Analyse de la résistance des chiffrements par blocs aux attaques linéaires et différentielles

Joëlle Roué

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Informatique

École doctorale Informatique, Télécommunications et Électronique (Paris)

Présentée par

Joëlle ROUÉ

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

Analyse de la résistance des chiffrements par blocs aux attaques linéaires et différentielles

soutenue le 14 octobre 2015

devant le jury composé de :

Anne CANTEAUT	Inria Paris-Rocquencourt	Directrice de thèse
Daniel AUGOT	Inria Saclay	Rapporteur
Thierry BERGER	Université de Limoges	Rapporteur
Joan DAEMEN	STMicroelectronics	Examinateur
Henri GILBERT	ANSSI	Examinateur
Antoine JOUX	UPMC	Examinateur
Marine MINIER	INSA Lyon	Examinatrice
María NAYA-PLASENCIA	Inria Paris-Rocquencourt	Examinatrice

Analyse de la résistance des chiffrements par blocs aux attaques linéaires et différentielles

Joëlle Roué

Sous la direction d'Anne Canteaut

Financé par l'Agence Nationale de la Recherche grâce au projet BLOC

Inria Paris-Rocquencourt
Équipe-Projet SECRET

Remerciements

Ces trois années de thèse, précédées d'un stage, ont été effectuées au sein du projet SECRET de l'Inria Paris-Rocquencourt. Un grand merci à toutes les personnes qui ont participé de près ou de loin au bon déroulement de ma thèse.

En particulier, je tiens à remercier Anne Canteaut pour m'avoir offert la possibilité de travailler sur un sujet aussi intéressant, pour sa disponibilité, pour m'avoir guidée et conseillée durant ces trois années. Elle a été la meilleure directrice de thèse que l'on puisse imaginer.

Je remercie les rapporteurs de ma thèse, Daniel Augot et Thierry Berger. Je vous suis très reconnaissante d'avoir accepté de relire mon manuscrit : vos suggestions et remarques m'ont permis d'améliorer sa qualité.

Merci à Joan Daemen, Henri Gilbert, Antoine Joux, Marine Minier et María Naya-Plasencia pour leur participation à ce jury de thèse, pour l'intérêt qu'ils ont portés à mes travaux et pour le temps qu'ils m'accordent en assistant à cette soutenance.

J'ai eu la chance de travailler au sein du projet SECRET, où règnent la bonne humeur, la convivialité ainsi que le partage de connaissances et d'idées. Merci aux membres et aux amis de ce projet qui veillent à maintenir cette ambiance. Je remercie les permanents du projet : Anne, Pascale (pour m'avoir accueillie au projet, pour avoir encadré mon stage et pour sa bienveillance), Nicolas, María, Anthony, André et Jean-Pierre (pour sa confiance en ma connaissance de la langue française).

Merci aux personnes avec qui j'ai partagé mon bureau : María, Rafael, Sébastien et Audrey (pour son amitié et pour toute l'aide qu'elle m'a apporté, par exemple en programmation ou pour trouver les teintes de bleu parfaites). Merci à Virginie pour son amitié, pour m'avoir aidé à exprimer certaines idées clairement et pour sa complicité durant l'école doctorale en Islande, les Journées C2 à Grenoble et durant EUROCRYPT 2015.

Pour leur gentillesse et les bons moments partagés, je remercie Aurélie, Christina, Denise, Dimitris, Gaëtan, Georgi, Grégory (pour son aide en programmation), Irene, Julia (pour son sourire et sa gaieté permanente), Kaushik, Marion, Nicky, Rodolfo, Valentin et Yann. Et bien sûr, merci à Christelle pour sa bonne humeur, sa disponibilité, son efficacité et son aide inestimable.

Je voudrais remercier tous les membres de l'ANR BLOC. Je ne me risquerai pas à tous les citer, j'ai bien trop peur d'oublier une de ces personnes grâce à qui j'ai beaucoup appris.

Je souhaite remercier mes amis, en particulier David et les Toulousains, pour m'avoir offert d'agréables moments de détente, parce qu'ils savent me faire rire même lorsque je suis d'humeur morose, ainsi que pour leurs conseils.

Enfin, merci beaucoup à toute ma famille et aux personnes qui me sont les plus chères (mes parents, mes sœurs, Florence, Laure, Max et Pascal) pour leur soutien, leur confiance et leurs encouragements. Je ne pense pas que je serais arrivée jusque là sans vous.

Pour finir, je tiens à vous dire que ces quelques mots sont bien faibles comparés à la gratitude que j'éprouve envers vous tous.

Table des matières

Introduction

Les travaux menés dans cette thèse se situent dans le domaine de la cryptographie symétrique et portent plus précisément sur l'analyse de la résistance aux cryptanalyses linéaires et différentielles des chiffrements par blocs de la famille des réseaux de substitution-permutation. Ce type de chiffrements, parmi lequel figure l'actuel standard de chiffrement symétrique AES, repose sur l'application successive d'une permutation hautement non-linéaire, appelée boîte-S, et d'une fonction linéaire assurant la diffusion.

Lors de la conception de l'AES, différents critères (rappelés dans le chapitre 2) ont été énoncés afin de choisir les composants ayant la meilleure résistance possible aux attaques linéaires et différentielles, comme l'uniformité différentielle et la non-linéarité pour la boîte-S, ainsi que le branch number de la fonction de diffusion. Toutefois, la modélisation des cryptanalyses qui a abouti à ces critères ne rend pas parfaitement compte de la complexité réelle des attaques. On constate par exemple que deux boîtes-S ayant la même uniformité différentielle (respectivement la même non-linéarité) ne présentent pas toujours la même résistance à la cryptanalyse différentielle (respectivement linéaire). Dans cette thèse, nous avons raffiné les critères classiques de résistance des réseaux de substitution-permutation aux attaques linéaires et différentielles, en considérant les quantités qui mesurent effectivement la résistance à ces deux familles d'attaques : la valeur maximale de l'espérance de la probabilité d'une différentielle (MEDP) et la valeur maximale du potentiel linéaire moyen (MELP).

Dans le chapitre 3, nous présentons une nouvelle borne sur le MEDP (resp. MELP) sur deux tours, qui ne dépend que de la boîte-S et du branch number de la fonction de diffusion, et qui n'est plus invariante par composition avec des permutations affines contrairement aux résultats précédents. Cette borne s'applique si la fonction de diffusion est linéaire sur $\mathbf{F}_{2^m}$ (l'alphabet sur lequel la boîte-S est définie), comme dans l'AES. De plus, pour toute boîte-S et toute valeur t, nous avons montré qu'il existe toujours au moins une permutation linéaire de $(\mathbf{F}_{2^m})^t$ de branch number maximal pour laquelle le MEDP_2 (resp. MELP_2) dépasse une certaine quantité. Ainsi, sous certaines conditions sur la boîte-S S et le branch number d, il est impossible de trouver une meilleure borne ne dépendant que de S et de d.

Dans le chapitre 4, nous introduisons une nouvelle propriété des boîtes-S qui simplifie le calcul de la borne, vérifiée par exemple par les fonctions puissance. Le cas où S et S^{-1} vérifient cette propriété est particulièrement intéressant car on peut prouver que notre borne inférieure est satisfaite pour toute fonction M de branch number maximal. En conséquence, si S est l'inversion dans $\mathbf{F}_{2^m}$ par exemple, alors la valeur exacte de MEDP_2 (resp. MELP_2) est toujours la plus grande possible parmi toutes les fonctions

de la même classe d'équivalence. Par contre, le fait de composer l'inversion avec une permutation affine permet en général de diminuer significativement la valeur de MEDP_2 (resp. MELP_2). C'est ce qui a été constaté dans le cas particulier de l'AES. Les travaux de ces deux chapitres ont donnés lieu à des présentations dans les conférences *11th International Conference on Finite Fields and their applications Fq11* [CR14] et *EUROCRYPT 2015* [CR15b].

Par ailleurs, nous avons cherché à infirmer l'idée communément admise que le MEDP_2 est atteint par une différentielle ayant le plus petit nombre possible de boîtes-S actives. Lorsque la boîte-S est fixée, le maximum de la probabilité d'une caractéristique sur r tours diminue lorsque le nombre de boîtes-S actives augmente. Ainsi, beaucoup d'analyses de sécurité s'intéressent au nombre minimal de boîtes-S actives sur r tours consécutifs d'un chiffrement lorsque r varie. Cependant, la complexité des attaques différentielles dépend de la probabilité d'une différentielle, *i.e.*, de la somme des probabilités de toutes les caractéristiques dont la différence en entrée et la différence en sortie sont fixées. De plus, sur deux tours d'un réseau de substitution-permutation, le nombre de caractéristiques dans une différentielle augmente généralement avec le poids de Hamming de la différentielle. Donc le MEDP_2 pourrait provenir d'une différentielle ayant un grand nombre de caractéristiques de petite probabilité : il n'y a pas de raison de croire que le MEDP_2 est atteint par une différentielle ayant le plus petit nombre possible de boîtes-S actives. Toutefois, en pratique, le MEDP_2 est atteint par une différentielle de plus petit poids possible pour un grand nombre de chiffrements. Dans le chapitre 5, nous avons réussi à démontrer cette observation pour certaines familles de boîtes-S lorsque la fonction de diffusion M a un branch number maximal. Cependant, nous montrons que ce phénomène n'est pas général : nous avons trouvé les premiers exemples de réseaux de substitution-permutation pour lesquels le MEDP_2 est atteint par une différentielle dont le nombre de boîtes-S actives est supérieur au branch number de M. Ces travaux ont donné lieu à une présentation lors de la conférence *C2SI 2015* [CR15a].

Tous les travaux sur la cryptanalyse différentielle présentés jusque là portent sur l'étude de la moyenne sur les clés de la probabilité d'une différentielle. Dans le chapitre 6, nous nous intéressons à cette probabilité lorsque la clé est fixée. Nous présentons les résultats de travaux en cours sur l'étude de la structure de l'ensemble des clés telles que la probabilité à clé fixée d'une caractéristique sur deux tours est non nulle, dans le but de pouvoir utiliser cette structure pour déterminer la probabilité à clé fixée d'une différentielle sur deux tours.

1

Les chiffrements par blocs et leurs composants

1.1 Introduction à la cryptographie symétrique

1.1.1 La cryptologie

La cryptologie est la science du secret. Cette science est née de la volonté des hommes de partager certaines informations uniquement avec quelques personnes. La cryptologie est composée de deux parties complémentaires. D'un côté, la cryptographie est l'ensemble des méthodes permettant de protéger les informations. Cela inclut la confidentialité (assurer que les données et les messages ne soient lisibles que par les personnes autorisées), l'authenticité (s'assurer de la provenance des messages) et l'intégrité (assurer que les messages n'ont pas été modifiés par un adversaire entre le moment où ils ont été envoyés et celui où ils ont été reçus). D'autre part, la cryptanalyse est l'ensemble des méthodes permettant à un adversaire de briser la confidentialité, l'authenticité ou l'intégrité des données.

Dans ce document, nous nous intéressons au chiffrement, qui est une technique permettant d'assurer la confidentialité des données ou des communications en transformant les données en une suite de symboles incompréhensible par toute personne ne possédant pas un certain secret, appelé une clé. Considérons deux personnes, Alice et Bob, qui souhaitent communiquer de manière confidentielle via un canal de communication non sécurisé. Il est donc possible qu'un adversaire, Ève, écoute le canal.

Pour qu'Alice puisse envoyer un message m à Bob de manière confidentielle, Alice et Bob se munissent d'une clé notée k_A pour Alice et k_B pour Bob, et de deux fonctions : une fonction de chiffrement E, qui transforme en fonction de la clé k_A le message m à envoyer en un message incompréhensible c, et une fonction de déchiffrement D, qui, à partir de la clé k_B et du message reçu c, retrouve le message m. Le message m est appelé *message clair*, c est appelé *message chiffré*.

De nos jours, la cryptographie est composée de deux familles qui diffèrent par l'utilisation par Alice et Bob d'une clé commune ou de deux clés distinctes. Les chiffrements suivent donc deux modèles différents. Le modèle le plus ancien s'appelle la cryptographie symétrique (puisque les clés utilisées pour chiffrer et déchiffrer sont les mêmes : $k_A = k_B = k$) ou encore la cryptographie à clé secrète.

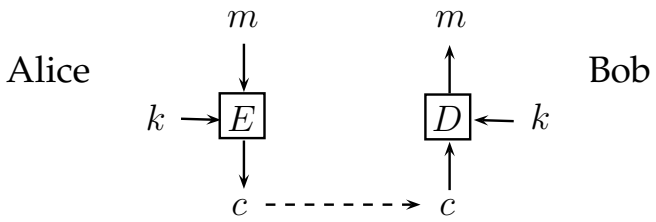


FIGURE 1.1 – Chiffrement symétrique.

L'inconvénient principal de la cryptographie symétrique est l'échange de clé. En effet, cet échange nécessite un canal sécurisé pour transmettre la clé ou un protocole pour qu'Alice et Bob créent une clé commune inconnue de toute autre personne. Le protocole d'échange de clé le plus célèbre est celui de Whitfield Diffie et Martin Hellman [DH76].

Dans l'article [DH76], les auteurs présentent aussi un deuxième modèle de chiffrements. Dans ce modèle, chaque clé est composée de deux parties : une clé publique, notée pk, qui est destinée à être connue de tous ; et une clé secrète, notée sk, qui n'est connue de son propriétaire. Comme les clés utilisées ne sont pas les mêmes, ce modèle est appelé cryptographie asymétrique, ou cryptographie à clé publique. Pour qu'Alice et Bob puissent communiquer, ils doivent posséder chacun un couple de clé (pk_A, sk_A) et (pk_B, sk_B). Pour envoyer un message m à Bob, Alice va chiffrer m avec la clé publique pk_B de Bob en un message chiffré c. Le message m est alors retrouvé à partir du message chiffré c et de la clé secrète sk_B, c'est-à-dire que n'importe qui peut chiffrer un message à l'attention de Bob, mais lui-seul est capable de déchiffrer un message chiffré avec sa clé publique pk_B. Bob utilise de manière similaire la clé publique d'Alice pour lui répondre.

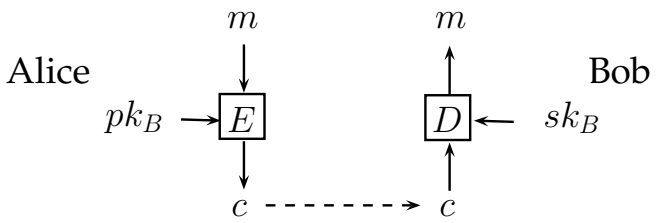


FIGURE 1.2 – Chiffrement asymétrique.

Cependant, les chiffrements asymétriques sont généralement beaucoup moins performants que les chiffrements symétriques. C'est pourquoi un système hybride est généralement utilisé : Alice et Bob utilisent un chiffrement asymétrique pour échanger une

clé k, qu'ils utilisent ensuite comme clé d'un chiffrement symétrique pour protéger leur communication. C'est pourquoi les chiffrements symétriques sont encore très utilisés aujourd'hui.

1.1.2 La cryptographie symétrique

Les chiffrements par blocs et les chiffrements à flot sont les deux grandes familles de chiffrement symétrique. Du fait de leur vitesse de chiffrement et de la possibilité de les implémenter sur des systèmes contraints, ils assurent aujourd'hui la confidentialité d'une grande partie des données et communications informatiques. En particulier, nous nous intéressons aux réseaux de substitution-permutation, qui forment une famille de chiffrements par blocs qui contient l'actuel standard de chiffrement symétrique, l'AES.

Dans ce chapitre, nous allons rappeler les notions nécessaires à la compréhension des travaux présentés dans la suite ce document. En particulier, nous présentons quelques résultats sur les codes correcteurs et les fonctions booléennes, qui permettent d'étudier les propriétés cryptographiques des composants des chiffrements par blocs.

1.2 Les réseaux de substitution-permutation

1.2.1 Chiffrements par blocs

Définition 1.1. *Soient κ et n deux entiers strictement positifs. Un* chiffrement par blocs *sur $\mathbf{F}_2^n$ est une famille de permutations E de $\mathbf{F}_2^n$, paramétrées par une clé $k \in \mathbf{F}_2^\kappa$. Une entrée de ces permutations est un message clair $P \in \mathbf{F}_2^n$ et l'image de P par la permutation de paramètre $k \in \mathbf{F}_2^\kappa$ est un message chiffré $C \in \mathbf{F}_2^n$:*

$$\begin{array}{rccc} E: & \mathbf{F}_2^\kappa \times \mathbf{F}_2^n & \longrightarrow & \mathbf{F}_2^n \\ & (k, P) & \longmapsto & C = E(k, P). \end{array}$$

Pour tout $k \in \mathbf{F}_2^\kappa$, la permutation $E_k : P \longmapsto E_k(P) := E(k, P)$ est appelée fonction de chiffrement.

La bijectivité des fonctions de chiffrement est indispensable pour que le déchiffrement soit possible.

Un chiffrement par blocs permet naturellement de chiffrer des messages de taille fixe, cette taille variant le plus souvent entre 64 et 256 bits. Pour chiffrer un message de taille quelconque, ce message est coupé en blocs de n bits et les blocs sont chiffrés les uns à la suite des autres et liés par un mode opératoire. Les plus utilisés sont les modes CBC et CTR [MvOV97].

La sécurité des chiffrements par blocs repose sur la clé, qui est la seule partie secrète. Il doit donc être impossible de trouver la clé à partir de couples (P, C) de messages clairs/chiffrés. Une façon simple de retrouver la clé est la recherche exhaustive. Cela consiste à chiffrer le message P avec toutes les clés possibles jusqu'à trouver celle qui vérifie $E_k(P) = C$. Pour se protéger de cette attaque universelle, il faut que l'espace des clés $\mathbf{F}_2^\kappa$ soit suffisamment grand pour que la recherche exhaustive prenne trop de

temps. D'un autre côté, pour des raisons pratiques, la taille de la clé ne doit pas être trop grande. Actuellement, on considère que κ doit être supérieur à 80 pour assurer une sécurité minimale, et 128 est la valeur communément admise pour une sécurité à long terme.

De plus, pour qu'un chiffrement par blocs E offre une sécurité adéquate, il ne doit pas exister d'algorithme qui puisse distinguer la fonction de chiffrement E_k, pour une clé k choisie aléatoirement, d'une permutation aléatoire en une complexité significativement inférieure à celle de la recherche exhaustive de la clé.

Concevoir des familles avec ces propriétés et de bonnes performances est assez difficile. La méthode la plus utilisée consiste à concevoir des permutations ayant une implémentation peu coûteuse et à itérer ces permutations plusieurs fois pour obtenir le niveau de sécurité souhaité. De tels chiffrements sont dits *itératifs*. Pour minimiser la taille du circuit implémentant le chiffrement, les fonctions qui composent un chiffrement itératif, appelées *fonctions de tour*, sont identiques (ou elles diffèrent simplement par l'addition d'une constante par exemple), chacune étant paramétrée par une quantité secrète. Cette quantité, appelée *sous-clé*, est dérivée de la clé k à l'aide d'un algorithme dit de cadencement de clé. Une fonction de chiffrement E_k d'un chiffrement itératif s'écrit :

$$E_k = f_{k_r} \circ f_{k_{r-1}} \circ \cdots \circ f_{k_1},$$

où $r \in \mathbb{N}^*$ est le nombre d'itérations de la fonction et f_{k_i} est la fonction itérée. Dans toute la suite, le nombre r de fonctions est appelé nombre de *tours* ou d'étages.

Selon les principes énoncés par Claude Shannon [Sha49], la fonction interne f_{k_i} doit assurer la *confusion* et la *diffusion* :

Confusion : faire disparaître les structures linéaires et algébriques du chiffrement ;

Diffusion : faire en sorte que chaque bit de la sortie soit influencé par le plus grand nombre possible de bits de l'entrée et de la clé.

Il y a deux familles de chiffrements itératifs par blocs : les réseaux de substitution-permutation et les schémas de Feistel.

1.2.2 Réseaux de substitution-permutation

Pour minimiser le coût d'implémentation d'un chiffrement itératif, la sous-clé est souvent insérée dans le chiffrement à la fin de chaque itération au moyen d'une simple addition dans $\mathbf{F}_2^n$: on parle alors de *chiffrement alternant clés et permutations* (*cf.* figure 1.3).

Une première sous-clé k_0 est additionnée au texte clair P au début du chiffrement, sinon la valeur $f(P)$ à la sortie du premier tour serait calculable sans connaissance de la clé, donc la première application de la fonction de tour f n'apporterait aucune sécurité supplémentaire mais diminuerait les performances du chiffrement.

Un cas particulier de chiffrement alternant clés et permutations est celui des *réseaux de substitution-permutation* (ou SPN), où la fonction de tour f suit à la lettre les critères de Shannon : elle est la composée d'une permutation linéaire M (assurant la diffusion

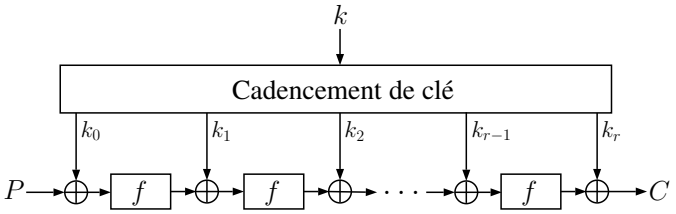


FIGURE 1.3 – Chiffrement itératif par blocs alternant clés et permutations.

de l'information) et d'une fonction de substitution Sub (assurant la confusion),

$$f(x) = M(\mathsf{Sub}(x)).$$

Cette structure ne garantit pas l'inversibilité, donc chaque composante de la fonction interne doit elle-même être inversible. Le déchiffrement est effectué en itérant les inverses de chacune des composantes.

L'implémentation d'une fonction de grande taille qui n'est pas linéaire est très coûteux, c'est pourquoi la fonction de substitution Sub consiste généralement en l'application en parallèle d'une fonction S opérant sur un ensemble plus petit, appelée *boîte-S*.

Dans ce document, nous utiliserons la notation suivante pour décrire les paramètres d'un réseau de substitution-permutation.

Notation 1.2. Soient m et t deux entiers strictement positifs. Soit S une permutation de $\mathbf{F}_2^m$ et M une permutation linéaire de $\mathbf{F}_2^{mt}$. On note $\mathrm{SPN}(m, t, S, M)$ le réseau de substitution-permutation défini sur $\mathbf{F}_2^{mt}$ dont la fonction de substitution est composée de t applications de S et dont la fonction de diffusion est M.

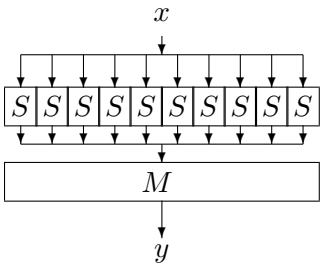


FIGURE 1.4 – Fonction interne d'un chiffrement SPN, ici de la forme $SPN(m, 10, S, M)$.

Le standard de chiffrement symétrique, l'AES, est un réseau de substitution-permutation.

Le chiffrement AES

L'AES a été conçu par Joan Daemen et Vincent Rijmen sous le nom de Rijndael [DR00]. En 2000, le chiffrement Rijndael a remporté la compétition publique organisée par le NIST pour définir un nouveau standard de chiffrement par blocs, puis a été normalisé [FIP01] et renommé AES (Advanced Encryption Standard).

La taille d'un bloc est $n = 128$ et il y a trois tailles de clé possibles : 128 bits (AES-128), 192 bits (AES-192) et 256 bits (AES-256). L'AES suit la construction $\mathrm{SPN}(8, 16, S, M)$ sur 10 tours pour AES-128, 12 tours pour AES-192 et 14 tours pour AES-256. Généralement, l'état interne de l'AES est représenté sous forme d'une matrice de 4×4 octets, notés de x_0 à x_{15} (voir figure 1.5).

x_0	x_4	x_8	x_{12}
x_1	x_5	x_9	x_{13}
x_2	x_6	x_{10}	x_{14}
x_3	x_7	x_{11}	x_{15}

FIGURE 1.5 – État interne de l'AES.

La fonction non-linéaire dans l'AES (l'étage de boîtes-S) est appelée SubBytes et notée Sub. La boîte-S s'applique à chaque octet.

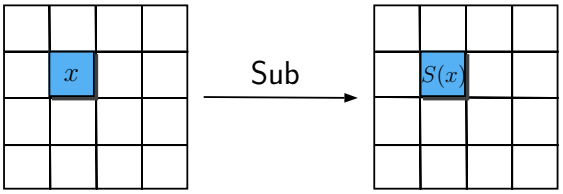


FIGURE 1.6 – La fonction SubBytes de l'AES.

Elle est obtenue à partir d'une fonction dans le corps fini $\mathbf{F}_{2^8}$. L'isomorphisme entre l'espace vectoriel $\mathbf{F}_2^8$ et le corps $\mathbf{F}_{2^8}$ est donné par

$$\begin{array}{rccc} \varphi : & \mathbf{F}_2^8 & \longrightarrow & \mathbf{F}_{2^8} \\ & (x_0, \ldots, x_7) & \longmapsto & \sum_{i=0}^{7} x_i X^i \end{array}$$

où les opérations sont effectuées modulo le polynôme irréductible

$$X^8 + X^4 + X^3 + X + 1.$$

Avec cette identification, la boîte-S de l'AES correspond à la composée de la fonction inverse du corps $\mathbf{F}_{2^8}$, c'est-à-dire la fonction

$$\begin{array}{ccc} \mathbf{F}_{2^8} & \longrightarrow & \mathbf{F}_{2^8} \\ x & \longmapsto & x^{254} \end{array},$$

avec une permutation affine sur $\mathbf{F}_2^8$: $S(x) = A \circ \varphi^{-1}([\varphi(x)]^{254})$, où la permutation A

est définie par

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} = \begin{bmatrix} 1\,0\,0\,0\,1\,1\,1\,1 \\ 1\,1\,0\,0\,0\,1\,1\,1 \\ 1\,1\,1\,0\,0\,0\,1\,1 \\ 1\,1\,1\,1\,0\,0\,0\,1 \\ 1\,1\,1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,1\,1\,0\,0 \\ 0\,0\,1\,1\,1\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1\,1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

La permutation linéaire M, vue comme une fonction de $(\mathbf{F}_2^8)^{16}$, est la composition de deux fonctions nommées ShiftRows et MixColumns. La fonction ShiftRows, notée SR, correspond à une permutation des 16 coordonnées de son entrée. Lorsque l'état interne est représenté par une matrice, la fonction ShiftRows correspond à une rotation des lignes de cette matrice : la ligne i est décalée de i positions vers la gauche, $0 \leqslant i \leqslant 3$.

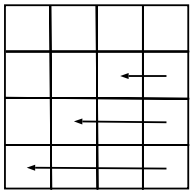


FIGURE 1.7 – La fonction ShiftRows de l'AES.

La fonction MixColumns, notée MC, transforme $(x_0, \ldots, x_{15})$ en appliquant en parallèle à chaque quadruplet $(x_{4i}, \ldots, x_{4i+3})$ une permutation M_c de $(\mathbf{F}_2^8)^4$. Lorsque l'état interne est représenté sous forme de matrice, la permutation M_c s'applique à chaque colonne de cette matrice. La permutation M_c correspond à une permutation de $(\mathbf{F}_{2^8})^4$ linéaire sur $\mathbf{F}_{2^8}$, où le corps $\mathbf{F}_{2^8}$ est identifié à $\mathbf{F}_2^8$ par l'isomorphisme φ défini précédemment, et dont la matrice est :

$$\begin{pmatrix} \alpha & \alpha+1 & 1 & 1 \\ 1 & \alpha & \alpha+1 & 1 \\ 1 & 1 & \alpha & \alpha+1 \\ \alpha+1 & 1 & 1 & \alpha \end{pmatrix},$$

où α est une racine du polynôme irréductible $X^8 + X^4 + X^3 + X + 1$.

L'ordre des fonctions dans un tour est le suivant : SubBytes, ShiftRows, MixColumns et l'addition de sous-clé AddKey. Il y a une addition de sous-clé supplémentaire avant le premier tour (comme dans tout chiffrement alternant clés et permutations) et la fonction MixColumns du dernier tour n'est pas effectuée puisqu'elle n'a aucun effet sur les propriétés cryptographiques du chiffrement.

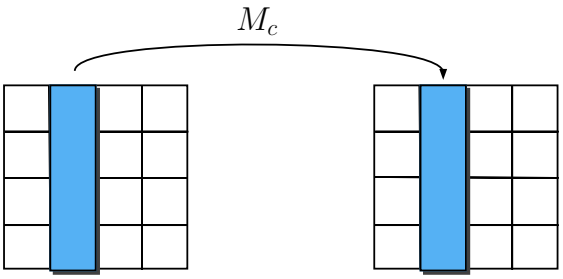


FIGURE 1.8 – La fonction MixColumns de l'AES.

Superboîte-S

Grâce aux propriétés des composants de l'AES, il existe une deuxième représentation de ce chiffrement comme réseau de substitution-permutation, qui conduit à la notion de *superboîte-S* [DR06]. Les fonctions ShiftRows et SubBytes agissant indépendamment sur les octets, elles commutent. En utilisant cette commutativité, deux tours de l'AES peuvent s'écrire :

$$\mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR} \circ \mathrm{SB} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SB} \circ \mathrm{SR}.$$

Les fonctions SubBytes et AddKey agissant sur des octets et MixColumns sur les quatre octets d'une colonne de l'état interne, la fonction composée SB ∘ AK ∘ MC ∘ SB agit indépendamment sur chaque groupe de 4 octets, c'est-à-dire sur chaque colonne de la matrice de l'état interne (*cf* figure 1.9). Cette fonction peut donc être vue comme l'application en parallèle de quatre permutations sur $(\mathbf{F}_2^{32})^4$, appelées superboîtes-S et paramétrées par des clés différentes (*cf* figure 1.10).

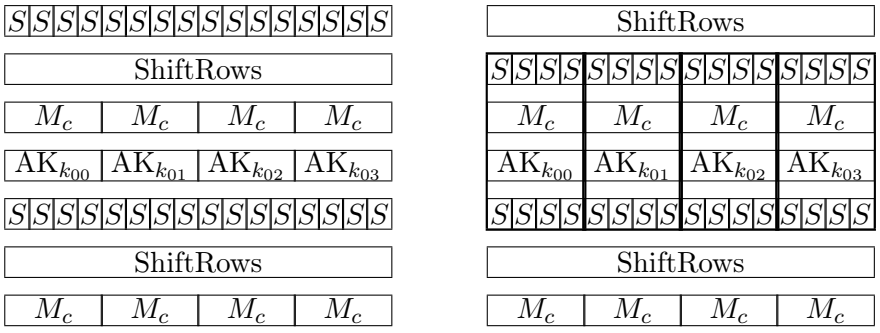


FIGURE 1.9 – Les deux représentations équivalentes de deux tours de l'AES (sans l'addition de sous-clé du deuxième tour) : la représentation usuelle à gauche et la représentation avec les superboîtes-S à droite.

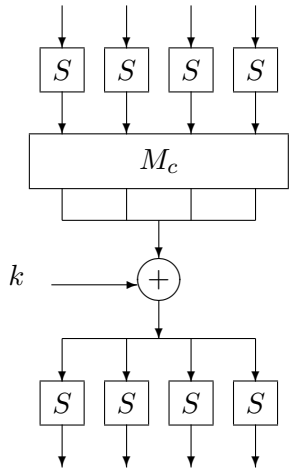


FIGURE 1.10 – La superboîte-S de l'AES.

Les fonctions qui sont appliquées après l'étage de superboîtes-S sont toutes linéaires. Donc l'AES est un réseau de substitution-permutation de la forme

$$\mathrm{SPN}(32, 4, \mathrm{SBS}, \mathrm{SR} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR}),$$

où la superboîte-S est notée SBS (la dernière fonction ShiftRows provient du tour suivant, cette fonction étant commutée avec la fonction SubBytes).

1.2.3 Schéma de Feistel

L'autre construction très utilisée pour un chiffrement par blocs, alternative au réseau de substitution-permutation, est la construction dite de Feistel introduite dans les années 1970 par Horst Feistel pour la conception chiffrement Lucifer [Fei74], le prédécesseur du DES (Data Encryption Standard) [FIP77] qui fut le standard de chiffrement symétrique qui a précédé l'AES. Dans un schéma de Feistel, les bits d'un bloc de message sont séparés en deux parties de même taille. La fonction de tour (*cf* figure 1.11) est la suivante :

$$\begin{array}{rcl} \mathbf{F}_2^n \times \mathbf{F}_2^n & \longrightarrow & \mathbf{F}_2^n \times \mathbf{F}_2^n \\ (L_i, R_i) & \longmapsto & (L_{i+1}, R_{i+1}) = (R_i + f(L_i, k_i), L_i). \end{array}$$

Cette structure étant bijective, la fonction f n'a donc pas besoin de l'être. La fonction f doit assurer la confusion et la diffusion, l'addition dans $\mathbf{F}_2^n$ (l'opération OU exclusif) et la transposition des deux moitiés de bloc assurant également la diffusion. La fonction de tour étant une involution (à la transposition des deux moitiés de bloc près), le déchiffrement consiste à appliquer les mêmes opérations que pour le chiffrement en inversant l'ordre des sous-clés k_i.

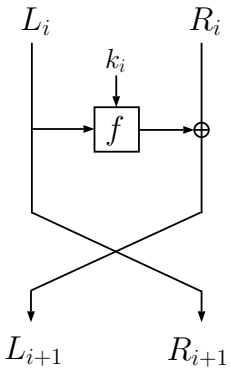


FIGURE 1.11 – La fonction de tour d'un schéma de Feistel.

1.3 Codes correcteurs

Nous rappelons ici quelques définitions et résultats classiques sur les codes correcteurs d'erreurs [MS77, Cha98] qui vont nous être utiles pour analyser les propriétés de diffusion dans les chiffrements par blocs.

1.3.1 Propriétés générales

Lorsqu'un message est envoyé à travers un canal, ce message peut être modifié suite à des interférences ou du « bruit » sur le canal. Pour pouvoir retrouver le message original malgré la présence d'erreurs, une idée est d'ajouter à ce message de la redondance, c'est-à-dire de répéter des informations contenues dans ce message. Si le nombre d'erreurs dues au passage par le canal n'est pas trop élevé, le récepteur doit pouvoir, grâce à cette redondance, repérer s'il y a des erreurs dans le message reçu et les corriger. Un exemple simple mais non optimal consiste à envoyer le message deux fois.

