAFER K/SK
- DEAN
Outline

Distinguishers between two sources

Projection-based distinguishers between two sources

Practical Implications for block ciphers

• Cryptanalysis of SAFER K/SK
• DEAN

[BJVa04] [BSVsac07] [BVicits08]

Distinguisher between two Sources
The Game

- \( P_0 \) and \( P_1 \) are two arbitrary distributions over a finite set \( \mathcal{Z} \).
The Game

• $P_0$ and $P_1$ are two arbitrary distributions over a finite set $\mathcal{Z}$.
The Game

- $P_0$ and $P_1$ are two arbitrary distributions over a finite set $\mathcal{Z}$.

\[
\text{Adv}_A(P_0, P_1) = |\Pr_{P_0}[A(Z_1, \ldots, Z_q) = 1] - \Pr_{P_1}[A(Z_1, \ldots, Z_q) = 1]| \]

- The ability of $A$ to distinguish $P_0$ from $P_1$ is its advantage:
Example: Biased Dice
Example: Biased Dice
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]

\[ P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]

\[ P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]

\[ P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]

\[ P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \]

\[ P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
An Optimal Distinguisher

Using maximum-likelihood techniques, the $q$-limited distinguisher $\mathcal{A}^*$ which outputs 1 when by

$$D(P\|P_1) \leq D(P\|P_0)$$

can be shown to be optimal.
An Optimal Distinguisher

Using maximum-likelihood techniques, the $q$-limited distinguisher $A^*$ which outputs 1 when by

$$D(P \parallel P_1) \leq D(P \parallel P_0)$$

can be shown to be optimal.

$$D(p\|q) = \sum_{a \in Z} p[a] \log \frac{p[a]}{q[a]}$$

always non-negative, 0 iff $p=q$, infinite iff $\text{Supp}(p) \not\subseteq \text{Supp}(q)$
Data Complexity Analysis

Using the theory of types & Sanov’s theorem asymptotic data complexity of $A^*$. 
Data Complexity Analysis

Using the theory of types & Sanov’s theorem as asymptotic data complexity of $A^*$. 
Data Complexity Analysis

Using the theory of types & Sanov’s theorem \( \Rightarrow \) asymptotic data complexity of \( A^* \).

**Theorem**

Let \( P_0 \) and \( P_1 \) be two distributions s.t. \( \text{Supp}(P_0) \cup \text{Supp}(P_1) = \mathcal{Z} \). The advantage of \( A^* \) verifies

\[
1 - \text{BestAdv}_q(P_0, P_1) = 2^{-qC(P_0, P_1)}
\]

where

\[
C(P_0, P_1) = - \inf_{0<\lambda<1} \log \sum_{a \in \text{Supp}(P_0) \cap \text{Supp}(P_1)} P_0[a]^{1-\lambda} P_1[a]^{\lambda}
\]

is the Chernoff information between \( P_0 \) and \( P_1 \).
Data Complexity Analysis

Using the theory of types & Sanov’s theorem, the asymptotic data complexity of $A^*$ is given by

Let $P_0$ and $P_1$ be two distributions s.t. $\text{Supp}(P_0) \cup \text{Supp}(P_1) = \mathcal{Z}$. The advantage of $A^*$ verifies

$$1 - \text{BestAdv}_q(P_0, P_1) \geq 2^{-qC(P_0, P_1)}$$

where

$$C(P_0, P_1) = -\inf_{0<\lambda<1} \log \sum_{a \in \text{Supp}(P_0) \cap \text{Supp}(P_1)} P_0[a]^{1-\lambda} P_1[a]^\lambda$$

is the Chernoff information between $P_0$ and $P_1$.

Notation: $f(q) \sim g(q)$ means that $f(q) = g(q)e^{o(q)}$, i.e.,

$$\lim_{q \to \infty} \frac{1}{q} \log \frac{f(q)}{g(q)} = 0.$$
Data Complexity Analysis

Using the theory of types & Sanov’s theorem \(\Rightarrow\) asymptotic data complexity of \(A^*\).

