Practical Decorrelation

La théorie de la décorrélation en pratique

Matthieu Finiasz

Mise au point par S. Vaudenay [Journal of Cryptology 03]

- Permet de prouver des résultats de sécurité sur des chiffrements par blocs.
- Utilisé pour prouver la sécurité de DFC contre :
  - la cryptanalyse linéaire,
  - la cryptanalyse différentielle.

Modèle d'attaque :

L'attaquant doit distinguer un chiffrement par bloc C (utilisant une clef inconnue) du chiffrement parfait $C^\star$ en n'ayant accès qu'à d couples de clair/chiffré.

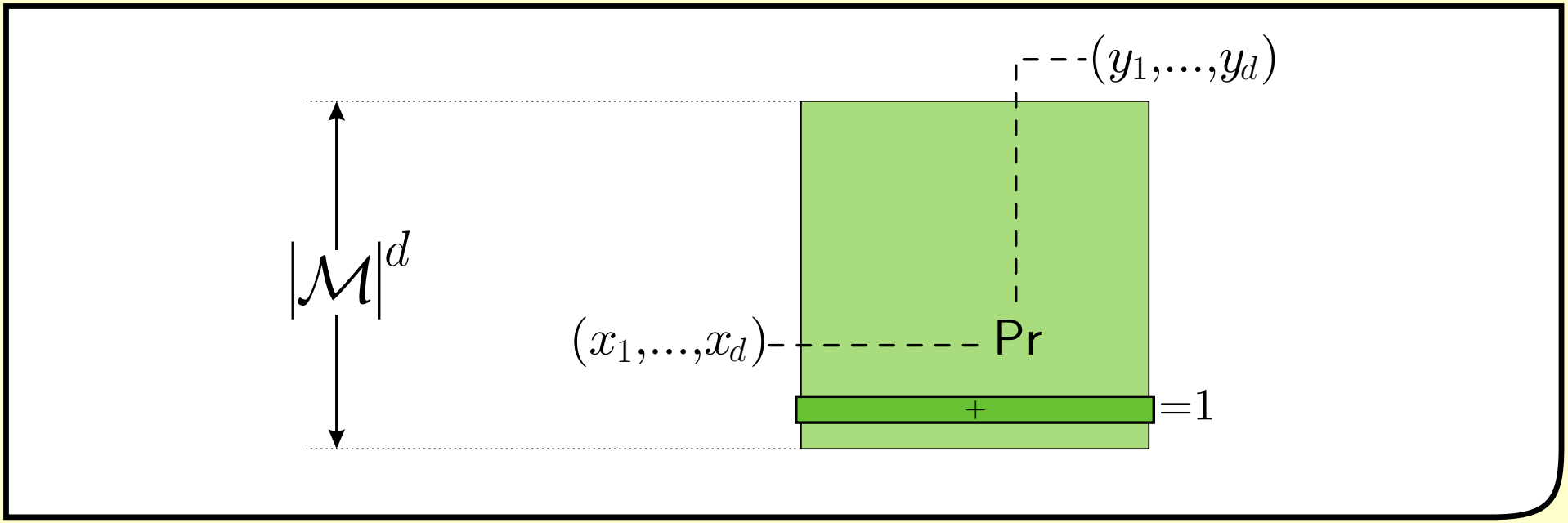
Matrices de distribution

L'outil de base

Soit $\mathsf{C} : \mathcal{M} \to \mathcal{M}$, une permutation aléatoire.

- $[\mathsf{C}]^d$ est la matrice de distribution d'ordre d définie par :

$$[\mathsf{C}]^d_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)} = \Pr_{c\in\mathsf{C}}\big(c(x_1) = y_1, \ldots, c(x_d) = y_d\big).$$



Soit $\mathsf{C} : \mathcal{M} \to \mathcal{M}$, une permutation aléatoire.

- $[\mathsf{C}]^d$ est la matrice de distribution d'ordre d définie par :

$$[\mathsf{C}]^d_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)} = \Pr_{c\in\mathsf{C}}\big(c(x_1) = y_1, \ldots, c(x_d) = y_d\big).$$

- Deux permutations aléatoires sont faciles à distinguer si leurs matrices sont très différentes.
  $\longrightarrow$ C est sûr si $[\mathsf{C}]^d$ est très proche de $[\mathsf{C}^\star]^d$.
- Deux normes existent pour mesurer cette distance.

Mesure de l'avantage

Les normes $|||\cdot|||_\infty$ et $\|\cdot\|_a$

- Triple norme infinie : $|||M|||_\infty = \max_x \sum_y |M_{x,y}|$.

L'avantage du meilleur distingueur non-adaptatif contre C est donné par : $\mathrm{Adv}^d = \frac{1}{2} \left|\left|\left| [\mathsf{C}]^d - [\mathsf{C}^\star]^d \right|\right|\right|_\infty$.