Le codage en blocs consiste à remplacer le message initial $m = (m_1, \ldots, m_k) \in X^k$, où X est un ensemble et k un entier positif, par un *mot de code* $x = (x_1, \ldots, x_n) \in X^n$, où n est un entier plus grand que k. Les mots de code sont tous composés du même nombre n d'éléments d'un ensemble X (appelé *alphabet*) et l'ensemble de ces mots est appelé *code*, de longueur n. Un code doit vérifier au mieux les deux propriétés suivantes :

- détection d'erreurs : déterminer si le message reçu est égal au message envoyé x ;
- correction d'erreurs : le message obtenu après détection de l'erreur et décodage doit correspondre au message original m.

Une notion très importante est la distance minimale d'un code.

Définition 1.3. *Soit $\mathcal{C}$ un code de longueur n sur un alphabet X.*

- *La* distance de Hamming *du mot a au mot b est le nombre de coordonnées de a et b qui sont distinctes :*

$$d_H(a, b) = \#\{i \in [1; n] \,|\, a_i \neq b_i\}.$$

La distance de Hamming est une distance sur X^n.

– *La* distance minimale d *de* $\mathcal{C}$ *est la plus petite distance de Hamming possible entre deux mots distincts du code :*

$$d = \min\{d_H(a,b) \mid a,b \in \mathcal{C}, a \neq b\}.$$

Le nombre d'erreurs qui peuvent être détectées et le nombre d'erreurs qui peuvent être corrigées dépendent de la distance minimale du code. Si $\mathcal{C}$ est un code de distance minimale d, alors $\mathcal{C}$ peut détecter $d-1$ erreurs et en corriger $\lfloor \frac{d-1}{2} \rfloor$.

1.3.2 Codes linéaires

Une façon courante de concevoir un code est d'utiliser un sous-espace vectoriel. Le code est alors dit linéaire. Notons $\mathbf{F}$ un corps fini et n un entier strictement positif.

Définition 1.4. *Un* code linéaire $\mathcal{C}$ *sur* $\mathbf{F}$ *est un* $\mathbf{F}$*-sous-espace vectoriel de* $\mathbf{F}^n$*. On dit que* $\mathcal{C}$ *est un code* $[n,k,d]$*, où* n *est sa longueur,* k *est sa dimension et* d *sa distance minimale.*

Le code $\mathcal{C}$ peut alors être indifféremment représenté par une des matrices suivantes :
– une *matrice génératrice* G est une matrice de taille $k \times n$ dont les lignes forment une base du code $\mathcal{C}$. Pour tout message $m \in \mathbf{F}^k$, le mot de code associé est $x = mG$.
On dit que G est sous *forme systématique* si $G = [I_k, A]$, où I_k est la matrice identité de taille $k \times k$ et A est une matrice de taille $k \times (n-k)$. Dans ce cas, le mot de code associé à un message m vérifie $x = mG = (m_1, \ldots, m_k, x_{k+1}, \ldots, x_n)$, c'est-à-dire qu'il est formé du message concaténé avec $n-k$ éléments dépendant du message.
– une *matrice de parité* (ou *matrice de contrôle*) H est une matrice $(n-k) \times n$ telle qu'un mot $x \in \mathbf{F}^n$ appartient au code si et seulement si $H^t x = 0$. Cette matrice vérifie : $H^t G = 0$.

Proposition 1.5. *[MS77, chap.1] Si la matrice génératrice* G *de* $\mathcal{C}$ *a pour forme systématique* $G = [I_k, A]$*, où* A *est une matrice* $k \times (n-k)$*, alors une matrice de parité* H *de* $\mathcal{C}$ *est* $H = [{}^tA, -I_{n-k}]$.

Définition 1.6. *Le* code dual *de* $\mathcal{C}$ *est défini comme l'ensemble des vecteurs qui sont orthogonaux aux mots du code* $\mathcal{C}$ *:*

$$\mathcal{C}^\perp = \{y \in \mathbf{F}^n \mid \forall x \in \mathcal{C},\ x \cdot y = 0\},$$

où $x \cdot y = \sum_{i=1}^{n} x_i y_i$ *est le produit scalaire canonique de* $\mathbf{F}^n$.

Le code $\mathcal{C}^\perp$ est un code linéaire $[n, n-k, d']$ dont une matrice génératrice est la matrice de parité H de $\mathcal{C}$.

La définition suivante présente des notions qui seront utilisées tout au long ce document, pour des codes linéaires et d'autres structures où l'utilisation du vocabulaire des codes est intéressante.

Définition 1.7. *Soit $\mathcal{C}$ un code linéaire de paramètres $[n, k, d]$.*

- *Le* support *d'un mot de code $x \in \mathcal{C}$ est l'ensemble des indices des symboles non nuls :*

$$\mathsf{Supp}(x) = \{i \in [1; n] \,|\, x_i \neq 0\}.$$

- *Le* poids *d'un mot de code $x \in \mathcal{C}$ est le nombre de coordonnées non nulles du mot x :*

$$wt(x) = \#\{i \in [1; n] \,|\, x_i \neq 0\}.$$

  Le poids d'un mot correspond au cardinal de son support.
- *La* distribution des poids *du code $\mathcal{C}$ est l'ensemble des valeurs*

$$A_i = \#\{x \in \mathcal{C} \,|\, wt(x) = i\}$$

  pour $0 \leqslant i \leqslant n$.

Proposition 1.8. *[MS77, chap.1] La distance minimale d'un code linéaire est égale au plus petit poids non nul de ce code.*

Remarque 1.9. La notion de poids d'un mot est aussi définie pour les codes non linéaires et la proposition 1.8 est valable pour une famille de codes plus grande que les codes linéaires : les codes additifs. Un code $\mathcal{C}$ inclus dans un ensemble muni d'une loi notée « + » est dit *additif* si pour tous mots c_1 et c_2 de $\mathcal{C}$, $c_1 + c_2$ est aussi un mot de $\mathcal{C}$. Pour un tel code, la distance minimale est égale au plus petit poids non nul de ce code.

Une autre information intéressante sur un code est le nombre de mots de code ayant un poids donné. La distribution des poids d'un code est souvent représentée par un polynôme à deux indéterminées.

Définition 1.10. *Le* polynôme des poids *d'un code $\mathcal{C}$ est :*

$$W_{\mathcal{C}}(x, y) = \sum_{i=0}^{n} A_i x^{n-i} y^i,$$

où $A_i = \#\{x \in \mathcal{C} \,|\, wt(x) = i\} \in \mathbb{N}$.

Les polynômes des poids de $\mathcal{C}$ et de son dual $\mathcal{C}^\perp$ sont liés par la relation suivante.

Théorème 1.11. *[Mac63, chap.5] Soit p un nombre premier et m un entier strictement positif. Soit $\mathcal{C}$ un code linéaire de longueur n sur $\mathbf{F}_{p^m}$. Alors les polynômes des poids de $\mathcal{C}$ et de son dual $\mathcal{C}^\perp$ vérifient :*

$$W_{\mathcal{C}^\perp}(x, y) = \frac{1}{\#\mathcal{C}} W_{\mathcal{C}}(x + (p^m - 1)y, x - y).$$

1.3.3 Codes MDS

La distance minimale d'un code permet de savoir combien d'erreurs peuvent être détectées et corrigées. Cette notion a donc été beaucoup étudiée. En particulier, la valeur maximale pour la distance minimale d'un code s'exprime en fonction des paramètres de ce code. Soit p un nombre premier et m un entier strictement positif. Dans la suite de cette section, on note $q = p^m$.

Théorème 1.12 (Borne de Singleton [Sin64])**.** *Si $\mathcal{C}$ est un code de longueur n et de taille q^k sur un alphabet de taille q, alors sa distance minimale d sur cet alphabet vérifie : $d \leqslant n - k + 1$.*

Les codes qui vérifient $d = n - k + 1$ sont appelés *Maximum Distance Separable* (MDS). Ce sont les codes qui ont la plus grande distance minimale possible parmi l'ensemble des codes ayant la même longueur et la même taille.

Lorsque le code est linéaire, plusieurs critères ont été énoncés pour montrer qu'un code est MDS.

Proposition 1.13. *[MS77, chap.11] Pour un code $\mathcal{C}$ linéaire, les assertions suivantes sont équivalentes :*

- *Le code $\mathcal{C}$ est MDS ;*
- *Chaque ensemble de $n - k$ colonnes de la matrice de parité H de $\mathcal{C}$ est de rang $n - k$;*
- *Le code dual $\mathcal{C}^\perp$ de $\mathcal{C}$ est MDS ;*
- *Chaque ensemble de k colonnes de la matrice génératrice G de $\mathcal{C}$ est de rang k ;*
- *Toute sous-matrice carrée de A est inversible, A étant la matrice $k \times (n-k)$ telle que la matrice génératrice G de $\mathcal{C}$ s'écrit $G = [I_k, A]$.*

Dans le cas où le code est MDS, la distribution des poids est connue.

Théorème 1.14. *[MS77, chap.11] Si un code $[n, k, d]$ $\mathcal{C}$ est MDS sur $\mathbf{F}_q$, alors le nombre de mots de $\mathcal{C}$ de poids w est :*

$$A_w = \binom{n}{w} (q-1) \sum_{j=0}^{w-d} (-1)^j \binom{w-1}{j} q^{w-d-j},$$

où $w \in [1; n]$ représente les différents poids possibles du code $\mathcal{C}$.

Les codes MDS n'existent que pour certaines classes de paramètres : lorsque $n \leqslant q + 1$, des familles de codes MDS sont connues pour tout n et pour tout $k \leqslant n$ (par exemple les codes GRS et GRS étendus [MS77]). Lorsque $n > q + 1$, il est conjecturé qu'il n'existe pas de code MDS, sauf exception.

Conjecture MDS. *[Seg55] Un code $[n, k, d]$ MDS sur $\mathbf{F}_{2^m}$ ($q = 2^m$) avec $k = 3$ ou $k = q - 1$ vérifie $n \leqslant q + 2$. Tout autre code $[n, k, d]$ MDS sur $\mathbf{F}_q$ (q quelconque) vérifie $n \leqslant q + 1$.*

Cette conjecture a été démontrée dans deux cas : lorsque q est un nombre premier et $k \leqslant q$, lorsque $q = p^m$, p étant un nombre premier et $m > 1$, et $k \leqslant 2p - 2$ [BDB12].

1.3.4 Codes GRS

Les codes GRS sont une famille infinie de codes MDS.

Définition 1.15. *Soit n un entier inférieur à $\#\mathbf{F}$ et k un entier compris entre 0 et n. Soient $\alpha_1, \ldots, \alpha_n$ n éléments distincts de $\mathbf{F}$ et $v_1, \ldots, v_n$ n éléments non nuls de $\mathbf{F}$. Le* code de Reed-Solomon Généralisé *$GRS_{n,k}(\alpha, v)$ est l'ensemble des vecteurs de la forme $(v_1 P(\alpha_1), \ldots, v_n P(\alpha_n))$ pour tout polynôme P de $\mathbf{F}[x]$ de degré strictement inférieur à k :*

$$GRS_{n,k} = \{(v_1 P(\alpha_1), \ldots, v_n P(\alpha_n)) \,|\, P(x) \in \mathbf{F}[x]_{k-1}\}.$$

Théorème 1.16. *[MS77, chap.10] Un code de Reed-Solomon Généralisé $GRS_{n,k}(\alpha, v)$ est un code $[n, k]$-linéaire sur $\mathbf{F}$. Si $k \neq 0$, ce code est MDS.*

Une matrice génératrice du code $\mathrm{GRS}_{n,k}(\alpha, v)$ est :

$$\begin{pmatrix} v_1 & v_2 & \ldots & v_n \\ v_1\alpha_1 & v_2\alpha_2 & \ldots & v_n\alpha_n \\ \vdots & \vdots & & \vdots \\ v_1\alpha_1^i & v_2\alpha_2^i & \ldots & v_n\alpha_n^i \\ \vdots & \vdots & & \vdots \\ v_1\alpha_1^{k-1} & v_2\alpha_2^{k-1} & \ldots & v_n\alpha_n^{k-1} \end{pmatrix}.$$

Théorème 1.17. *[MS77, chap.10] Le dual d'un code GRS est un code GRS. Plus exactement, le code dual du code $GRS_{n,k}(\alpha, v)$ est le code $GRS_{n,n-k}(\alpha, u)$ avec*

$$u_i^{-1} = v_i \prod_{j \neq i} (\alpha_i - \alpha_j).$$

1.4 Fonctions booléennes

Les fonctions booléennes apparaissent également dans l'étude des chiffrements par blocs. Nous rappelons ici leurs principales propriétés. Ces résultats sont connus et leurs démonstrations se trouvent par exemple dans [Car10].

1.4.1 Propriétés générales

Soit n un nombre entier strictement positif.

Définition 1.18. *On appelle* fonction booléenne à n variables *toute application f de $\mathbf{F}_2^n$ à valeurs dans $\mathbf{F}_2$. L'ensemble des fonctions booléennes à n variables est noté $\mathbf{B}_n$. Son cardinal est 2^{2^n}.*

Une fonction booléenne peut être représentée de manière unique par 2^n bits : son vecteur des valeurs.

Définition 1.19. *Le* vecteur des valeurs *d'une fonction booléenne f à n variables est un vecteur dont les coordonnées sont les images $f(x)$ des 2^n éléments x de $\mathbf{F}_2^n$ par la fonction f.*

Cette représentation est unique lorsque l'ordre des éléments de $\mathbf{F}_2^n$ est choisi. L'ensemble des fonctions booléennes à n variables est donc en bijection avec $\mathbf{F}_2^{2^n}$. Au vu de ce lien, il est intéressant d'utiliser le vocabulaire des codes.

Définition 1.20. *Soient f et g deux fonctions booléennes à n variables.*
- *Le* support *de f est l'ensemble des vecteurs de $\mathbf{F}_2^n$ qui ont une image non nulle par la fonction f :* $\mathsf{Supp}(f) = \{x \in \mathbf{F}_2^n \mid f(x) \neq 0\}$.
- *Le* poids de Hamming *de f est le cardinal du support de f :* $wt(f) = \#\mathsf{Supp}(f)$.
- *La* distance de Hamming *de f à g est le poids de Hamming de la somme des vecteurs de valeurs de f et g :* $d_H(f,g) = \{x \in \mathbf{F}_2^n \mid f(x) \neq g(x)\}$.

Notons le produit scalaire canonique de deux éléments $x = (x_0, \ldots, x_{n-1})$ et $y = (y_0, \ldots, y_{n-1})$ de $\mathbf{F}_2^n$ par

$$x \cdot y = \sum_{i=0}^{n-1} x_i y_i,$$

où le signe somme correspond à l'addition dans $\mathbf{F}_2$.

Une deuxième représentation d'une fonction booléenne, par un polynôme multivarié, est appelée forme algébrique normale.

Théorème 1.21. *Toute fonction booléenne f à n variables peut être représentée par un unique polynôme multivarié dans $\mathbf{F}_2[x_0, \ldots, x_{n-1}]/(x_0^2 + x_0, \ldots, x_{n-1}^2 + x_{n-1})$ défini par :*

$$f(x) = \sum_{u \in \mathbf{F}_2^n} f_u x^u = \sum_{u \in \mathbf{F}_2^n} f_u \left(\prod_{i=0}^{n-1} x_i^{u_j} \right).$$

Cette représentation est appelée forme algébrique normale *(ANF) de f.*

Notation 1.22. L'inclusion $\mathsf{Supp}(u) \subseteq \mathsf{Supp}(x)$ est notée $u \preceq x$, où u et x appartiennent à $\mathbf{F}_2^n$. Cette inclusion est vérifiée lorsque $u_i \leqslant x_i$ pour tout $i \in \{0, \ldots, n-1\}$.

Définition 1.23. *La* transformée de Möbius *de f est définie par*

$$\begin{array}{rccc} f^\circ : & \mathbf{F}_2^n & \longrightarrow & \mathbf{F}_2 \\ & u & \longmapsto & \displaystyle\sum_{v \preceq u} f(v) \ . \end{array}$$

Proposition 1.24 (Relation entre l'ANF et la transformée de Möbius de f)**.** *Soit f une fonction booléenne à n variables dont l'ANF est $f(x) = \sum_{u \in \mathbf{F}_2^n} f_u x^u$ et dont la fonction de Möbius est f°. Alors on a :*

$$f_u = f^\circ(u) = \sum_{v \preceq u} f(v) \text{ et } f(v) = \sum_{u \preceq v} f^\circ(u) = \sum_{u \preceq v} f_u.$$

Avec cette représentation, il est possible de définir le degré algébrique d'une fonction booléenne.

Définition 1.25. *Le degré de l'ANF est noté* $\deg(f)$ *et appelé* degré algébrique *de f :*

$$\deg(f) = \max_{u \in \mathbf{F}_2^n} \{wt(u) | f^\circ(u) \neq 0\}.$$

La fonction booléenne f est dite *constante* si $\deg(f) = 0$, *affine* si $\deg(f) = 1$ et *quadratique* si $\deg(f) = 2$. Les fonctions booléennes affines à n variables sont de la forme $f(x) = a \cdot x + c$, où $a \in \mathbf{F}_2^n$ et $c \in \mathbf{F}_2$. Pour tout $a \in \mathbf{F}_2^n$, la fonction linéaire $x \to a \cdot x$ est notée φ_a.

Définition 1.26. *La* corrélation $\mathrm{C}(f)$ *d'une fonction booléenne à n variables f est définie par :*

$$\mathrm{C}(f) = \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)}.$$

Lemme 1.27. *Soit $a \in \mathbf{F}_2^n$. On a :*

$$\mathrm{C}(\varphi_a) = \sum_{x \in \mathbf{F}_2^n} (-1)^{a \cdot x} = \begin{cases} 2^n & \text{si } a = 0 \\ 0 & \text{sinon.} \end{cases}$$

Démonstration. Pour $a = 0$, le résultat est évident. Pour $a \neq 0$, l'application $x \mapsto a \cdot x$ est une forme linéaire sur $\mathbf{F}_2^n$. Son noyau est un hyperplan de $\mathbf{F}_2^n$, c'est-à-dire un sous-espace de $\mathbf{F}_2^n$ de dimension $n-1$. Il y a donc 2^{n-1} éléments x de $\mathbf{F}_2^n$ tels que $a \cdot x = 0$ et $2^n - 2^{n-1} = 2^{n-1}$ éléments x de $\mathbf{F}_2^n$ tels que $a \cdot x = 1$. D'où $\sum_{x \in \mathbf{F}_2^n} (-1)^{a \cdot x} = 0$. □

Définition 1.28. *Une fonction booléenne à n variables est dite* équilibrée *si son image contient autant de* 0 *que de* 1*, c'est-à-dire si $wt(f) = 2^{n-1}$ ou, de manière équivalente, si* $\mathrm{C}(f) = 0$.

1.4.2 Non-linéarité

La notion de distance (au sens de Hamming) entre une fonction booléenne et l'ensemble des fonctions booléennes affines est importante en cryptographie symétrique. Cette notion indique s'il existe une fonction affine qui est une bonne approximation de la fonction booléenne considérée. La cryptanalyse linéaire, une des attaques les plus célèbres sur les chiffrements par blocs, utilise l'existence de ce type d'approximations. La distance d'une fonction booléenne aux fonctions affines correspond à sa transformée de Walsh.

Définition 1.29. *Soit f une fonction booléenne à n variables. La* transformée de Walsh *de f est la fonction*

$$\begin{array}{rccl} \mathcal{W}^f : & \mathbf{F}_2^n & \to & \mathbb{Z} \\ & u & \mapsto & \mathcal{W}^f(u) = \displaystyle\sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)+u \cdot x}. \end{array}$$

Pour un élément $u \in \mathbf{F}_2^n$, $\mathcal{W}^f(u)$ est appelé coefficient de Walsh *de f au point u. L'amplitude maximale des coefficients de Walsh de f est appelée* linéarité *et est notée $\mathcal{L}(f)$:*

$$\mathcal{L}(f) = \max_{u \neq 0} \left| \mathcal{W}^f(u) \right|.$$

De façon équivalente, on définit souvent la non-linéarité d'une fonction booléenne.

Définition 1.30. *Soit f une fonction booléenne à n variables. La* non-linéarité *de f est définie par :*

$$\mathcal{NL}(f) = 2^{n-1} - \frac{1}{2}\,\mathcal{L}(f).$$

La transformée de Walsh étant une transformée de Fourier discrète, elle vérifie l'ensemble des propriétés mathématiques d'une transformée de Fourier, en particulier la relation de Parseval.

Théorème 1.31 (Relation de Parseval)**.** *Soit f une fonction booléenne à n variables. Alors les coefficients de Walsh vérifient :*

$$\sum_{u \in \mathbf{F}_2^n} (\mathcal{W}^f(u))^2 = 2^{2n}.$$

Pour toute fonction booléenne, on a $\mathcal{L}(f) \geqslant 2^{n/2}$ [Rot76]. Les fonctions qui atteignent la borne, donc qui ont la plus petite linéarité possible, sont appelées fonctions *courbes*. Elles n'existent que lorsque n est pair. Si une fonction booléenne est courbe, sa transformée de Walsh ne prend que deux valeurs : $\pm 2^{n/2}$ [Rot76]. La corrélation $\mathrm{C}(f)$, qui est égale au coefficient de Walsh de f au point 0, $\mathcal{W}^f(0)$, est non nulle, c'est-à-dire que les fonctions courbes ne sont pas équilibrées. Elles sont donc rarement utilisées en cryptographie. Pour qu'une fonction booléenne ait de bonnes propriétés cryptographiques, il faut que sa linéarité soit aussi proche de cette borne que possible.

Définition 1.32. *[ZZ99] Une fonction f est dite* plateau *si ses coefficients de Walsh prennent au plus trois valeurs : 0 et $\pm\mathcal{L}(f)$.*

Proposition 1.33. *[ZZ99] Si f est une fonction plateau, alors $\mathcal{L}(f) = 2^s$ avec $s \geqslant n/2$.*

Donc les coefficients de Walsh d'une fonction plateau sont divisibles par $2^{n/2}$ si n est pair et par $2^{(n+1)/2}$ si n est impair. Les fonctions courbes sont les fonctions plateaux qui vérifient $s = n/2$.

1.5 Fonctions booléennes vectorielles

1.5.1 Propriétés générales

Définition 1.34. *Une* fonction booléenne vectorielle *à n variables et m coordonnées est un produit cartésien de m fonctions booléennes :*

$$\begin{array}{rccc} F : & \mathbf{F}_2^n & \to & \mathbf{F}_2^m \\ & (x_0, \ldots, x_{n-1}) & \mapsto & (f_0(x_0, \ldots, x_{n-1}), \ldots, f_{m-1}(x_0, \ldots, x_{n-1})), \end{array}$$

où les fonctions $f_0, \ldots, f_{m-1}$ sont des fonctions booléennes à n variables. Elles sont appelées fonctions coordonnées *de F.*

Les fonctions booléennes correspondent au cas où $m = 1$.

Définition 1.35. *Soit F une fonction booléenne vectorielle à n variables et m coordonnées. Pour tout $\lambda \in \mathbf{F}_2^m$, les fonctions de la forme*

$$\begin{array}{rccc} F_\lambda : & \mathbf{F}_2^n & \to & \mathbf{F}_2 \\ & x & \mapsto & \lambda \cdot F(x) \end{array}$$

sont appelées fonctions composantes *de F.*

Les fonctions composantes correspondent aux combinaisons linéaires dans $\mathbf{F}_2^m$ des fonctions coordonnées de F. La fonction composante F_0 (cas où $\lambda = 0$) est la fonction nulle.

Définition 1.36. *Le* degré algébrique *d'une fonction booléenne vectorielle à n variables et m coordonnées F est égal au maximum des degrés algébriques de ces fonctions coordonnées $f_0, \ldots, f_{m-1}$:*

$$\deg(F) = \max_{0 \leqslant i < m} \deg(f_i).$$

La définition suivante présente une relation d'équivalence sur l'ensemble des fonctions booléennes vectorielles. L'intérêt de cette relation, dite équivalence affine, est qu'une partie des propriétés cryptographiques des fonctions booléennes vectorielles sont invariantes sur les classes d'équivalence affine.

Définition 1.37. *Soient F_1 et F_2 deux fonctions booléennes vectorielles à n variables et n coordonnées. Les fonctions F_1 et F_2 sont dites* affinement équivalentes *s'il existe deux permutations affines A_1 et A_2 telles que : $F_2 = A_2 \circ F_1 \circ A_1$.*

En particulier, nous allons voir dans les sections suivantes que c'est le cas pour deux des principales propriétés cryptographiques des fonctions booléennes vectorielles : l'uniformité différentielle et la non-linéarité de deux permutations affinement équivalentes sont égales.

1.5.2 Uniformité différentielle

Un des deux composants d'un réseau de substitution-permutation est la fonction S, qui est une fonction booléenne vectorielle non linéaire à n variables et n coordonnées. Son rôle est d'assurer la confusion, c'est-à-dire de faire disparaître les liens entre les bits du texte clair et ceux du texte chiffré. C'est donc principalement de cette boîte-S que dépend la résistance à différentes attaques. Dans cette section, nous présentons un des critères qui permet d'estimer si une boîte-S a de bonnes propriétés cryptographiques, appelé uniformité différentielle, qui a été introduit par Kaisa Nyberg [Nyb94].

Définition 1.38. *Soit F une fonction booléenne vectorielle de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^m$, n et m étant deux entiers strictement positifs. Soit a un élément non nul de $\mathbf{F}_2^n$. La* dérivée *de F dans la direction a est la fonction D_aF définie par :*

$$\begin{array}{rccc} D_aF : & \mathbf{F}_2^n & \to & \mathbf{F}_2^m \\ & x & \mapsto & F(x) + F(x+a). \end{array}$$

Définition 1.39 (Uniformité différentielle)**.** *Soit F une fonction de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^m$. Pour tout α et β dans $\mathbf{F}_2^n$, on définit*

$$\delta(\alpha, \beta) = \#\{x \in \mathbf{F}_2^n, F(x+\alpha) + F(x) = \beta\} .$$

Le multi-ensemble $\{\delta(\alpha,\beta),\ \alpha \in \mathbf{F}_2^n \setminus \{0\}, \beta \in \mathbf{F}_2^m\}$ est appelé spectre différentiel *de F, et son maximum*

$$\Delta(F) = \max_{\alpha \neq 0, \beta} \delta(\alpha, \beta)$$

est appelé uniformité différentielle *de F.*

Dans le cas où il pourrait y avoir ambiguïté (plusieurs fonctions sont considérées), la notation $\delta^F(\alpha,\beta)$ (avec en exposant le nom de la fonction) sera utilisée.

Dans la suite, nous nous intéressons à des fonctions booléennes vectorielles à n variables et n coordonnées (*i.e.* $m = n$).

Remarquons que si x est solution de l'équation $F(x+\alpha)+F(x) = \beta$, alors $x+\alpha$ l'est aussi. Les coefficients $(\delta(\alpha,\beta))_{\alpha,\beta\in\mathbf{F}_2^n}$ sont donc pairs et la plus petite valeur possible pour l'uniformité différentielle est 2.

Théorème 1.40. *[NK93] Pour toute fonction $F : \mathbf{F}_2^n \longrightarrow \mathbf{F}_2^n$, $\Delta(F) \geqslant 2$.*

Les fonctions qui ont la plus petite uniformité différentielle possible sont appelées APN.

Définition 1.41. *Une fonction $F : \mathbf{F}_2^n \to \mathbf{F}_2^n$ est dite* APN (Almost Perfect Nonlinear) *si son uniformité différentielle vérifie $\Delta(F) = 2$.*

Les fonctions APN ont fait (et font encore) l'objet de beaucoup de recherches, en particulier lorsqu'elles sont bijectives. Quelques familles de permutations APN sont connues pour des valeurs impaires de n, mais lorsque n est pair, il n'y a qu'une permutation connue (à équivalence affine près) qui soit APN, pour $n = 6$ [BDMW10]. Pour des raisons d'implémentation, n est généralement pair et les boîtes-S utilisées dans les réseaux de substitution-permutation ont donc souvent une uniformité différentielle égale à 4, comme pour l'AES.

Les valeurs $(\delta(\alpha,\beta))_{\alpha,\beta\in\mathbf{F}_2^n}$ sont présentées dans un tableau appelé *table des différences* de F, les lignes correspondant aux valeurs $\alpha \in \mathbf{F}_2^n$, les colonnes correspondant aux valeurs $\beta \in \mathbf{F}_2^n$.

Exemple 1.42. Notons S la fonction de $\mathbf{F}_2^4$ correspondant à la fonction $x \mapsto x^{13}$ dans le corps $\mathbf{F}_{2^4}$, lorsque l'isomorphisme entre l'espace vectoriel $\mathbf{F}_2^4$ et le corps $\mathbf{F}_{2^4}$ est donné par

$$\varphi : \begin{array}{ccc} \mathbf{F}_2^4 & \longrightarrow & \mathbf{F}_{2^4} \\ (x_0, \ldots, x_3) & \longmapsto & \sum_{i=0}^{3} x_i X^i \end{array}$$

où les opérations sont effectuées modulo le polynôme irréductible $X^4 + X + 1$:

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	1	13	11	14	9	6	7	10	4	15	2	8	3	5	12

où les mots de 4 bits sont identifiés aux entiers par leur décomposition en base 2. La figure 1.12 correspond à la table des différences de cette fonction S.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	4	0	0	0	0	2	2	0	2	0	2	0	2	2	0
2	0	0	0	0	0	2	2	0	2	0	2	0	0	4	2	2
3	0	0	0	0	2	0	2	0	2	2	0	4	2	0	0	2
4	0	0	2	0	0	0	0	2	2	0	2	2	2	0	4	0
5	0	0	0	2	0	0	0	2	0	4	2	0	2	2	0	2
6	0	2	2	2	0	0	4	2	2	0	0	0	0	0	0	2
7	0	2	0	0	2	2	2	4	0	0	2	0	2	0	0	0
8	0	0	2	2	0	2	2	0	0	2	4	2	0	0	0	0
9	0	2	2	0	4	0	0	0	0	0	2	2	0	2	0	2
10	0	0	0	2	2	2	0	2	0	0	0	2	0	0	2	4
11	0	2	4	0	0	2	0	0	0	2	0	0	2	0	2	2
12	0	0	2	0	2	2	0	2	4	2	0	0	0	2	0	0
13	0	2	0	4	2	0	0	0	2	2	2	0	0	0	2	0
14	0	2	0	2	0	4	0	0	2	0	0	2	2	2	0	0
15	0	0	2	2	2	0	2	0	0	0	0	0	4	2	2	0

FIGURE 1.12 – Table des différences de la fonction S.

Dans cette table, nous remarquons que $\delta(0,0) = 2^n$. En effet, Pour toute fonction $F : \mathbf{F}_2^n \to \mathbf{F}_2^n$, tout élément $x \in \mathbf{F}_2^n$ est solution de l'équation $F(x+0) + F(x) = 0$. De plus, l'équation $F(x+0)+F(x) = \beta$ n'a pas de solution si β est non nul, donc pour tout élément non nul $\beta \in \mathbf{F}_2^n$, nous avons $\delta(0,\beta) = 0$. Par ailleurs, si F est une permutation, alors pour tout $\alpha \neq 0$, l'équation $F(x+\alpha) + F(x) = 0$ n'a pas de solutions et nous avons $\delta(\alpha, 0) = 0$.

Si F est une fonction booléenne vectorielle, les coefficients du spectre différentiel de F^{-1} et des fonctions affinement équivalentes à F se déduisent de ceux de F.

Proposition 1.43. *[DR06] Soit $F : \mathbf{F}_2^n \to \mathbf{F}_2^n$ une permutation. Alors*

$$\delta^{F^{-1}}(a,b) = \delta^F(b,a).$$

Proposition 1.44. *[DR06] Soient F_1 et F_2 deux fonctions affinement équivalentes, c'est-à-dire qu'il existe deux permutations de $\mathbf{F}_2^n$, $A_1 : x \mapsto L_1(x) + c_1$ et $A_2 : x \mapsto L_2(x) + c_2$, telles que*

$$F_2 = A_2 \circ F_1 \circ A_1.$$

Alors F_1 et F_2 vérifient :

$$\delta^{F_2}(a,b) = \delta^{F_1}(L_1(a), L_2^{-1}(b)).$$

Démonstration. Soient a, b deux éléments de $\mathbf{F}_2^n$. On a :

$$\begin{aligned} F_2(x+a)+F_2(x)=b &\Leftrightarrow A_2\circ F_1(A_1(x+a))+A_2\circ F_1(A_1(x))=b \\ &\Leftrightarrow L_2\circ F_1(L_1(x+a)+c_1)+L_2\circ F_1(L_1(x)+c_1)=b \\ &\Leftrightarrow F_1(L_1(x)+L_1(a)+c_1)+F_1(L_1(x)+c_1)=L_2^{-1}(b) \\ &\Leftrightarrow F_1(y+L_1(a))+F_1(y)=L_2^{-1}(b). \end{aligned}$$

Donc le nombre $\delta^{F_2}(a,b)$ de solutions à l'équation $F_2(x+a)+F_2(x)=b$ est égal au nombre $\delta^{F_1}(L_1(a), L_2^{-1}(b))$ de solutions à l'équation $F_1(y+L_1(a))+F_1(y)=L_2^{-1}(b)$. □

1.5.3 Non-linéarité

Nous présentons ici un deuxième critère pour savoir si une fonction peut être utilisée dans un réseau de substitution-permutation : la non-linéarité. Cette notion est déjà définie pour les fonctions booléennes ; la non-linéarité pour une fonction booléenne vectorielle correspond à la non-linéarité maximale de ses fonctions composantes.

Définition 1.45. *Soit F une fonction de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. Pour tout u et v dans $\mathbf{F}_2^n$, on définit*

$$\mathcal{W}^F(u,v) = \sum_{x\in\mathbf{F}_2^n} (-1)^{u\cdot x+v\cdot F(x)} .$$

L'ensemble $\{\mathcal{W}^F(u,v),\ u\in\mathbf{F}_2^n, v\in\mathbf{F}_2^n\backslash\{0\}\}$ est appelé spectre de Walsh *de F, et son amplitude maximale $\mathcal{L}(F)=\max_{u,v\neq 0}\left|\mathcal{W}^F(u,v)\right|$ est appelée* linéarité *de F.*

Remarquons que la fonction $u\mapsto\mathcal{W}^F(u,v)$ est la transformée de Walsh de la fonction composante F_v de F. Lorsqu'il n'y a pas d'ambiguïté, l'exposant indiquant la fonction qui est considérée sera omis.

La définition de la non-linéarité des fonctions booléennes vectorielles est similaire à celle des fonctions booléennes.