**Theorem**

Let \(P_0\) and \(P_1\) be two distributions s.t. \(\text{Supp}(P_0) \cup \text{Supp}(P_1) = \mathcal{Z}\). The advantage of \(A^*\) verifies

\[
1 - \text{BestAdv}_q(P_0, P_1) \doteq 2^{-qC(P_0, P_1)}
\]

where

\[
C(P_0, P_1) \approx \frac{\|P_1 - P_0\|_2^2}{8 \ln 2}
\]

is the Chernoff information between \(P_0\) and \(P_1\).

Notation: \(f(q) \doteq g(q)\) means that \(f(q) = g(q)e^{o(q)}\), i.e., \(\lim_{q \to \infty} \frac{1}{q} \log \frac{f(q)}{g(q)} = 0\).
Data Complexity Analysis

Using the theory of types & Sanov’s theorem, asymptotic data complexity of $A^*$. 

**Theorem**

Let $P_0$ and $P_1$ be two distributions s.t. $\text{Supp}(P_0) \cup \text{Supp}(P_1) = \mathcal{Z}$. The advantage of $A^*$ verifies

$$1 - \text{BestAdv}_q(P_0, P_1) \approx 2^{-qC(P_0, P_1)}$$

where

$$C(P_0, P_1) \approx \frac{||P_1 - P_0||^2}{8 \ln 2}$$

is the Chernoff information between $P_0$ and $P_1$. 
Data Complexity Analysis

Using the **theory of types & Sanov’s theorem** asymptotic data complexity of $\mathcal{A}^*$. 

**Theorem**

Let $P_0$ and $P_1$ be two distributions s.t. $\text{Supp}(P_0) \cup \text{Supp}(P_1) = Z$. The advantage of $A^*$ verifies

$$\text{BestAdv}(P_0, P_1) \leq 1 - C(P_0, P_1)$$

where $C(P_0, P_1)$ is the Chernoff information between $P_0$ and $P_1$.

Heuristic: $q \approx 1/C(P_0, P_1)$ allows $A^*$ to reach a non-negligible advantage.
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \quad P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \quad P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]

\[ C(P_0, P_1) = \max_{0 < \lambda < 1} \log \left( \frac{6}{2^\lambda + 4} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \quad P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6} \right) \]

\[ C(P_0, P_1) = \max_{0 < \lambda < 1} \log \left( \frac{6}{2^\lambda + 4} \right) \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \quad P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6} \right) \]

\[ C(P_0, P_1) = \max_{0 < \lambda < 1} \log \left( \frac{6}{2\lambda + 4} \right) \]

\[ \approx 0.263 \]
Example: Biased Dice

\[ P_0 = \left( \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right) \quad \text{and} \quad P_1 = \left( \frac{1}{6}, \frac{1}{6}, \frac{2}{6}, 0, \frac{1}{6}, \frac{1}{6} \right) \]

\[ C(P_0, P_1) = \max_{0 < \lambda < 1} \log \left( \frac{6}{2\lambda + 4} \right) \approx 0.263 \]

\[ \approx \frac{1}{0.263} \approx 3.8 \text{ queries (rolls) are sufficient to distinguish one dice from the other.} \]

\[ \text{This is the proof that all this theory has a practical application...} \]
Example: Biased Coin

\[ P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2}(1 - \epsilon), \frac{1}{2}(1 + \epsilon) \right) \]
Example: Biased Coin

\[ P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2}(1 - \epsilon), \frac{1}{2}(1 + \epsilon) \right) \]
Example: Biased Coin

\[ P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2} (1 - \epsilon), \frac{1}{2} (1 + \epsilon) \right) \]

\[ C(P_0, P_1) = -\inf_{0 < \lambda < 1} \log \frac{1}{2} ((1 - \epsilon)^\lambda + (1 + \epsilon)^\lambda) \]
Example: Biased Coin

\[ P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2} (1 - \epsilon), \frac{1}{2} (1 + \epsilon) \right) \]

\[ C(P_0, P_1) = - \inf_{0 < \lambda < 1} \log \frac{1}{2} \left( (1 - \epsilon)\lambda + (1 + \epsilon)\lambda \right) \]
Example: Biased Coin

\[ P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2}(1 - \epsilon), \frac{1}{2}(1 + \epsilon) \right) \]

\[ C(P_0, P_1) = - \inf_{0 < \lambda < 1} \log \frac{1}{2} \left( (1 - \epsilon)^\lambda + (1 + \epsilon)^\lambda \right) \]