- Norme "a" : $\|M\|_a = \max_{x_1} \sum_{y_1} \ldots \max_{x_d} \sum_{y_d} |M_{x,y}|$.

L'avantage du meilleur distingueur adaptatif contre C est donné par : $\mathrm{Adv}_a^d = \frac{1}{2} \left\| [\mathsf{C}]^d - [\mathsf{C}^\star]^d \right\|_a$.

⚠ Attaques limitées à d clairs choisis.

Extension aux attaques itérées

- Attaque itérées d'ordre d et complexité n :
  - l'adversaire fait n requêtes indépendantes de d clairs aléatoires,
  - pour chaque requête il reçoit les chiffrés et conserve un bit d'information,
  - à la fin il doit choisir s'il s'agit de C ou de $C^\star$.

- Cryptanalyse linéaire $\longrightarrow$ attaque itérée d'ordre 1.

- Cryptanalyse différentielle $\longrightarrow$ attaque itérée d'ordre 2.

Extension aux attaques itérées

- Si $\left|\!\left|\!\left|[\mathsf{C}]^{2d} - [\mathsf{C}^\star]^{2d}\right|\!\right|\!\right|_\infty \leq \varepsilon$ alors :

$$\mathsf{Adv}^d_{\mathrm{iter}} \leq \sqrt[3]{n^2 \left(\frac{d^2}{|\mathcal{M}|} + \varepsilon\right)}$$

  - ▷ il faut $n \simeq \frac{1}{\sqrt{\varepsilon}}$ pour bien distinguer.

- Pour les cryptanalyses linéaires et différentielles il existe un résultat est un peu plus fort :
  - ▷ il faut $n \simeq \frac{1}{\left|\!\left|\!\left|[\mathsf{C}]^2 - [\mathsf{C}^\star]^2\right|\!\right|\!\right|_\infty}$ pour distinguer.

Début de conclusion

- La théorie de la décorrélation permet de prouver la résistance à de grandes classes d'attaques :
  - Une grande partie des attaques utilisées contre les chiffrement par blocs sont des attaques itérées.

- Problèmes :
  - Comment calculer les matrices de distribution ?
  - Comment gérer des matrices de taille $2^{256} \times 2^{256}$?

La décorrélation en pratique

Combinaison de permutations

Une propriété essentielle

- A et B sont deux permutation aléatoires indépendantes :
$$[\mathsf{B} \circ \mathsf{A}]^d = [\mathsf{A}]^d \times [\mathsf{B}]^d.$$

- Si C est composé de 10 tours T utilisant des clefs indépendantes :
$$[\mathsf{C}]^d = \left([\mathsf{T}]^d\right)^{10}.$$
  - Valable aussi pour des transformations constantes.

On peut réussir à calculer $[\mathsf{C}]^d$ sans avoir à calculer C sur toutes les entrées possibles.

$C^\star$ est "absorbant"

- Sur l'ensemble $\{1,2,3\}$, $[C^\star]^2$ vaut :

$$
\begin{array}{c}
 \\ (1,1) \\ (1,2) \\ (1,3) \\ (2,1) \\ (2,2) \\ (2,3) \\ (3,1) \\ (3,2) \\ (3,3)
\end{array}
\begin{array}{c}
\begin{array}{ccccccccc} (1,1) & (1,2) & (1,3) & (2,1) & (2,2) & (2,3) & (3,1) & (3,2) & (3,3) \end{array} \\
\left[\begin{array}{ccccccccc}
\frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
\frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
\frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3}
\end{array}\right]
\end{array}
$$

  - les 0 sont là pour toute permutation,
  - la somme par colonnes vaut 1 (pour toute permutation),
  - une ligne/colonne ne contient qu'une valeur.

$$\rightarrow [C^\star]^d \times [C]^d = [C]^d \times [C^\star]^d = [C^\star]^d.$$

Les modules de décorrélation

Définition, exemple

- Un module de décorrélation est une permutation aléatoire C telle qu'il existe d pour lequel $[\mathsf{C}]^d = [\mathsf{C}^\star]^d$.
  - Aussi valable pour tout $i \leq d$.
  - Pour tout A et B : $[\mathsf{B} \circ \mathsf{C} \circ \mathsf{A}]^d = [\mathsf{C}^\star]^d$. [COCONUT98]

- Exemples :
  - ordre 1 : $X \mapsto A \oplus X$, où A est une variable aléatoire équidistribuée.
  - ordre 2 : $X \mapsto AX + B$, où $A \neq 0$ et B sont équidistribuées.
  - au-delà : rien de très simple ?