Définition 1.46. *Soit F une fonction de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. La* non-linéarité *de F est définie par :*

$$\mathcal{NL}(F) = 2^{n-1} - \frac{1}{2}\mathcal{L}(F).$$

Il a été démontré que la non-linéarité d'une fonction est bornée et que lorsque n est impair, cette borne est atteinte.

Proposition 1.47. *[CV94] Soit n un entier. Toute fonction de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$ vérifie*

$$\mathcal{NL}(F) \leqslant 2^{n-1} - 2^{\frac{n-1}{2}}.$$

Les fonctions atteignant cette borne sont appelées fonctions presque-courbes *et n'existent que si n est impair.*

Les fonctions presque-courbes ont les meilleures propriétés cryptographiques : en plus d'avoir la plus grande non-linéarité, ces fonctions sont des fonctions APN, c'est-à-dire qu'elles ont aussi la plus petite uniformité différentielle.

Théorème 1.48. *[CV94] Toute fonction presque-courbe est APN.*

La réciproque de ce théorème n'est pas vraie.

Cependant, ces fonctions existent uniquement pour n impair, alors que les fonctions booléennes vectorielles utilisées dans les chiffrements sont souvent définies avec n pair.

Les coefficients de Walsh $(\mathcal{W}^F(u,v))_{u,v\in\mathbf{F}_2^n}$ sont regroupés dans une table appelée *table des biais linéaires* ou *table de Walsh*. La figure 1.13 correspond à la table de Walsh de la fonction S définie à l'exemple 1.42.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	8	8	4	-4	4	-4	4	-4	-4	4
2	0	0	4	4	0	8	-4	4	8	0	-4	4	0	0	-4	-4
3	0	0	4	4	8	0	-4	4	-4	-4	0	0	-4	4	0	8
4	0	0	8	0	4	4	4	-4	-4	4	-4	-4	8	0	0	0
5	0	0	0	8	4	4	4	-4	0	0	8	0	-4	-4	4	-4
6	0	8	-4	-4	4	4	0	8	-4	4	0	0	0	0	4	-4
7	0	8	4	4	-4	-4	8	0	0	0	-4	4	-4	4	0	0
8	0	4	-4	8	0	-4	-4	0	4	8	0	-4	4	0	0	4
9	0	-4	-4	0	0	4	4	0	0	4	-4	8	0	-4	4	8
10	0	4	0	-4	8	-4	0	-4	4	0	4	8	4	0	-4	0
11	0	-4	0	4	0	-4	0	4	0	-4	0	4	8	4	8	-4
12	0	4	-4	0	-4	8	0	-4	0	-4	4	0	4	8	0	4
13	0	-4	4	0	-4	0	0	4	-4	8	8	4	0	4	-4	0
14	0	-4	0	-4	4	0	4	0	8	4	0	-4	-4	8	4	0
15	0	4	8	-4	-4	0	-4	0	4	0	4	0	0	-4	8	4

FIGURE 1.13 – Table de Walsh de la fonction S définie à l'exemple 1.42.

Dans cette table, nous remarquons que $\mathcal{W}^F(0,0) = 2^n$. En effet, pour toute fonction F de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$, tout élément $x \in \mathbf{F}_2^n$ vérifie $0 \cdot x + 0 \cdot F(x) = 0$. De plus, d'après le lemme 1.27, $\mathcal{W}^F(u,0) = \sum_{x\in\mathbf{F}_2^n}(-1)^{u\cdot x} = 0$ si u est non nul. Par ailleurs, si F est une permutation, alors pour tout $v \neq 0$, nous avons $\mathcal{W}^F(0,v) = \sum_{x\in\mathbf{F}_2^n}(-1)^{v\cdot F(x)} = \sum_{y\in\mathbf{F}_2^n}(-1)^{v\cdot y} = 0$.

Si F est une fonction booléenne vectorielle, les coefficients de Walsh des fonctions affinement équivalentes à F se déduisent de ceux de F. Ce lien fait intervenir l'adjoint d'une fonction linéaire, qui est défini ci-dessous.

Définition 1.49. *Soit F une fonction linéaire de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. L'*adjoint *de F, noté F^*, est l'unique fonction linéaire de $\mathbf{F}_2^n$ qui vérifie : pour tous éléments x, y de $\mathbf{F}_2^n$,*

$x \cdot F(y) = F^*(x) \cdot y$. La matrice de l'adjoint de F correspond à la transposée de la matrice de F.

Proposition 1.50. *Soient F_1 et F_2 deux fonctions affinement équivalentes, c'est-à-dire qu'il existe deux permutations affines de $\mathbf{F}_2^n$, $A_1 : x \mapsto L_1(x)+c_1$ et $A_2 : x \mapsto L_2(x)+c_2$, telles que*

$$F_2 = A_2 \circ F_1 \circ A_1.$$

Alors F_1 et F_2 vérifient : pour tous éléments $u, v \in \mathbf{F}_2^n$,

$$\mathcal{W}^{F_2}(u,v) = (-1)^{\varepsilon}\, \mathcal{W}^{F_1}((L_1^{-1})^*(u), L_2^*(v)),$$

avec $\varepsilon = u \cdot L_1^{-1}(c_1) + v \cdot c_2$ et où L_1^* et L_2^* sont les adjoints de L_1 et L_2.

Démonstration. Soient u, v deux éléments de $\mathbf{F}_2^n$. On a :

$$\begin{aligned} u \cdot x + v \cdot F_2(x) &= u \cdot x + v \cdot [A_2 \circ F_1 \circ A_1(x)] \\ &= u \cdot x + v \cdot [L_2 \circ F_1(L_1(x) + c_1)] + v \cdot c_2 \\ &= u \cdot x + L_2^*(v) \cdot F_1(L_1(x) + c_1) + v \cdot c_2 \\ &= u \cdot [L_1^{-1}(y) + L_1^{-1}(c_1)] + L_2^*(v) \cdot F_1(y) + v \cdot c_2 \\ &= (L_1^{-1})^*(u) \cdot y + u \cdot L_1^{-1}(c_1) + L_2^*(v) \cdot F_1(y) + v \cdot c_2, \end{aligned}$$

où $y = L_1(x) + c_1$. Donc

$$\begin{aligned} \mathcal{W}^{F_2}(u,v) &= (-1)^{u \cdot L_1^{-1}(c_1) + v \cdot c_2} \sum_{y \in \mathbf{F}_2^n} (-1)^{(L_1^{-1})^*(u) \cdot y + L_2^*(v) \cdot F_1(y)} \\ &= (-1)^{u \cdot L_1^{-1}(c_1) + v \cdot c_2}\, \mathcal{W}^{F_1}((L_1^{-1})^*(u), L_2^*(v)). \end{aligned}$$

□

Les fonctions booléennes vectorielles plateau sont définies de manière similaire aux fonctions booléennes plateau.

Définition 1.51. *Une fonction F est dite* plateau *si ses coefficients de Walsh prennent au plus trois valeurs : 0 et $\pm\mathcal{L}(F)$.*

Dans [BN13], les auteurs ont démontré que les coefficients du spectre différentiel et les coefficients de Walsh sont liés par la relation suivante.

Proposition 1.52. *[BN13] Soit S une fonction de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. Pour tous éléments u et v non nuls de $\mathbf{F}_2^n$, nous avons*

$$(\mathcal{W}(u,v))^2 = \sum_{a,b \in \mathbf{F}_2^n} (-1)^{a \cdot u + b \cdot v} \delta(a,b) = 2^n + \sum_{a,b \in \mathbf{F}_2^n, a \neq 0} (-1)^{a \cdot u + b \cdot v} \delta(a,b).$$

1.6 Le corps à 2^n éléments

Nous avons vu dans la description de l'AES (section 1.2.2) qu'une identification de l'espace vectoriel $\mathbf{F}_2^8$ au corps $\mathbf{F}_{2^8}$ était utilisée. En effet, lorsque la fonction de diffusion d'un réseau de substitution-permutation est une permutation linéaire sur $\mathbf{F}_{2^n}$ lorsque $\mathbf{F}_2^n$ est identifié à $\mathbf{F}_{2^n}$, il est pratique de la décrire sur $\mathbf{F}_{2^n}$. Alors, en utilisant le même isomorphisme pour décrire la boîte-S, comme dans le cas de l'AES, le chiffrement tout entier peut se décrire dans le corps $\mathbf{F}_{2^n}$. C'est pourquoi nous rappelons dans cette section quelques notions sur les corps finis qui seront utilisées dans la suite de ce document. Nous renvoyons le lecteur aux ouvrages de R. Lidl et H. Niederreiter [LN83] ou R. McEliece [McE87] (entre autres) pour une description plus complète de la théorie des corps finis.

Soit n un entier strictement positif. Une fonction très importante de $\mathbf{F}_{2^n}$ dans $\mathbf{F}_2$ est appelée *trace absolue*.

Définition 1.53. *La fonction* trace absolue *de* $\mathbf{F}_{2^n}$ *dans* $\mathbf{F}_2$ *est notée* Tr *et est définie pour tout élément* x *de* $\mathbf{F}_{2^n}$ *par*

$$\mathrm{Tr}(x) = \sum_{i=0}^{n-1} x^{2^i}.$$

Dans ce document, cette fonction sera simplement appelée trace. Une de ses propriétés intéressante est d'être linéaire.

Proposition 1.54. *[LN83, chap. 2] Soient* x, y *deux éléments de* $\mathbf{F}_{2^n}$ *et* c *un élément de* $\mathbf{F}_2$. *La fonction trace vérifie :*
- $\mathrm{Tr}(x+y) = \mathrm{Tr}(x) + \mathrm{Tr}(y)$;
- $\mathrm{Tr}(cx) = c\mathrm{Tr}(x)$;
- $\mathrm{Tr}(c) = nc$;
- $\mathrm{Tr}(x^2) = \mathrm{Tr}(x)$.

De plus, la fonction trace permet de décrire toutes les fonctions linéaires de $\mathbf{F}_{2^n}$ dans $\mathbf{F}_2$, autrement dit, toutes les fonctions booléennes linéaires.

Proposition 1.55. *[LN83, chap. 2] Les fonctions linéaires de* $\mathbf{F}_{2^n}$ *dans* $\mathbf{F}_2$ *sont exactement les fonctions* $x \mapsto \mathrm{Tr}(ax)$ *pour tout* $a \in \mathbf{F}_{2^n}$.

Le corps fini $\mathbf{F}_{2^n}$ peut être vu comme un espace vectoriel de dimension n sur $\mathbf{F}_2$. Nous pouvons donc utiliser la notion de base sur le corps $\mathbf{F}_{2^n}$.

Définition 1.56. *Une famille de* n *éléments* $\{b_0, \ldots, b_{n-1}\}$ *de* $\mathbf{F}_{2^n}$ *est une* base *de ce corps si tout élément* $x \in \mathbf{F}_{2^n}$ *peut s'écrire de manière unique* $x = x_0 b_0 + \ldots + x_{n-1} b_{n-1}$ *avec* $x_i \in \mathbf{F}_2$ *pour tout* i *compris entre* 0 *et* $n-1$.

La notion de base duale peut aussi être définie.

Définition 1.57. *Deux bases* $\{a_0, \ldots, a_{n-1}\}$ *et* $\{b_0, \ldots, b_{n-1}\}$ *de* $\mathbf{F}_{2^n}$ *sont dites* duales *si pour tous les entiers* i *et* j *compris entre* 0 *et* $n-1$*, nous avons :*

$$Tr(a_i b_j) = \begin{cases} 0 & si\ i \neq j \\ 1 & si\ i = j. \end{cases}$$

Proposition 1.58. *[LN83, chap. 2] Chaque base de* $\mathbf{F}_2^n$ *admet une unique base duale.*

Le corps fini $\mathbf{F}_{2^n}$ étant un espace vectoriel sur $\mathbf{F}_2$, pour chaque base $\{b_0, \ldots, b_{n-1}\}$ de $\mathbf{F}_{2^n}$, il existe un isomorphisme d'espace vectoriel φ entre $\mathbf{F}_2^n$ et $\mathbf{F}_{2^n}$:

$$\begin{array}{rccc} \varphi : & \mathbf{F}_2^n & \longrightarrow & \mathbf{F}_{2^n} \\ & (x_0, \ldots, x_{n-1}) & \longmapsto & x_0 b_0 + \ldots + x_{n-1} b_{n-1}. \end{array}$$

Alors le produit scalaire sur le corps correspondant au produit scalaire euclidien de $\mathbf{F}_2^n$ (noté « $\cdot$ ») est défini par la fonction trace.

Proposition 1.59. *Soit* $\{b_0, \ldots, b_{n-1}\}$ *une base de* $\mathbf{F}_{2^n}$ *et* φ *l'isomorphisme associé. Alors pour tout* $(x, y) \in \mathbf{F}_2^n$*, nous avons*

$$x \cdot y = \mathrm{Tr}(\varphi(x)\psi(y)),$$

où ψ *est l'isomorphisme correspondant à la base* $\{a_0, \ldots, a_{n-1}\}$ *duale de* $\{b_0, \ldots, b_{n-1}\}$*.*

Démonstration. D'après la définition de la base duale, nous avons :

$$\begin{aligned} \mathrm{Tr}(\varphi(x)\psi(y)) &= \mathrm{Tr}\left(\left(\sum_{i=0}^{n-1} x_i b_i\right)\left(\sum_{j=0}^{n-1} y_j a_j\right)\right) \\ &= \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} x_i y_j \mathrm{Tr}(a_j b_i) \\ &= \sum_{i=0}^{n-1} x_i y_i = x \cdot y. \end{aligned}$$

□

Toutes les définitions et résultats sur les fonctions booléennes vectorielles de la partie précédente peuvent donc s'énoncer pour des fonctions définies sur le corps $\mathbf{F}_{2^n}$.

1.7 Cryptanalyses différentielle et linéaire

Deux des plus célèbres familles d'attaques sur les chiffrements par blocs sont la cryptanalyse différentielle et la cryptanalyse linéaire. Elles ont été présentées en 1990 et 1992 respectivement, avec comme cible principale le chiffrement DES, le standard de chiffrement symétrique utilisé entre 1977 et 2000. Aujourd'hui, tout chiffrement doit se prémunir contre ces attaques, puisqu'elles permettent de retrouver la clé utilisée à partir d'un certain nombre de couples de messages clairs/chiffrés.

1.7.1 Principe des attaques statistiques

Les cryptanalyses différentielle et linéaire sont des attaques statistiques, c'est-à-dire qu'elles utilisent des relations entre les entrées et les sorties d'une fonction apparaissant dans le chiffrement qui arrivent avec une certaine probabilité pour la distinguer d'une fonction aléatoire. Un distingueur $\mathcal{D}$ pour une famille de fonctions $(F_k)_k$ est un algorithme qui prend en entrée N couples (x_i, y_i), $1 \leqslant i \leqslant N$, et qui retourne 0 ou 1. Il doit décider si les couples (x_i, y_i) sont des paires d'entrées/sorties pour une fonction choisie aléatoirement parmi la famille $(F_k)_k$ (dans ce cas, $\mathcal{D}$ retourne 1) ou pour une fonction choisie aléatoirement parmi toutes les fonctions de $\mathbf{F}_2^n$ (dans ce cas, $\mathcal{D}$ retourne 0). L'avantage du distingueur est défini par

$$\mathrm{Adv}(\mathcal{D}) = |\Pr(\mathcal{D}^\pi = 1 | \pi \in_R \{F_k, k \in \mathbf{F}_2^\kappa\}) - \Pr(\mathcal{D}^\pi = 1 | \pi \in_R \mathrm{Fonc}_n)|$$

où Fonc_n est l'ensemble des fonctions de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. On définit de manière similaire un distingueur sur une famille de permutations.

Pour qu'un chiffrement itératif soit sûr, il ne doit pas exister de distingueur sur un grand nombre de tours consécutifs du chiffrement qui ait un avantage non négligeable. Pour un chiffrement itératif $(E_k)_{k \in \mathbf{F}_2^\kappa}$ de fonction de tour f_{k_i}, c'est-à-dire

$$E_k = f_{k_r} \circ \ldots \circ f_{k_1},$$

l'attaquant peut par exemple considérer le chiffrement réduit, c'est-à-dire la famille $(G_k)_k$ obtenue en enlevant le dernier tour de la fonction de chiffrement E_k : pour tout $k \in \mathbf{F}_2^\kappa$,

$$G_k = f_{k_{r-1}} \circ \ldots \circ f_{k_1}.$$

S'il existe un distingueur $\mathcal{D}$ sur le chiffrement réduit $(G_k)_k$, il est possible de retrouver la sous-clé k_r par une attaque sur le dernier tour. Cette attaque consiste à utiliser le distingueur sur une famille $(H_{\widehat{k}_r})$ définie par :

$$H_{\widehat{k}_r} = f_{\widehat{k}_r}^{-1} \circ E_k = f_{\widehat{k}_r}^{-1} \circ f_{k_r} \circ \ldots \circ f_{k_1}$$

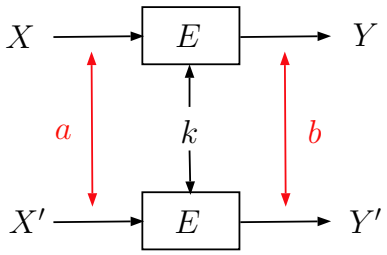
où $\widehat{k}_r$ prend toutes les valeurs possibles pour la sous-clé du dernier tour. Si la clé essayée est la bonne, c'est-à-dire si $\widehat{k}_r = k_r$, alors $H_{\widehat{k}_r}$ est une fonction de la famille $(G_k)_k$. Sinon, la fonction $H_{\widehat{k}_r}$ est supposée avoir le même comportement qu'une fonction aléatoire de $\mathbf{F}_2^n$ (hypothèse de répartition aléatoire par fausse clé) [HKM95, Har96]. La connaissance de la sous-clé k_r (ou d'une partie de cette sous-clé) permet d'avoir des informations sur la clé maître k. Pour déterminer k entièrement, il est possible d'itérer cette attaque ou alors de faire une recherche exhaustive sur les derniers bits inconnus de k.

Les attaques statistiques sur le dernier tour supposent généralement que le comportement des fonctions de la famille $(G_k)_k$ est similaire pour toutes les clés (hypothèse d'équivalence stochastique pour la cryptanalyse différentielle et hypothèse d'équivalence de clé fixée pour la cryptanalyse linéaire)[LMM91]. Toutefois, ces hypothèses ne sont pas toujours vérifiées, en particulier pour un chiffrement ne contenant que peu de tours [DR02, DR07b, BBL13]. Nous reviendrons sur ces hypothèses dans la suite de ce document.

1.7.2 Cryptanalyse différentielle

Principe

Introduite par Eli Biham et Adi Shamir dans [BS90, BS91], la cryptanalyse différentielle consiste à exploiter un biais statistique dans la distribution de la différence entre les images de deux entrées dont la différence est fixée. Dans tout ce document, les opérations sont définies en caractéristique 2, donc une différence sera toujours représentée par un signe « + ». Toutefois, il est possible d'utiliser une autre loi de groupe pour définir la différence (*cf* [LMM91]).



Pour une permutation aléatoire π de $\mathbf{F}_2^n$, la probabilité qu'une différence non nulle $a \in \mathbf{F}_2^n$ en entrée mène à une différence non nulle $b \in \mathbf{F}_2^n$ en sortie est

$$\Pr_X[\pi(X+a)+\pi(X)=b] = \frac{1}{2^n-1}.$$

Si on peut trouver un couple (a,b) (appelé une *différentielle*) tel que, pour un grand nombre de clés k, la probabilité $p = \Pr_X[E_k(X+a)+E_k(X)=b]$ est significativement supérieure à $(2^n-1)^{-1}$, alors il est possible de distinguer la fonction de chiffrement E_k, où k est une clé choisie aléatoirement, d'une permutation aléatoire. Le nombre de couples de messages clairs/chiffrés (la complexité en données) nécessaire pour distinguer la fonction de chiffrement d'une permutation aléatoire est $\mathcal{O}\left[p\ln(p(2^n-1)\right]^{-1}$ [BGT11].

Notons $(E_k)_k$ un chiffrement itératif opérant sur des blocs de n bits et composé de r tours, X_0 et X_0' un couple de textes clairs et X_i (respectivement X_i') l'image de X_0 (respectivement X_0') après i tours de la fonction de chiffrement E_k, pour i compris entre 1 et r.

Définition 1.60. *Une* différentielle *sur r tours est un couple $(a,b) \in \mathbf{F}_2^n$ de différences. Une différentielle est caractérisée par sa probabilité pour la clé k, définie par :*

$$\mathrm{DP}_r^{E_k}(a,b) = \Pr_X[E_k(X)+E_k(X+a)=b];$$

où l'indice correspond au nombre de tours et l'exposant au chiffrement pour lequel la probabilité est calculée. Lorsqu'il n'y a aucune ambiguïté, l'exposant peut être omis.

Remarque 1.61. Il est souvent utile d'étudier la probabilité d'un chiffrement sur un nombre de tours inférieur au nombre total de tours du chiffrement. Dans ce cas, la probabilité d'une différentielle sera notée $\mathrm{DP}_i^{E_k}(a,b)$, où $1 \leqslant i \leqslant r$.

Une différentielle est définie par une différence en entrée et une différence en sortie du chiffrement. Les différences sur les états intermédiaires peuvent donc prendre différentes valeurs et il y a généralement un grand nombre de suites de différences sur les états internes successifs partant d'une différence fixée en entrée du chiffrement et menant à une différence fixée à la sortie. Une telle suite de différences fixées sur les états internes successifs d'un chiffrement est appelée caractéristique différentielle.

Définition 1.62. *Une* caractéristique différentielle Q *sur r tours est un $(r+1)$-uplet $(a_0, \ldots, a_r)$ tel que $a_0 = X_0 + X'_0$, $a_1 = X_1 + X'_1$, $\ldots$, $a_r = X_r + X'_r$.*

Une caractéristique différentielle $(a_0, \ldots, a_r)$ est dite dans *une différentielle (a, b) si $a_0 = a$ et $a_r = b$. Une différentielle (a, b) est* composée *des caractéristiques $(a_0, \ldots, a_r)$ vérifiant $a_0 = a$ et $a_r = b$.*

Définition 1.63. *Pour la clé k, la* probabilité de la caractéristique différentielle $Q = (a_0, \ldots, a_r)$ *sur r tours est :*

$$\mathrm{DCP}_r^{E_k}(Q) = \Pr_{X_0}(X_1 + X'_1 = a_1; \ldots; X_r + X'_r = a_r \,|\, X_0 + X'_0 = a_0) \,.$$

Une caractéristique différentielle de probabilité non nulle est appelée chemin différentiel.

Distribution des caractéristiques et différentielles quand la clé varie

La probabilité d'une caractéristique (donc d'une différentielle) peut beaucoup varier en fonction de la clé, ce qui peut rendre la distribution de la probabilité de la différentielle très difficile à calculer en général. Dans certains cas particuliers, il a été possible de déterminer cette distribution. Par exemple, dans [DR09], les auteurs ont présenté une condition sur les composants d'un réseau de substitution-permutation qui assure que la détermination de la distribution de la probabilité d'une caractéristique sur deux tours de ce chiffrement est assez simple. Une partie de ces travaux est détaillée dans le chapitre 6.

Une autre méthode pour approximer la distribution de la probabilité d'une caractéristique (resp. d'une différentielle) est de calculer la moyenne sur les clés de la probabilité de cette caractéristique (resp. différentielle).

Définition 1.64. *Soit $(E_k)_k$ un chiffrement itératif de r tours et de taille de clé κ. La probabilité en moyenne sur les clés ("Expected Differential Probability" en anglais) de la différentielle (a, b) sur r tours est*

$$\mathrm{EDP}_r^E(a, b) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \Pr_X[E_k(X) + E_k(X + a) = b] \,,$$

où l'exposant E correspond au chiffrement pour lequel la probabilité est calculée. Il sera omis lorsqu'il n'y a pas d'ambiguïté. Le nombre de tours considéré r, noté en indice, ne sera pas indiqué dans le cas où $r = 1$.

De même, la probabilité en moyenne sur les clés de la caractéristique différentielle Q sur r tours est

$$\mathrm{EDCP}_r^E(Q) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \mathrm{DCP}_r^{E_k}(Q) \,.$$

Dans [DR05, Théorèmes 13, 14], les auteurs ont montré que, sous certaines hypothèses, la distribution de la probabilité d'une caractéristique Q (resp. d'une différentielle (a, b)) à *clé fixée* peut être approximée par une distribution de Poisson.

Définition 1.65. *Une variable aléatoire dénombrable X, à valeurs entières et positives, suit une* loi de Poisson de paramètre λ ($\lambda > 0$) *si et seulement si, pour tout entier naturel s,*

$$\Pr(X = s) = \frac{\lambda^s}{s!}\, e^{-\lambda}.$$

Cette loi est notée $\mathcal{P}(s)$. Son espérance est $E(X) = \lambda$, sa variance est $V(X) = \lambda$ et son écart-type est $\sigma(X) = \sqrt{\lambda}$.

Théorème 1.66. *[DR05] La distribution de la probabilité d'une caractéristique Q (resp. d'une différentielle (a, b)) à* clé fixée *peut être approximée par :*

$$\Pr(2^{n-1}\mathrm{DCP}^{E_k}(Q) = i) \approx \mathcal{P}(2^{n-1}\mathrm{EDCP}^{E}(Q)),$$
$$\textit{resp. } \Pr(2^{n-1}\mathrm{DP}^{E_k}(a, b) = i) \approx \mathcal{P}(2^{n-1}\mathrm{EDP}^{E}(a, b)),$$

où $\mathcal{P}(s)$ désigne la loi de Poisson de paramètre s.

Cette distribution est à comparer avec la distribution obtenue sur toutes les fonctions de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$.

Théorème 1.67. *[DR07b] Soit $(a, b) \in (\mathbf{F}_2^n)^2$ une différentielle. La distribution de $2^{n-1}\mathrm{DP}^F(a, b)$ sur toutes les fonctions F de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$ est binomiale :*

$$\Pr_F(2^{n-1}\,\mathrm{DP}^F(a, b) = i) = 2^{-in}(1 - 2^{-n})^{2^{n-1}-i} \binom{2^{n-1}}{i}.$$

Cette distribution est approximée par une loi de Poisson $\mathcal{P}$ de paramètre $\frac{1}{2}$:

$$\Pr_F(2^{n-1}\,\mathrm{DP}^F(a, b) = i) \approx \mathcal{P}\left(\frac{1}{2}\right).$$

Si $2^{n-1}\mathrm{EDP}^E(a, b)$ est proche de $1/2$, c'est-à-dire si $\mathrm{EDP}^E(a, b)$ est proche de 2^{-n}, alors la distribution de la probabilité d'une différentielle à clé fixée est proche de la distribution de la probabilité de cette différentielle sur toutes les fonctions de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. C'est pourquoi la moyenne sur les clés de la probabilité d'une différentielle joue un rôle important, en particulier son écart à 2^{-n}.

Probabilité d'une caractéristique et d'une différentielle en moyenne sur les clés

Calculer la probabilité d'une caractéristique différentielle est plus simple que calculer celle d'une différentielle, en particulier pour un chiffrement de Markov [LMM91]. Pour ce type de chiffrement, la probabilité en moyenne sur les clés d'une caractéristique est égale au produit des probabilités des différentielles sur un tour.

Définition 1.68. *[LMM91] Un chiffrement itératif $(E_k)_k$ possède la propriété de Markov si la probabilité de toute différentielle sur sa fonction de tour f_i est indépendante du choix du texte clair : si la sous-clé k_i est choisie aléatoirement selon la loi uniforme, on a pour tout X_0 :*

$$\begin{aligned}\Pr_{X,k_i}(f_i(X;k_i)+f_i(X';k_i)=b \,|\, X+X'=a,\ X=X_0) = \\ \Pr_{X,k_i}(f_i(X;k_i)+f_i(X';k_i)=b \,|\, X+X'=a)\ .\end{aligned}$$

Les réseaux de substitution-permutation dont l'opération d'insertion des sous-clés est une addition dans $\mathbf{F}_2^n$ sont des chiffrements de Markov [LMM91].

Théorème 1.69. *[LMM91] Si $(E_k)_k$ est un chiffrement de Markov dont les sous-clés sont indépendantes et uniformément distribuées, la probabilité en moyenne sur les clés d'une caractéristique différentielle sur r tours est égale au produit des probabilités en moyenne des r différentielles sur un tour qui la composent :*

$$\mathrm{EDCP}_r^E(a_0,\ldots,a_r)=\prod_{i=1}^{r}\mathrm{EDP}_1^{f_i}(a_{i-1},a_i),$$

où f_i est la fonction de tour du chiffrement. Dans ce cas, la probabilité en moyenne d'une différentielle (a_0,a_r) sur r tours du chiffrement est :

$$\mathrm{EDP}_r^E(a_0,a_r)=\sum_{a_1\in\mathbf{F}_2^n}\sum_{a_2\in\mathbf{F}_2^n}\cdots\sum_{a_{r-1}\in\mathbf{F}_2^n}\prod_{i=1}^{r}\mathrm{EDP}_1^{f_i}(a_{i-1},a_i).$$

En pratique, dans les analyses de sécurité sur les réseaux de substitution-permutation, la quantité évaluée est le maximum de la probabilité en moyenne sur les clés d'une différentielle.

Définition 1.70. *Soit $(E_k)_k$ un chiffrement itératif de r tours et de taille de clé κ. Le maximum de la probabilité en moyenne sur les clés d'une différentielle sur r tours est*

$$\mathrm{MEDP}_r^E=\max_{a\neq 0,b}\mathrm{EDP}_r^E(a,b)\ .$$

L'exposant E et l'indice r seront omis lorsqu'il n'y a pas d'ambiguïté.

Il y a généralement une différence entre la valeur MEDP et, pour une clé fixée, la probabilité maximale des différentielles. Cette différence a été mise en avant par exemple pour l'AES dans [DR07b, Section 5]. Cependant, la valeur MEDP est considérée comme une bonne estimation pour déterminer si un chiffrement est résistant à la cryptanalyse différentielle, cela a été justifié par exemple dans [BBL13] dans le cas de deux tours d'un réseau de substitution-permutation.

1.7.3 Cryptanalyse linéaire

Principe

La cryptanalyse linéaire a été introduite dans [TCG91], puis dans [Mat94, Mat95] sur le DES, le standard de chiffrement par blocs qui a précédé l'AES. Elle consiste à

exploiter les approximations affines du chiffrement qui font intervenir des bits du texte clair, du chiffré et de la clé et ayant une très grande ou très petite probabilité. Ces approximations affines sont de la forme $x \mapsto u \cdot x + v \cdot E_k(x) + c \cdot k$, où $(u, v) \in (\mathbf{F}_2^n)^2$ et $c \in \mathbf{F}_2^\kappa$. Le couple (u, v) est appelé *masque linéaire*. Pour une permutation aléatoire π de $\mathbf{F}_2^n$, la probabilité que $u \cdot x + v \cdot \pi(x)$ soit égal à 0 est $\frac{1}{2}$. Donc s'il existe u et v non nuls tels que $u \cdot x = v \cdot E_k(x)$ avec une probabilité éloignée de $\frac{1}{2}$, alors il est possible de distinguer la fonction de chiffrement d'une permutation aléatoire. On s'intéresse donc à la différence entre la probabilité d'une approximation affine et la valeur $\frac{1}{2}$, qui est appelée *biais*. Pour qu'un chiffrement résiste à la cryptanalyse linéaire, il faut que le biais de toute approximation affine soit proche de 0.

Notons $(E_k)_k$ un chiffrement itératif de r tours et n la taille d'un bloc de E_k. Le biais ε d'une approximation affine pour ce chiffrement est défini par :

$$\varepsilon = \frac{\#\{x \in \mathbf{F}_2^n \mid u \cdot x + v \cdot E_k(x) = 0\}}{2^n} - \frac{1}{2}.$$

Lorsque l'équation $u \cdot x + v \cdot E_k(x) = 0$ a un grand nombre de solutions, le biais ε est proche de $\frac{1}{2}$ et $c \cdot k = 0$ avec une grande probabilité. Si, au contraire, l'équation $u \cdot x + v \cdot E_k(x) = 0$ a très peu de solutions, le biais ε est proche de $-\frac{1}{2}$ et $c \cdot k = 1$, avec une grande probabilité. Dans ces deux cas, il est possible de distinguer le chiffrement d'une permutation aléatoire.

Pour simplifier les calculs, la notion de corrélation d'un masque linéaire à clé fixée, qui correspond au double du biais, est utilisée.

Définition 1.71. *La* corrélation du masque linéaire (u, v) sur r tours pour la clé k *est la valeur définie par :*

$$\mathrm{C}_r^{E_k}(u, v) = 2^{-n} \sum_{x \in \mathbf{F}_2^n} (-1)^{u \cdot x + v \cdot E_k(x)},$$

où l'exposant E_k correspondant à la permutation pour laquelle la probabilité est calculée est omis lorsqu'il n'y a pas d'ambiguïté.

En général, c'est la corrélation d'un masque linéaire pour un nombre de tours inférieur au nombre de tours du chiffrement qui est calculée. Cette corrélation est notée $\mathrm{C}_i^{E_k}(u, v)$, où le nombre de tours considéré i, noté en indice, ne sera pas indiqué dans le cas où $i = 1$.

Distribution de la corrélation d'un chemin et d'un masque linéaire quand la clé varie

De même que pour la cryptanalyse différentielle, il est plus facile de calculer la corrélation d'un masque sur r tours lorsque les masques intermédiaires sont fixés. Dans ce cas, on parle de *chemin linéaire* sur r tours.

Définition 1.72. *Soit $(E_k)_k$ un chiffrement itératif de r tours de taille de bloc n et de fonction de tour f_i, c'est-à-dire*

$$E_k = f_r \circ f_{r-1} \circ \cdots \circ f_1.$$

Un chemin linéaire *sur r tours pour la clé k est un $(r+1)$-uplet $(u_0, \ldots, u_r)$ tel que pour tout i compris entre 1 et r, (u_{i-1}, u_i) est un masque linéaire sur le tour i du chiffrement.*

La corrélation d'un chemin linéaire *sur r tours $(u_0, \ldots, u_r)$ est définie comme étant le produit des corrélations des masques linéaires (u_i, u_{i+1}) :*

$$\mathrm{TC}_r^{E_k}(u_0, \ldots, u_r) = \prod_{i=1}^{r} \mathrm{C}_1^{f_i}(u_{i-1}, u_i).$$

La notation TC pour la corrélation d'un chemin linéaire ("linear trail correlation" en anglais) n'est pas standard mais permet de rendre plus visible la distinction entre un chemin et un masque linéaire.