Minimum reached for \( \lambda \approx \frac{1}{2} \)

\[ C(P_0, P_1) \approx - \log \left( 1 - \frac{\epsilon^2}{8} \right) \approx \frac{\epsilon^2}{8 \ln 2} \]
Example: Biased Coin

$$P_0 = \left( \frac{1}{2}, \frac{1}{2} \right) \quad P_1 = \left( \frac{1}{2}(1 - \epsilon), \frac{1}{2}(1 + \epsilon) \right)$$

$$C(P_0, P_1) = -\inf_{0 < \lambda < 1} \log \frac{1}{2} \left( (1 - \epsilon)\lambda + (1 + \epsilon)\lambda \right)$$

Minimum reached for $\lambda \approx \frac{1}{2}$

$$C(P_0, P_1) \approx -\log \left( 1 - \frac{\epsilon^2}{8} \right) \approx \frac{\epsilon^2}{8 \ln 2}$$

$$q \approx \frac{8 \ln 2}{\epsilon^2}$$ allow to reach a non-negligible advantage.
Possible Extensions

- Case where the distributions are “close” to each other
- Case where one of the hypotheses is composite
- Case where one of the two distributions is unknown
- etc.

Projection Based Distinguishers
On the Need for Projection-Based Distinguishers

- If $|\mathcal{Z}|$ is too large, the best distinguisher cannot be implemented.
On the Need for Projection-Based Distinguishers

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

\[
\text{Distinguish in } \mathcal{G} \text{ instead of } \mathcal{Z}.
\]

\[\text{This reduces the power of the distinguisher.}\]
Example: Linear Distinguishers

- \( Z = \{0, 1\}^n \)  \( G = \{0, 1\} \)  \( P_0 = U \)  \( P_1 = P \)  \( h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n \)

- This is a **linear distinguisher** based on the mask \( a \).
Example: Linear Distinguishers

- \( Z = \{0, 1\}^n \) \( \mathcal{G} = \{0, 1\} \) \( P_0 = U \) \( P_1 = P \) \( h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n \)

- This is a linear distinguisher based on the mask \( a \).
**Example: Linear Distinguishers**

- \( Z = \{0, 1\}^n \) \( G = \{0, 1\} \) \( P_0 = U \) \( P_1 = P \) \( h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n \)

- This is a **linear distinguisher** based on the **mask** \( a \).

- By implementing the optimal strategy (after the linear compression), the advantage of this linear distinguisher verifies:

\[
1 - \text{Adv}(U, P) \doteq 2^{-q_{C(U, P)}}
\]
Example: Linear Distinguishers

- \( \mathcal{Z} = \{0, 1\}^n \)  \( \mathcal{G} = \{0, 1\} \)  \( \mathcal{P}_0 = \mathcal{U} \)  \( \mathcal{P}_1 = \mathcal{P} \)  \( h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n \)

- This is a linear distinguisher based on the mask \( a \).

- By implementing the optimal strategy (after the linear compression), the advantage of this linear distinguisher verifies:

\[
1 - \text{Adv}(\mathcal{U}, \mathcal{P}) = 2^{-q(C(\overline{\mathcal{U}}, \overline{\mathcal{P}}))}
\]

\[
a \cdot Z \sim \overline{\mathcal{P}} \iff Z \sim \mathcal{P}
\]

\[
a \cdot Z \sim \overline{\mathcal{U}} \iff Z \sim \mathcal{U}
\]
Example: Linear Distinguishers

- $Z = \{0, 1\}^n$, $G = \{0, 1\}$, $P_0 = U$, $P_1 = P$, $h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n$

- This is a linear distinguisher based on the mask $a$.