Les modules de décorrélation

Avantages / inconvénients

- ◇ simples à comprendre
- ◇ parfaits jusqu'à l'ordre d

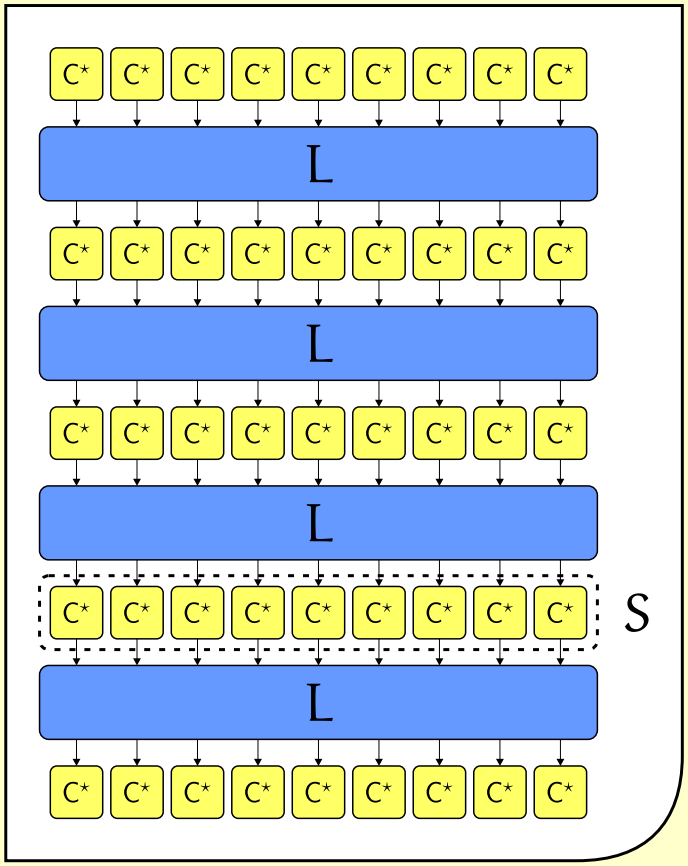
- ◇ sur le bloc complet ⟶ assez lents (sauf à l'ordre 1)
- ◇ a priori, très mauvais au-delà de l'ordre d
  - ▷ attaque "boomerang" sur COCONUT.

- ▶ Utiles pour des preuves, nécessitent une intégration dans une construction déjà sûre.

Utiliser des symétries :
Dial C for Cipher

Description de C

- On reprend la structure de l'AES, et on remplace le XOR de clef et les boites S par 16 permutations parfaitement aléatoires et indépendantes
  - 16 petits $C^\star$ sur 8 bits.



- S est parfaite à l'ordre 1.
- On s'intéresse à la décorrélation d'ordre 2 :
  - à quoi $[S]^2$ ressemble-t-elle ?
  - comment $[S]^2$ et $[L]^2$ se combinent-elle ?

Propriétés de $[S]^2$

- On note $\textsc{Supp}(x, x')$ le support de $x \oplus x'$: l'élément de $\{0,1\}^{16}$ nul sur les coefficients où $x_i = x'_i$.
- On note $w(x, x')$ le poids de Hamming de $\textsc{Supp}(x, x')$.
- On note $q = 2^8$, pour simplifier.

$$[\mathrm{S}]^2_{(x,x'),(y,y')} = \frac{\mathbf{1}_{\textsc{Supp}(x,x')=\textsc{Supp}(y,y')}}{q^{16} \cdot (q-1)^{w(x,x')}}.$$

Propriétés de $[S]^2$

- On note $\textsc{Supp}(x, x')$ le support de $x \oplus x'$: l'élément de $\{0,1\}^{16}$ nul sur les coefficients où $x_i = x'_i$.
- On note $w(x, x')$ le poids de Hamming de $\textsc{Supp}(x, x')$.
- On note $q = 2^8$, pour simplifier.

$$[\mathrm{S}]^2_{(x,x'),(y,y')} = \frac{\mathbf{1}_{\textsc{Supp}(x,x')=\textsc{Supp}(y,y')}}{q^{16} \cdot (q-1)^{w(x,x')}}.$$

- On définit deux matrices de conversion Paire $\leftrightarrow$ Support :