Remarque 1.73. Dans le cas d'un chiffrement $(E_k)_k$ alternant clés et permutations, c'est-à-dire tel que la fonction de tour est $f_i(x) = f(x) + k_i$, k_i étant la sous-clé dérivée de k pour le tour i, la corrélation d'un chemin linéaire $(u_0, \ldots, u_r)$ sur r tours vérifie :

$$\begin{aligned}
\mathrm{TC}_r^{E_k}(u_0, \ldots, u_r) &= \prod_{i=1}^{r} \mathrm{C}_1^{f_i}(u_{i-1}, u_i) \\
&= \prod_{i=1}^{r} \left(2^{-n} \sum_{x \in \mathbf{F}_2^n} (-1)^{u_{i-1} \cdot x + u_i \cdot f(x) + u_i \cdot k_i} \right) \\
&= \prod_{i=1}^{r} (-1)^{u_i \cdot k_i} \left(2^{-n} \sum_{x \in \mathbf{F}_2^n} (-1)^{u_{i-1} \cdot x + u_i \cdot f(x)} \right) \\
&= (-1)^{u_1 \cdot k_1 + \ldots u_r \cdot k_r} \prod_{i=1}^{r} \mathrm{C}^f(u_{i-1}, u_i).
\end{aligned}$$

Donc la valeur absolue de la corrélation d'un chemin linéaire est indépendante de la clé, seul le signe de la corrélation du chemin est déterminé par la clé.

Il est donc plus facile de déterminer la corrélation d'un chemin linéaire à clé fixée que de déterminer la probabilité d'une caractéristique différentielle à clé fixée. Cependant, la difficulté apparait lorsqu'on s'intéresse non plus à un chemin linéaire mais à la corrélation d'un masque linéaire. La corrélation d'un masque linéaire (u, v) est égale à la somme des corrélations des chemins linéaires $(u_0, \ldots, u_r)$ vérifiant $u = u_0$ et $v = u_r$.

Théorème 1.74. *[DR02, Th. 7.8.1] Soit $(E_k)_k$ un chiffrement itératif de r tours de taille de bloc n et de la forme*

$$E_k = f_r \circ f_{r-1} \circ \cdots \circ f_1.$$

La corrélation du masque linéaire (u_0, u_r) sur r tours est égale à la somme des corrélations des chemins linéaires tels que le masque à l'entrée du chiffrement est u_0 et le masque après les r tours est u_r :

$$\mathrm{C}_r^{E_k}(u_0, u_r) = \sum_{u_1,\ldots,u_{r-1}\in\mathbf{F}_2^n} \prod_{i=1}^{r} \mathrm{C}_1^{f_i}(u_{i-1}, u_i).$$

Démonstration. Nous allons faire la démonstration pour trois tours d'un chiffrement puisque cette démonstration se généralise très simplement à un nombre quelconque de tours. Notons A la valeur $\sum_{u_1,u_2\in\mathbf{F}_2^n} \mathrm{C}_1^{f_1}(u_0,u_1)\,\mathrm{C}_1^{f_2}(u_1,u_2)\,\mathrm{C}_1^{f_3}(u_2,u_3)$. En écrivant la définition de la corrélation d'un masque linéaire, nous obtenons :

$$\begin{aligned} A &= \sum_{u_1,u_2\in\mathbf{F}_2^n} \left(2^{-n}\sum_{x_1\in\mathbf{F}_2^n}(-1)^{u_0\cdot x_1+u_1\cdot f_1(x_1)}\right)\left(2^{-n}\sum_{x_2\in\mathbf{F}_2^n}(-1)^{u_1\cdot x_2+u_2\cdot f_2(x_2)}\right)\left(2^{-n}\sum_{x_3\in\mathbf{F}_2^n}(-1)^{u_2\cdot x_3+u_3\cdot f_3(x_3)}\right) \\ &= 2^{-3n}\sum_{x_1,x_2,x_3\in\mathbf{F}_2^n}(-1)^{u_0\cdot x_1+u_3\cdot f_3(x_3)}\left(\sum_{u_1\in\mathbf{F}_2^n}(-1)^{u_1\cdot(x_2+f_1(x_1))}\right)\left(\sum_{u_2\in\mathbf{F}_2^n}(-1)^{u_2\cdot(x_3+f_2(x_2))}\right). \end{aligned}$$

D'après le théorème 1.27, nous avons pour $i = 1, 2$:

$$\sum_{u_i\in\mathbf{F}_2^n}(-1)^{u_i\cdot(x_{i+1}+f_i(x_i))} = \begin{cases} 2^n & \text{si } x_{i+1} = f_i(x_i) \\ 0 & \text{sinon.} \end{cases}$$

En appliquant ce résultat pour $i = 2$, puis pour $i = 1$, nous obtenons :

$$\begin{aligned} A &= 2^{-2n}\sum_{x_1,x_2\in\mathbf{F}_2^n}(-1)^{u_0\cdot x_1+u_3\cdot f_3(f_2(x_2))}\left(\sum_{u_1\in\mathbf{F}_2^n}(-1)^{u_1\cdot(x_2+f_1(x_1))}\right) \\ &= 2^{-n}\sum_{x_1\in\mathbf{F}_2^n}(-1)^{u_0\cdot x_1+u_3\cdot f_3(f_2(f_1(x_1)))} \\ &= \mathrm{C}_3^{E_k}(u_0, u_3). \end{aligned}$$

□

Comme dans le cas de la cryptanalyse différentielle, les variations de la corrélation d'un masque linéaire selon la clé peuvent rendre le calcul de la distribution de la corrélation assez complexe. Par exemple, dans le cas d'un chiffrement alternant clés et permutations, on déduit de la remarque 1.73 que la corrélation sur r tours du masque (u_0, u_r) est donnée par :

$$\mathrm{C}_r^{E_k}(u_0, u_r) = \sum_{u_1,\ldots,u_{r-1}\in\mathbf{F}_2^n}(-1)^{u_1\cdot k_1+\ldots u_r\cdot k_r}\prod_{i=1}^{r}\mathrm{C}_1^{f}(u_{i-1}, u_i),$$

où $(k_1, \ldots, k_r)$ est la suite des sous-clés. En fonction de la clé, certains chemins linéaires peuvent donc apporter une contribution positive ou négative à la corrélation

totale. C'est donc ici que la clé joue un rôle essentiel, alors que pour la cryptanalyse différentielle, la clé apparaît dans le calcul des contributions des différentes caractéristiques, mais celles-ci sont ensuite simplement sommées.

Cependant, pour les chiffrements alternant clés et permutations où les sous-clés sont indépendantes et identiquement distribuées, la moyenne sur les clés de la corrélation de tout masque linéaire est nulle.

Proposition 1.75. *[DR02, Section 7.9][AÅBL12, Prop. 1] Soit $(E_k)_k$ un chiffrement alternant clés et permutations dont les sous-clés $k_1, \ldots, k_r$ sont indépendantes et identiquement distribuées. Alors la moyenne sur les clés de la corrélation de tout masque linéaire est nulle,* i.e. *pour tout masque $(u_0, u_r) \neq (0, 0)$, nous avons*

$$\sum_{k_1,\ldots,k_r \in \mathbf{F}_2^n} \mathrm{C}_r^{E_k}(u_0, u_r) = 0.$$

Démonstration. En utilisant la remarque 1.73, nous pouvons écrire que la corrélation d'un masque linéaire (u_0, u_r) sur r tours est :

$$\begin{aligned} \mathrm{C}_r^{E_k}(u_0, u_r) &= \sum_{u_1,\ldots,u_{r-1} \in \mathbf{F}_2^n} (-1)^{u_1 \cdot k_1 + \ldots u_r \cdot k_r} \prod_{i=1}^{r} \mathrm{C}_1^f(u_{i-1}, u_i) \\ &= \sum_{u_1,\ldots,u_{r-1} \in \mathbf{F}_2^n} (-1)^{U \cdot K} \prod_{i=1}^{r} \mathrm{C}_1^f(u_{i-1}, u_i) \end{aligned}$$

où $U = (u_1, \ldots, u_r)$ et $K = (k_1, \ldots, k_r)$. Lorsque nous considérons la somme sur toutes les sous-clés $K \in (\mathbf{F}_2^n)^r$ possibles, nous obtenons :

$$\begin{aligned} \sum_{K \in (\mathbf{F}_2^n)^r} \mathrm{C}_r^{E_k}(u_0, u_r) &= \sum_{K \in (\mathbf{F}_2^n)^r} \sum_{u_1,\ldots,u_{r-1} \in \mathbf{F}_2^n} (-1)^{U \cdot K} \prod_{i=1}^{r} \mathrm{C}_1^f(u_{i-1}, u_i) \\ &= \sum_{u_1,\ldots,u_{r-1} \in \mathbf{F}_2^n} \prod_{i=1}^{r} \mathrm{C}_1^f(u_{i-1}, u_i) \sum_{K \in (\mathbf{F}_2^n)^r} (-1)^{U \cdot K}. \end{aligned}$$

D'après le lemme 1.27, $\sum_{K \in (\mathbf{F}_2^n)^r} (-1)^{U \cdot K} = 0$ si $U = (u_1, \ldots, u_r) \neq 0$. Donc

$$\sum_{K \in (\mathbf{F}_2^n)^r} \mathrm{C}_r^{E_k}(u_0, u_r) = 2^{nr} \mathrm{C}_1^f(u_0, 0) \left(\prod_{i=2}^{r-1} \mathrm{C}_1^f(0, 0) \right) \mathrm{C}_1^f(0, u_r).$$

Or $\mathrm{C}_1^f(u_0, 0) = 0$ sauf si $u_0 = 0$ et $\mathrm{C}_1^f(0, u_r) = 0$ sauf si $u_r = 0$. D'où

$$\sum_{K \in (\mathbf{F}_2^n)^r} \mathrm{C}_r^{E_k}(u_0, u_r) = 0$$

car $(u_0, u_r) \neq (0, 0)$. □

La valeur qui est utilisée pour estimer la corrélation d'un masque linéaire est donc la variance sur les clés de la corrélation (aussi appelée *potentiel linéaire moyen*) d'un masque linéaire.

Définition 1.76. *[Nyb95] Soit $(E_k)_k$ un chiffrement itératif de r tours, de taille de bloc n et de taille de clé κ. Le* potentiel linéaire moyen d'un masque (u_0, u_r) *sur r tours est*

$$\mathrm{ELP}_r^E(u_0,u_r) = 2^{-\kappa} \sum_{k\in\mathbf{F}_2^\kappa} (\mathrm{C}_r^{E_k}(u_0,u_r))^2 = 2^{-2n-\kappa} \sum_{k\in\mathbf{F}_2^\kappa} \left(\sum_{x\in\mathbf{F}_2^n} (-1)^{u_0\cdot x + u_r\cdot E_k(x)} \right)^2 .$$

Le potentiel linéaire d'un chemin $(u_0, \ldots, u_r)$ *sur r tours est*

$$\mathrm{ELTP}_r^E(u_0,\ldots,u_r) = 2^{-\kappa} \sum_{k\in\mathbf{F}_2^\kappa} (\mathrm{TC}_r^{E_k}(u_0,\ldots,u_r))^2.$$

Dans [DR05], il est montré que sous certaines hypothèses, le comportement de la distribution de la corrélation d'un masque linéaire *à clé fixée*, pour un masque (u_0, u_r) donné, peut être approximée par une distribution normale centrée de variance $\mathrm{ELP}_r^E(u_0, u_r)$.

Définition 1.77. *Une variable aléatoire X suit une loi normale de paramètres μ et σ^2 si sa densité de probabilité est*

$$D(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Cette loi est notée $\mathcal{N}(\mu, \sigma^2)$. Son espérance est $E(X) = \mu$ et sa variance est $V(X) = \sigma^2$.

Théorème 1.78. *[DR05] Soit (u_0, u_r) un masque linéaire tel qu'il existe un grand nombre de chemin ayant pour masque d'entrée u_0 et pour masque de sortie u_r. Alors la distribution de la corrélation du masque (u_0, u_r) peut être approximée par :*

$$\Pr(\mathrm{C}_r^{E_k}(u_0,u_r) = z) \approx \mathcal{N}(0, \mathrm{ELP}_r^E(u_0,u_r))$$

si $z > 0$ et par 0 sinon.

Ce comportement est à comparer avec la distribution de la corrélation d'un masque linéaire sur les permutations de $\mathbf{F}_2^n$.

Théorème 1.79. *[DR05] Soit $(u_0, u_r) \in (\mathbf{F}_2^m)^2$, $u_r \neq 0$, un masque linéaire. La distribution de $2^{n-1}\mathrm{C}(u_0, u_r)$ sur toutes les permutations F de $\mathbf{F}_2^n$ est :*

$$\Pr_F(2^{n-1}\,\mathrm{C}(u_0,u_r) = 2x) = \frac{\binom{2^{n-1}}{2^{n-2}+x}^2}{\binom{2^n}{2^{n-1}}}.$$

La distribution de la corrélation d'un masque linéaire sur toutes les permutations de $\mathbf{F}_2^n$ est approximée par une loi normale centrée de variance 2^{-n} :

$$\Pr_F(\mathrm{C}(u,v) = i) \approx \mathcal{N}\left(0; 2^{-n}\right).$$

Si $\mathrm{ELP}_r^E(u_0, u_r)$ est proche de 2^{-n}, alors la distribution de la corrélation d'un masque à clé fixée est proche de la distribution de la corrélation de ce masque sur toutes les fonctions de $\mathbf{F}_2^n$ dans $\mathbf{F}_2^n$. C'est pourquoi la variance de la corrélation d'un masque linéaire joue un rôle important, en particulier son écart à 2^{-n}.

Un chiffrement ne possédant pas le comportement attendu a cependant été présenté dans [AÅBL12]. Mais, dans les analyses de sécurité des chiffrements itératifs, la quantité évaluée pour analyser la résistance à la cryptanalyse linéaire d'un chiffrement itératif est le maximum du potentiel linéaire moyen d'un masque.

Définition 1.80. *Le maximum du potentiel linéaire moyen pour un masque sur r tours est*

$$\mathrm{MELP}_r^E = \max_{u,v\neq 0} \mathrm{ELP}_r^E(u, v).$$

Dans la suite de ce document, nous nous intéressons à l'évaluation précise des probabilités des différentielles et du carré des corrélations des masques linéaires, notamment en moyenne sur les clés, c'est-à-dire aux valeurs MEDP et MELP.

2

Critères classiques de résistance aux cryptanalyses différentielle et linéaire

Un réseau de substitution-permutation est résistant à la cryptanalyse différentielle si, pour toute clé, la probabilité maximale des différentielles est suffisamment petite, de l'ordre de 2^{-n}. Dans la majeure partie des analyses de sécurité sur les réseaux de substitution-permutation, plutôt que de déterminer la distribution sur les clés de la probabilité maximale des différentielles, c'est la moyenne sur les clés de la probabilité d'une différentielle sur r tours (a, b) qui est utilisée :

$$\mathrm{EDP}_r^E(a,b) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \Pr_X[E_k(X) + E_k(X+a) = b] \ ,$$

où $(E_k)_k$ est un chiffrement itératif de r tours et de taille de clé κ.

Le maximum sur les différentielles de la valeur EDP, appelé probabilité maximale en moyenne sur les clés d'une différentielle sur r tours, défini par

$$\mathrm{MEDP}_r^E = \max_{a \neq 0, b} \mathrm{EDP}_r^E(a,b) \ ,$$

est considéré comme une bonne estimation pour déterminer si un chiffrement est résistant à la cryptanalyse différentielle [BBL13].

De même, la résistance à la cryptanalyse linéaire est en général estimée par la variance sur les clés de la corrélation (aussi appelée potentiel linéaire moyen) d'un masque linéaire (u, v) sur r tours :

$$\mathrm{ELP}_r^E(u,v) = 2^{-2n-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \left(\sum_{x \in \mathbf{F}_2^n} (-1)^{u \cdot x + v \cdot E_k(x)} \right)^2 \ ,$$

où n et κ désignent respectivement la taille de bloc et la taille de clé du chiffrement $(E_k)_k$. En particulier, nous nous intéressons au maximum sur les masques linéaires sur

r tours de cette valeur, c'est-à-dire au maximum du potentiel linéaire d'un masque sur r tours, qui est défini par

$$\mathrm{MELP}_r^E = \max_{u,v\neq 0} \mathrm{ELP}_r^E(u,v).$$

Dans ce chapitre et les deux suivants, nous allons chercher à déterminer les valeurs MEDP et MELP pour deux tours d'un réseau de substitution-permutation, puisque ces deux valeurs fournissent une bonne estimation de la sécurité d'un tel chiffrement vis-à-vis des cryptanalyses linéaire et différentielle. Nous nous concentrons sur deux tours mais le MEDP_2 et le MELP_2 permettent également d'obtenir une estimation du MEDP et du MELP sur un plus grand nombre de tours.

Considérons un chiffrement $(E_k)_k$ de la forme $\mathrm{SPN}(m,t,S,M)$ de r tours, donc de taille de bloc $n = mt$. La fonction de diffusion M étant linéaire, la probabilité de la différentielle (a,b) sur le chiffrement E_k est égale à la probabilité de la différentielle $(a,M(b))$ sur le chiffrement $M\circ E_k$. La dernière fonction de diffusion linéaire d'un réseau de substitution-permutation n'a donc aucune influence sur la valeur MEDP. Il en est de même pour la valeur MELP. C'est pourquoi les réseaux de substitution-permutation de la forme $\mathrm{SPN}(m,t,S,M)$ et à r tours, $r \geqslant 1$, considérés dans le reste de ce document ne comportent pas de fonction de diffusion M dans le dernier tour.

Dans ce chapitre, nous rappelons les critères permettant d'évaluer la sécurité des réseaux de substitution-permutation, ainsi que les premières estimations du MEDP et du MELP. Nous rappelons dans la première partie les définitions d'uniformité différentielle d'une boîte-S et de branch number différentiel d'une fonction de diffusion. Ces critères ont été établis il y a une quinzaine d'années pour calculer la probabilité maximale des caractéristiques différentielles sur deux tours de l'AES et de n'importe quel réseau de substitution-permutation [Dae95, DR01, DR02]. Cette probabilité donne une première idée du niveau de sécurité d'un tel chiffrement. Dans la deuxième partie, nous énonçons les critères analogues pour la cryptanalyse linéaire : la non-linéarité d'une boîte-S et le branch number linéaire d'une fonction de diffusion [Dae95, DR02]. Dans la troisième partie, nous présentons les premières bornes établies sur les valeurs MEDP_2 et MELP_2 [HLL$^+$00, DR02], ainsi que la borne présentée en 2003 à la conférence FSE [PSLL03, CKL$^+$03].

2.1 Critères de résistance à la cryptanalyse différentielle

La valeur MEDP_2 étant difficile à calculer, la probabilité maximale des caractéristiques sur deux tours d'un réseau de substitution-permutation est souvent considérée comme suffisante pour décider si un chiffrement est résistant à la cryptanalyse différentielle. De plus, comme nous allons le rappeler maintenant, le calcul de la probabilité d'une caractéristique différentielle est assez simple puisque toutes les différences intermédiaires sont déterminées. Pour fixer les notations, commençons par calculer la probabilité d'une différentielle sur un tour d'un réseau de substitution-permutation.

2.1.1 Différentielle sur un tour

Considérons un tour du réseau de substitution-permutation E_k. Comme la dernière fonction de diffusion est ignorée pour l'analyse (puisqu'elle ne modifie pas le MEDP), le calcul de la probabilité d'une différentielle sur un tour revient à calculer la probabilité d'un étage de boîtes-S, noté Sub.

Une différence $a \in \mathbf{F}_2^n$ peut aussi être vue comme un élément $(a_1, \ldots, a_t) \in (\mathbf{F}_2^m)^t$. On définit le support d'une différence comme étant l'ensemble des indices des coordonnées non nulles de a, lorsque a est considérée comme un vecteur de $(\mathbf{F}_2^m)^t$:

$$\mathsf{Supp}(a) = \{i \in \{1, \ldots, t\} \,|\, a_i \neq 0\},$$

et le poids d'une différence comme étant le nombre de coordonnées non nulles de a :

$$wt(a) = \#\mathsf{Supp}(a).$$

De même, on définit le support (respectivement le poids) d'une différentielle $(a, b) = (a_1, \ldots, a_t, b_1, \ldots, b_t) \in (\mathbf{F}_2^m)^{2t}$, où a_i correspond à la différence à l'entrée de la i-ème boîte-S du premier tour du réseau de substitution-permutation et b_j correspond à la différence à la sortie de la j-ème boîte-S du dernier tour, par l'ensemble des indices (respectivement le nombre) des coordonnées non nulles de (a, b), vu comme un vecteur de $(\mathbf{F}_2^m)^{2t}$.

Remarque 2.1. La notion de poids d'une différentielle définie ici et utilisée dans tout ce document correspond au poids au sens de Hamming. Dans [DR07a] et [DR09], les auteurs nomment poids d'une différentielle une notion totalement différente de celle utilisée ici.

La probabilité d'une différentielle (a, b) sur un étage de boîtes-S est égale au produit des probabilités des différentielles (a_i, b_i) sur chacune des boîtes-S :

$$\mathrm{DP}^{\mathsf{Sub}}(a, b) = \prod_{i=1}^{t} \mathrm{DP}^{S}(a_i, b_i).$$

La quantité qui intervient dans le calcul de la probabilité d'une différentielle sur une boîte-S correspond à la notion de spectre différentiel (voir chapitre 1, partie 1.5.2). Si S est une fonction de $\mathbf{F}_2^m$ dans $\mathbf{F}_2^m$, nous avons

$$\mathrm{DP}^{S}(\alpha, \beta) = 2^{-m}\delta(\alpha, \beta).$$

Nous pouvons donc donner la probabilité d'une différentielle sur un tour d'un réseau de substitution-permutation en fonction du spectre différentiel de la boîte-S de ce chiffrement.

Lemme 2.2. *[DR02] Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$. La probabilité d'une différentielle (a, b) pour la fonction de substitution Sub est :*

$$\mathrm{DP}^{\mathsf{Sub}}(a, b) = \begin{cases} 0 & \text{si } \mathsf{Supp}(a) \neq \mathsf{Supp}(b) \\ 2^{-mwt(a)} \displaystyle\prod_{i \in \mathsf{Supp}(a)} \delta(a_i, b_i) & \text{sinon.} \end{cases}$$

Démonstration. Si les supports de a et de b ne coïncident pas, il existe au moins un indice $i \in \{1, \ldots, t\}$ tel que a_i soit nul et b_i ne le soit pas (ou l'inverse). Comme la boîte-S est une permutation, on a $\delta(a_i, b_i) = 0$ donc $\mathrm{DP}^{\mathsf{Sub}}(a, b) = 0$.

Si les supports de a et b coïncident, nous avons :

$$\begin{aligned} \mathrm{DP}^{\mathsf{Sub}}(a,b) &= \prod_{i \in \mathsf{Supp}(a)} \mathrm{DP}^S(a_i, b_i) \times \prod_{i \notin \mathsf{Supp}(a)} \mathrm{DP}^S(0,0) \\ &= \prod_{i \in \mathsf{Supp}(a)} (2^{-m}\delta(a_i, b_i)) \times 1. \end{aligned}$$

□

Le paramètre essentiel ici est donc l'uniformité différentielle de la boîte-S, définie par

$$\Delta(S) = \max_{\alpha \neq 0, \beta} \delta(\alpha, \beta) = \max_{\alpha \neq 0, \beta} \#\{x \in \mathbf{F}_2^m, S(x+\alpha) + S(x) = \beta\}.$$

Naturellement, plus l'uniformité différentielle de la boîte-S est petite, plus la probabilité d'une différentielle sur un tour est petite. Ce critère a été complété dans [DR02], où la probabilité d'une caractéristique différentielle sur deux tours est étudiée.

2.1.2 Caractéristique différentielle sur deux tours

Considérons deux tours du réseau de substitution-permutation E_k de la forme $\mathrm{SPN}(m, t, S, M)$ et de taille de bloc $n = mt$. Rappelons qu'une caractéristique différentielle Q sur deux tours est un triplet de différences $Q = (a, c, b) \in (\mathbf{F}_2^n)^3$ tel que a est la différence en entrée de $(E_k)_k$, c est la différence en sortie du premier étage de boîtes-S et b la différence en sortie du deuxième étage de boîtes-S.

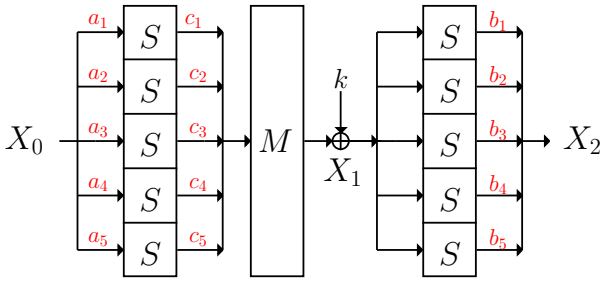


FIGURE 2.1 – Une caractéristique sur deux tours d'un chiffrement de la forme $\mathrm{SPN}(m, 5, S, M)$.

L'espérance sur les clés de la probabilité d'une caractéristique différentielle sur deux tours est :

$$\mathrm{EDCP}_2^E(Q) = 2^{-\kappa} \sum_{k \in \mathbf{F}_2^\kappa} \Pr_{X_0}(X_1 + X_1' = M(c); X_2 + X_2' = b \,|\, X_0 + X_0' = a),$$

où κ est la taille de la clé du chiffrement, X_1 (resp. X'_1) est l'image de X_0 (resp. X'_0) après le premier tour du chiffrement et X_2 (resp. X'_2) est l'image de X_0 (resp. X'_0) après le deuxième étage de boîtes-S.

Lorsque le chiffrement est un chiffrement de Markov, comme dans le cas des réseaux de substitution-permutation où les sous-clés sont indépendantes et uniformément distribuées, la différence en sortie du i-ème tour dépend uniquement de la différence en sortie du $(i-1)$-ème tour, donc la probabilité d'une caractéristique sur r tours est égale au produit des probabilités des différentielles pour chaque tour. Nous avons donc :

$$\mathrm{EDCP}_2(Q) = \mathrm{EDP}_1^E(a, M(c)) \times \mathrm{EDP}_1^E(M(c), b).$$

Comme les fonctions de tour d'un chiffrement alternant clés et permutations sont indépendantes de la clé, la probabilité en moyenne sur les clés correspond à la probabilité de la fonction de tour privée de l'addition de clé. De plus, nous avons supposé que le deuxième et dernier tour ne comporte pas de fonction de diffusion. Donc :

$$\begin{aligned} \mathrm{EDCP}_2(Q) &= \mathrm{DP}^{M\circ\mathsf{Sub}}(a, M(c)) \times \mathrm{DP}^{\mathsf{Sub}}(M(c), b) \\ &= \mathrm{DP}^{\mathsf{Sub}}(a, c) \times \mathrm{DP}^{\mathsf{Sub}}(M(c), b) \end{aligned}$$

où Sub est la fonction qui correspond à un étage de boîtes-S.

D'après le lemme 2.2, la probabilité de la caractéristique Q est nulle si les supports de a et c ou ceux de $M(c)$ et b ne coïncident pas. Lorsque ces supports coïncident, c'est-à-dire $\mathsf{Supp}(a) = \mathsf{Supp}(c)$ et $\mathsf{Supp}(M(c)) = \mathsf{Supp}(b)$, nous avons :

$$\begin{aligned} \mathrm{EDCP}_2(Q) &= \left(2^{-m\,wt(c)} \prod_{i\in\mathsf{Supp}(c)} \delta(a_i, c_i)\right)\left(2^{-m\,wt(M(c))} \prod_{j\in\mathsf{Supp}(M(c))} \delta((M(c))_j, b_j)\right) \\ &= 2^{-m(wt(c)+wt(M(c)))}\left(\prod_{i\in\mathsf{Supp}(c)} \delta(a_i, c_i)\right)\left(\prod_{j\in\mathsf{Supp}(M(c))} \delta((M(c))_j, b_j)\right) \end{aligned} \quad (2.1)$$

En majorant les coefficients $\delta(\alpha, \beta)$, α, $\beta \in \mathbf{F}_2^m$, par leur valeur maximale, l'uniformité différentielle $\Delta(S)$, nous obtenons une majoration de la probabilité d'une caractéristique différentielle sur deux tours [Dae95] :

$$\mathrm{EDCP}_2(Q) \leq \left(2^{-m}\Delta(S)\right)^{wt(c)+wt(M(c))}. \quad (2.2)$$

Cette majoration dépend de l'uniformité différentielle de la boîte-S et du poids des vecteurs de la forme $(c, M(c))$, $c \in (\mathbf{F}_2^m)^t$: plus le poids du vecteur $(c, M(c))$ est grand et plus l'uniformité différentielle est petite, plus la probabilité maximale de la caractéristique différentielle est petite. Cette remarque a amené à la définition de deux critères de résistance à la cryptanalyse différentielle concernant respectivement la boîte-S et la fonction de diffusion : l'uniformité différentielle et le branch number différentiel. Le premier critère pour construire un réseau de substitution-permutation résistant à la cryptanalyse différentielle est de choisir une boîte-S avec la plus petite uniformité différentielle possible. Le deuxième critère est de choisir la fonction de diffusion M du

chiffrement avec le plus grand branch number différentiel possible.

Comme la fonction de diffusion du chiffrement est linéaire sur $\mathbf{F}_2$, le code formé des mots $(c, M(c))$, $c \in (\mathbf{F}_2^m)^t$ est un code additif : si $(c_1, M(c_1))$ et $(c_2, M(c_2))$ sont deux mots du code, alors $(c_1 + c_2, M(c_1) + M(c_2))$ l'est aussi. Donc le critère qui nous intéresse, le plus petit poids pour un mot non nul du code, correspond à la distance minimale de ce code.

Définition 2.3. *[Dae95, p.107] Soit M une permutation linéaire de $(\mathbf{F}_2^m)^t$. Nous associons à M le code $\mathcal{C}_M$ défini par*

$$\mathcal{C}_M = \{(c, M(c)),\ c \in (\mathbf{F}_2^m)^t\}\ .$$

Le code $\mathcal{C}_M$ est linéaire sur $\mathbf{F}_2$, de longueur $2t$ et de taille 2^{mt} sur $\mathbf{F}_2^m$. On appelle branch number différentiel de M relativement à $\mathbf{F}_2^m$ la distance minimale du code $\mathcal{C}_M$.

Pour une caractéristique différentielle sur deux tours d'un réseau de substitution-permutation, le branch number différentiel correspond au nombre minimal de *boîtes-S actives* dans cette caractéristique, c'est-à-dire au nombre de boîtes-S dont les différences en entrée et en sortie sont non nulles.

Définition 2.4. *Une boîte-S est dite* active *dans une caractéristique différentielle si les différences en entrée et en sortie de cette boîte-S sont non nulles.*

D'après la borne de Singleton, la plus grande valeur possible pour le branch number différentiel d'une permutation de $(\mathbf{F}_2^m)^t$ est $(t+1)$. Il s'agit du cas où le code associé à M est MDS (voir partie 1.3.3). Les fonctions de diffusion dont le code associé est MDS sont donc celles qui ont la meilleure résistance à la cryptanalyse différentielle.

Le lemme suivant résume les résultats que nous venons d'exposer : il donne une borne supérieure pour la probabilité d'une caractéristique différentielle sur deux tours d'un réseau de substitution-permutation. Cette borne dépend des deux critères précédemment énoncés : l'uniformité différentielle de la boîte-S et le branch number différentiel de la fonction de diffusion.

Lemme 2.5. *[DR02] Soit $(E_k)_k$ un chiffrement de la forme* $\mathrm{SPN}(m, t, S, M)$*. Notons d le branch number différentiel de M. Alors la probabilité de toute caractéristique différentielle Q sur deux tours vérifie :*

$$\mathrm{EDCP}_2^E(Q) \leq \left(2^{-m}\Delta(S)\right)^d\ .$$

En particulier, cette borne est atteinte si pour tout élément non nul α de $\mathbf{F}_2^m$, il existe deux éléments β et γ tels que $\delta(\alpha, \beta) = \delta(\gamma, \alpha) = \Delta(S)$.

Démonstration. La borne du lemme est obtenue grâce à l'équation (2.2) et à la définition de branch number différentiel. Soit $(c, M(c))$ un mot de poids minimal de $\mathcal{C}_M$. Pour tout $i \in \mathsf{Supp}(c)$, il existe par hypothèse un élément a_i de $\mathbf{F}_2^m$ tel que $\delta(a_i, c_i) = \Delta(S)$. De même, pour tout $j \in \mathsf{Supp}(M(c))$, il existe un élément b_j de $\mathbf{F}_2^m$ tel que $\delta((M(c))_j, b_j) = \Delta(S)$. Si i (resp. j) n'appartient pas au support de c (resp.

$M(c)$), alors on choisit $a_i = 0$ (resp. $b_j = 0$). Les éléments a_i, $i = 1, \ldots, t$, définissent un mot a de même support que c, et les éléments b_j, $j = 1, \ldots, t$, définissent un mot b de même support que $M(c)$. Ces deux mots vérifient donc

$$\mathrm{EDCP}_2^E(a, c, b) = \left(2^{-m}\Delta(S)\right)^d \ .$$

□

La condition sur la table des différences pour avoir égalité dans le lemme 2.5 est d'avoir un coefficient égal à $\Delta(S)$ sur chaque ligne et chaque colonne de la table. Cette condition est par exemple vérifiée pour les boîtes-S affinement équivalentes à des fonctions puissances (comme S présentée dans l'exemple 1.42 de la section 1.5.2).

L'AES a été conçu de façon à minimiser la probabilité d'une caractéristique différentielle sur deux tours. Rappelons que deux tours de l'AES peuvent être vus comme l'application en parallèle de quatre permutations de $\mathbf{F}_2^{32}$ appelées superboîtes-S (*cf* page 10). La probabilité d'une caractéristique sur deux tours sera minimale si la différence d'entrée est nulle sur trois des quatre superboîtes-S. La probabilité minimale d'une caractéristique différentielle sur deux tours de l'AES est donc égale à la probabilité minimale d'une caractéristique différentielle sur une superboîte-S. Or une superboîte-S correspond à deux tours d'un réseau de substitution-permutation de la forme $\mathrm{SPN}(8, 4, S, M_c)$, où S et M_c sont les composants de l'AES définis à la page 7. L'uniformité différentielle de la boîte-S de l'AES est la plus petite connue ($\Delta(S) = 4$, puisqu'aucune permutation APN n'a été trouvée sur $\mathbf{F}_2^8$) et M_c, la fonction de diffusion sur $\mathbf{F}_2^{32}$, est de branch number différentiel maximal ($d = 5$, le code associé est un code de longueur 8 et de taille 2^{32} qui est donc MDS sur $\mathbf{F}_2^8$). En appliquant le lemme à une superboîte-S, on obtient que l'espérance maximale de la probabilité d'une caractéristique sur deux tours de l'AES est au plus égale à 2^{-30} (puisque la boîte-S de l'AES, qui est affinement équivalente à une fonction puissance, vérifie les hypothèses du lemme 2.5). De plus, il existe des caractéristiques différentielles qui atteignent cette valeur : par exemple, l'espérance de la probabilité de la caractéristique $Q = (a, c, b)$ avec $a = (\texttt{0x75}, 0, 0, 0)$, $c = (\texttt{0xfe}, 0, 0, 0)$ et $b = (\texttt{0xf7}, \texttt{0xd8}, \texttt{0xd8}, \texttt{0xb7})$ sur une superboîte-S est égale à 2^{-30}.