- By implementing the optimal strategy (after the linear compression), the advantage of this linear distinguisher verifies:

  \[
  1 - \text{Adv}(U, P) \overset{!}{=} 2^{-q_{C(U,P)}}
  \]

  $a \cdot Z \sim \overline{P} \iff Z \sim P$
  
  $a \cdot Z \sim \overline{U} \iff Z \sim U$

- Definition: linear probability of $P$: 
  
  \[
  \text{LP}_a(P) = \left( E_P \left( (-1)^{a \cdot Z} \right) \right)^2
  \]
Example: Linear Distinguishers

- \( Z = \{0, 1\}^n \quad G = \{0, 1\} \quad P_0 = U \quad P_1 = P \quad h(Z) = a \cdot Z = a_1 Z_1 \oplus \cdots \oplus a_n Z_n \)

- This is a linear distinguisher based on the mask \( a \).

- By implementing the optimal strategy (after the linear compression), the advantage of this linear distinguisher verifies:

\[
1 - \text{Adv}(U, P) \doteq 2^{-qC(U, P)}
\]

\[
a \cdot Z \sim \overline{P} \iff Z \sim P
\]

\[
a \cdot Z \sim \overline{U} \iff Z \sim U
\]

- Definition: linear probability of \( P \):

\[
LP_a(P) = \left( E_P \left( (-1)^{a \cdot Z} \right) \right)^2
\]

- Roughly:

\[
C(U, P) \approx \frac{LP_a(P)}{8 \ln 2}
\]

\[
q \approx \frac{8 \ln 2}{LP_a(P)}
\]

are enough (well known...)

Thomas Baignères

PhD Defense
Extending the Notion of Linear Probability

• The previous example only works for sets of the form $Z = \{0, 1\}^n$.

• We at least need to generalize the notion of linear probability to arbitrary sets.
Extending the Notion of Linear Probability

• The previous example only works for sets of the form $\mathcal{Z} = \{0, 1\}^n$.

• We at least need to generalize the notion of linear probability to arbitrary sets.
Extending the Notion of Linear Probability

• The previous example only works for sets of the form $\mathcal{Z} = \{0, 1\}^n$.

• We at least need to generalize the notion of linear probability to arbitrary sets.

Definition

The linear probability of $P$ over the group $\mathcal{Z}$ with respect to the character $\chi$ is

$$\text{LP}_\chi(P) = |\mathbb{E}_P(\chi(Z))|^2$$
Extending the Notion of Linear Probability

• The previous example only works for sets of the form $\mathcal{Z} = \{0, 1\}^n$.

• We at least need to generalize the notion of linear probability to arbitrary sets.

**Definition**

The linear probability of $P$ over the group $\mathcal{Z}$ with respect to the character $\chi$ is

$$LP_\chi(P) = |E_\chi(P)\chi(\mathcal{Z})|^2$$

• A character of $\mathcal{Z}$ is a homomorphism $\chi : \mathcal{Z} \rightarrow \mathbb{C}^\times$

• Example: when $\mathcal{Z} = \{0, 1\}^n$ we have $\chi(a) = (-1)^{u \cdot a}$ for some $u$
Extending the Notion of Linear Probability

- The previous example only works for sets of the form $\mathcal{Z} = \{0, 1\}^n$.
- We at least need to generalize the notion of linear probability to arbitrary sets.

**Definition**

The linear probability of $P$ over the group $\mathcal{Z}$ with respect to the character $\chi$ is

$$LP_{\chi}(P) = |E_{\chi}(\chi(\mathcal{Z}))|^2$$

- A character of $\mathcal{Z}$ is a homomorphism $\chi : \mathcal{Z} \rightarrow \mathbb{C}^\times$
- Example: when $\mathcal{Z} = \{0, 1\}^n$ we have $\chi(a) = (-1)^{u \cdot a}$ for some $u$
- Consequence: when $\mathcal{Z} = \{0, 1\}^n$ this new definition corresponds to the old one!
Lin. Dist. for Sources overs Arbitrary Sets

We have wonderful lemma...
Lin. Dist. for Sources overs Arbitrary Sets

We have wonderful lemma...