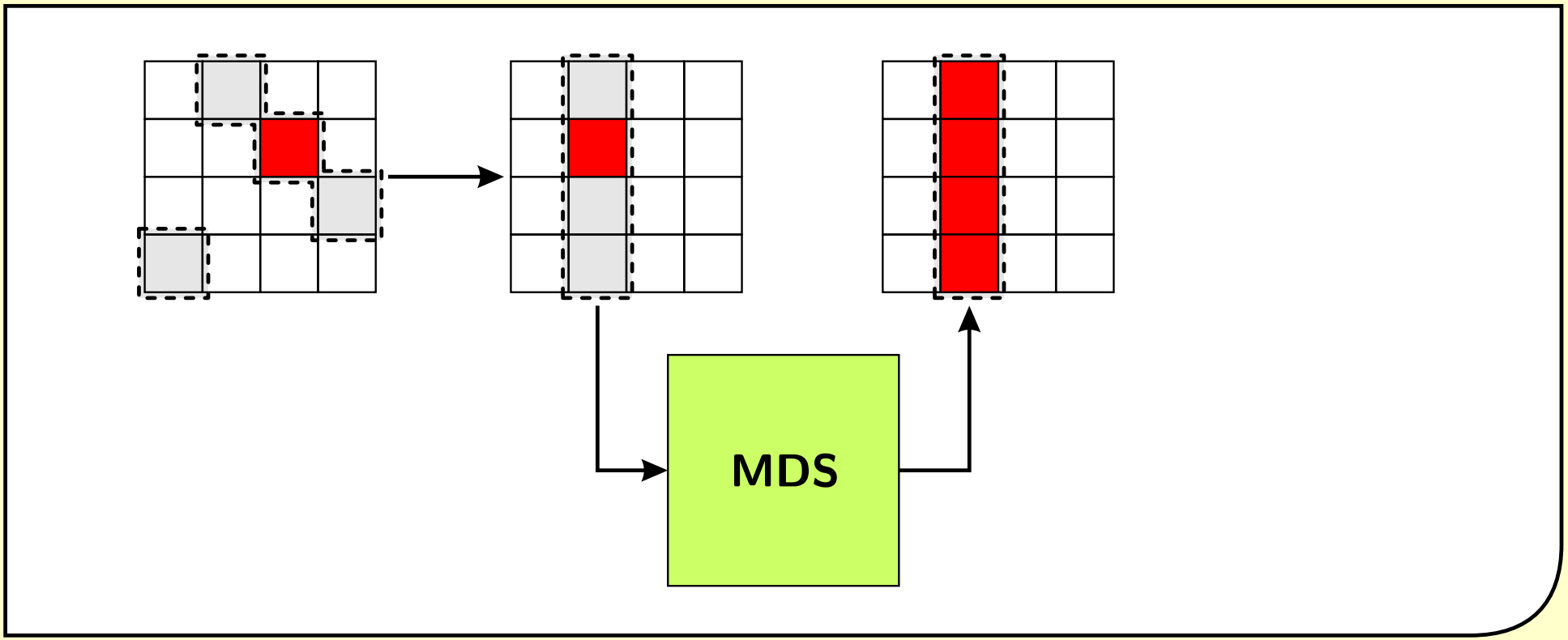
$$PS_{(x,x'),\gamma} = \mathbf{1}_{\gamma=\textsc{Supp}(x,x')},$$

$$SP_{\gamma',(y,y')} = \mathbf{1}_{\gamma'=\textsc{Supp}(y,y')} q^{-16} \cdot (q-1)^{-w(\gamma')}.$$