Cette analyse se généralise naturellement à un plus grand nombre de tours : l'espérance de la probabilité d'une caractéristique différentielle sur r tours est inférieure ou égale à $(2^{-m}\Delta(S))^{d_r}$, où d_r est le nombre minimal de boîtes-S actives sur r tours. Pour estimer l'évolution de d_r lorsque le nombre de tours r augmente, des propriétés combinatoires ou des algorithmes sont utilisés. En particulier, dans [MWGP12], les auteurs présentent un algorithme qui calcule d_r et qui s'applique à un grand nombre de chiffrements. Cet algorithme utilise la programmation linéaire, c'est-à-dire l'optimisation d'une fonction linéaire sous certaines conditions sur les variables. Une fois le chiffrement décrit sous forme d'un problème de programmation linéaire, les techniques de résolution de ce domaine sont utilisées pour retrouver le nombre minimum de boîtes-S actives du chiffrement.

Pour quatre tours de l'AES, il y a $d_4 = 25$ boîtes-S actives au minimum. Ce résultat vient de la construction de l'AES et s'obtient par exemple avec la représentation du chiffrement en superboîtes-S (*cf* page 10). Avec cette représentation, quatre tours de

l'AES correspondent à deux tours du chiffrement de la forme $\mathrm{SPN}(32, 4, \mathrm{SBS}, \mathrm{SR} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR})$, où la fonction de diffusion a un branch number égal à 5. Il y a donc au minimum 5 superboîtes-S actives, chacune contenant au minimum 5 boîtes-S actives. Cette construction a été généralisée dans [ADK+14] par la construction par entrelacement.

2.1.3 Différentielle sur deux tours

Cependant, la complexité des attaques est déterminée par la probabilité des différentielles et non des caractéristiques. La différence entre ces deux notions est que dans une caractéristique, les différences intermédiaires sont fixées, tandis que ces différences sont libres dans une différentielle. Or les différences intermédiaires ne sont pas connues par l'attaquant. La probabilité d'une différentielle (a, b) correspond donc à la somme des probabilités de toutes les caractéristiques ayant a comme différence en entrée du chiffrement et b comme différence de sortie. Ainsi, même si les probabilités des caractéristiques sont très petites, il pourrait exister une différentielle telle que la somme des probabilités des caractéristiques qui la composent (c'est-à-dire ayant les mêmes différences en entrée et sortie du chiffrement) soit grande. Donc la quantité qui permet effectivement de s'assurer de la résistance d'un chiffrement à la cryptanalyse différentielle est la probabilité des différentielles.

En particulier, pour deux tours d'un réseau de substitution-permutation, nous cherchons à estimer le maximum de la probabilité d'une différentielle, cette probabilité vérifiant :

$$\mathrm{EDP}_2(a, b) = \sum_{x \in \mathbf{F}_2^{mt}} \mathrm{EDCP}_2^E(a, x, b) \; .$$

D'après la borne (2.2), pour tout élément x de $(\mathbf{F}_2^m)^t$, la probabilité de la caractéristique (a, x, b) est non nulle si et seulement si les supports de x et a et ceux de $M(x)$ et b coïncident, c'est-à-dire si l'élément $c = (x, M(x)) \in (\mathbf{F}_2^m)^{2t}$ est un mot du code $\mathcal{C}_M$ associé à la fonction de diffusion de même support que la différentielle (a, b). La probabilité d'une différentielle sur deux tours est donc donnée dans le lemme suivant.

Lemme 2.6. *Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$. L'espérance de la probabilité de la différentielle (a, b) sur deux tours vérifie :*

$$\mathrm{EDP}_2(a, b) = 2^{-m\, wt(a,b)} \sum_{\substack{c \in \mathcal{C}_M : \\ \mathsf{Supp}(c) = \mathsf{Supp}(a,b)}} \left(\prod_{i \in \mathsf{Supp}(a)} \delta^S(a_i, c_i) \right) \left(\prod_{j \in \mathsf{Supp}(b)} \delta^S(c_{t+j}, b_j) \right) .$$

2.2 Critères de résistance à la cryptanalyse linéaire

Nous présentons ici les critères de sécurité développés classiquement pour étudier la résistance aux attaques linéaires de l'AES et des réseaux de substitution-permutation. Le premier critère, la linéarité, apparaît lorsqu'un tour est étudié et ne dépend que de la boîte-S utilisée dans le chiffrement. Le deuxième critère dépend de la fonction de diffusion et est similaire au branch number différentiel.

2.2.1 Masque linéaire sur un tour

Considérons un tour du chiffrement $(E_k)_k$ de la forme $\mathrm{SPN}(m,t,S,M)$. Le support et le poids d'un masque linéaire (u,v) sont définis comme pour une différentielle, c'est-à-dire qu'ils correspondent au support et au poids de (u,v) vu comme un élément de $(\mathbf{F}_2^m)^{2t}$: $(u,v)=(u_1,\ldots,u_t,v_1,\ldots,v_t)\in(\mathbf{F}_2^m)^{2t}$.

Comme la dernière fonction de diffusion est ignorée pour l'analyse, le calcul du potentiel linéaire sur un tour revient à calculer le carré de la corrélation d'un étage de boîtes-S, noté Sub.

La corrélation (respectivement le potentiel linéaire moyen) d'un masque (u,v) sur un étage de boîtes-S est égal au produit des corrélations (respectivement des potentiels linéaires moyens) (u_i,v_i), $1\leqslant i\leqslant t$, sur chaque boîte-S. En effet, si $u=(u_1,u_2)$ et $v=(v_1,v_2)$ sont deux éléments de $(\mathbf{F}_2^m)^2$ et qu'un étage de boîtes-S est composé de deux boîtes-S, alors nous avons par bilinéarité du produit scalaire :

$$u\cdot x+v\cdot\mathsf{Sub}(x)=(u_1,0)\cdot(x_1,0)+(0,u_2)\cdot(0,x_2)+(v_1,0)\cdot(S(x_1),0)+(0,v_2)\cdot(0,S(x_2)).$$

Donc

$$\begin{aligned}\mathrm{C}^{\mathsf{Sub}}(u,v) &= 2^{-2m}\sum_{x_1,x_2\in\mathbf{F}_2^m}(-1)^{(u_1,0)\cdot(x_1,0)+(v_1,0)\cdot(S(x_1),0)}\,(-1)^{(0,u_2)\cdot(0,x_2)+(0,v_2)\cdot(0,S(x_2))}\\ &= \left(2^{-m}\sum_{x_1\in\mathbf{F}_2^m}(-1)^{u_1\cdot x_1+v_1\cdot S(x_1)}\right)\left(2^{-m}\sum_{x_2\in\mathbf{F}_2^m}(-1)^{u_2\cdot x_2+v_2\cdot S(x_2)}\right)\\ &= \mathrm{C}^S(u_1,v_1)\,\mathrm{C}^S(u_2,v_2).\end{aligned}$$

La quantité qui intervient dans le calcul de la corrélation (respectivement du potentiel linéaire moyen) d'un masque $(\alpha,\beta)\in\mathbf{F}_2^m$ sur une boîte-S correspond à la notion de spectre de Walsh de la boîte-S (voir partie 1.7.3). Comme les coefficients du spectre de Walsh vérifient $\mathcal{W}^S(u,0)=0$ pour tout élément $u\in\mathbf{F}_2^m$ et $\mathcal{W}^S(0,v)=0$ pour tout élément $v\in\mathbf{F}_2^m$ si et seulement si S est une permutation, nous en déduisons le lemme suivant.

Lemme 2.7. *Le potentiel linéaire d'un masque (u,v) sur un étage de boîte-S vérifie :*

$$\mathrm{ELP}^{\mathsf{Sub}}(u,v)=\begin{cases}0 & \text{si } \mathsf{Supp}(u)\neq\mathsf{Supp}(v)\\ 2^{-2m\,wt(u)}\displaystyle\prod_{i\in\mathsf{Supp}(u)}\mathcal{W}^S(u_i,v_i)^2 & \text{sinon.}\end{cases}$$

Le paramètre essentiel ici est la linéarité, définie pour une fonction S de $\mathbf{F}_2^m$ dans $\mathbf{F}_2^m$ par :

$$\mathcal{L}(S)=\max_{\alpha,\beta\neq0}\left|\mathcal{W}^S(\alpha,\beta)\right|=\max_{\alpha,\beta\neq0}\left|\sum_{x\in\mathbf{F}_2^m}(-1)^{\alpha\cdot x+\beta\cdot S(x)}\right|,$$

où $\cdot$ est le produit scalaire dans $\mathbf{F}_2^m$ et $|x|$ est la valeur absolue de x.

Nous allons voir que de façon similaire à l'espérance de la probabilité d'une différentielle sur deux tours, le potentiel linéaire moyen d'un masque sur deux tours dépend du potentiel linéaire maximal sur les boîtes-S et de la distance minimale d'un code linéaire défini par la fonction de diffusion.

2.2.2 Chemin linéaire sur deux tours

Considérons deux tours du réseau de substitution-permutation E_k de la forme $\mathrm{SPN}(m,t,S,M)$ et de taille de bloc $n = mt$. Le potentiel linéaire moyen d'un chemin (u,c,v) sur deux tours d'un réseau de substitution-permutation est :

$$\begin{aligned}\mathrm{ELTP}_2^E(u,c,v) &= 2^{-\kappa}\sum_{k\in\mathbf{F}_2^\kappa}\left[(-1)^{c\cdot k_1+v\cdot k_2}\mathrm{C}^{M\circ\mathsf{Sub}}(u,c)\mathrm{C}^{\mathsf{Sub}}(c,v)\right]^2\\ &= \left[\mathrm{C}^{M\circ\mathsf{Sub}}(u,c)\mathrm{C}^{\mathsf{Sub}}(c,v)\right]^2,\end{aligned}$$

où Sub est la fonction correspondant à l'étage de boîtes-S. Rappelons que l'adjoint de M est noté M^* et vérifie : pour tous x, y éléments de $\mathbf{F}_2^n$, $x\cdot M(y) = M^*(x)\cdot y$. En utilisant la définition de la corrélation d'un masque linéaire et celle de l'adjoint d'une fonction linéaire, nous avons :

$$\begin{aligned}\mathrm{C}^{M\circ\mathsf{Sub}}(u,c) &= 2^{-n}\sum_{x\in\mathbf{F}_2^n}(-1)^{u\cdot x+c\cdot M\circ\mathsf{Sub}(x)}\\ &= 2^{-n}\sum_{x\in\mathbf{F}_2^n}(-1)^{u\cdot x+M^*(c)\cdot\mathsf{Sub}(x)}\\ &= \mathrm{C}^{\mathsf{Sub}}(u,M^*(c)).\end{aligned}$$

Pour tous les éléments $u=(u_1,\ldots,u_t)$ et $v=(v_1,\ldots,v_t)$ de $(\mathbf{F}_2^m)^t$, la corrélation d'un masque linéaire (u,v) sur un étage de boîtes-S est égal au produit des corrélations des masques (u_i,v_i) sur chacune des boîtes-S. Nous obtenons donc :

$$\begin{aligned}\mathrm{ELTP}_2^E(u,c,v) &= \left(\mathrm{C}^{\mathsf{Sub}}(u,M^*(c))\mathrm{C}^{\mathsf{Sub}}(c,v)\right)^2\\ &= \left(\prod_{i=0}^{t}\mathrm{C}^S(u_i,(M^*(c))_i)\right)^2\left(\prod_{j=0}^{t}\mathrm{C}^S(c_j,v_j)\right)^2\end{aligned}$$

et en appliquant le lemme 2.7, si $\mathsf{Supp}(u)=\mathsf{Supp}(M^*(c))$ et $\mathsf{Supp}(c)=\mathsf{Supp}(v)$, alors :

$$\mathrm{ELTP}_2^E(u,c,v) = 2^{-2m(wt(M^*(c))+wt(c))}\left(\prod_{i\in\mathsf{Supp}(M^*(c))}\mathcal{W}^S(u_i,(M^*(c))_i)\right)^2\left(\prod_{j\in\mathsf{Supp}(c)}\mathcal{W}^S(c_j,v_j)\right)^2. \tag{2.3}$$

Le code linéaire qui intervient dans le calcul du potentiel linéaire moyen d'un masque sur deux tours est donc le suivant.

Définition 2.8. *[Dae95] Soit M une permutation linéaire de $(\mathbf{F}_2^m)^t$. Nous associons à M le code $\mathcal{C}_M^\perp$ défini par*

$$\mathcal{C}_M^\perp = \{(M^*(c),c),\ c\in(\mathbf{F}_2^m)^t\}\ .$$

Le code $\mathcal{C}_M^\perp$ est linéaire sur $\mathbf{F}_2$, de longueur $2t$ et de taille 2^{mt} sur $\mathbf{F}_2^m$. On appelle branch number linéaire de M la distance minimale du code $\mathcal{C}_M^\perp$.

Le code $\mathcal{C}_M^\perp$ et sa distance minimale permettent d'estimer la résistance d'un chiffrement à la cryptanalyse linéaire, de façon similaire au code $\mathcal{C}_M$ pour la cryptanalyse différentielle. Ces codes ont la même longueur et la même taille. De plus, ces codes sont duaux : si c_1 et c_2 sont deux éléments de $(\mathbf{F}_2^m)^t$, on a

$$(c_1, M(c_1)) \cdot (M^*(c_2), c_2) = c_1 \cdot M^*(c_2) + M(c_1) \cdot c_2 = 0.$$

C'est pourquoi ce code est noté $\mathcal{C}_M^\perp$. D'après la borne de Singleton, la valeur maximale des distances minimales sur $\mathbf{F}_2^m$ de ces deux codes est $t+1$, les codes atteignant cette valeur maximale sont dits MDS. Comme le code dual d'un code MDS est MDS (*cf* proposition 1.13), les fonctions de diffusion qui ont un branch number différentiel maximal ont aussi un branch number linéaire maximal. Donc les fonctions de diffusion qui résistent le mieux à la cryptanalyse différentielle sont aussi celles qui résistent le mieux à la cryptanalyse linéaire.

Lemme 2.9. *[DR02] Soit $(E_k)_k$ un chiffrement de la forme* $\mathrm{SPN}(m, t, S, M)$*. Notons $d^\perp$ le branch number linéaire de M. Alors le potentiel linéaire d'un chemin sur deux tours vérifie :*

$$\mathrm{ELTP}_2^E(u, c, v) \leqslant \left(2^{-m}\mathcal{L}(S)\right)^{2d^\perp} .$$

En particulier, cette borne est atteinte si pour tout élément non nul α de $\mathbf{F}_2^m$, il existe deux éléments β et γ tels que $|\mathcal{W}^S(\alpha, \beta)| = |\mathcal{W}^S(\gamma, \alpha)| = \mathcal{L}(S)$.

Le cas où la borne est atteinte est obtenu de façon similaire à ce qui est fait dans la démonstration du lemme 2.5.

La condition sur la table de Walsh pour avoir égalité dans le lemme 2.9 est d'avoir un coefficient égal à $\pm\mathcal{L}(S)$ sur chaque ligne et chaque colonne de la table. Cette condition est par exemple vérifiée pour les boîtes-S affinement équivalentes à des fonctions puissances.

Les chiffrements qui résistent le mieux à la cryptanalyse linéaire sont donc ceux dont la boîte-S a la plus petite linéarité possible et dont la fonction de diffusion a un branch number maximal.

La boîte-S de l'AES a la plus petite linéarité connue pour une permutation de $\mathbf{F}_2^8$: $\mathcal{L}(S) = 32$ et sa fonction de diffusion est de branch number linéaire maximal $d^\perp = 5$ (puisque $d^\perp = d$). En appliquant le lemme précédent à l'AES, nous obtenons donc que le potentiel linéaire maximal d'un chemin linéaire sur deux tours est égal à 2^{-30}.

2.2.3 Masque linéaire sur deux tours

La complexité des attaques linéaires est déterminée par le potentiel linéaire moyen d'un masque sur deux tours. Le potentiel linéaire moyen du masque (u, v) se calcule à partir du potentiel linéaire moyen des chemins dont le masque d'entrée est u et celui de sortie est v :

$$\mathrm{ELP}_2^E(u, v) = \sum_{c \in \mathbf{F}_2^{mt}} \mathrm{ELTP}_2^E(u, c, v).$$

D'après l'équation (2.3), le potentiel linéaire moyen d'un masque sur deux tours d'un réseau de substitution-permutation se calcule en fonction des mots du code $\mathcal{C}_M^\perp$ et des coefficients du spectre de Walsh de la boîte-S comme le montre le lemme suivant.

Lemme 2.10. *Soit $(E_k)_k$ un chiffrement de la forme* $\mathrm{SPN}(m,t,S,M)$*. Le potentiel linéaire moyen d'un masque (u,v) sur deux tours vérifie :*

$$\mathrm{ELP}_2(u,v) = 2^{-2m\,wt(u,v)} \sum_{\substack{c\in\mathcal{C}_M^\perp:\\ \mathsf{Supp}(c)=\mathsf{Supp}(u,v)}} \left(\prod_{i\in\mathsf{Supp}(u)} \mathcal{W}^S(u_i,c_i)\right)^2 \left(\prod_{j\in\mathsf{Supp}(v)} \mathcal{W}^S(c_{t+j},v_j)\right)^2.$$

Les formules donnant l'espérance de la probabilité d'une différentielle sur deux tours et le potentiel linéaire moyen d'un masque sur deux tours (lemmes 2.6 et 2.10) sont similaires : elles ne diffèrent que par l'utilisation des coefficients du spectre différentiel dans un cas, et du carré des coefficients du spectre de Walsh dans l'autre. Nous utiliserons dans la suite de ce document une notation générique, en remplaçant les $2^m \times 2^m$ coefficients $2^{-m}\delta(\alpha,\beta)$, ou les coefficients $2^{-2m}\mathcal{W}(\alpha,\beta)^2$, $\alpha,\ \beta \in \mathbf{F}_2^m$, par une matrice $(\Lambda(\alpha,\beta))_{\alpha,\beta\in\mathbf{F}_2^m}$ dont les coefficients $\Lambda(\alpha,\beta)$ vérifient les trois propriétés suivantes :

$$\begin{aligned}
&- \forall(\alpha,\beta) \in (\mathbf{F}_{2^m})^2,\ \Lambda(\alpha,\beta) \in [0;1];\\
&- \forall\alpha \neq 0,\ \Lambda(\alpha,0) = \Lambda(0,\alpha) = 0;\\
&- \forall\alpha \in \mathbf{F}_{2^m},\ \sum_{\beta\in\mathbf{F}_{2^m}} \Lambda(\alpha,\beta) = \sum_{\beta\in\mathbf{F}_{2^m}} \Lambda(\beta,\alpha) = 1.
\end{aligned} \tag{2.4}$$

Il est assez simple de voir que les coefficients $2^{-m}\delta(\alpha,\beta)$ et les coefficients $2^{-2m}\mathcal{W}(\alpha,\beta)^2$, $\alpha,\ \beta \in \mathbf{F}_2^m$, vérifient ces trois propriétés. Tous les résultats que nous allons obtenir dans les chapitres suivants sont donc valides pour toute matrice Λ vérifiant ces trois conditions. Pour une matrice de ce type, nous allons donc étudier la quantité

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{c\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i,c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j},b_j)\right)\\
&= \sum_{\substack{c\in\mathcal{C}:\\ \mathsf{Supp}(c)=\mathsf{Supp}(a,b)}} \left(\prod_{i\in\mathsf{Supp}(a)} \Lambda(a_i,c_i)\right)\left(\prod_{j\in\mathsf{Supp}(b)} \Lambda(c_{t+j},b_j)\right).
\end{aligned}$$

2.3 Bornes sur le MEDP_2 et le MELP_2

Dans [HLL+00] et [DR02, Section B.2], les auteurs déduisent du lemme 2.6 une première borne supérieure sur la valeur du MEDP_2 dans le cas où le branch number différentiel de la fonction de diffusion est maximal. Une borne similaire sur la valeur du MELP_2 dans le cas où le branch number linéaire de la fonction de diffusion est maximal est déduite du lemme 2.10.

Théorème 2.11. *[HLL$^+$00, DR02] Soit $(E_k)_k$ un réseau de substitution-permutation de la forme $\mathrm{SPN}(m,t,S,M)$ avec M de branch number maximal ($d = d^\perp = t+1$). Alors :*

$$\mathrm{MEDP}_2 \leqslant \left(2^{-m}\Delta(S)\right)^t \quad et \quad \mathrm{MELP}_2 \leqslant \left(2^{-m}\mathcal{L}(S)\right)^{2t} .$$

Remarque 2.12. Ce théorème est vrai pour toute fonction de diffusion : si $(E_k)_k$ est un réseau de substitution-permutation de la forme $\mathrm{SPN}(m,t,S,M)$, alors :

$$\mathrm{MEDP}_2 \leqslant \left(2^{-m}\Delta(S)\right)^{d-1} \quad \text{et} \quad \mathrm{MELP}_2 \leqslant \left(2^{-m}\mathcal{L}(S)\right)^{2(d^\perp-1)} ,$$

où d est le branch number différentiel de M et $d^\perp$ est le branch number linéaire de M.

Dans la suite de ce document, lorsque nous nous référerons à ce théorème, nous considérerons le cas où les branch number différentiel et linéaire de la fonction de diffusion sont quelconques. La démonstration présentée ici reprend les idées de [HLL$^+$00, DR02] mais nous utilisons des notations différentes. Ces notations seront reprises dans la suite de ce document.

Pour démontrer ces bornes, les auteurs ont étudié les motifs formés par certains mots des codes $\mathbf{F}_2$-linéaires. Le lemme ci-dessous présente la propriété de ces codes qui a permis de démontrer le théorème 2.11.

Notation 2.13. Soit c un mot de longueur n et I un sous-ensemble de $\{1,\ldots,n\}$. La décomposition de c selon I est notée $(x,y)_I$: x correspond à la restriction de c à I, et y correspond à la restriction de c au sous-ensemble complémentaire $\bar{I}$ de I. Pour simplifier la notation, les $\#I$ coordonnées de x (resp. les coordonnées de y) ont pour indice les éléments de I (resp. de $\bar{I}$), *i.e.*, $x_i = c_i$ pour tout $i \in I$ et $y_j = c_j$ pour tout $j \in \bar{I}$.

Lemme 2.14. *Soit $\mathcal{C}$ un code $\mathbf{F}_2$-linéaire de longueur n, de dimension mk et de distance minimum d sur $\mathbf{F}_2^m$. Pour tout sous-ensemble $I \subset \{1,\ldots,n\}$ de taille $(n-d)$, et pour tout $x \in (\mathbf{F}_2^m)^{n-d}$, notons*

$$Z(I,x) = \{y \ : (x,y)_I \in \mathcal{C}\} .$$

Alors, pour tout I de taille $(n-d)$, pour tout $x \in (\mathbf{F}_2^m)^{n-d}$:

- *soit $Z(I,x)$ est vide ;*
- *soit $Z(I,x)$ contient un unique élément ;*
- *soit $\#Z(I,x) \leqslant 2^m$ et pour tout $j \in \bar{I}$, pour tous éléments distincts y et y' de $Z(I,x)$, $y_j \neq y'_j$.*

Démonstration. Supposons que $Z(I,x)$ contient au moins deux éléments distincts y et y'. Supposons qu'il existe $j \in \bar{I}$ tel que $y_j = y'_j$. Comme (x,y) et (x,y') sont des mots du code $\mathcal{C}$ qui est $\mathbf{F}_2$-linéaire, $(x,y)+(x,y')$ est aussi un mot de $\mathcal{C}$. Or les $(n-d)$ coordonnées $i \in I$ du mot $(x,y)+(x,y')$ ainsi que la coordonnée j de ce mot sont nulles, c'est-à-dire que $(x,y)+(x,y')$ est un mot de $\mathcal{C}$ de poids au plus $n-(n-d+1) = d-1$, ce qui est en contradiction avec le fait que la distance minimale de $\mathcal{C}$ est d. Donc pour tout $j \in \bar{I}$, les valeurs y_j et y'_j sont distinctes. Comme il y a au maximum 2^m valeurs distinctes possibles pour une coordonnée d'un mot de $Z(I,x)$, nous avons $\#Z(I,x) \leqslant 2^m$.

□

Nous pouvons donc démontrer la version plus générale du théorème 2.11, présentée dans le théorème ci-dessous. Cette version s'applique quel que soit le branch number de M, et également pour toute matrice Λ vérifiant les conditions (2.4). Cette généralisation permet de démontrer le théorème dans le cas différentiel et linéaire, puisque nous avons vu que l'espérance de la probabilité d'une différentielle sur deux tours et le potentiel linéaire d'un masque sur deux tours ont une forme similaire.

Théorème 2.15. *Soient m et t deux entiers positifs. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients $\Lambda(\alpha,\beta),(\alpha,\beta) \in (\mathbf{F}_2^m)^2$, vérifient les conditions (2.4). Alors, pour tout code $\mathcal{C}$ sur $\mathbf{F}_2^m$ de longueur $(2t)$ et de distance minimum d, pour tous a non nul et b éléments de $(\mathbf{F}_2^m)^t$, nous avons :*

$$\Lambda_{a,b} = \sum_{c\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i,c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j},b_j)\right) \leqslant \left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*} \Lambda(\alpha,\beta)\right)^{d-1}.$$

Démonstration. Soient a, b deux éléments non nuls de $(\mathbf{F}_2^m)^t$. Pour tout mot c du code $\mathcal{C}$ tel que $\mathsf{Supp}(c) \neq \mathsf{Supp}(a,b)$, il existe $\ell \in \{1,\ldots,t\}$ tel que $\Lambda(a_\ell,c_\ell) = 0$ ou $\Lambda(c_{t+\ell},b_\ell) = 0$. Donc nous obtenons :

$$\Lambda_{a,b} = \sum_{c\in\mathcal{C}:\mathsf{Supp}(c)=\mathsf{Supp}(a,b)} \left(\prod_{i=1}^{t} \Lambda(a_i,c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j},b_j)\right).$$

Si $wt(a)+wt(b) \leqslant d$, alors $\Lambda_{a,b} = 0$ puisqu'il n'existe pas de mot c du code tel que $\mathsf{Supp}(c) = \mathsf{Supp}(a,b)$. Supposons donc que $wt(a)+wt(b) \geqslant d$. Nous pouvons choisir deux ensembles I_1 et I_2 de $\{1,\ldots,t\}$ tels que $I_1 \subseteq \mathsf{Supp}(a)$, $I_2 \subseteq \mathsf{Supp}(b)$ et $(\#I_1)+(\#I_2) = d$. Décomposons tout mot c de $\mathcal{C}$ dont le support est égal à celui de $\mathsf{Supp}((a,b))$ en deux parties : $c = (y,x)_I$ où $I = (\{1,\ldots,t\} \setminus I_1) \cup \{t+j, j \notin I_2\}$, c'est-à-dire que y correspond à la restriction de c aux positions en dehors de I_1 et I_2, tandis que x correspond aux d autres coordonnées. Rappelons que, conformément à la définition 2.13, les coordonnées de y (resp. de x) sont indexées par les éléments de I (resp. de $I_1 \cup \{t+j, j \in I_2\}$). Alors, pour $Z(I,y) = \{x \ : (y,x)_I \in \mathcal{C}\}$,

$$\Lambda_{a,b} = \sum_{y\in(\mathbf{F}_2^m)^{n-d}} \left(\prod_{i\notin I_1} \Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2} \Lambda(y_{t+j},b_j)\right) \mathcal{Q}_{a,b}(I,y) \tag{2.5}$$

$$\text{où } \mathcal{Q}_{a,b}(I,y) = \sum_{x\in Z(I,y)} \left(\prod_{i\in I_1} \Lambda(a_i,x_i)\right)\left(\prod_{j\in I_2} \Lambda(x_{t+j},b_j)\right).$$

Nous allons maintenant déterminer une majoration de la valeur $\mathcal{Q}_{a,b}(I,y)$. Pour tout $i \in I_1$ et pour tout $j \in I_2$, les valeurs $\Lambda(a_i,x_i)$ et $\Lambda(x_{t+j},b_j)$ sont inférieures ou égales

à $\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)$. Choisissons un élément $i'\in I_1$. Alors :

$$\begin{aligned}\mathcal{Q}_{a,b}(I,y) &\leqslant \sum_{x\in Z(I,y)}\Lambda(a_{i'},x_{i'})\left(\prod_{i\in I_1;i\neq i'}\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)\left(\prod_{j\in I_2}\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)\\ &\leqslant \left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)^{(\#I_1)-1+(\#I_2)}\sum_{x\in Z(I,y)}\Lambda(a_{i'},x_{i'}).\end{aligned}$$

Or $(\#I_1)-1+(\#I_2)=d-1$. Par ailleurs, d'après le lemme 2.14 et puisque les coefficients $\Lambda(\alpha,\beta)$ sont positifs, nous avons :

$$\sum_{x\in Z(I,y)}\Lambda(a_{i'},x_{i'})\leqslant\sum_{\gamma\in\mathbf{F}_2^m}\Lambda(a_{i'},\gamma).$$

Par hypothèse du théorème, $\sum_{\gamma\in\mathbf{F}_2^m}\Lambda(a_{i'},\gamma)=1$. Donc

$$\mathcal{Q}_{a,b}(I,y)\leqslant\left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)^{d-1}.$$

En reportant cette majoration dans l'équation (2.5), nous avons :

$$\begin{aligned}\Lambda_{a,b} &\leqslant \left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)^{d-1}\sum_{y\in(\mathbf{F}_2^m)^{n-d}}\left(\prod_{i\notin I_1}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\Lambda(y_{t+j},b_j)\right)\\ &\leqslant \left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)^{d-1}\left(\prod_{i\notin I_1}\sum_{y_i\in\mathbf{F}_2^m}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\sum_{y_{t+j}\in\mathbf{F}_2^m}\Lambda(y_{t+j},b_j)\right).\end{aligned}$$

Par hypothèse, pour tout $i\notin I_1$ et pour tout $j\notin I_2$,

$$\sum_{y_i\in\mathbf{F}_2^m}\Lambda(a_i,y_i)=1\quad\text{ et }\quad\sum_{y_{t+j}\in\mathbf{F}_2^m}\Lambda(y_{t+j},b_j)=1$$

donc

$$\Lambda_{a,b}\leqslant\left(\max_{\alpha,\beta\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\beta)\right)^{d-1}.$$

□

En appliquant ce théorème avec $t=d-1$ et avec $\Lambda(\alpha,\beta)=2^{-m}\delta(\alpha,\beta)$ d'une part et $(\Lambda(\alpha,\beta)=2^{-m}\mathcal{W}^S(\alpha,\beta))^2$ d'autre part, nous obtenons le théorème 2.11.

Il est intéressant d'observer que la borne du théorème 2.11 sur le MEDP (resp. le MELP) est similaire à celle du lemme 2.5 sur les caractéristiques différentielles (resp. du lemme 2.9 sur les chemins linéaires) mais que l'exposant a été décrémenté.

Bornes de FSE 2003

En 2003, deux nouvelles bornes, une sur le MEDP_2 et l'autre sur le MELP_2, ont été présentées dans l'article [CKL+03] et à FSE [PSLL03]. Dans la suite de ce document, ces bornes sont appelées bornes de FSE 2003. Elles ont été obtenues en utilisant le lemme 2.14 et une version généralisée de l'inégalité de Hölder présentée ci-dessous.

Lemme 2.16 (Inégalité de Hölder [HLP52])**.** *Soit* $\{x_i^{(j)}\}_{i=1}^n$, $1 \leqslant j \leqslant p$, p *suites de* n *nombres réels. Alors*

$$\sum_{i=1}^{n} \left| \prod_{j=1}^{p} x_i^{(j)} \right| \leqslant \prod_{j=1}^{p} \left(\sum_{i=1}^{n} |x_i^{(j)}|^p \right)^{\frac{1}{p}}.$$

Théorème 2.17 (Borne de FSE 2003 [PSLL03, CKL+03])**.** *Soit* $(E_k)_k$ *un chiffrement par bloc de la forme* $\mathrm{SPN}(m, t, S, M)$*, où* M *est une permutation linéaire de branch number différentiel* d *et de branch number linéaire* $d^\perp$*. Alors, nous avons :*

$$\mathrm{MEDP}_2^E \leqslant 2^{-md} \max \left(\max_{a \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \delta^S(a,\gamma)^d, \max_{b \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \delta^S(\gamma,b)^d \right),$$

$$\mathrm{MELP}_2^E \leqslant 2^{-2md^\perp} \max \left(\max_{u \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \mathcal{W}^S(u,\gamma)^{2d^\perp}, \max_{v \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \mathcal{W}^S(\gamma,v)^{2d^\perp} \right).$$

De nouveau, nous allons démontrer une version générique de ce théorème. Après cette démonstration, nous donnerons un exemple pour expliquer la signification de cette borne.