**Lemma 7.5** Let $P_0$ be the uniform distribution on a finite subgroup $H$ of $\mathbb{C}^\times$ of order $d$. Let $D = \{P_u : u \in H\}$ be a set of $d$ distributions on $H$ defined by (7.10). The $q$-limited distinguisher between the null hypothesis $H_0 : P = P_0$ and the alternate hypothesis $H_1 : P \in D$ defined by the distribution acceptance region $\Pi^*_q = \Pi^* \cap P_q$, where

$$\Pi^* = \left\{ P \in \mathcal{P} : \|P\|_\infty \geq \frac{\log(1 - \epsilon)}{\log(1 - \epsilon) - \log(1 + (d-1)\epsilon)} \right\},$$

is asymptotically optimal and its advantage $\text{BestAdv}_q$ is such that

$$1 - \text{BestAdv}_q(H_0, H_1) \geq 2^{q \inf_{\lambda > 0} \log \frac{1}{d}((1+(d-1)\epsilon)\lambda+(d-1)(1-\epsilon)^d)}.$$
We have wonderful lemma...

**Lemma 7.5** Let $P_0$ be the uniform distribution on a finite subgroup $H$ of $\mathbb{C}^\times$ of order $q$. Let $D = \{P_u : u \in H\}$ be a set of distributions on $\mathbb{C}^\times$ defined by (7.10). The distribution over $H$ is balanced when $q = P_0$ and the alternate hypothesis $H_1 : P \in D$ defined by the best distinguishing acceptance region $\Pi^*_q = \Pi^*_r \cap \Pi^*_q$ is asymptotically optimal and its advantage $\text{BestAdv}_q(H_0, H_1)$ needs $q \approx \frac{8 \ln 2}{(d - 1)(1 - (1 - \epsilon)\lambda)}$ to reach a good advantage.

Which shows how to use the generalized LP to build a linear distinguisher over arbitrary sets...
Practical Implications for Block Ciphers
Applications on SAFER K/SK

• We attack SAFER with a \( \oplus \)-linear cryptanalysis.

• Use the toolbox to find characteristics within SAFER K/SK.

• To compute the complexities 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.
Applications on SAFER K/SK

• We attack SAFER with a $\boxplus$-linear cryptanalysis.

• Use the toolbox to find characteristics within SAFER K/SK.

• To compute the complexities 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.

<table>
<thead>
<tr>
<th>Nbr Rounds</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$2^{23}/2^{31}$</td>
</tr>
<tr>
<td>3</td>
<td>$2^{38}$</td>
</tr>
<tr>
<td>4</td>
<td>$2^{49}$</td>
</tr>
<tr>
<td>5</td>
<td>$2^{56}$</td>
</tr>
</tbody>
</table>
Other Applications

- Two new Digital Encryption Algorithm for Numbers (based on the AES): DEAN18 and DEAN27 which respectively encrypts blocks made of 18 and 27 decimal digits.

- Resistance against our generalization of linear cryptanalysis.

- New attacks on TOY100 (toy cipher that encrypts blocks of 32 decimal digits).

- Break 9 (10 ?) rounds out of 12.
Part II: Designs and Security Proofs
Outline

- Block Ciphers
- Dial C for Cipher
- KFC: the Krazy Feistel Cipher
Outline

Block Ciphers

Dial C for Cipher

KFC: the Krazy Feistel Cipher

• The Luby-Rackoff Model

• Vaudenay’s decorrelation theory
Outline

Block Ciphers

Dial C for Cipher

KFC: the Krazy Feistel Cipher

Round 1

Round 2

Round 3

Round 10
Outline

Block Ciphers

Dial C for Cipher

KFC: the Krazy Feistel Cipher
Outline

Block Ciphers

Dial C for Cipher

KFC: the Krazy Feistel Cipher

[BVsac05]  [BFsac06]  [BFa06]
Part II: Designs and Security Proofs

_block_ciphers_
A block cipher on a finite set is a family of permutations on that set, indexed by a parameter called the key.
A Typical Iterated Block Cipher

• A block cipher on a finite set is a family of permutations on that set, indexed by a parameter called the key.

• Such a cipher is usually iterated, i.e., made of several rounds.