$$\rightarrow [\mathrm{S}]^2 = PS \times SP$$

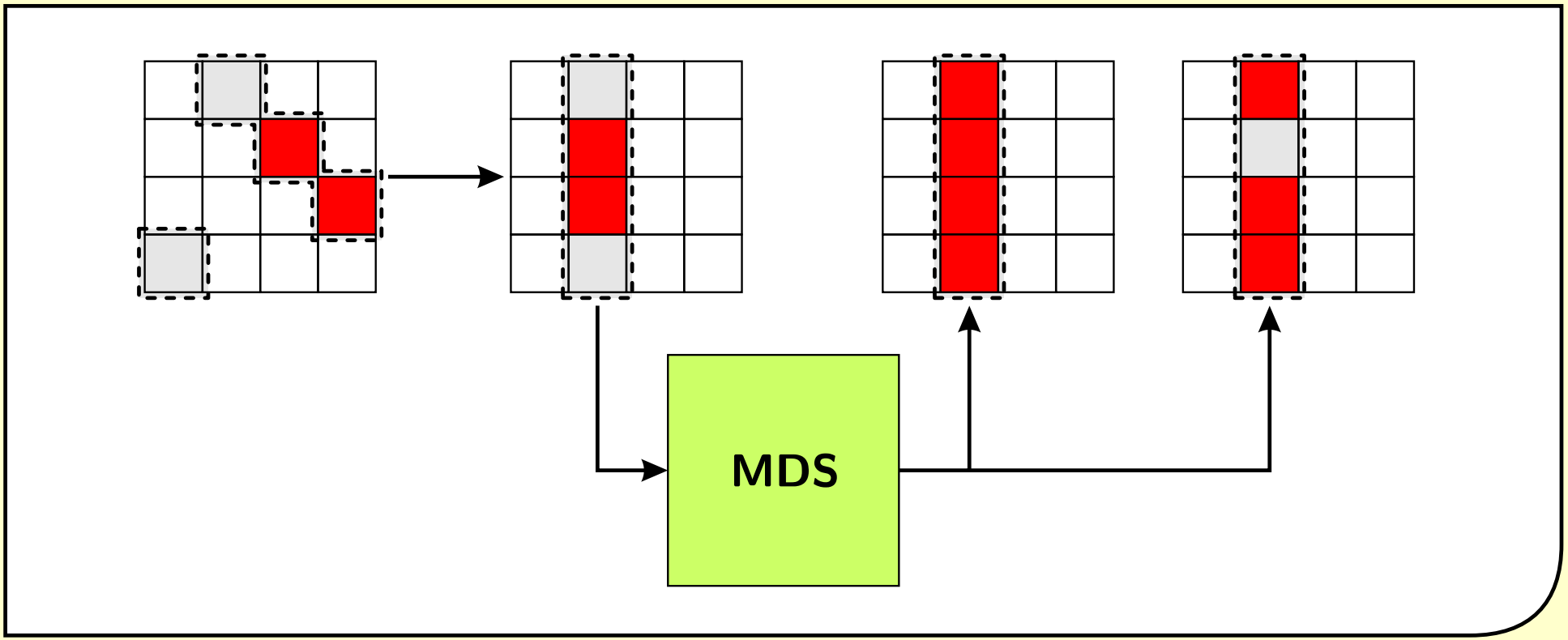
Propriétés de $[\mathrm{L}]^2$

- $[\mathrm{L}]^2$ est une permutation :
  - impossible à simplifier a priori,
  - entourée par $[\mathrm{S}]^2$ de chaque côté
    - → on regarde comment simplifier $SP \times [\mathrm{L}]^2 \times PS$.



Propriétés de $[\mathrm{L}]^2$

- $[\mathrm{L}]^2$ est une permutation :
  - impossible à simplifier a priori,
  - entourée par $[\mathrm{S}]^2$ de chaque côté
    - → on regarde comment simplifier $SP \times [\mathrm{L}]^2 \times PS$.



Propriétés de $[\mathrm{L}]^2$

- Ne dépend pas du support complet, juste des poids sur les diagonales d'entrée et sur les colonnes de sortie.
  - On introduit deux nouvelles matrices Support $\leftrightarrow$ Poids

$$SW_{\gamma,(w_1,w_2,w_3,w_4)} = \mathbf{1}_{\mathsf{weights}_{\mathsf{diag}}(\gamma)=(w_1,w_2,w_3,w_4)}$$

$$WS_{(w'_1,w'_2,w'_3,w'_4),\gamma'} = \frac{\mathbf{1}_{\mathsf{weights}_{\mathsf{col}}(\gamma')=(w'_1,w'_2,w'_3,w'_4)}}{\binom{4}{w'_1}\binom{4}{w'_2}\binom{4}{w'_3}\binom{4}{w'_4}}$$

$$\longrightarrow SP \times [\mathrm{L}]^2 \times PS = SW \times \bar{\mathrm{L}} \times WS$$

- $\bar{\mathrm{L}}$ est une matrice 625×625 facile à calculer.

Calcul exact de $\|[\mathrm{C}]^2 - [\mathrm{C}^\star]^2\|_\infty$

- On se ramène entièrement à un calcul de matrices 625×625 :

$$[\mathrm{C}]^2 = PS \times SW \times \bar{\mathrm{L}} \times WS \times SW \times \cdots \times \bar{\mathrm{L}} \times WS \times SP$$
$$= PW \times (\bar{\mathrm{L}} \times \bar{\mathrm{W}})^{r-2} \times \bar{\mathrm{L}} \times WP$$

- On peut calculer exactement l'avantage du meilleur distingueur d'ordre 2 (adaptatif ou non) :

r	2	3	4	5	6	7	8	9	10	11	12
Adv	1	$2^{-4.0}$	$2^{-23.4}$	$2^{-45.8}$	$2^{-71.0}$	$2^{-126.3}$	$2^{-141.3}$	$2^{-163.1}$	$2^{-185..5}$	$2^{-210.8}$	$2^{-238.9}$

- Résiste aux attaques itérées d'ordre 1.

Autres résultats

- 4 tours : cryptanalyse linéaire et différentielle
- 5 tours : différentielles impossibles
- 10 tours : toutes les attaques futures ?

⚠ If faut des boites aléatoires et indépendantes.

Autres résultats

- ◇ 4 tours : cryptanalyse linéaire et différentielle
- ◇ 5 tours : différentielles impossibles
- ◇ 10 tours : toutes les attaques futures ?