Théorème 2.18. *Soient* m *et* t *deux entiers positifs. Soit* Λ *une matrice de taille* $2^m \times 2^m$ *dont les coefficients* $\Lambda(\alpha,\beta), (\alpha,\beta) \in (\mathbf{F}_2^m)^2$*, vérifient les conditions (2.4). Alors, pour tout code* $\mathcal{C}$ $\mathbf{F}_2$*-linéaire de longueur* $(2t)$ *et de distance minimum* d *sur* $\mathbf{F}_2^m$*, pour tout* a *non nul et* b *éléments de* $(\mathbf{F}_2^m)^t$*, nous avons :*

$$\begin{aligned} \Lambda_{a,b} &= \sum_{c \in \mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i) \right) \left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j) \right) \\ &\leqslant \max \left(\max_{\alpha \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \Lambda(\alpha,\gamma)^d, \max_{\beta \in (\mathbf{F}_2^m)^*} \sum_{\gamma \in (\mathbf{F}_2^m)^*} \Lambda(\gamma,\beta)^d \right). \end{aligned}$$

Démonstration. Soient a, b deux éléments non nuls de $(\mathbf{F}_2^m)^t$. De façon similaire à la démonstration du théorème 2.15, nous obtenons l'égalité (2.5) :

$$\Lambda_{a,b} = \sum_{y \in (\mathbf{F}_2^m)^{n-d}} \left(\prod_{i \notin I_1} \Lambda(a_i, y_i) \right) \left(\prod_{j \notin I_2} \Lambda(y_{t+j}, b_j) \right) \mathcal{Q}_{a,b}(I, y)$$

$$\text{où } \mathcal{Q}_{a,b}(I,y) = \sum_{x\in Z(I,y)} \left(\prod_{i\in I_1} \Lambda(a_i,x_i)\right)\left(\prod_{j\in I_2} \Lambda(x_{t+j},b_j)\right) .$$

Il s'agit de nouveau de déterminer une majoration de la valeur $\mathcal{Q}_{a,b}(I,y)$. Pour cela, nous allons appliquer l'inégalité de Hölder à $\mathcal{Q}_{a,b}(I,y)$. Nous avons $\#I_1$ suites $\{\Lambda(a_i,x_i)\}_{x\in Z(I,y)}$ et $\#I_2$ suites $\{\Lambda(x_{t+j},b_j)\}_{x\in Z(I,y)}$, donc $\#I_1+\#I_2 = d$ suites au total. Nous obtenons :

$$\begin{aligned}\mathcal{Q}_{a,b}(I,y) &= \sum_{x\in Z(I,y)} \left(\prod_{i\in I_1} \Lambda(a_i,x_i)\right)\left(\prod_{j\in I_2} \Lambda(x_{t+j},b_j)\right)\\ &\leqslant \prod_{i\in I_1}\left(\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d\right)^{\frac{1}{d}} \prod_{j\in I_2}\left(\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d\right)^{\frac{1}{d}} .\end{aligned}$$

En majorant chacune des sommes $\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d$ et $\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d$, $i\in I_1$, $j\in I_2$, par la plus grande d'entre elles, nous obtenons :

$$\mathcal{Q}_{a,b}(I,y) \leqslant \left(\max_{i\in I_1}\left(\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d\right)^{\frac{1}{d}}\right)^{\#I_1} \left(\max_{j\in I_2}\left(\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d\right)^{\frac{1}{d}}\right)^{\#I_2} .$$

Si nous prenons le maximum entre les deux valeurs $\max_{i\in I_1}\left(\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d\right)^{\frac{1}{d}}$ et $\max_{j\in I_2}\left(\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d\right)^{\frac{1}{d}}$, nous obtenons :

$$\mathcal{Q}_{a,b}(I,y) \leqslant \max\left(\max_{i\in I_1}\left(\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d\right)^{\frac{1}{d}\times d}, \max_{j\in I_2}\left(\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d\right)^{\frac{1}{d}\times d}\right).$$

La somme $\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d$ ne dépend pas de tout le vecteur $x\in Z(I,y)$ mais seulement de sa coordonnée i, donc d'après le lemme 2.14, pour tout $i\in I_1$:

$$\sum_{x\in Z(I,y)} \Lambda(a_i,x_i)^d \leqslant \sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(a_i,\gamma)^d.$$

De même, pour tout $j\in I_2$, nous avons :

$$\sum_{x\in Z(I,y)} \Lambda(x_{t+j},b_j)^d \leqslant \sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(\gamma,b_j)^d.$$

Donc la valeur $\mathcal{Q}_{a,b}(I,y)$ est majorée par :

$$\begin{aligned}\mathcal{Q}_{a,b}(I,y) &\leqslant \max\left(\max_{i\in I_1}\sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(a_i,\gamma)^d, \max_{j\in I_2}\sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(\gamma,b_j)^d\right)\\ &\leqslant \max\left(\max_{\alpha\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(\alpha,\gamma)^d, \max_{\beta\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*} \Lambda(\gamma,\beta)^d\right).\end{aligned}$$

Notons

$$\mathcal{M} = \max\left(\max_{\alpha\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\gamma)^d, \max_{\beta\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\Lambda(\gamma,\beta)^d\right).$$

Cette valeur est indépendante des mots de code $c = (y,x)_I$, donc nous avons :

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{y\in(\mathbf{F}_2^m)^{n-d}}\left(\prod_{i\notin I_1}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\Lambda(y_{t+j},b_j)\right)\mathcal{Q}_{a,b}(I,y)\\
&\leqslant \mathcal{M}\times\sum_{y\in(\mathbf{F}_2^m)^{n-d}}\left(\prod_{i\notin I_1}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\Lambda(y_{t+j},b_j)\right)\\
&\leqslant \mathcal{M}\times\left(\prod_{i\notin I_1}\sum_{y_i\in\mathbf{F}_2^m}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\sum_{y_{t+j}\in\mathbf{F}_2^m}\Lambda(y_{t+j},b_j)\right).
\end{aligned}$$

Par hypothèse, pour tout $i \notin I_1$ et pour tout $j \notin I_2$,

$$\sum_{y_i\in\mathbf{F}_2^m}\Lambda(a_i,y_i) = 1 \quad \text{et} \quad \sum_{y_{t+j}\in\mathbf{F}_2^m}\Lambda(y_{t+j},b_j) = 1$$

donc

$$\Lambda_{a,b} \leqslant \max\left(\max_{\alpha\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\Lambda(\alpha,\gamma)^d, \max_{\beta\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\Lambda(\gamma,\beta)^d\right).$$

□

Il est possible de montrer que les bornes de FSE 2003 sont toujours inférieures ou égales à celles du théorème 2.11.

Corollaire 2.19. *[PSLL03, CKL+03] Soit E un chiffrement par blocs de la forme* $\mathrm{SPN}(m,t,S,M)$*, où M est une permutation linéaire de branch number différentiel d et de branch number linéaire $d^\perp$. Alors la borne de FSE 2003 (théorème 2.17) sur le* MEDP_2 *est majorée par :*

$$\left(2^{-m}\Delta(S)\right)^{d-1},$$

avec égalité par exemple lorsque le spectre différentiel de S ne prend que deux valeurs. La borne de FSE 2003 sur le MELP_2 *est majorée par :*

$$\left(2^{-m}\mathcal{L}(S)\right)^{2(d^\perp-1)},$$

avec égalité par exemple lorsque S est une fonction plateau (les carrés de ses coefficients de Walsh prennent au plus deux valeurs).

Démonstration. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients vérifient les conditions (2.4). Soit d un entier strictement positif. Pour tout élément a non nul de $\mathbf{F}_2^m$, nous avons :

$$\sum_{\gamma \in (\mathbf{F}_2^m)^*} \Lambda(a,\gamma)^d \leq \left(\max_{\alpha,\beta \in (\mathbf{F}_2^m)^*} \Lambda(\alpha,\beta) \right)^{d-1} \left(\sum_{\gamma \in (\mathbf{F}_2^m)^*} \Lambda(a,\gamma) \right) = \left(\max_{\alpha,\beta \in (\mathbf{F}_2^m)^*} \Lambda(\alpha,\beta) \right)^{d-1},$$

avec égalité si et seulement si pour tout $\gamma \in (\mathbf{F}_2^m)^*$, $\Lambda(a,\gamma) = 0$ ou $\Lambda(a,\gamma) = \max_{\alpha,\beta \in (\mathbf{F}_2^m)^*} \Lambda(\alpha,\beta)$. La borne est identique pour $\sum_{\gamma \in (\mathbf{F}_2^m)^*} \Lambda(\gamma,b)^d$. En appliquant ce résultat avec $\Lambda(\alpha,\beta) = 2^{-m}\delta(\alpha,\beta)$ d'une part, $\Lambda(\alpha,\beta) = (2^{-m}\mathcal{W}^S(\alpha,\beta))^2$ d'autre part, nous obtenons la borne du corollaire.

Le cas d'égalité entre le MEDP_2 et la borne a lieu lorsque les coefficients du spectre différentiel ne prennent que deux valeurs distinctes. Le cas d'égalité entre le MELP_2 et la borne a lieu lorsque le carré des coefficients de Walsh ne prennent que deux valeurs distinctes, par exemple lorsque les coefficients de Walsh prennent les valeurs 0 et $\pm\mathcal{L}(S)$, ce qui correspond à une fonction plateau. □

Exemple 2.20. Pour comprendre la signification de la borne de FSE 2003 (théorème 2.17), il nous faut comprendre la signification de la formule de l'espérance de la probabilité d'une différentielle sur deux tours (lemme 2.6) et celle du potentiel linéaire moyen d'un masque sur deux tours (lemme 2.10). Ces deux formules sont similaires, c'est pourquoi nous allons nous intéresser à une seule d'entre elles dans cet exemple. Rappelons donc la formule de l'espérance de la probabilité d'une différentielle $(a,b) \in (\mathbf{F}_2^m)^{2t}$ sur deux tours d'un réseau de substitution-permutation de la forme $\mathrm{SPN}(m,t,S,M)$ (lemme 2.6) :

$$\mathrm{EDP}_2(a,b) = 2^{-m\,wt(a,b)} \sum_{\substack{c \in \mathcal{C}_M : \\ \mathsf{Supp}(c) = \mathsf{Supp}(a,b)}} \left(\prod_{i \in \mathsf{Supp}(a)} \delta(a_i, c_i) \right) \left(\prod_{j \in \mathsf{Supp}(b)} \delta(c_{t+j}, b_j) \right).$$

Lorsque le mot c du code $\mathcal{C}_M$ varie parmi les mots de même support que (a,b), les valeurs des coordonnées du support de c varient dans $(\mathbf{F}_2^m)^*$. Lorsque γ parcourt $(\mathbf{F}_2^m)^*$, les coefficients $\delta(\alpha,\gamma)$ parcourent la ligne α de la table des différences de la boîte-S et les coefficients $\delta(\gamma,\beta)$ parcourent la colonne β de la table des différences de la boîte-S. Donc, pour $(a,b) \in (\mathbf{F}_2^m)^{2t}$, le calcul de $\mathrm{EDP}_2(a,b)$ fait intervenir les coefficients des $wt(a)$ lignes de la table des différences de la boîte-S correspondant aux valeurs a_i, $i \in \mathsf{Supp}(a)$, et les coefficients des $wt(b)$ colonnes de cette table correspondant aux valeurs b_j, $j \in \mathsf{Supp}(b)$.

Prenons comme boîte-S la permutation S de $\mathbf{F}_2^4$ définie au chapitre 1, exemple 1.42, $t = 2$ et $(a,b) = (3,8,0,2) \in (\mathbf{F}_2^4)^4$, où les mots de 4 bits sont identifiés aux entiers par leur décomposition en base 2 (*cf* figure 2.2).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	4	0	0	0	0	2	2	0	2	0	2	0	2	2	0
2	0	0	0	0	2	2	0	2	0	2	0	0	4	2	2
3	0	0	0	2	0	2	0	2	2	0	4	2	0	0	2
4	0	2	0	0	0	0	2	2	0	2	2	2	0	4	0
5	0	0	2	0	0	0	2	0	4	2	0	2	2	0	2
6	2	2	2	0	0	4	2	2	0	0	0	0	0	0	2
7	2	0	0	2	2	2	4	0	0	2	0	2	0	0	0
8	0	2	2	0	2	2	0	0	2	4	2	0	0	0	0
9	2	2	0	4	0	0	0	0	0	2	2	0	2	0	2
10	0	0	2	2	2	0	2	0	0	0	2	0	0	2	4
11	2	4	0	0	2	0	0	0	2	0	0	2	0	2	2
12	0	2	0	2	2	0	2	4	2	0	0	0	2	0	0
13	2	0	4	2	0	0	0	2	2	2	0	0	0	2	0
14	2	0	2	0	4	0	0	2	0	0	2	2	2	0	0
15	0	2	2	2	0	2	0	0	0	0	0	4	2	2	0

FIGURE 2.2 – Table des différences de S ; les lignes et colonnes colorées sont celles intervenant dans le calcul de $\mathrm{EDP}_2(3, 8, 0, 2)$.

Le calcul de $\mathrm{EDP}_2(a, b)$ consiste à additionner, pour chaque mot de $\mathcal{C}_M$ de même support que (a, b), un produit de $wt(a) + wt(b)$ coefficients de la table des différences. En fonction du mot c de même support que (a, b), un coefficient de chacune des $wt(a)$ lignes et $wt(b)$ colonnes sélectionnées est utilisé. Par exemple, la figure 2.3 correspond au calcul de $\mathrm{EDP}_2(a, b)$, en supposant que les mots de code de même support que (a, b) sont listés dans la première colonne de cette figure.

c	$\delta(a_1, c_1)$	$\delta(a_2, c_2)$	$\delta(c_4, b_2)$	$\delta(a_1, c_1)\delta(a_2, c_2)\delta(c_4, b_2)$
$(1, 2, 0, 1)$	0	2	0	0
$(2, 4, 0, 2)$	0	0	0	0
$(3, 6, 0, 3)$	0	2	0	0
$(4, 8, 0, 4)$	2	0	2	0
$(5, 10, 0, 5)$	0	4	0	0
$(6, 12, 0, 6)$	2	0	2	0
$(7, 14, 0, 7)$	0	0	0	0
$(8, 3, 0, 8)$	2	2	2	8
$(9, 1, 0, 9)$	2	0	2	0
$(10, 7, 0, 10)$	0	0	0	0
$(11, 5, 0, 11)$	4	2	4	32
$(12, 11, 0, 12)$	2	2	2	8
$(13, 9, 0, 13)$	0	2	0	0
$(14, 15, 0, 14)$	0	0	0	0
$(15, 13, 0, 15)$	2	0	2	0

FIGURE 2.3 – Calcul de $\mathrm{EDP}_2(3, 8, 0, 2) = (8 + 32 + 8) \times 2^{-4\times 3}$.

Comme il y a un grand nombre de coefficients nuls dans la table des différences d'une boîte-S (et dans la table de Walsh), la valeur $\mathrm{EDP}_2(a,b)$ sera grande lorsque les zéros seront multipliés entre eux, ainsi que lorsque les coefficients avec les plus grandes valeurs seront multipliés entre eux. La valeur maximale possible est obtenue lorsque les coefficients qui sont multipliés entre eux sont égaux, c'est-à-dire lorsque $\mathrm{EDP}_2(a,b) = \max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta(a,\gamma)^d$ ou $\mathrm{EDP}_2(a,b) = \max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta(\gamma,b)^d$, ce qui correspond à la borne de FSE 2003 .

La question est de savoir si les bornes de FSE 2003 (théorème 2.17) sont éloignées des valeurs réelles du MEDP_2 et du MELP_2. Ces bornes ont une propriété intéressante : elles sont invariantes par équivalence affine de la boîte-S.

Proposition 2.21. *Soient S_1 et S_2 deux permutations de $\mathbf{F}_2^m$ telles que $S_2 = A_1 \circ S_1 \circ A_2$, où A_1 et A_2 sont des permutations affines de $\mathbf{F}_2^m$. Alors :*

$$\begin{aligned}
\max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta^{S_1}(a,\gamma)^d &= \max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta^{S_2}(a,\gamma)^d,\\
\max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta^{S_1}(\gamma,b)^d &= \max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta^{S_2}(\gamma,b)^d,\\
\max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \mathcal{W}^{S_1}(a,\gamma)^{2d^\perp} &= \max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \mathcal{W}^{S_2}(a,\gamma)^{2d^\perp},\\
\max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \mathcal{W}^{S_1}(\gamma,b)^{2d^\perp} &= \max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \mathcal{W}^{S_2}(\gamma,b)^{2d^\perp}.
\end{aligned}$$

Donc les bornes de FSE 2003 pour S_1 et S_2 sont égales.

Démonstration. L'ensemble des éléments de chaque ligne ou colonne de la table des différences (resp. de la table de Walsh) est invariant par équivalence affine. Pour chaque ligne (resp. chaque colonne) a de la table des différences d'une fonction S, il existe, pour toute fonction S' affinement équivalente à S, une ligne (resp. une colonne) a' de la table des différences de S' telle que les ensembles des éléments des lignes (resp. colonnes) a et a' sont égaux. Il en est de même pour la table de Walsh d'une fonction. □

Par exemple, comme la boîte-S de l'AES correspond à la fonction inverse du corps $\mathbf{F}_{2^8}$ composée par une permutation affine, la borne de FSE 2003 sur le MEDP_2 est identique pour la boîte-S de l'AES et pour la boîte-S correspondant à la fonction inverse dans $\mathbf{F}_{2^8}$.

Corollaire 2.22. *Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(8,4,S,M)$ où M est la matrice de diffusion de l'AES. Si S est une permutation de $\mathbf{F}_2^8$ affinement équivalente à la fonction inverse dans le corps $\mathbf{F}_{2^8}$, alors la borne de FSE 2003 (théorème 2.17) sur le MEDP_2 est*

$$2^{-8\times 5} \max\left(\max_{a\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta(a,\gamma)^5, \max_{b\in(\mathbf{F}_2^m)^*} \sum_{\gamma\in(\mathbf{F}_2^m)^*} \delta(\gamma,b)^5 \right) = 79\times 2^{-34}$$

et la borne de FSE 2003 sur le MELP_2 *est*

$$2^{-2\times 40}\max\left(\max_{u\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\mathcal{W}(u,\gamma)^{2\times 5},\max_{v\in(\mathbf{F}_2^m)^*}\sum_{\gamma\in(\mathbf{F}_2^m)^*}\mathcal{W}(\gamma,v)^{2\times 5}\right) = 48\,193\,409\times 2^{-52}$$
$$\approx 2,873\times 2^{-28}.$$

Cependant, les valeurs exactes du MEDP_2 et du MELP_2 pour deux réseaux de substitution-permutation utilisant deux boîtes-S affinement équivalentes peuvent être différentes. Il a été démontré par J. Daemen et V. Rijmen [DR06] que la valeur exacte du MEDP_2 pour l'AES où la boîte-S est remplacée par la fonction inverse de $\mathbf{F}_{2^8}$ est égale à la borne de FSE 2003 : ils ont présenté une différentielle sur deux tours de probabilité 79×2^{-34}.

Par ailleurs, L. Keliher et J. Sui [KS07] ont proposé un algorithme pour calculer les valeurs exactes du MEDP_2 et du MELP_2 et l'ont appliqué à l'AES avec la boîte-S originale. La valeur obtenue pour le MEDP_2 est plus petite que la valeur pour la fonction inverse. À notre connaissance, la valeur exacte du MELP_2 pour l'AES avec la fonction inverse comme boîte-S n'a jamais été calculée, mais la borne de FSE 2003 pour ce chiffrement est égale à $48\,193\,409\times 2^{-52}\approx 2,873\times 2^{-28}$, et les résultats présentés dans la suite de ce document montrent que la borne est atteinte dans ce cas. La valeur exacte du MELP_2 pour l'AES avec la boîte-S originale, calculée avec l'algorithme de Keliher et Sui, est aussi plus petite que la valeur pour la fonction inverse.

Proposition 2.23. *[KS07] La valeur exacte du* MEDP_2 *pour l'AES est*

$$\mathrm{MEDP}_2 = 53\times 2^{-34}.$$

La valeur exacte du MELP_2 *pour l'AES est*

$$\mathrm{MELP}_2 = 109\,953\,193\times 2^{-54}\approx 1,638\times 2^{-28}.$$

La boîte-S de l'AES résiste donc mieux aux cryptanalyses différentielle et linéaire que la fonction inverse bien que la seule différence entre les deux soit une permutation affine, qui est composée à la fonction inverse pour obtenir la boîte-S de l'AES.

L'algorithme de Keliher et Sui pour obtenir la valeur exacte du MEDP_2 consiste à calculer $\mathrm{EDP}(a,b)$ pour une partie des différentielles possibles, sachant que la méthode de calcul de l'algorithme permet de vérifier que la valeur $\mathrm{EDP}(a,b)$ pour les autres différentielles sera inférieure au maximum (la méthode est similaire pour le calcul de la valeur exacte du MELP_2). Ainsi, l'algorithme de Keliher et Sui demande moins de $(2^m)^{2t}$ calculs d'espérances de la probabilité d'une différentielle, mais sa complexité reste assez élevée. Ceci le rend inutilisable pour étudier un grand nombre de boîtes-S de $\mathbf{F}_2^8$ (de même taille que celle de l'AES : il y a environ 2^{64} différentielles) afin de trouver celles qui ont la meilleure résistance aux attaques linéaires et différentielles. Nous avons donc cherché à comprendre pourquoi deux boîtes-S de même spectre différentiel et spectre de Walsh ont un comportement différent et si ce comportement dépendait du choix de la fonction de diffusion utilisée.

	Borne de FSE 2003 (théorème 2.17)	Valeur exacte pour l'AES avec la fonction inverse	Valeur exacte pour l'AES avec la boîte-S normale
MEDP_2	79×2^{-34}	79×2^{-34} [DR06]	53×2^{-34} [KS07]
MELP_2	$2,873 \times 2^{-28}$	$2,873 \times 2^{-28}$ [CR15b]	$1,638 \times 2^{-28}$ [KS07]

FIGURE 2.4 – Valeurs exactes et bornes sur les valeurs de MEDP_2 et MELP_2 pour l'AES avec la boîte-S normale et la boîte-S naïve (la borne de FSE 2003 est identique pour les deux boîtes-S).

3

Nouvelles bornes

Dans ce chapitre sont présentées de nouvelles bornes supérieures et inférieures sur le MEDP et sur le MELP de deux tours d'un réseau de substitution-permutation. Ces bornes ne sont pas invariantes sur les classes d'équivalence affine et sont valides lorsque la fonction de diffusion est linéaire sur le corps $\mathbf{F}_{2^m}$, où m est la taille des boîtes-S, comme c'est le cas dans l'AES, LED [GPPR11], KLEIN [GNL12], mCrypton [LK06], PRØST [KLL+14]... Dans cette situation, les codes associés à la matrice de diffusion du chiffrement ont une structure plus forte, ce qui permet d'affiner les bornes de FSE 2003. Ces travaux ont donnés lieu à des présentations dans les conférences *11th International Conference on Finite Fields and their applications Fq11* [CR14] et *EUROCRYPT 2015* [CR15b].

Nous avons expliqué au chapitre précédent la signification de la formule de la valeur exacte de l'espérance de la probabilité d'une différentielle (respectivement du potentiel linéaire) sur deux tours d'un réseau de substitution-permutation, donnée dans le lemme 2.6 (respectivement le lemme 2.10). Le calcul consiste à additionner, pour chaque mot de $\mathcal{C}_M$ (respectivement $\mathcal{C}_M^\perp$) de même support que la différentielle (a, b) (respectivement que le masque linéaire (a, b)), un produit de $wt(a) + wt(b)$ coefficients de la table des différences (respectivement la table de Walsh). La valeur maximale possible est obtenue lorsque les coefficients qui sont multipliés entre eux sont égaux, c'est ainsi qu'est obtenue la borne de FSE 2003.

Cependant, il n'est pas toujours possible de multiplier les coefficients égaux entre eux : dans le calcul exact de $\mathrm{EDP}_2(a, b)$ (respectivement de $\mathrm{ELP}_2(a, b)$), le produit fait intervenir des coefficients des lignes et des colonnes. Or il n'y a aucune raison pour que la table des différences (respectivement la table de Walsh) d'une boîte-S contienne une ligne et une colonne identiques. Pour trouver une approximation plus fine, nous avons donc cherché à multiplier des coefficients d'au moins une ligne et une colonne.

De plus, les coefficients utilisés dans la multiplication dépendent des mots du code $\mathcal{C}_M$ (respectivement $\mathcal{C}_M^\perp$). Pour un code linéaire sur le corps $\mathbf{F}_{2^m}$, nous pouvons expliciter la forme des mots de même support de ce code, et l'utiliser pour affiner la borne.

Dans la première partie de ce chapitre, nous définissons et posons les notations des réseaux de substitution-permutation définis sur un corps correspondant à l'alphabet de la boîte-S. Puis nous établissons de nouvelles bornes sur les valeurs MEDP_2 et MELP_2, des bornes supérieures dans la deuxième partie et inférieures dans la troisième. Enfin, nous verrons que si les définitions des composants du réseau de substitution-permutation n'utilisent pas la même représentation, c'est-à-dire si la boîte-S est définie sur l'espace vectoriel $\mathbf{F}_2^m$ et la fonction de diffusion sur le corps $\mathbf{F}_{2^m}$, alors la représentation du corps $\mathbf{F}_{2^m}$ choisie peut avoir une influence sur les valeurs MEDP_2 et MELP_2.

3.1 Réseaux de substitution-permutation définis sur un corps

Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$ tel que la fonction de diffusion M de $(\mathbf{F}_2^m)^t$ est linéaire sur $\mathbf{F}_{2^m}$. Alors le chiffrement tout entier peut se décrire dans le corps $\mathbf{F}_{2^m}$. Pour cela, nous allons identifier l'espace vectoriel $\mathbf{F}_2^m$ au corps $\mathbf{F}_{2^m}$ grâce à un isomorphisme φ associé à une base $(\alpha_0, \ldots, \alpha_{m-1})$, c'est-à-dire :

$$\begin{array}{rcl} \varphi : \mathbf{F}_2^m & \to & \mathbf{F}_{2^m} \\ (x_0, \ldots, x_{m-1}) & \mapsto & \sum_{i=0}^{m-1} x_i \alpha_i \ . \end{array}$$

Alors la boîte-S et la fonction de diffusion peuvent être représentées comme des fonctions du corps $\mathbf{F}_{2^m}$ par

$$\mathcal{S} = \varphi \circ S \circ \varphi^{-1} \text{ et } \mathcal{M} = \tilde{\varphi} \circ M \circ \tilde{\varphi}^{-1} \ ,$$

où $\tilde{\varphi}$ est la concaténation de t copies de φ. Il s'ensuit que r tours d'un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$ peuvent s'écrire comme indiqué dans [DR02, Section A.5] :

$$\tilde{\varphi}^{-1} \circ \mathrm{AK}_{k_r} \circ \ldots \circ \mathcal{F} \circ \mathrm{AK}_{k_1} \circ \mathcal{F} \circ \mathrm{AK}_{k_0} \circ \tilde{\varphi} \tag{3.1}$$

où la fonction de tour $\mathcal{F} = \mathcal{M} \circ \tilde{\mathcal{S}}$ est une permutation de $(\mathbf{F}_{2^m})^t$ et AK_x correspond à l'addition de x dans $(\mathbf{F}_{2^m})^t$. Nous pouvons donc définir le chiffrement sur $\mathbf{F}_{2^m}$ de la manière suivante.

Définition 3.1. *Soient m et t deux entiers positifs. Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$ et $\mathcal{M}$ une permutation de $(\mathbf{F}_{2^m})^t$ qui est linéaire sur $\mathbf{F}_{2^m}$. Nous noterons*

$$\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$$

un réseau de substitution-permutation défini sur $(\mathbf{F}_{2^m})^t$ dont la fonction de substitution est la concaténation de t copies de $\mathcal{S}$ et dont la fonction de diffusion correspond à $\mathcal{M}$.

Pour éviter toute ambiguïté, les quantités correspondant à la représentation dans le corps $\mathbf{F}_{2^m}$ auront un indice F, et toutes les fonctions définies sur $\mathbf{F}_{2^m}$ seront notées avec des lettres calligraphiques.

Ainsi les chiffrements $\mathrm{SPN}(m,t,S,M)$ et $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ ne diffèrent que par l'application de $\tilde{\varphi}$ en entrée et $\tilde{\varphi}^{-1}$ en sortie, comme le montre l'équation (3.1). Il est évident que composer le chiffrement au début par $\tilde{\varphi}$ et à la fin par $\tilde{\varphi}^{-1}$ ne modifie pas les valeurs MEDP et MELP. Donc MEDP_r^E et MELP_r^E dépendent uniquement de $\mathcal{M}$ et $\mathcal{S}$, c'est-à-dire de la représentation de la boîte-S et de la fonction de diffusion sur $\mathbf{F}_{2^m}$. En particulier, il est expliqué dans [BB02] que le choix du polynôme irréductible dans l'AES est arbitraire et qu'il n'y a pas d'avantage à choisir un polynôme qui serait primitif plutôt que celui qui est actuellement utilisé. Dans l'essentiel de la suite du document, nous étudierons donc des réseaux de substitution-permutation définis sur $\mathbf{F}_{2^m}$. Les résultats détaillés au chapitre précédent portent sur le chiffrement binaire, mais ils peuvent également s'exprimer à partir des propriétés du chiffrement sur $\mathbf{F}_{2^m}$. Pour une fonction de $\mathbf{F}_{2^m}$, les coefficients du spectre différentiel et ceux du spectre de Walsh sont définis comme suit.

Définition 3.2. *Soit $\mathcal{S}$ une fonction de $\mathbf{F}_{2^m}$ dans $\mathbf{F}_{2^m}$. Alors, pour tout $(a,b) \in \mathbf{F}_{2^m}$, on définit*

$$\begin{aligned} \delta_F^{\mathcal{S}}(a,b) &= \#\{x \in \mathbf{F}_{2^m}, \mathcal{S}(x+a)+\mathcal{S}(x) = b\} \\ \mathcal{W}_F^{\mathcal{S}}(a,b) &= \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(ax+b\mathcal{S}(x))} \end{aligned}$$

Le lien entre le spectre différentiel et le spectre de Walsh d'une boîte-S de $\mathbf{F}_2^m$ et ceux de sa représentation sur le corps $\mathbf{F}_{2^m}$ est donné dans la proposition suivante. Il fait intervenir la notion de base duale définie dans la section 1.6.

Proposition 3.3. *[DR11] Soit $(\alpha_0,\ldots,\alpha_{m-1})$ une base de $\mathbf{F}_{2^m}$ et φ l'isomorphisme de $\mathbf{F}_2^m$ dans $\mathbf{F}_{2^m}$ correspondant. Soit S une permutation de $\mathbf{F}_2^m$ et $\mathcal{S} = \varphi \circ S \circ \varphi^{-1}$. Alors, pour tout $(a,b) \in \mathbf{F}_{2^m}$,*

$$\begin{aligned} \delta_F^{\mathcal{S}}(a,b) &= \delta^S(\varphi^{-1}(a),\varphi^{-1}(b)) \\ \mathcal{W}_F^{\mathcal{S}}(a,b) &= \mathcal{W}^S(\psi^{-1}(a),\psi^{-1}(b)) \end{aligned}$$

où ψ est l'isomorphisme de $\mathbf{F}_2^m$ dans $\mathbf{F}_{2^m}$ défini par la base duale de $(\alpha_0,\ldots,\alpha_{m-1})$.

Démonstration. Pour la première égalité, nous avons :

$$\begin{aligned} \delta_F^{\mathcal{S}}(a,b) &= \#\{x \in \mathbf{F}_{2^m}, \varphi\circ S\circ\varphi^{-1}(x+a)+\varphi\circ S\circ\varphi^{-1}(x) = b\} \\ &= \#\{x \in \mathbf{F}_{2^m}, \varphi[S(\varphi^{-1}(x)+\varphi^{-1}(a))+S(\varphi^{-1}(x))] = b\} \\ &= \#\{y \in \mathbf{F}_2^m, S(y+\varphi^{-1}(a))+S(y) = \varphi^{-1}(b)\} = \delta^S(\varphi^{-1}(a),\varphi^{-1}(b)) \,. \end{aligned}$$

Pour la deuxième égalité, nous avons :

$$\begin{aligned} \mathcal{W}_F^{\mathcal{S}}(a,b) &= \sum_{x \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(ax+b\,\varphi\circ S\circ\varphi^{-1}(x))} \\ &= \sum_{y \in \mathbf{F}_2^m} (-1)^{\mathrm{Tr}(a\,\varphi(y))+\mathrm{Tr}(b\,\varphi[S(y)])} \,. \end{aligned}$$

Or, d'après la proposition 1.59, pour deux éléments x et y de $\mathbf{F}_2^m$, nous avons $x \cdot y = \mathrm{Tr}(\psi(x)\varphi(y))$, et pour deux éléments a et b de $\mathbf{F}_{2^m}$, nous avons

$$\mathrm{Tr}(ab) = \psi^{-1}(a) \cdot \varphi^{-1}(b) \ .$$

Nous en déduisons :

$$\mathrm{Tr}(a\,\varphi(y)) = \psi^{-1}(a) \cdot y \text{ et } \mathrm{Tr}(b\,\varphi\,[S(y)]) = \psi^{-1}(b) \cdot S(y) \ .$$

Donc

$$\mathcal{W}_F^{\mathcal{S}}(a,b) = \sum_{y \in \mathbf{F}_2^m} (-1)^{\psi^{-1}(a)\cdot y + \psi^{-1}(b)\cdot S(y)} = \mathcal{W}^S(\psi^{-1}(a), \psi^{-1}(b)) \ .$$

□

L'étude faite dans ce chapitre et le suivant concernent les chiffrements décrits sur le corps $\mathbf{F}_{2^m}$, c'est-à-dire avec la représentation $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$.

3.2 Une nouvelle borne supérieure

Le théorème 3.6 présente une nouvelle borne supérieure sur le MEDP_2 et le MELP_2 d'un chiffrement qui suit la construction $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$. Cette nouvelle borne utilise la forme des mots des codes linéaires sur $\mathbf{F}_{2^m}$. Pour décrire ces mots, nous allons à nouveau utiliser la définition 2.13.

Rappel (Définition 2.13)**.** *Soit c un mot de longueur n et I un sous-ensemble de $\{1,\ldots,n\}$. La décomposition de c selon I est notée $(x,y)_I$: x correspond à la restriction de c à I, et y correspond à la restriction de c au sous-ensemble complémentaire $\bar{I}$ de I. Pour simplifier la notation, les $\#I$ coordonnées de x (resp. les coordonnées de y) ont pour indice les éléments de I (resp. de $\bar{I}$),* i.e.*, $x_i = c_i$ pour tout $i \in I$ et $y_j = c_j$ pour tout $j \in \bar{I}$.*

Lemme 3.4. *Soit $\mathcal{C}$ un code linéaire de longueur n, de dimension k et de distance minimum d sur $\mathbf{F}_{2^m}$. Pour tout sous-ensemble $I \subset \{1,\ldots,n\}$ de taille $(n-d)$, et pour tout $x \in (\mathbf{F}_{2^m})^{n-d}$, notons*

$$Z(I,x) = \{y \ : (x,y)_I \in \mathcal{C}\} \ .$$

Alors, pour tout I de taille $(n-d)$,

- *soit $Z(I,0)$ est vide, soit il existe $y_0 \in (\mathbf{F}_{2^m}^*)^d$ tel que $Z(I,0) = \{\gamma y_0,\ \gamma \in \mathbf{F}_{2^m}\}$;*
- *Pour tout $x \neq 0$, soit $Z(I,x)$ est vide, soit il existe $y_0 \in (\mathbf{F}_{2^m}^*)^d$ et $y_1 \in (\mathbf{F}_{2^m})^d$ tel que $Z(I,x) \subseteq \{y_1 + \gamma y_0,\ \gamma \in \mathbf{F}_{2^m}\}$.*

Démonstration.