• Each round is parameterized by a key derived from the main secret key by means of a Key Schedule.
A Typical Iterated Block Cipher

- A block cipher on a finite set is a family of permutations on that set, indexed by a parameter called the key.

- Such a cipher is usually iterated, i.e., made of several rounds.

- Each round is parameterized by a key derived from the main secret key by means of a Key Schedule.

- Usually, the rounds all share the same design, e.g., a round key addition followed by a fixed (nonlinear) transformation.
What Should we Expect from a Block Cipher?

It should be fast and secure!
What Should we Expect from a Block Cipher?

It should be fast and secure!
What Should we Expect from a Block Cipher?

It should be fast and secure!

\[ P_1, P_2, \ldots, P_q \rightarrow P_1^*, P_2^*, \ldots, P_q^* \rightarrow C_1, C_2, \ldots, C_q \]

I’m bad
What Should we Expect from a Block Cipher?

It should be **fast** and **secure!**

\[ P_1, P_2, \ldots, P_q \rightarrow C_1, C_2, \ldots, C_q \]

My guess is...
We consider a $q$-limited adversary $\mathcal{A}$ in the Luby-Rackoff Model:
The Luby-Rackoff Model

We consider a $q$-limited adversary $A$ in the Luby-Rackoff Model:

The block cipher $C$ is secure if the advantage of $A$ is negligible for all $A$’s.

Advantage of the $q$-limited adversary $A$ between $C$ and $C^*$

$\text{Adv}_A(C, C^*) = \left| \Pr[A(C) = 1] - \Pr[A(C^*) = 1] \right|$
The Luby-Rackoff Model

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

$\mathcal{A}$ is non-adaptive if the $q$ plaintexts are chosen “at once”.

$\mathcal{O}(p_1), \ldots, \mathcal{O}(p_q) \\ p_1, \ldots, p_q \\ C \text{ or } C^*$

$\mathcal{A}$

$\text{0 or 1}$
We consider a $q$-limited adversary $\mathcal{A}$ in the Luby-Rackoff Model:

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

---

**The Luby-Rackoff Model**

$\mathcal{O}$ or $\mathcal{C}^*$

$\mathcal{O}(p_1)$

$\mathcal{O}(p_q)$

$p_1$

$p_q$

$\mathcal{A}$

0 or 1
Computing $\text{Adv}_A(C, C^*)$

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

$$\max_A \text{Adv}_A(C, C^*) = \frac{1}{2} \| [C]^q - [C^*]^q \|$$

$[C]^q = |M|^q$

$|M|^q = 2^{128\cdot q}$ for a 128-bits block cipher
Computing $\text{Adv}_A(C, C^*)$

- Computing the advantage is not a trivial task in general.

- Possible solution: use Vaudenay’s Decorrelation Theory.

$$\max_A \text{Adv}_A(C, C^*) = \frac{1}{2} \| [C]^q - [C^*]^q \|$$

$[C]^q = |M|^q$  $|M|^q = 2^{128 \cdot q}$ for a 128-bits block cipher
Tricks for Computing $\text{Adv}_A(C, C^*)$

To deal with the size of the distribution matrices:

$$[C_2 \circ C_1]^q = [C_1]^q \times [C_2]^q$$
Tricks for Computing $\text{Adv}_A(C, C^*)$

To deal with the size of the distribution matrices:

$$[C_2 \circ C_1]^q = [C_1]^q \times [C_2]^q$$

Take advantage of the symmetries of the block cipher in order to compute the distribution matrix of each round.
Part II: Designs and Security Proofs

Dial C for Cipher
Description of $C$

$C$ corresponds to the AES where “addRoundKeys $\rightarrow$ SubBytes” is replaced by mutually independent random permutations.