⚠ If faut des boites aléatoires et indépendantes.

- ▶ Sur 8 bits il faut $\log_2(256!) \approx 1684$ bits de clef par boite
  - ▷ générer 269440 bits à partir d'une clef de 128,
  - ▷ utiliser un générateur prouvé sûr (Blum-Blum-Shub...)

⟶▷ attaque sur $\mathrm{C}_{\mathsf{key}} \Rightarrow$ attaque sur C ou sur BBS.

Si on veut aller plus loin

À l'ordre 3...

- La structure de $[S]^3$ n'est pas simple à comprendre :
  - indexer par un triplet de supports
    - → certains n'existent pas
  - une autre idée ?
- Sur la partie linéaire $[\mathrm{L}]^3$:
  - sans une simplification de $[S]^3$ on ne peut rien faire...
- J'ai pas trop d'autres idées :
  - trouver un L qui fonctionne bien à l'ordre 3 ?
  - faire le calcul avec les triplets de supports possibles ?

Aux ordres supérieurs :
KFC - the Krazy Feistel Cipher

La magie des fonctions aléatoires

- On note $\mathsf{F}^\star$ la fonction parfaitement aléatoire
  - si $x \neq x'$ alors $\mathsf{F}^\star(x)$ et $\mathsf{F}^\star(x')$ sont indépendants.
- On note F la couche de 16 boites $\mathsf{F}^\star$ sur 8 bits
  - si $w(\textsc{Supp}(x, x')) = 16$ ($x \neq x'$ sur chaque boite) alors $\mathsf{F}(x)$ et $\mathsf{F}(x')$ sont indépendants.
  - si parmi d entrées, l'une est partout différente de toutes les autres, sa sortie est indépendante.

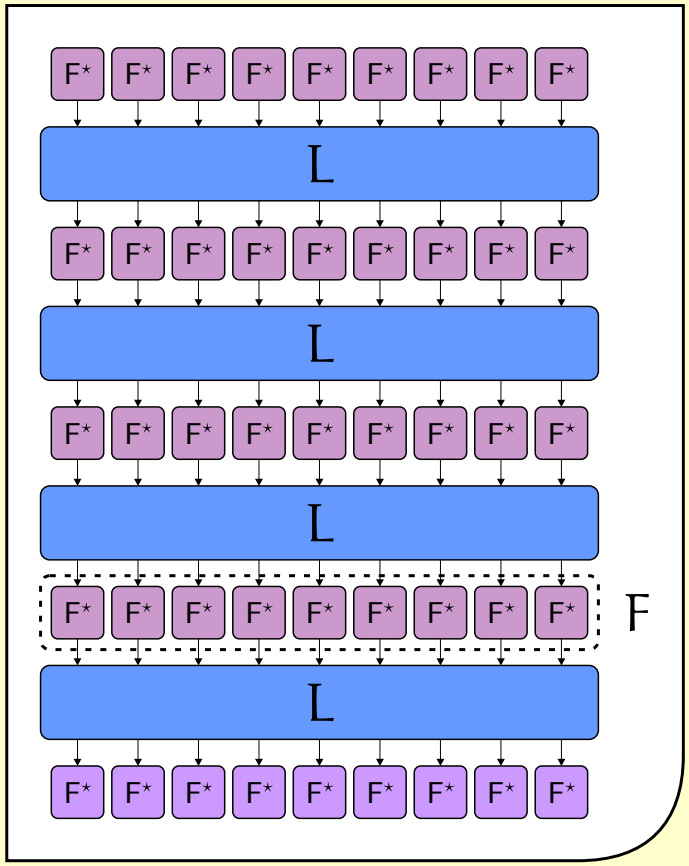
    $\longrightarrow$ un attaquant d'ordre d n'est pas meilleur qu'un attaquant d'ordre $d-1$.
- Si on peut borner la probabilité d'être différent, on peut borner l'avantage du meilleur attaquant d'ordre d.

Construction proposée

- F_{KFC} : alternance de F et de diffusion linéaire L
  - pour la diffusion on prend directement un code MDS.

- On sait calculer $[F_{\mathsf{KFC}}]^2$
  - calcul de matrices 17×17.
- Ça ne marche pas bien :
  - $\left\| [F_{\mathsf{KFC}}]^2 - [\mathrm{F}^\star]^2 \right\|_\infty \simeq \frac{1}{q}$

- Il ne faut pas commencer par F
  - on ajoute des couches S.