- Supposons que $Z(I,0)$ n'est pas vide. Puisque $\mathcal{C}$ est linéaire sur $\mathbf{F}_{2^m}$, pour tout $y_0 \in Z(I,0)$, $(0,y_0)_I$ appartient à $\mathcal{C}$ et pour tout $\gamma \in \mathbf{F}_{2^m}$, $\gamma(0,y_0)_I$ appartient aussi à $\mathcal{C}$.

– Soit $x \neq 0$. Comme le résultat est évident quand $\#Z(I,x) \leqslant 1$, supposons que $\#Z(I,x) \geqslant 2$. Soient y et y' deux éléments distincts de $Z(I,x)$, les mots $c = (x,y)_I$ et $c' = (x,y')_I$ appartiennent à $\mathcal{C}$, donc $(y+y') \in Z(I,0)$. D'après la première partie du lemme, il existe y_0 tel que $y + y' = \gamma y_0$ avec $\gamma \in \mathbf{F}_{2^m}$. Donc y' s'écrit $y' = y + \gamma y_0$. Comme le poids $wt(c+c') = wt(y+y')$ ne peut pas être inférieur à d, toutes les coordonnées de y_0 sont non nulles.

□

Les nouvelles bornes sont décrites en fonction des quantités définies ci-dessous.

Notation 3.5. Soient m et d deux entiers strictement positifs. Pour tout élément μ de $\mathbf{F}_{2^m}$ et pour tout entier u strictement positif, notons

$$\begin{aligned}
\mathcal{B}_u(\mu) &= \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\alpha,\gamma)^u \delta_F(\gamma\lambda+\mu,\beta)^{(d-u)} \ , \\
\mathcal{B}_u^{\perp}(\mu) &= \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha,\gamma)^{2u} \mathcal{W}_F(\gamma\lambda+\mu,\beta)^{2(d^{\perp}-u)} \ , \\
\mathcal{B}(\mu) &= \max_{1 \leq u < d} \mathcal{B}_u(\mu) \ , \\
\mathcal{B}^{\perp}(\mu) &= \max_{1 \leq u < d^{\perp}} \mathcal{B}_u^{\perp}(\mu) \ .
\end{aligned}$$

Remarquons qu'avec cette notation, la borne de FSE 2003 sur le MEDP_2^E correspond à $\max_{\mu \in \mathbf{F}_{2^m}} \max(\mathcal{B}_0(\mu), \mathcal{B}_d(\mu))$ et la borne sur le MELP_2^E correspond à $\max_{\mu \in \mathbf{F}_{2^m}} \max(\mathcal{B}_0^{\perp}(\mu), \mathcal{B}_d^{\perp}(\mu))$.

Théorème 3.6. *Soit E un chiffrement de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$, où $\mathcal{M}$ est linéaire sur $\mathbf{F}_{2^m}$, de branch number différentiel d et de branch number linéaire $d^{\perp}$. Alors,*

$$\mathrm{MEDP}_2^E \leqslant 2^{-md} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}(\mu) \text{ et } \mathrm{MELP}_2^E \leqslant 2^{-2md^{\perp}} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}^{\perp}(\mu) \ .$$

La démonstration de ce théorème est proche de celle des bornes de FSE 2003. Nous utiliserons l'inégalité de Hölder généralisée (lemme 2.16). La démonstration est faite dans le cas général, en remplaçant les $2^m \times 2^m$ coefficients $2^{-m}\delta(\alpha,\beta)$, ou les coefficients $2^{-2m}\mathcal{W}(\alpha,\beta)^2$, $\alpha,\ \beta \in \mathbf{F}_2^m$, par une matrice $(\Lambda(\alpha,\beta))_{\alpha,\beta \in \mathbf{F}_2^m}$ vérifiant les conditions (2.4).

Théorème 3.7. *Soient m et t deux entiers positifs. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients $\Lambda(\alpha,\beta), (\alpha,\beta) \in (\mathbf{F}_{2^m})^2$, vérifient les conditions (2.4). Alors, pour tout code $\mathcal{C}$ linéaire sur $\mathbf{F}_{2^m}$ de longueur $2t$ et de distance minimum d, pour tous éléments a non nul et b de $\mathbf{F}_{2^m}^t$, nous avons :*

$$\Lambda_{a,b} = \sum_{c \in \mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i) \right) \left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j) \right) \leqslant \max_{1 \leqslant u < d} \max_{\mu \in \mathbf{F}_{2^m}} \mathcal{B}_u(\mu)$$

où

$$\mathcal{B}_u(\mu) = \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma\lambda+\mu,\beta)^{d-u}.$$

Démonstration. Soient a, b deux éléments non nuls de $(\mathbf{F}_{2^m})^t$. D'après la démonstration du théorème 2.15, pour tous sous-ensembles $I_1 \subseteq \mathsf{Supp}(a)$ et $I_2 \subseteq \mathsf{Supp}(b)$ tels que $(\#I_1) + (\#I_2) = d$, la valeur $\Lambda_{a,b}$ vérifie :

$$\Lambda_{a,b} = \sum_{y \in \mathbf{F}_{2^m}^{n-d}} \left(\prod_{i \notin I_1} \Lambda(a_i, y_i) \right) \left(\prod_{j \notin I_2} \Lambda(y_{t+j}, b_j) \right) \mathcal{Q}_{a,b}(I, y)$$

où

$$\mathcal{Q}_{a,b}(I, y) = \sum_{x \in Z(I,y)} \left(\prod_{i \in I_1} \Lambda(a_i, x_i) \right) \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) .$$

La démonstration consiste à nouveau à déterminer une majoration de la valeur $\mathcal{Q}_{a,b}(I, y)$. Pour cela, notons $u = \#I_1$. D'après le lemme 2.16,

$$\begin{aligned} \mathcal{Q}_{a,b}(I, y) &= \sum_{x \in Z(I,y)} \prod_{i \in I_1} \left[\Lambda(a_i, x_i) \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right)^{\frac{1}{u}} \right] \\ &\leqslant \prod_{i \in I_1} \left[\sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) \right]^{\frac{1}{u}} . \end{aligned}$$

Pour tout $i \in I_1$, appliquons à nouveau le lemme 2.16 :

$$\begin{aligned} \sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \left(\prod_{j \in I_2} \Lambda(x_{t+j}, b_j) \right) &= \sum_{x \in Z(I,y)} \prod_{j \in I_2} \left(\Lambda(a_i, x_i)^{\frac{u}{d-u}} \Lambda(x_{t+j}, b_j) \right) \\ &\leqslant \prod_{j \in I_2} \left(\sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \Lambda(x_{t+j}, b_j)^{d-u} \right)^{\frac{1}{d-u}} . \end{aligned}$$

Nous savons d'après le lemme 3.4 que si $Z(I, y) \neq \emptyset$, alors il existe $\alpha \in (\mathbf{F}_{2^m}^*)^d$ et $\beta \in (\mathbf{F}_{2^m})^d$ tels que $Z(I, y) \subseteq \{\gamma\alpha + \beta, \gamma \in \mathbf{F}_{2^m}\}$. Donc, pour tout couple $(i, j) \in I_1 \times I_2$, nous pouvons écrire :

$$\begin{aligned} \sum_{x \in Z(I,y)} \Lambda(a_i, x_i)^u \Lambda(x_{t+j}, b_j)^{d-u} &\leqslant \sum_{\gamma \in \mathbf{F}_{2^m}} \Lambda(a_i, \gamma\alpha_i + \beta_i)^u \Lambda(\gamma\alpha_{t+j} + \beta_{t+j}, b_j)^{d-u} \\ &= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(a_i, \gamma')^u \Lambda(\gamma'\lambda + \mu, b_j)^{d-u} , \end{aligned}$$

où la dernière égalité est obtenue en remplaçant $\gamma\alpha_i + \beta_i$ par γ' puisque $\alpha_i \neq 0$, et en posant $\lambda = \alpha_{t+j}\alpha_i^{-1}$ et $\mu = \beta_{t+j} + \alpha_{t+j}\alpha_i^{-1}\beta_i$. De plus, la somme peut être effectuée sur les éléments non nuls γ' de $\mathbf{F}_{2^m}$ car $\Lambda(a_i, \gamma') = 0$ pour $\gamma' = 0$. Notons

$$\mathcal{B}_u = \max_{a,b,\lambda \in \mathbf{F}_{2^m}^*} \max_{\mu \in \mathbf{F}_{2^m}} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a, \gamma)^u \Lambda(\gamma\lambda + \mu, b)^{d-u} .$$

Alors, nous avons

$$\begin{aligned}\mathcal{Q}_{a,b}(I,y) &\leqslant \prod_{i\in I_1}\left[\prod_{j\in I_2}\left(\sum_{x\in Z(I,y)}\Lambda(a_i,x_i)^u\Lambda(x_{t+j},b_j)^{d-u}\right)^{\frac{1}{(d-u)}}\right]^{\frac{1}{u}}\\ &\leqslant \prod_{i\in I_1}\left[(\mathcal{B}_u)^{\frac{d-u}{d-u}}\right]^{\frac{1}{u}}\\ &\leqslant (\mathcal{B}_u)^{\frac{u}{u}}=\mathcal{B}_u\end{aligned}$$

car $\#I_1 = u$ et $\#I_2 = d-u$.

En utilisant l'équation (2.5) et le fait que $\sum_{\beta\in\mathbf{F}_{2^m}}\Lambda(\alpha,\beta)=\sum_{\alpha\in\mathbf{F}_{2^m}}\Lambda(\alpha,\beta)=1$, nous obtenons finalement

$$\begin{aligned}\Lambda_{a,b} &= \sum_{y\in\mathbf{F}_{2^m}^{n-d}}\left(\prod_{i\notin I_1}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\Lambda(y_{t+j},b_j)\right)\mathcal{Q}_{a,b}(I,y)\\ &\leqslant \mathcal{B}_u\sum_{y\in\mathbf{F}_{2^m}^{n-d}}\left(\prod_{i\notin I_1}\Lambda(a_i,y_i)\right)\left(\prod_{j\notin I_2}\Lambda(y_{t+j},b_j)\right)\leqslant\mathcal{B}_u\ .\end{aligned}$$

□

On observe en particulier que cette nouvelle borne n'est à priori pas invariante par équivalence affine.

Grâce à un exemple, nous allons comprendre la signification de cette borne. Nous allons de nouveau nous intéresser uniquement à la borne sur le MEDP_2, la borne sur le MELP_2 étant similaire.

Exemple 3.8. La borne sur le MEDP_2 est le maximum sur les éléments non nuls α, β et λ de $\mathbf{F}_2^m$, sur les éléments μ de $\mathbf{F}_2^m$ et sur les entiers u compris entre 1 et d de

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*}\delta_F(\alpha,\gamma)^u\,\delta_F(\gamma\lambda+\mu,\beta)^{(d-u)}. \tag{3.2}$$

Soit un couple $(\alpha,\beta)\in(\mathbf{F}_2^m)^{2t}$ et une valeur u comprise entre 1 et d. Les coefficients de la table des différences qui interviennent dans le calcul de la somme (3.2) sont ceux de la ligne α élevés à la puissance u et ceux de la colonne β élevés à la puissance $d-u$. La différence par rapport au calcul exact de $\mathrm{EDP}_2(a,b)$ décrit dans l'exemple 2.20 est que nous supposons dans le calcul de la borne que les lignes a_i utilisées sont toutes identiques, ou que $a_i=\alpha$ pour tout $i\in\mathsf{Supp}(a)$ (de même pour les colonnes).

Ensuite, le calcul de la borne consiste à additionner des produits d'un coefficient de la ligne α avec un coefficient de la colonne β, ces coefficients étant déterminés par les valeurs de λ et μ.

Prenons comme exemple la fonction $x\mapsto x^{13}$ de $\mathbf{F}_{2^4}$ dans $\mathbf{F}_{2^4}$. Notons g un élément primitif de $\mathbf{F}_{2^4}$. La table des différences de $\mathcal{S}$ est présentée dans la figure 3.1.

	1	g	g^2	g^3	g^4	g^5	g^6	g^7	g^8	g^9	g^{10}	g^{11}	g^{12}	g^{13}	g^{14}
1	4	0	0	0	0	2	0	2	0	0	2	2	0	2	2
g	0	0	0	2	0	2	0	0	2	2	0	2	2	4	0
g^2	0	2	0	2	0	0	2	2	0	2	2	4	0	0	0
g^3	0	2	0	0	2	2	0	2	2	4	0	0	0	0	2
g^4	0	0	2	2	0	2	2	4	0	0	0	0	2	0	2
g^5	2	2	0	2	2	4	0	0	0	0	2	0	2	0	0
g^6	0	2	2	4	0	0	0	0	2	0	2	0	0	2	2
g^7	2	4	0	0	0	0	2	0	2	0	0	2	2	0	2
g^8	0	0	0	0	2	0	2	0	0	2	2	0	2	2	4
g^9	0	0	2	0	2	0	0	2	2	0	2	2	4	0	0
g^{10}	2	0	2	0	0	2	2	0	2	2	4	0	0	0	0
g^{11}	2	0	0	2	2	0	2	2	4	0	0	0	0	2	0
g^{12}	0	2	2	0	2	2	4	0	0	0	0	2	0	2	0
g^{13}	2	0	2	2	4	0	0	0	0	2	0	2	0	0	2
g^{14}	2	2	4	0	0	0	0	2	0	2	0	0	2	2	0

FIGURE 3.1 – Table des différences de la fonction $x \mapsto x^{13}$ du corps $\mathbf{F}_{2^4}$.

Prenons $d = 3$ et $u = 2$, $\alpha = g^3$, $\beta = g^7$, $\lambda = g^2$, $\mu = 0$. Alors le calcul de la somme (3.2) est décrit dans la figure 3.2.

γ	1	g	g^2	g^3	g^4	g^5	g^6	g^7	g^8	g^9	g^{10}	g^{11}	g^{12}	g^{13}	g^{14}
$\delta_F(g^3,\gamma)^2$	0	4	0	0	4	4	0	4	4	16	0	0	0	0	4
	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
$\delta_F(g^2\,\gamma, g^7)$	0	0	0	0	2	0	2	0	0	2	2	0	2	2	4
	0	+0	+0	+0	+8	+0	+0	+0	+0	+32	+0	+0	+0	+0	+16

FIGURE 3.2 – Calcul de la somme $\sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(g^3,\gamma)^2\, \delta_F(g^2\,\gamma, g^7) = 8 + 32 + 16 = 56$.

En faisant ce calcul pour toutes les valeurs possibles pour u, nous parcourons tous les supports possibles pour une différentielle du chiffrement. En faisant ce calcul pour toutes les valeurs possibles pour α et β, nous parcourons toutes les différentielles dont le support est déterminé par u. En faisant ce calcul pour toutes les valeurs possibles pour λ et μ, nous parcourons au moins tous les mots du code $\mathcal{C}_M$ de même support que la différentielle déterminée par les valeurs de u, α et β. En effet, le terme $\lambda\gamma + \mu$ provient de la forme des mots d'un code $\mathbf{F}_{2^m}$-linéaire déterminée dans le lemme 3.4.

Le calcul de cette nouvelle borne est clairement plus long que le calcul de la borne de FSE 2003, puisque les calculs pour toutes les valeurs λ et μ s'ajoutent à ceux pour les valeurs α et β. Cependant, nous verrons au chapitre suivant que cette borne peut parfois se simplifier, en particulier quand la boîte-S est une fonction puissance com-

posée avec une permutation affine, comme c'est le cas pour l'AES. Nous verrons aussi qu'il est parfois possible de déduire de cette borne la valeur exacte du MEDP_2 (resp. du MELP_2), de façon plus efficace que l'algorithme de Keliher et Sui.

Alors que la borne de FSE 2003 utilise indépendamment les coefficients d'une ligne ou d'une colonne de la table des différences (resp. la table de Walsh), la borne du théorème 3.6 mélange une ligne et une colonne de cette table, c'est-à-dire qu'elle dépend des liens entre les dérivées (resp. les transformées de Walsh) de la boîte-S $\mathcal{S}$ et celles de son inverse $\mathcal{S}^{-1}$. Le rôle de la permutation inverse de $\mathcal{S}$ est plus évident lorsque les coefficients de la table des différences (resp. le carré des coefficients de la table de Walsh) de $\mathcal{S}$ ne prennent que deux valeurs distinctes.

Corollaire 3.9. *Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$. Si les coefficients de la table des différences de $\mathcal{S}$ ne prennent que deux valeurs distinctes, alors*

$$\mathcal{B}(0) = \Delta(\mathcal{S})^d \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Im}(D_\alpha \mathcal{S}) \cap \left[\lambda \mathsf{Im}(D_\beta \mathcal{S}^{-1})\right]\right)$$

et

$$\max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}(\mu) = \Delta(\mathcal{S})^d \max_{\alpha,\beta,\lambda,\mu \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Im}(D_\alpha \mathcal{S}) \cap \left[\lambda \mathsf{Im}(D_\beta \mathcal{S}^{-1}) + \mu\right]\right)$$

où $D_\alpha \mathcal{S}$ est la dérivée de $\mathcal{S}$ au point α.

Si $\mathcal{S}$ est une fonction plateau, c'est-à-dire si les coefficients de la table de Walsh de $\mathcal{S}$ ne prennent que les valeurs 0 et $\pm\mathcal{L}(S)$, alors nous avons

$$\mathcal{B}^\perp(0) = \mathcal{L}(\mathcal{S})^{2d^\perp} \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\alpha^{-1}}) \cap \left[\lambda \mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta})\right]\right) ,$$

et

$$\max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}^\perp(\mu) = \mathcal{L}(\mathcal{S})^{2d^\perp} \max_{\alpha,\beta,\lambda,\mu \in \mathbf{F}_{2^m}^*} \#\left(\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\alpha^{-1}}) \cap \left[\lambda \mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta}) + \mu\right]\right) ,$$

où $\mathcal{S}_\alpha$ est une fonction composante de $\mathcal{S}$ et $\mathcal{W}^f$ est la transformée de Walsh de la fonction booléenne f.

Démonstration. Si les coefficients de la table des différences de $\mathcal{S}$ ne prennent que deux valeurs distinctes (0 et $\Delta(\mathcal{S})$), notons

$$I_{\alpha,\beta,\lambda} = \{\gamma \in \mathbf{F}_{2^m}^* \mid \delta_F(\alpha,\gamma) \neq 0\} \cap \{\gamma \in \mathbf{F}_{2^m}^* \mid \delta_F(\lambda\gamma + \mu, \beta) \neq 0\}.$$

Donc pour tout $\gamma \in I_{\alpha,\beta,\lambda}$, nous avons $\delta_F(\alpha,\gamma) = \delta_F(\lambda\gamma+\mu,\beta) = \Delta(\mathcal{S})$. Nous obtenons :

$$\begin{aligned} \mathcal{B}(\mu) &= \max_{1 \leqslant u < d} \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in I_{\alpha,\beta,\lambda}} \Delta(\mathcal{S})^u \times \Delta(\mathcal{S})^{d-u} \\ &= \Delta(\mathcal{S})^d \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \#I_{\alpha,\beta,\lambda}. \end{aligned}$$

Or, d'après la définition du coefficient $\delta_F(\alpha,\gamma)$ et de la dérivée de $\mathcal{S}$ au point α (*cf* section 1.5.2), nous avons :

$$\begin{aligned} \delta_F(\alpha,\gamma) \neq 0 &\Leftrightarrow \#\{x \in \mathbf{F}_{2^m} \mid D_\alpha(x) = \gamma\} \neq 0 \\ &\Leftrightarrow \gamma \in \mathsf{Im}(D_\alpha \mathcal{S}). \end{aligned}$$

De même, pour le coefficient $\delta_F(\lambda\gamma+\mu,\beta)$, nous avons :

$$\begin{aligned}\delta_F(\lambda\gamma+\mu,\beta)\neq 0 &\Leftrightarrow \lambda\gamma+\mu\in \mathsf{Im}(D_\beta\mathcal{S}^{-1}))\\ &\Leftrightarrow \gamma\in\lambda^{-1}(\mathsf{Im}(D_\beta\mathcal{S}^{-1})+\mu).\end{aligned}$$

Le résultat s'ensuit.

Si $\mathcal{S}$ est une fonction plateau, c'est-à-dire si les coefficients de la table de Walsh de $\mathcal{S}$ ne prennent que les valeurs 0 et $\pm\mathcal{L}(S)$ (*cf* section 1.5.3), alors nous avons

$$\begin{aligned}\mathcal{B}^{\perp}(\mu) &= \max_{1\leqslant u<d}\max_{\alpha,\beta,\lambda\in\mathbf{F}_{2^m}^*}\sum_{\gamma\in J_{\alpha,\beta,\lambda}}\mathcal{L}(\mathcal{S})^{2u}\times\mathcal{L}(\mathcal{S})^{2(d^{\perp}-u)}\\ &= \mathcal{L}(\mathcal{S})^{2d^{\perp}}\max_{\alpha,\beta,\lambda\in\mathbf{F}_{2^m}^*}\#J_{\alpha,\beta,\lambda},\end{aligned}$$

où $J_{\alpha,\beta,\lambda}=\{\gamma\in\mathbf{F}_{2^m}^*\,|\,\mathcal{W}_F(\alpha,\gamma)\neq 0\}\cap\{\gamma\in\mathbf{F}_{2^m}^*\,|\,\mathcal{W}_F(\lambda\gamma+\mu,\beta)\neq 0\}$. Or d'après la définition des coefficients de Walsh, nous avons :

$$\begin{aligned}\mathcal{W}_F(\alpha,\gamma) &= \sum_{x\in\mathbf{F}_{2^m}}(-1)^{\mathrm{Tr}(\alpha x+\gamma\mathcal{S}(x))}\\ &= \sum_{y\in\mathbf{F}_{2^m}}(-1)^{\mathrm{Tr}(\alpha\mathcal{S}^{-1}(y)+\gamma y)}\\ &= \sum_{y\in\mathbf{F}_{2^m}}(-1)^{\mathrm{Tr}(\mathcal{S}_\alpha^{-1}(y)+\gamma y)}\\ &= \mathcal{W}^{\mathcal{S}_\alpha^{-1}}(\gamma)\end{aligned}$$

où $\mathcal{S}_\alpha^{-1}$ est une fonction composante de $\mathcal{S}^{-1}$ (*cf* section 1.5) et $\mathcal{W}^f$ est la transformée de Walsh de la fonction booléenne f (*cf* section 1.4.2). Donc

$$\mathcal{W}_F(\alpha,\gamma)\neq 0\Leftrightarrow\gamma\in\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\alpha^{-1}}).$$

De même, pour le coefficient $\mathcal{W}_F(\lambda\gamma+\mu,\beta)$, nous avons :

$$\begin{aligned}\mathcal{W}_F(\lambda\gamma+\mu,\beta)\neq 0 &\Leftrightarrow \lambda\gamma+\mu\in\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta})\\ &\Leftrightarrow \gamma\in\lambda^{-1}(\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta})+\mu).\end{aligned}$$

Nous en déduisons le résultat. □

De façon évidente, le cardinal de l'ensemble $\mathsf{Im}(D_\alpha\mathcal{S})\cap\left[\lambda\mathsf{Im}(D_\beta\mathcal{S}^{-1})+\mu\right]$ (resp. $\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\alpha^{-1}})\cap\left[\lambda\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta})+\mu\right]$) décrit dans le corollaire ne peut pas être supérieur aux cardinaux des ensembles qui le composent, c'est-à-dire à $2^m/\Delta(\mathcal{S})$ (resp. $2^{2m}/\mathcal{L}^2(\mathcal{S})$). La valeur maximale est obtenue lorsque $\mathcal{S}$ est une involution, *i.e.* lorsque $\mathcal{S}=\mathcal{S}^{-1}$. Mais lorsque la boîte-S est composée avec une permutation affine choisie de façon aléatoire, les deux ensembles $\mathsf{Im}(D_\alpha\mathcal{S})$ et $\lambda\mathsf{Im}(D_\beta\mathcal{S}^{-1})+\mu$ (resp. $\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\alpha^{-1}})$ et $\lambda\mathsf{Supp}(\mathcal{W}^{\mathcal{S}_\beta})+\mu$) peuvent être considérés comme indépendants. La valeur moyenne pour le cardinal de leur intersection est $2^m\pi_\Delta^2=2^m/\Delta(\mathcal{S})^2$ (resp. $2^m\pi_{\mathcal{L}}^2=2^{3m}/\mathcal{L}(\mathcal{S})^4$) où $\pi_\Delta=1/\Delta(\mathcal{S})$ est la proportion d'éléments non nuls dans une ligne ou une colonne

de la table des différences (resp. $\pi_{\mathcal{L}} = 2^m/\mathcal{L}(\mathcal{S})^2$ est la proportion d'éléments non nuls dans une ligne ou une colonne de la table de Walsh). Par exemple, pour une boîte-S presque courbe, c'est-à-dire pour m impair, $\Delta(\mathcal{S}) = 2$ et $\mathcal{L}(\mathcal{S}) = 2^{(m+1)/2}$, le cardinal moyen des deux ensembles considérés dans le corollaire est 2^{m-2}, alors qu'il est égal à 2^{m-1} lorsque $\mathcal{S}$ est une involution.

Plus généralement, la proposition 3.10 montre que nos nouvelles bornes sont toujours inférieures ou égales à celles de FSE 2003. En particulier, les deux bornes sont égales lorsque la boîte-S est une involution.

Proposition 3.10. *Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$ et d un entier strictement positif. Alors chacune des deux bornes du théorème 3.6 est inférieure ou égale à la borne de FSE 2003 correspondante (théorème 2.17). De plus, il y a égalité lorsque $\mathcal{S}$ est une involution, car dans ce cas, pour tout entier $u < d$,*

$$\max_{\mu\in\mathbf{F}_{2^m}} \mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \max_{a\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta_F(a,\gamma)^d = \max_{b\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta_F(\gamma,b)^d$$

et

$$\max_{\mu\in\mathbf{F}_{2^m}} \mathcal{B}_u^{\perp}(\mu) = \mathcal{B}_u^{\perp}(0) = \max_{a\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}_F(a,\gamma)^{2d} = \max_{b\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}_F(\gamma,b)^{2d}.$$

Nous allons montrer cette proposition dans le cas générique, c'est-à-dire avec la matrice Λ. La proposition 3.10 s'écrit alors de la façon suivante.

Proposition 3.11. *Soient m et d deux entiers positifs. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients vérifient les conditions (2.4). Alors, pour tout $1 \leqslant u < d$ et tout $\mu \in \mathbf{F}_{2^m}$, nous avons :*

$$\mathcal{B}_u(\mu) \leqslant \max\left(\max_{a\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(a,\gamma)^d, \max_{b\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\gamma,b)^d \right),$$

où

$$\mathcal{B}_u(\mu) = \max_{\alpha,\beta,\lambda\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma\lambda+\mu,\beta)^{d-u}.$$

De plus, si $\Lambda(\alpha,\beta) = \Lambda(\beta,\alpha)$ pour tout $(\alpha,\beta) \in (\mathbf{F}_{2^m})^2$, alors, pour tout $1 \leqslant u < d$,

$$\max_{\mu\in\mathbf{F}_{2^m}} \mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \max_{a\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(a,\gamma)^d = \max_{b\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\gamma,b)^d .$$

Démonstration. Le lemme 2.16 implique que, pour tout ensemble de p suites $\{x_i^{(j)}\}_{i=1}^n$, $1 \leqslant j \leqslant p$,

$$\sum_{i=1}^{n} \left| \prod_{j=1}^{p} x_i^{(j)} \right| \leqslant \max_{1\leqslant j\leqslant p} \sum_{i=1}^{n} |x_i^{(j)}|^p .$$

En utilisant cette inégalité avec $p = d$, nous obtenons la majoration suivante : pour tous $1 \leqslant u < d$, $\alpha, \beta, \lambda \in \mathbf{F}_{2^m}^*$ et $\mu \in \mathbf{F}_{2^m}$

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma\lambda + \mu)^u \Lambda(\gamma, \beta)^{d-u} \leqslant \max\left(\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^d, \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma\lambda + \mu, \beta)^d \right) .$$

Puisque $\lambda \neq 0$, nous avons $\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\gamma\lambda + \mu, \beta)^d = \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\gamma', \beta)^d$. Nous en déduisons l'inégalité.

Maintenant, supposons que $\Lambda(a, b) = \Lambda(b, a)$ pour tout couple (a, b). Alors,

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\beta, \gamma\lambda + \mu)^{d-u} .$$

Pour $\mu = 0$, le maximum de cette valeur sur tous les triplets non nuls α, β, λ est supérieur ou égal à la valeur obtenue pour $\beta = \alpha$ et $\lambda = 1$, ce qui implique que

$$\mathcal{B}_u(0) \geqslant \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\alpha, \gamma)^{d-u} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^d .$$

Alors $\max_{a \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a, \gamma)^d$ est une borne inférieure de $\mathcal{B}_u(0)$, et donc de $\max_\mu \mathcal{B}_u(\mu)$. Puisque nous avons aussi montré que c'est une borne supérieure de $\max_\mu \mathcal{B}_u(\mu)$, les deux quantités sont égales. □

Les involutions ne sont pas les seules permutations telles que la borne de FSE 2003 et la borne du théorème 3.6 sont égales. Par exemple, pour la boîte-S utilisée dans les chiffrements `LED` [GPPR11] et PRESENT [BKL+07], définie à la figure 3.3, les deux bornes sur le MEDP_2 (resp. le MELP_2) sont égales et leur valeur est $\mathrm{MEDP}_2 \leqslant 2^{-8}$ (resp. $\mathrm{MELP}_2 \leqslant 2^{-8}$).

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	12	5	6	11	9	0	10	13	3	14	15	8	4	7	1	2

FIGURE 3.3 – Boîte-S utilisée dans les chiffrements `LED` et PRESENT, où les mots de 4 bits sont identifiés aux entiers par leur décomposition en base 2.

Pour l'AES, nous obtenons les bornes supérieures suivantes :

$$\mathrm{MEDP}_2 \leqslant 55,5 \times 2^{-34} \text{ et } \mathrm{MELP}_2 \leqslant 31\,231\,767 \times 2^{-52},$$

(au lieu de $\mathrm{MEDP}_2 \leqslant 79 \times 2^{-34}$ et $\mathrm{MELP}_2 \leqslant 48\,193\,441 \times 2^{-52}$ pour la borne de FSE 2003).

De plus, lorsque la boîte-S de l'AES est remplacée par la fonction inverse de $\mathbf{F}_{2^8}$, nous déduisons de la proposition 3.10 que la borne donnée par le théorème 3.6 est égale à la borne de FSE 2003.

3.3 Optimalité de la nouvelle borne

Nous avons vu que les bornes du théorème 3.6 sont meilleures que les bornes précédemment connues. Nous allons montrer que, sous certaines conditions, ces bornes sont optimales, au sens où il existe au moins une fonction de diffusion pour laquelle la borne est atteinte. Pour cela, nous allons présenter des fonctions $\mathcal{M}$ avec un branch number maximal telles que la valeur EDP (resp. ELP) de certaines différentielles de poids minimal (resp. masques linéaires) sur deux tours est liée à la borne du théorème 3.6. La borne inférieure ainsi obtenue résulte de la forme des mots de poids minimum de même support d'un code MDS linéaire sur $\mathbf{F}_{2^m}$: pour c un mot de poids minimum, les mots de même support que c forment un ensemble $\{\lambda c, \lambda \in \mathbf{F}_{2^m}^*\}$ (*cf* lemme 3.4). Un tel ensemble est appelé *paquet* dans [DR06]. Remarquons que $\mathcal{B}(0)$ (resp. $\mathcal{B}^\perp(0)$) correspond à la valeur maximale de EDP (resp. de ELP) pour certains paquets. De plus, pour n'importe lequel de ces paquets particuliers, une fonction $\mathcal{M}$ telle que $\mathcal{C}_\mathcal{M}$ (resp. $\mathcal{C}_\mathcal{M}^\perp$) contient ce paquet peut être construite à partir d'un code GRS (*cf* section 1.3.4).

Proposition 3.12. *Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$ et t un entier tel que $t \leqslant 2^{m-1}$. Alors il existe deux fonctions de diffusion $\mathcal{M}_1$ et $\mathcal{M}_2$ de $\mathbf{F}_{2^m}^t$ linéaires sur $\mathbf{F}_{2^m}$ de branch number maximal $d = t+1$ telles que tout chiffrement E_1 de la forme $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_1)$ et E_2 de la forme $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M}_2)$ vérifient*

$$\mathrm{MEDP}_2^{E_1} \geqslant 2^{-m(t+1)}\mathcal{B}(0) \text{ et } \mathrm{MELP}_2^{E_2} \geqslant 2^{-2m(t+1)}\mathcal{B}^\perp(0)$$

où $\mathcal{B}(0)$ et $\mathcal{B}^\perp(0)$ sont définies comme dans la notation 3.5.

Nous allons donner ici la démonstration dans le cas générique, c'est-à-dire pour la borne du théorème 3.7.