AES

![AES Diagram]

Thomas Baignères

PhD Defense
Description of $\mathcal{C}$

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

- $\mathcal{C}$ is made of 9 identical rounds, followed by a layer of substitution boxes.
- $\mathcal{C}$ uses $16 \cdot 10 = 160$ mutually independent random 8-bits substitution boxes.
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:
Computing $[\mathbf{C}]^2$

We consider a version of $\mathbf{C}$ reduced to 3 rounds:

\[
\begin{align*}
S_1^{(1)} & \quad S_2^{(1)} & \quad S_3^{(1)} & \quad S_{16}^{(1)} \\
S_1^{(2)} & \quad S_2^{(2)} & \quad S_3^{(2)} & \quad S_{16}^{(2)} \\
S_1^{(3)} & \quad S_2^{(3)} & \quad S_3^{(3)} & \quad S_{16}^{(3)}
\end{align*}
\]

\[
\begin{align*}
\{ & \; [S]^2 \} \\
\{ & \; [L]^2 \} \\
\{ & \; [S]^2 \} \\
\{ & \; [L]^2 \} \\
\{ & \; [S]^2 \}
\end{align*}
\]
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:

$$[C]^2 = [S]^2 \times [L]^2 \times [S]^2 \times [L]^2 \times [S]^2$$
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:

$[C]^2 \times [L]^2 \times [S]^2 \times [L]^2 \times [S]^2$

$[S]^2 = \text{PS} \times \text{SP}$
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:

$$[C]^2 = [S]^2 \times [L]^2 \times [S]^2 \times [L]^2 \times [S]^2$$
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:

$[C]^2 = [S]^2 \times [L]^2 \times [S]^2 \times [L]^2 \times [S]^2$
Computing $[C]^2$

We consider a version of $C$ reduced to 3 rounds:

$$[C]^2 = [S]^2 \times [L]^2 \times [S]^2 \times [L]^2 \times [S]^2$$

$$= \begin{array}{c}
\text{PS} \\
\text{SP}
\end{array} \times \begin{array}{c}
[L]^2 \\
\text{PS}
\end{array} \times \begin{array}{c}
\text{SP}
\end{array} \times \begin{array}{c}
[L]^2 \\
\text{PS}
\end{array} \times \begin{array}{c}
\text{SP}
\end{array}$$

$$= \begin{array}{c}
\text{PS} \\
\text{L} \times \text{L} \times \text{SP}
\end{array}$$
Computing $\text{Adv}_A(C, C^*)$

For a $r$-round version of $C$ we have:

$$[C]^2 = PS \times (\overline{L})^{r-1} \times SP$$

where $\overline{L}$ is a $2^{16} \times 2^{16}$ matrix.
Computing $\text{Adv}_A(C, C^*)$

For a $r$-round version of $C$ we have:

$$[C]^2 = PS \times (\overline{L})^{r-1} \times SP$$

where $\overline{L}$ is a $2^{16} \times 2^{16}$ matrix.

$$\max_{A} \text{Adv}_A(C, C^*) = \frac{1}{2} ||(\overline{L})^{r-1} - C^*||_\infty$$
Computing $\text{Adv}_A(\mathbf{C}, \mathbf{C}^*)$

For a $r$-round version of $\mathbf{C}$ we have:

$$[\mathbf{C}]^2 = \mathbf{PS} \times (\mathbf{L})^{r-1} \times \mathbf{SP}$$

where $\mathbf{L}$ is a $2^{16} \times 2^{16}$ matrix.

$$\max_A \text{Adv}_A(\mathbf{C}, \mathbf{C}^*) = \frac{1}{2} \| (\mathbf{L})^{r-1} - \mathbf{C}^* \|_\infty$$

Can we reduce the computational complexity even further?

Yes! But the diffusion has to be chosen with care...
Computing $\text{Adv}_{\mathcal{A}}(C, C^*)$

For a $r$-round version of $C$ we have:

$$[C]^2 = PS \times (\overline{L})^{r-1} \times SP$$

where $\overline{L}$ is a $2^{16} \times 2^{16}$ matrix.

$$\max_{\mathcal{A}} \text{Adv}_{\mathcal{A}}(C, C^*) = \frac{1}{2} \|\|(\overline{L})^{r-1} - C^*\|\|_{\infty}$$

Can we reduce the computational complexity even further?