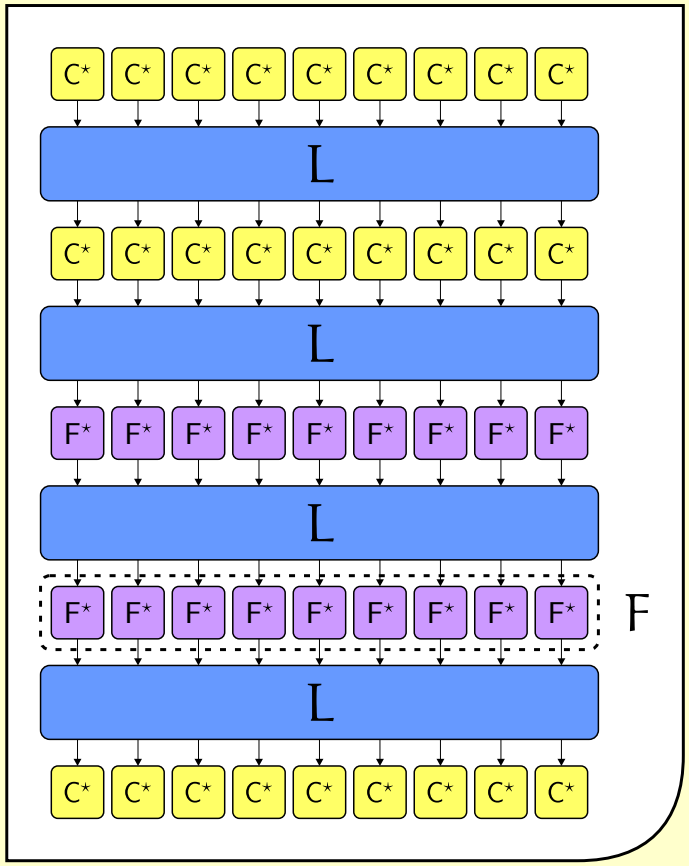


Construction proposée

- F_{KFC} : alternance de F et de diffusion linéaire L
  - ▷ pour la diffusion on prend directement un code MDS.

- On sait calculer $[F_{\mathsf{KFC}}]^2$
  - ▷ calcul de matrices 17×17.
- Ça ne marche pas bien :
  - ▷ $\left\| [F_{\mathsf{KFC}}]^2 - [\mathsf{F}^\star]^2 \right\|_\infty \simeq \frac{1}{q}$

- Il ne faut pas commencer par F
  - ▷ on ajoute des couches S.



Résultats obtenus

Décorrélation à l'ordre 2

- On obtient après calculs :

	$N=8$ et $q=2^8$				$N=8$ et $q=2^{16}$				$N=16$ et $q=2^8$			
$r_1 \backslash r_2$	0	1	10	100	0	1	10	100	0	1	10	100
0	1	2^{-5}	2^{-8}	2^{-8}	1	2^{-13}	2^{-16}	2^{-16}	1	2^{-4}	2^{-8}	2^{-8}
1	2^{-5}	2^{-50}	2^{-52}	2^{-49}	2^{-13}	2^{-114}	2^{-116}	2^{-113}	2^{-4}	2^{-95}	2^{-104}	2^{-103}
2	2^{-46}	2^{-53}	2^{-52}	2^{-49}	2^{-110}	2^{-117}	2^{-116}	2^{-113}	2^{-87}	2^{-104}	2^{-104}	2^{-103}
3	2^{-62}	2^{-53}	2^{-52}	2^{-49}	2^{-128}	2^{-117}	2^{-116}	2^{-113}	2^{-120}	2^{-104}	2^{-104}	2^{-103}

- Chaque couche F fait un peu diminuer $[F_{\mathsf{KFC}}]^2$
  - la décorrélation en sortie est une borne inférieure de la décorrélation en entrée de chaque F.
  - les entrées de F sont indépendantes deux à deux.

Récurrence entre les ordres

On note $F_{\mathsf{KFC}[i]}$, la fonction F_{KFC} avec i couches F.

- Soit α_i l'événement :

  Une entrée de la dernière couche F *de* $F_{\mathsf{KFC}[i]}$ *est différentes des* $d-1$ *autres sur les* N *coordonnées.*

- On peut écrire :

$$\mathsf{Adv}^{d}(F_{\mathsf{KFC}[i]}, \mathsf{F}^{\star}) \leq \mathsf{Adv}^{d-1}(F_{\mathsf{KFC}[i]}, \mathsf{F}^{\star}) + \mathsf{Pr}(\bar{\alpha}_i).$$

- On cherche à borner $\mathsf{Pr}(\bar{\alpha}_i)$:
  - ▷ on utilise l'indépendance deux à deux.

Borne sur $\Pr(\bar{\alpha}_i)$

Entrée : $(X_1, ..., X_d)$ avec $X_i = (X_{i,1}, ..., X_{i,N})$.
Soit $\lambda \leq d$ le nombre de X_i différents de tous les autres.