Yes! But the diffusion has to be chosen with care...

$$\max_{\mathcal{A}} \text{Adv}_{\mathcal{A}}(C, C^*) = \frac{1}{2} \|\|(\overline{L} \times W)^{r-2} \times \overline{L} - C^*\|\|_{\infty}$$

Computing the advantage of the best distinguisher (either adaptive or not) only requires operations on $625 \times 625$ matrices (instead of $2^{256} \times 2^{256}$ initially).
### Values of $\text{Adv}_A(C, C^*)$

<table>
<thead>
<tr>
<th>$r$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv($C, C^*$)</td>
<td>1</td>
<td>1</td>
<td>$2^{-4.0}$</td>
<td>$2^{-23.4}$</td>
<td>$2^{-45.8}$</td>
<td>$2^{-71.0}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$r$</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv($C, C^*$)</td>
<td>$2^{-126.3}$</td>
<td>$2^{-141.3}$</td>
<td>$2^{-163.1}$</td>
<td>$2^{-185.5}$</td>
<td>$2^{-210.8}$</td>
<td>$2^{-238.9}$</td>
</tr>
</tbody>
</table>
Values of $\text{Adv}_A(C, C^*)$

<table>
<thead>
<tr>
<th>$r$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv($C, C^*$)</td>
<td>1</td>
<td>1</td>
<td>$2^{-4.0}$</td>
<td>$2^{-23.4}$</td>
<td>$2^{-45.8}$</td>
<td>$2^{-71.0}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$r$</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adv($C, C^*$)</td>
<td>$2^{-126.3}$</td>
<td>$2^{-141.3}$</td>
<td>$2^{-163.1}$</td>
<td>$2^{-185.5}$</td>
<td>$2^{-210.8}$</td>
<td>$2^{-238.9}$</td>
</tr>
</tbody>
</table>

7 rounds of $C$ are enough to obtain provable security against 2-limited adversaries
Part II: Designs and Security Proofs

KFC: the Krazy Feistel Cipher
What about Higher Orders?

We did not manage to prove the security of $C$ against higher $q$-limited adversaries for $q > 2$. 
What about Higher Orders?

We did not manage to prove the security of $C$ against higher $q$-limited adversaries for $q > 2$.

Idea: try to bound the advantage of the best $q$-limited adversary by that of the best $(q-1)$-limited adversary.

Perfectly random permutation vs. Perfectly random function

- $S^*$: different inputs, different outputs
- $F^*$: different inputs, independent outputs
Rand. Permutations vs. Rand. Functions

2 correlated inputs distinct on each box input

2 correlated outputs

2 independent outputs
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 the order 70 (for an unreasonable set of parameters).

• The bounds are 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 the order 70 (for an unreasonable set of parameters).

• The bounds are not tight at all it is certainly possible to improve our results.
Conclusion
“[…] the methodology of provable security has become unavoidable in designing and evaluating new schemes”

[JSe03]
“[...] the methodology of provable security has become unavoidable in designing and evaluating new schemes”

[JSe03]
“[...] the methodology of provable security has become unavoidable in designing and evaluating new schemes”

[JSe03]

We hope to have made a significant step towards its extension to block ciphers!
Thank you for your attention! 😊
Publications

[BVicits08] *The Complexity of Distinguishing Distributions*
Joint work with Serge Vaudenay
Published in the proceedings of ICITS 08 (Calgary, Canada)

[BSVsac07] *Linear Cryptanalysis of Non Binary Ciphers (with an application to SAFER)*
Joint work with Jacques Stern & Serge Vaudenay
Published in the proceedings of SAC 07 (Ottawa, Canada)

[BFa06] *KFC - The Krazy Feistel Cipher*
Joint work with Matthieu Finiasz
Published in the proceedings of Asiacrypt 06 (Shangai, China)

[BFsac06] *Dial C for Cipher*
Joint work with Matthieu Finiasz
Published in the proceedings of SAC 06 (Montreal, Canada)

[BVsac05] *Proving the Security of the AES Substitution-Permutation Network*
Joint work with Serge Vaudenay
Published in the proceedings of SAC 05 (Kingston, Canada)

[BJVa04] *How Far Can We Go Beyond Linear Cryptanalysis?*
Joint work with Pascal Junod & Serge Vaudenay
Published in the proceedings of Asiacrypt 04 (Jeju Island, Korea)