- $E(\lambda) = d \cdot \mathcal{P}_d^N$ avec $\mathcal{P}_d = \Pr(X_{1,1} \notin \{X_{2,1}, ..., X_{d,1}\})$.

De plus :

$$\begin{aligned} \mathcal{P}_{d+1} &= \mathcal{P}_d - \Pr(X_{1,1} \notin \{X_{2,1}, ..., X_{d,1}\}, X_{1,1} = X_{d+1,1}) \\ &\geq \mathcal{P}_d - \Pr(X_{1,1} = X_{d+1,1}) = \mathcal{P}_d - \tfrac{1}{q} \end{aligned}$$

- Donc : $E(\lambda) \geq d \cdot \left(1 - \frac{d-1}{q}\right)^N$.

  Or $E(\lambda) = \sum_{k=1}^{d} \Pr(\lambda = k) \leq d \cdot \Pr(\lambda \neq 0) = d \cdot \Pr(\alpha_i)$

  $\rightarrow \Pr(\bar{\alpha}_i) \leq 1 - \frac{E(\lambda)}{d} \leq 1 - \left(1 - \frac{d-1}{q}\right)^N$.

En empilant les couches...

- Si on regarde t couches F successives on trouve :

$$\Pr(\bar{\alpha}_1, ..., \bar{\alpha}_t) \leq \left(1 - \left(1 - \tfrac{d-1}{q}\right)^N\right)^t$$

- Par récurrence on peut dire :

Pour tout d, et tout ensemble d'entiers $\{t_3, ..., t_d\}$ tels que $\sum_{i=3}^{d} t_i \leq r$, si $\left\| [F_{\mathsf{KFC}[r]}]^2 - [\mathsf{F}^\star]^2 \right\|_\infty \leq \varepsilon$ on a :

$$\mathsf{Adv}^d(F_{\mathsf{KFC}[r]}, \mathsf{F}^\star) \leq \varepsilon + \sum_{i=3}^{d} \left(1 - \left(1 - \tfrac{d-1}{q}\right)^N\right)^{t_i}$$

Décorrélation d'ordre d

On obtient, en choisissant bien les t_i :

	$N = 8$ et $q = 2^8$						
r \ d	2	3	4	8	16	32	64
10	2^{-52}	2^{-40}	2^{-17}	2^{-2}	1	1	1
100	2^{-49}	2^{-49}	2^{-49}	2^{-46}	2^{-11}	1	1
250	2^{-48}	2^{-48}	2^{-48}	2^{-48}	2^{-33}	2^{-5}	1
1000	2^{-46}	2^{-46}	2^{-46}	2^{-46}	2^{-46}	2^{-35}	2^{-2}

	$N = 8$ et $q = 2^{16}$						
r \ d	2	3	4	8	16	32	64
10	2^{-116}	2^{-116}	2^{-57}	2^{-11}	1	1	1
100	2^{-113}	2^{-113}	2^{-113}	2^{-113}	2^{-66}	2^{-23}	2^{-5}
250	2^{-112}	2^{-112}	2^{-112}	2^{-112}	2^{-112}	2^{-69}	2^{-25}
1000	2^{-110}	2^{-110}	2^{-110}	2^{-110}	2^{-110}	2^{-110}	2^{-110}

	$N = 16$ et $q = 2^8$						
r \ d	2	3	4	8	16	32	64
10	2^{-104}	2^{-31}	2^{-12}	1	1	1	1
100	2^{-103}	2^{-103}	2^{-103}	2^{-31}	2^{-5}	1	1
250	2^{-103}	2^{-103}	2^{-103}	2^{-81}	2^{-18}	1	1
1000	2^{-102}	2^{-102}	2^{-102}	2^{-102}	2^{-82}	2^{-12}	1

- Avec $N = 8$, $q = 2^{16}$ et $r = 1000$ on prouve la résistance aux attaques d'ordre 70.

Pour finir

- Il suffit d'utiliser F_{KFC} dans un Feistel :
  - on obtient un chiffrement par blocs,
  - conserve les propriétés de décorrélation.
- On reprends BBS pour générer la clef.

- La borne est très large :
  - on résiste sûrement à des attaques d'ordre plus grand,
  - on peut l'améliorer en regardant $\mathsf{Pr}(\lambda) \leq i$.

Conclusion

Conclusion

- La théorie de la décorrélation permet de prouver des résultats forts.
- On peut l'utiliser sans modules de décorrélation.
- Les constructions sont utilisables en pratique :
  - le chiffrement rapide, mais le temps de génération de clef est un peu long,
  - on peut remplacer les $\mathrm{F}^\star$ ou $\mathrm{C}^\star$ par des boites décorrélées à l'ordre d (2 pour C) et cela suffit,
  - il faut remplacer BBS par quelque chose de léger.