Proposition 3.13. *Soient m et t deux entiers positifs, avec $t \leqslant 2^{m-1}$. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients vérifient les conditions (2.4). Alors il existe un code $\mathcal{C}$ de $\mathbf{F}_{2^m}^t$ linéaire sur $\mathbf{F}_{2^m}$ et MDS (*i.e. *de distance minimale $t+1$) tel que :*

$$\max_{a,b \in \mathbf{F}_{2^m}^*} \Lambda_{a,b} \geqslant \mathcal{B}(0)$$

où

$$\mathcal{B}(0) = \max_{1 \leqslant u \leqslant t} \max_{\alpha,\beta,\lambda \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma\lambda,\beta)^{t+1-u}.$$

Démonstration. Soient $\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda} \in \mathbf{F}_{2^m}^*$ et $1 \leqslant \widehat{u} \leqslant t$ des valeurs telles que

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha},\gamma)^{\widehat{u}} \Lambda(\gamma\widehat{\lambda},\widehat{\beta})^{t+1-\widehat{u}} = \mathcal{B}(0).$$

Soit $a \in \mathbf{F}_{2^m}^t$ le vecteur dont les $\widehat{u}$ premières coordonnées sont égales à $\widehat{\alpha}$ et dont les $(t-\widehat{u})$ dernières coordonnées sont nulles. De même, $b \in \mathbf{F}_{2^m}^t$ est le vecteur dont les $(t+1-\widehat{u})$ premières coordonnées sont égales à $\widehat{\beta}$ et dont les $(\widehat{u}-1)$ dernières coordonnées sont nulles. Puisque

$$\Lambda_{a,b} = \sum_{c \in \mathcal{C}_\mathcal{M}} \left(\prod_{i=1}^{t} \Lambda(a_i, c_i) \right) \left(\prod_{j=1}^{t} \Lambda(c_{t+j}, b_j) \right) ,$$

$\Lambda_{a,b}$ est égal à $\mathcal{B}(0)$ si les mots de la forme

$$\gamma(\underbrace{1,\ldots,1}_{\widehat{u}},\underbrace{0,\ldots,0}_{t-\widehat{u}},\underbrace{\widehat{\lambda},\ldots,\widehat{\lambda}}_{t+1-\widehat{u}},\underbrace{0,\ldots,0}_{\widehat{u}-1}) \tag{3.3}$$

sont des mots du code $\mathcal{C}$ de même support que (a,b). Nous cherchons donc un code $\mathcal{C}$ MDS contenant ces mots. Comme $t \leqslant 2^{m-1}$, nous pouvons choisir $2t$ éléments distincts $x_1,\ldots,x_{2t}$ de $\mathbf{F}_{2^m}$. Pour tout choix de $2t$ éléments $v_1,\ldots,v_{2t}$, soit la matrice R de taille $t \times t$ définie par

$$R = \begin{bmatrix} 1 & 1 & ... & 1 \\ x_1v_1 & x_2v_2 & ... & x_tv_t \\ x_1^2v_1 & x_2^2v_2 & ... & x_t^2v_t \\ & & \cdots & \\ x_1^{t-1}v_1 & x_2^{t-1}v_2 & ... & x_t^{t-1}v_t \end{bmatrix}^{-1} \times \begin{bmatrix} 1 & 1 & ... & 1 \\ x_{t+1}v_{t+1} & x_{t+2}v_{t+2} & ... & x_{2t}v_{2t} \\ x_{t+1}^2v_{t+1} & x_{t+2}^2v_{t+2} & ... & x_{2t}^2v_{2t} \\ & & \cdots & \\ x_{t+1}^{t-1}v_{t+1} & x_{t+2}^{t-1}v_{t+2} & ... & x_{2t}^{t-1}v_{2t} \end{bmatrix}.$$

Alors $\mathcal{C} = \{(x, xR), x \in \mathbf{F}_{2^m}^t\}$ est le code de Reed-Solomon généralisé $\mathsf{GRS}_t(x_1,\ldots,x_{2t};v)$. Ce code est MDS et est composé de tous les mots de la forme $(v_1F(x_1),\ldots,v_{2t}F(x_{2t}))$ où F parcourt tous les polynômes de $\mathbf{F}_{2^m}[X]$ de degré strictement inférieur à t (*cf* section 1.3.4). Donc les mots de $\mathcal{C}$ de même support que (a,b) correspondent aux polynômes de degré au plus $(t-1)$ qui s'annulent en les $(t-1)$ points x_i pour $i \notin \mathsf{Supp}((a,b))$. Ces polynômes peuvent s'écrire $\gamma\widehat{F}(x)$, $\gamma \in \mathbf{F}_{2^m}^*$, et $\widehat{F}(x_i) \neq 0$ pour $i \in \mathsf{Supp}((a,b))$ car $\widehat{F}$ ne peut pas avoir plus de $(t-1)$ racines. Nous pouvons alors choisir le vecteur v tel que $v_i = 1/\widehat{F}(x_i)$ pour $1 \leqslant i \leqslant \widehat{u}$ et $v_i = \widehat{\lambda}/\widehat{F}(x_i)$ pour $t+1 \leqslant i \leqslant 2t+1-\widehat{u}$. Cela nous garantit que les mots de $\mathcal{C}$ de même support que (a,b) sont des mots de la forme (3.3). Nous obtenons donc

$$\begin{aligned} \Lambda_{a,b} &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{\widehat{u}} \Lambda(\widehat{\alpha}, \gamma v_i \widehat{F}(x_i))\right)\left(\prod_{j=1}^{t+1-\widehat{u}} \Lambda(\gamma v_{t+j}\widehat{F}(x_{t+j}), \widehat{\beta})\right) \\ &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha},\gamma)^{\widehat{u}}\Lambda(\gamma\widehat{\lambda},\widehat{\beta})^{t+1-\widehat{u}} = \mathcal{B}(0)\ . \end{aligned}$$

□

Remarque 3.14. Dans certains cas particuliers, nous pouvons trouver un code de Reed-Solomon généralisé correspondant à la fonction de diffusion $\mathcal{M}$ tel que les bornes sur MEDP et MELP sont atteintes en même temps. Pour cela, considérons les valeurs $\widehat{\alpha},\widehat{\beta},\widehat{\lambda} \in \mathbf{F}_{2^m}^*$ et $1 \leqslant \widehat{u} \leqslant t$ telles que

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta(\widehat{\alpha},\gamma)^{\widehat{u}}\delta(\gamma\widehat{\lambda},\widehat{\beta})^{t+1-\widehat{u}} = \mathcal{B}(0)$$

et les valeurs $\bar{\alpha},\bar{\beta},\bar{\lambda} \in \mathbf{F}_{2^m}^*$ et $1 \leqslant \bar{u} \leqslant t$ telles que

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}(\bar{\alpha},\gamma)^{2\bar{u}}\mathcal{W}(\gamma\bar{\lambda},\bar{\beta})^{2(t+1-\bar{u})} = \mathcal{B}^{\perp}(0).$$

En utilisant la démonstration de la proposition 3.13 avec $\Lambda_{a,b} = 2^{-m}\delta(a,b)$ et $\Lambda_{a,b} = 2^{-2m}\mathcal{W}(a,b)^2$, construire un code de Reed-Solomon généralisé correspondant à la fonction de diffusion $\mathcal{M}$ tel que les bornes sur MEDP et MELP sont atteintes en même temps revient à construire un code $\mathcal{C}_{\mathcal{M}}$ contenant les mots de la forme

$$\gamma(\underbrace{1,\ldots,1}_{\widehat{u}},\underbrace{0,\ldots,0}_{t-\widehat{u}},\underbrace{\widehat{\lambda},\ldots,\widehat{\lambda}}_{t+1-\widehat{u}},\underbrace{0,\ldots,0}_{\widehat{u}-1})$$

et dont le code dual $\mathcal{C}_{\mathcal{M}}^{\perp}$ contient les mots de la forme

$$\gamma(\underbrace{0,\ldots,0}_{t-\bar{u}},\underbrace{\bar{\lambda},\ldots,\bar{\lambda}}_{\bar{u}},\underbrace{0,\ldots,0}_{\bar{u}-1},\underbrace{1,\ldots,1}_{t+1-\bar{u}}),$$

pour tout $\gamma \in \mathbf{F}_{2^m}^*$. Or le dual du code $\mathsf{GRS}_t(x_1,\ldots,x_{2t};v)$ est un autre code de Reed-Solomon généralisé, $\mathsf{GRS}_t(x_1,\ldots,x_{2t};w)$ avec $w_i^{-1} = v_i\prod_{j\neq i}(x_i+x_j)$. En particulier, si $\bar{u}+\widehat{u}=t$, nous pouvons trouver un vecteur $(v_1,\ldots,v_{2t})$ tel que les deux conditions sont vérifiées en même temps. Cela se produit par exemple lorsque $\mathcal{S}$ est une involution, puisque $\mathcal{B}(0)$ (resp. $\mathcal{B}^{\perp}(0)$) est atteint pour tout $\widehat{u} < d$ (resp. pour tout $\bar{u} < d^{\perp}$).

Une situation intéressante correspond au cas où le maximum sur $\mu \in \mathbf{F}_{2^m}$ de $\mathcal{B}(\mu)$ (resp. de $\mathcal{B}^{\perp}(\mu)$) est atteint pour $\mu = 0$. Alors il existe des fonctions $\mathcal{M}$ telles que la borne supérieure du théorème 3.6 est atteinte pour $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$. Ceci implique qu'il est impossible de trouver une meilleure borne générale dépendant uniquement de $\mathcal{S}$ et de t. Cette situation se produit en particulier pour toute boîte-S involutive. En effet, en utilisant les propositions 3.10 et 3.12, nous obtenons que pour toute boîte-S involutive et pour tout $t \leqslant 2^{m-1}$, il existe des fonctions de diffusion linéaires de $\mathbf{F}_{2^m}^t$ telles que les valeurs exactes du MEDP_2 et du MELP_2 sont égales aux bornes de FSE 2003.

Corollaire 3.15. *Soit $\mathcal{S}$ une involution de $\mathbf{F}_{2^m}$ et t un entier tel que $t \leqslant 2^{m-1}$. Alors il existe une fonction de diffusion $\mathcal{M}$ de $\mathbf{F}_{2^m}^t$ linéaire sur $\mathbf{F}_{2^m}$ et de branch number maximal telle que tout chiffrement de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ vérifie*

$$\mathrm{MEDP}_2 = 2^{m(t+1)}\max_{a\in\mathbf{F}_{2^m}^*}\sum_{\gamma\in\mathbf{F}_{2^m}^*}\delta_F(a,\gamma)^{(t+1)} = 2^{m(t+1)}\max_{b\in\mathbf{F}_{2^m}^*}\sum_{\gamma\in\mathbf{F}_{2^m}^*}\delta_F(\gamma,b)^{(t+1)}$$

et

$$\mathrm{MELP}_2 = 2^{2m(t+1)}\max_{a\in\mathbf{F}_{2^m}^*}\sum_{\gamma\in\mathbf{F}_{2^m}^*}\mathcal{W}_F(a,\gamma)^{2(t+1)} = 2^{m(t+1)}\max_{b\in\mathbf{F}_{2^m}^*}\sum_{\gamma\in\mathbf{F}_{2^m}^*}\mathcal{W}_F(\gamma,b)^{2(t+1)}.$$

Démonstration. D'après la proposition 3.10, pour tout u, $\max_{\mu\in\mathbf{F}_{2^m}}\mathcal{B}_u(\mu) = \mathcal{B}_u(0) = \mathcal{B}(0)$, et $\max_{\mu\in\mathbf{F}_{2^m}}\mathcal{B}_u^{\perp}(\mu) = \mathcal{B}_u^{\perp}(0) = \mathcal{B}^{\perp}(0)$. De plus, toutes ces valeurs sont égales aux bornes de FSE 2003. En combinant le théorème 3.6 et la proposition 3.12, nous déduisons l'existence de fonctions de diffusion telles que le MEDP_2 (resp. le MELP_2) est majoré et minoré par $\mathcal{B}(0)$ (resp. $\mathcal{B}^{\perp}(0)$). Nous avons prouvé dans la proposition 3.10 que $\mathcal{B}(0)$ (resp. $\mathcal{B}^{\perp}(0)$) est atteint pour toute valeur de u. Ceci correspond au cas où nous pouvons construire un code GRS satisfaisant les conditions pour MEDP_2 et MELP_2 en même temps. □

Exemple 3.16. La permutation PRØST, définie sur $\mathbf{F}_2^{16d}$, $d \geqslant 1$, est utilisée dans plusieurs schémas de chiffrement authentifié soumis à la compétition CAESAR [KLL+14]. Cette permutation est de la forme SPN$(4, 4d, S, M)$ où S est une involution sur 4 bits appelée `SubRows` et M correspond à la composition de deux permutations linéaires, `MixSlices` et `ShiftPlanes`. L'addition de la constante de tour est omise dans cet exemple car elle n'a aucun impact sur l'étude. De façon similaire à l'AES, deux tours consécutifs de la permutation PRØST peuvent être vus comme l'application en parallèle de d copies d'une superboîte-S sur $\mathbf{F}_{16}$. Cette superboîte-S correspond à deux couches de la fonction `SubRows` séparées par la fonction `MixSlices`. De plus, même si ce n'est pas mentionné dans [KLL+14], nous pouvons vérifier que `MixSlices` est linéaire sur $\mathbf{F}_{16}$ si $\mathbf{F}_2^4$ est identifié à $\mathbf{F}_{16}$ par l'isomorphisme suivant :

$$\varphi : (x_0, \ldots, x_3) \mapsto x_1 + \alpha x_2 + \alpha^2 x_3 + \alpha^3 x_0$$

où α est une racine de $X^4 + X^3 + 1$. En effet, la fonction définie sur $\mathbf{F}_{16}^4$ par $\mathcal{M} = \tilde{\varphi} \circ \texttt{MixSlices} \circ \tilde{\varphi}^{-1}$ correspond à la multiplication par

$$\begin{pmatrix} 1 & \alpha & \alpha+\alpha^2 & \alpha^2 \\ \alpha & 1 & \alpha^2 & \alpha+\alpha^2 \\ \alpha+\alpha^2 & \alpha^2 & 1 & \alpha \\ \alpha^2 & \alpha+\alpha^2 & \alpha & 1 \end{pmatrix}.$$

Les résultats précédents s'appliquent donc pour un réseau de substitution-permutation avec les mêmes paramètres et la même boîte-S que la permutation PRØST, mais pas directement sur cette permutation puisque dans celle-ci, la clé est fixée, donc les notions de MEDP et MELP ne sont pas définies. Puisque la boîte-S est une involution, les nouvelles bornes sont égales aux bornes de FSE 2003 (proposition 3.10) : pour toute fonction de diffusion `MixSlices` linéaire sur $\mathbf{F}_2$ et MDS, le chiffrement de la forme SPN$(4, 4d, \texttt{SubRows}, \texttt{MixSlices})$ vérifie :

$$\mathrm{MEDP}_2 \leqslant 2^{-8} \text{ et } \mathrm{MELP}_2 \leqslant 2^{-8}.$$

Ces bornes sont atteintes (corollaire 3.15) : en utilisant la construction présentée dans la démonstration de la proposition 3.12, nous obtenons que la matrice $\mathcal{M}$ à coefficients dans $\mathbf{F}_{16}$ (avec l'identification décrite précédemment) présentée ci-dessous donne une variante de la permutation PRØST telle que $\mathrm{MEDP}_2 = \mathrm{MELP}_2 = 2^{-8}$:

$$\begin{pmatrix} \alpha^2+\alpha+1 & \alpha^3+\alpha & \alpha^3+\alpha+1 & 1 \\ \alpha+1 & \alpha^3+\alpha^2+\alpha & \alpha^2+\alpha+1 & 1 \\ \alpha^2+1 & \alpha^3+\alpha^2+1 & \alpha^3 & 1 \\ \alpha^2 & \alpha^3+\alpha^2 & \alpha^3+1 & 1 \end{pmatrix}.$$

Cela implique que, pour cette boîte-S particulière, la fonction `MixSlices` doit être choisie avec attention pour assurer que le MEDP_2 et le MELP_2 sont petits, tandis que pour une moitié des boîtes-S dans la même classe d'équivalence affine que `SubRows`, comme par exemple la boîte-S définie par :

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	6	11	7	2	4	5	9	8	12	3	13	1	15	14	10

où les mots de 4 bits sont identifiés aux entiers par leur décomposition en base 2, le théorème 3.6 nous montre que $\mathrm{MEDP}_2 \leqslant 3 \times 2^{-10}$ pour toute fonction `MixSlices` linéaire sur $\mathbf{F}_{16}$ de branch number maximal.

Cela ne fait pas une grande différence dans le cas de Prøst puisque l'alphabet est petit : les valeurs exactes du MEDP_2 et du MELP_2 se calculent facilement. Par exemple, pour la fonction `MixSlices` choisie par les concepteurs, nous avons $\mathrm{MEDP}_2 = 3 \times 2^{-11}$ et $\mathrm{MELP}_2 = 81 \times 2^{-16}$. Cependant, pour des boîtes-S définies sur $\mathbf{F}_{2^8}$, calculer ces valeurs exactes est très coûteux et obtenir de meilleures bornes est très utile.

3.4 Influence de la représentation du corps

Il n'y a pas de raisons pour que le MEDP ou le MELP soient les mêmes pour toute boîte-S d'une même classe d'équivalence affine. Il est donc normal que nos bornes ne soient pas invariantes par équivalence affine. Ce qui est plus étonnant, c'est qu'en combinant la borne supérieure du théorème 3.6 avec la borne inférieure de la proposition 3.12, nous pouvons présenter des exemples où la représentation du corps $\mathbf{F}_{2^m}$, *i.e.*, le choix de l'isomorphisme φ entre $\mathbf{F}_2^m$ et $\mathbf{F}_{2^m}$, semble influencer les valeurs du MEDP et du MELP.

Exemple 3.17. Considérons deux tours d'un chiffrement de la forme $\mathrm{SPN}(4, 4, S, M)$, où S est la permutation de $\mathbf{F}_2^4$ nommée S_6 dans [BCG+12a, Table 3] et qui est utilisée dans la famille de chiffrements incluant PRINCE [BCG+12b] :

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S_6(x)$	0	1	2	13	4	7	15	6	8	14	11	10	9	3	12	5

où chaque élément de $\mathbf{F}_2^4$ est représenté comme un entier entre 0 et 15. L'uniformité différentielle de cette boîte-S est égale à $\Delta(S_6) = 4$ et sa linéarité est égale à $\mathcal{L}(S_6) = 8$. Pour cette boîte-S, la borne de FSE 2003 donne $\mathrm{MEDP}_2^E \leqslant 34 \times 2^{-14}$ pour toute fonction de diffusion M de $\mathbf{F}_2^{16}$ $\mathbf{F}_2$-linéaire de branch number égal à 5, *i.e.* MDS. Si maintenant nous considérons une fonction de diffusion $\mathcal{M}$ de branch number égal à 5 et qui est linéaire sur $\mathbf{F}_{2^4}$ où $\mathbf{F}_{2^4}$ est identifié à $\mathbf{F}_2^4$ grâce à la base $\{1, \alpha, \alpha^2, \alpha^3\}$, α étant une racine du polynôme irréductible $X^4 + X^3 + X^2 + X + 1$, nous obtenons, d'après le théorème 3.6, que

$$\mathrm{MEDP}_2^E \leqslant 33 \times 2^{-14} ,$$

et cette inégalité est vérifiée par toute fonction $\mathcal{M}$ de ce type. Cependant, nous pouvons aussi considérer une permutation $\mathcal{M}'$ linéaire sur $\mathbf{F}_{2^4}$, mais où l'isomorphisme entre $\mathbf{F}_2^4$ et $\mathbf{F}_{2^4}$ est défini par une base différente, par exemple $\{1, \beta, \beta^2, \beta^3\}$ où β est une racine du polynôme primitif $X^4 + X + 1$. Alors la valeur $\mathcal{B}(0)$ qui intervient dans la proposition 3.12 est égale à 17×2^7, ce qui implique qu'il existe une permutation $\mathcal{M}'$ linéaire sur $\mathbf{F}_{2^4}$ de branch number différentiel 5 telle que

$$\mathrm{MEDP}_2^E = 34 \times 2^{-14} ,$$

ce qui est strictement plus grand que la borne supérieure que nous avons pour toute fonction de diffusion MDS $\mathbf{F}_{2^4}$-linéaire où $\mathbf{F}_{2^4}$ est défini avec la base $\{1, \alpha, \alpha^2, \alpha^3\}$.

Il n'y a pas de contradiction entre ces deux dernières affirmations puisqu'elles s'appliquent aux représentations de la boîte-S et de la fonction de diffusion sur $\mathbf{F}_{2^4}$ seulement. Ici, nous avons montré qu'il existe une fonction $\mathcal{M}'$ particulière telle que tout chiffrement de la forme $\mathrm{SPN}(4, 4, S, \widetilde{\psi}^{-1} \circ \mathcal{M}' \circ \widetilde{\psi})$ vérifie $\mathrm{MEDP}_2 = 34 \times 2^{-14}$, où $\widetilde{\psi}$ est la concaténation de 4 copies de l'isomorphisme ψ de $\mathbf{F}_2^4$ dans $\mathbf{F}_{2^4}$ défini par la base $\{1, \beta, \beta^2, \beta^3\}$. Mais si nous considérons la base $\{1, \alpha, \alpha^2, \alpha^3\}$ et l'isomorphisme correspondant φ, le théorème 3.6 fournit une borne pour tout chiffrement de la forme $\mathrm{SPN}_F(4, 4, \varphi \circ S \circ \varphi^{-1}, \mathcal{M})$, ce qui n'inclut pas le cas précédent car la permutation définie par $\mathcal{M} = \widetilde{\varphi} \circ \widetilde{\psi}^{-1} \circ \mathcal{M}' \circ \widetilde{\psi} \circ \widetilde{\varphi}^{-1}$ n'est pas linéaire sur $\mathbf{F}_{2^4}$, puisque $(\psi \circ \varphi^{-1})$ n'est pas un isomorphisme d'anneau.

Ce qui est plus inattendu est que le choix de l'isomorphisme entre $\mathbf{F}_{2^4}$ et $\mathbf{F}_2^4$ semble influencer la valeur du MEDP_2 alors que nous avons vu que les propriétés différentielles et linéaires d'un chiffrement vu sur le corps $\mathbf{F}_{2^m}$ étaient les mêmes que celles du chiffrement binaire et donc qu'elles étaient indépendantes de la base de $\mathbf{F}_{2^m}$ utilisée.

Cette apparente contradiction vient du fait que les définitions de la boîte-S et de la fonction de diffusion n'utilisent pas la même représentation : la boîte-S est définie sur $\mathbf{F}_2^4$ alors que la fonction de diffusion est définie sur $\mathbf{F}_{2^4}$. C'est pourquoi le choix de la base influence la valeur du MEDP alors que ce n'est évidement pas le cas lorsque les deux fonctions sont définies sur le même alphabet. Autrement dit, ces deux chiffrements sur $\mathbf{F}_{2^m}$ ne correspondent pas au même chiffrement binaire : pour qu'ils correspondent au même chiffrement binaire, il aurait fallu que le premier chiffrement utilise la fonction $\mathcal{S}_1 = \varphi \circ S \circ \varphi^{-1}$ et que le deuxième utilise la fonction $\mathcal{S}_2 = \psi \circ S \circ \psi^{-1}$.

Bien que cela ne corresponde pas à une description mathématique naturelle, il peut être utile d'utiliser la représentation binaire pour décrire la boîte-S, qui est choisie pour minimiser le nombre de portes logiques par exemple, alors que la représentation sur le corps est utilisée pour la couche de diffusion puisque la fonction assurant ce rôle est linéaire sur $\mathbf{F}_{2^m}$. C'est le cas par exemple pour le chiffrement `LED`.

Exemple 3.18. Dans [GPPR11, Section 3.2], les concepteurs de `LED` fournissent une borne supérieure sur les valeurs MEDP et MELP sur quatre tours d'un chiffrement de la forme $\mathrm{SPN}_F(4, 4, \mathcal{S}, \mathcal{M})$ où $\mathcal{M}$ est une fonction $\mathbf{F}_{2^4}$-linéaire de branch number égal à 5, $\mathcal{S}$ est la boîte-S définie dans la figure 3.3 et $\mathbf{F}_{2^4}$ est défini avec la base $(1, \alpha, \alpha^2, \alpha^3)$, où α est une racine du polynôme $X^4 + X + 1$. Pour cela, ils utilisent la borne de FSE 2003, qui donne $\mathrm{MEDP}_2 \leqslant 2^{-8}$ et $\mathrm{MELP}_2 \leqslant 2^{-8}$, ce qui implique que $\mathrm{MEDP}_4 \leqslant 2^{-32}$ et $\mathrm{MELP}_4 \leqslant 2^{-32}$. Pour ce chiffrement, la nouvelle borne est égale à celle de FSE 2003 et n'améliore donc pas le résultat. Cependant, si nous considérons la même boîte-S, mais que nous changeons la représentation de $\mathbf{F}_{2^4}$ et que nous choisissons la base définie par $X^4 + X^3 + 1$, nous obtenons d'après le théorème 3.6 : $\mathrm{MEDP}_2 \leqslant 3 \times 2^{-10}$. Ainsi, grâce à cette légère modification, la borne supérieure sur MEDP_4 est améliorée d'un facteur $(3/4)^4 = 0,3164$ (et la borne sur MELP_4 est inchangée). De même que pour Prøst, ce résultat ne s'applique pas directement à `LED` (bien qu'il soit utilisé par les concepteurs [GPPR11]) car les sous-clés ne sont insérées que tous les quatre tours.

Remarquons que la situation précédente n'est pas liée au fait que l'une des représentations du corps est définie avec un polynôme non primitif. En effet, dans l'exemple suivant, nous pouvons observer que changer le polynôme primitif utilisé pour représenter le corps $\mathbf{F}_{2^m}$ peut modifier les valeurs du MEDP et du MELP sur deux tours.

Exemple 3.19. Considérons deux tours d'un chiffrement de la forme $\mathrm{SPN}(5, 4, S, M)$ où S est la permutation de $\mathbf{F}_2^5$ définie par :

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	1	18	20	25	16	6	27	17	3	22	15	31	7	30	26
x	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
$S(x)$	4	23	29	21	9	10	24	2	14	5	13	8	28	19	12	11

Lorsque $\mathbf{F}_{2^5}$ est identifié à $\mathbf{F}_2^5$ avec la base $\{1, \alpha, \alpha^2, \alpha^3, \alpha^4\}$, où α est une racine du polynôme primitif $X^5 + X^2 + 1$, nous avons d'après le théorème 3.6 que toute fonction de diffusion MDS et linéaire sur $\mathbf{F}_{2^5}$ avec cette représentation vérifie

$$\mathrm{MEDP}_2^E \leqslant 13 \times 2^{-20} \text{ et } \mathrm{MELP}_2^E \leqslant 8407 \times 2^{-27} .$$

Lorsque $\mathbf{F}_{2^5}$ est construit avec le polynôme primitif $X^5 + X^3 + 1$, les bornes supérieures et inférieures du théorème 3.6 et de la proposition 3.12 sont égales, et il existe des fonctions de diffusion MDS linéaire sur $\mathbf{F}_{2^5}$ avec cette deuxième représentation telles que

$$\mathrm{MEDP}_2^E = 14 \times 2^{-20} \text{ et } \mathrm{MELP}_2^E = 8663 \times 2^{-27} .$$

Donc l'utilisation du premier polynôme primitif permet d'avoir des valeurs pour le MEDP et le MELP sur deux tours inférieures aux valeurs obtenues en utilisant le second polynôme.

4

Invariance multiplicative d'une boîte-S

Du fait de leur coût d'implémentation peu élevé en général, les permutations puissance, c'est-à-dire les permutations de $\mathbf{F}_{2^m}$ de la forme $x \mapsto x^s$, sont souvent utilisées comme boîtes-S. De plus, leur spectre différentiel et leur spectre de Walsh sont assez faciles à déterminer. Plus intéressant encore, les lignes et les colonnes de leur table des différences (resp. de leur table de Walsh) peuvent être obtenues à partir d'une seule d'entre elles par permutation des coefficients. Ceci vient du fait que toute permutation $\mathcal{S}$ est un endomorphisme sur le groupe multiplicatif $\mathbf{F}_{2^m}^*$, *i.e.*, $\mathcal{S}(xy) = \mathcal{S}(x)\mathcal{S}(y)$ pour tout couple d'éléments non nuls (x, y). Cette propriété permet de simplifier les bornes du chapitre précédent et surtout d'avoir une meilleure approximation des valeurs exactes du MEDP_2 et du MELP_2, cette approximation ne dépendant que de la boîte-S et du branch number.

Il y a malheureusement peu de chance de déterminer une famille plus large de permutations avec cette propriété car toute fonction $\mathcal{S}$ vérifiant $\mathcal{S}(xy) = \mathcal{S}(x)\mathcal{S}'(y)$ pour une certaine fonction $\mathcal{S}'$ est de la forme $\mathcal{S}(x) = cx^s$. Cependant nous allons présenter dans ce chapitre une propriété, dite d'invariance multiplicative, sur la table des différences (resp. sur la table de Walsh) de $\mathcal{S}$.

Après avoir défini cette propriété, nous verrons dans la deuxième partie que lorsque les lignes ou colonnes de la table des différences (resp. de la table de Walsh) d'une boîte-S peuvent être déduites de l'une d'entre elle, la borne supérieure sur le MEDP_2 (resp. MELP_2) du théorème 3.6 se simplifie. De plus, pour une telle boîte-S, la borne inférieure présentée dans la proposition 3.12 est valide pour toute fonction de diffusion MDS linéaire sur $\mathbf{F}_{2^m}$. Les bornes présentées dans ce chapitre sont donc indépendantes de la fonction de diffusion $\mathbf{F}_{2^m}$-linéaire utilisée : elles ne dépendent que du branch number. Ces résultats sont présentés dans la troisième partie. La dernière partie de ce chapitre est consacrée à la démonstration d'un résultat lié à la conjecture faite dans [DLP+09] : au sein de sa classe d'équivalence affine, la fonction inverse est la fonction qui mène à la plus grande valeur possible pour le MEDP_2 et le MELP_2.

Dans tout ce chapitre, les réseaux de substitution-permutation sont représentés sur le corps $\mathbf{F}_{2^m}$ (ils sont de la forme $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$).

4.1 Propriété d'invariance multiplicative

Soit s un entier strictement positif. La propriété sur la table des différences (resp. la table de Walsh) d'une permutation est définie ci-dessous.

Définition 4.1. *Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$.*

- *$\mathcal{S}$ est dite* à dérivées invariantes par multiplication *si, pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation π_λ de $\mathbf{F}_{2^m}^*$ telle que*

$$\delta_F(\alpha, \lambda\beta) = \delta_F(\pi_\lambda(\alpha), \beta) \quad \forall \beta \in \mathbf{F}_{2^m}^*.$$

- *$\mathcal{S}$ est dite* à transformée de Walsh invariante par multiplication *si, pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation ψ_λ de $\mathbf{F}_{2^m}^*$ telle que*

$$\mathcal{W}_F(\alpha, \lambda\beta)^2 = \mathcal{W}_F(\psi_\lambda(\alpha), \beta)^2 \quad \forall \beta \in \mathbf{F}_{2^m}^*.$$

Ces propriétés signifient que, quand les colonnes de la table des différences (resp. de Walsh) sont disposées dans l'ordre des puissances successives d'un élément primitif, alors les décalages circulaires d'une ligne donnée correspondent à l'ensemble des autres lignes de la table. Nous pouvons observer cette propriété sur la table des différences de la fonction $x \mapsto x^{13}$ de $\mathbf{F}_{2^4}$, présentée dans l'exemple 3.8 et rappelée ci-dessous.

	1	g	g^2	g^3	g^4	g^5	g^6	g^7	g^8	g^9	g^{10}	g^{11}	g^{12}	g^{13}	g^{14}
1	4	0	0	0	0	2	0	2	0	0	2	2	0	2	2
g	0	0	0	2	0	2	0	0	2	2	0	2	2	4	0
g^2	0	2	0	2	0	0	2	2	0	2	2	4	0	0	0
g^3	0	2	0	0	2	2	0	2	2	4	0	0	0	0	2
g^4	0	0	2	2	0	2	2	4	0	0	0	0	2	0	2
g^5	2	2	0	2	2	4	0	0	0	0	2	0	2	0	0
g^6	0	2	2	4	0	0	0	0	2	0	2	0	0	2	2
g^7	2	4	0	0	0	0	2	0	2	0	0	2	2	0	2
g^8	0	0	0	0	2	0	2	0	0	2	2	0	2	2	4
g^9	0	0	2	0	2	0	0	2	2	0	2	2	4	0	0
g^{10}	2	0	2	0	0	2	2	0	2	2	4	0	0	0	0
g^{11}	2	0	0	2	2	0	2	2	4	0	0	0	0	2	0
g^{12}	0	2	2	0	2	2	4	0	0	0	0	2	0	2	0
g^{13}	2	0	2	2	4	0	0	0	0	2	0	2	0	0	2
g^{14}	2	2	4	0	0	0	0	2	0	2	0	0	2	2	0

FIGURE 4.1 – Table des différences de la fonction $x \mapsto x^{13}$ du corps $\mathbf{F}_{2^4}$.

Comme nous l'avons déjà dit, toutes les permutations puissance sont à dérivées et à transformée de Walsh invariantes par multiplication. Ces définitions incluent de plus toutes les fonctions obtenues à partir des permutations puissance en les composant à droite par une permutation $\mathbf{F}_{2^m}$-linéaire.

Proposition 4.2. *Soit $\mathcal{S} = \mathcal{S}' \circ A$, où A est une permutation de $\mathbf{F}_{2^m}$ affine sur $\mathbf{F}_2$ et $\mathcal{S}' : x \mapsto x^s$ est une permutation puissance de $\mathbf{F}_{2^m}$. Alors les dérivées de $\mathcal{S}$ et sa transformée de Walsh sont invariantes par multiplication.*

Démonstration. D'après les propositions 1.44 et 1.50, nous savons que

$$\delta_F^{\mathcal{S}}(\alpha,\beta) = \delta_F^{\mathcal{S}'}(L(\alpha),\beta) \text{ et } \mathcal{W}_F^{\mathcal{S}}(\alpha,\beta)^2 = \mathcal{W}_F^{\mathcal{S}'}((L^{-1})^*(\alpha),\beta)^2 ,$$

où $L : x \mapsto A(x) + A(0)$ est la partie linéaire de A. Puisque $\mathcal{S}'(x) = x^s$, nous avons

$$\begin{aligned} \delta_F^{\mathcal{S}'}(\alpha,\beta\lambda) &= \#\{x \in \mathbf{F}_{2^m}, (x+\alpha)^s + x^s = \beta\lambda\} \\ &= \#\{x \in \mathbf{F}_{2^m}, \left(\lambda^{-e}x + \lambda^{-e}\alpha\right)^s + (\lambda^{-e}x)^s = \beta\} = \delta_F^{\mathcal{S}'}(\lambda^{-e}\alpha,\beta) \end{aligned}$$

où $x \mapsto x^e$ est la fonction inverse de $\mathcal{S}'$ pour la composition, *i.e.*, e est l'inverse de s modulo $(2^m - 1)$. Nous en déduisons que

$$\delta_F^{\mathcal{S}}(\alpha,\beta\lambda) = \delta_F^{\mathcal{S}'}(\lambda^{-e}L(\alpha),\beta) = \delta_F^{\mathcal{S}}(\pi_\lambda(\alpha),\beta)$$

avec $\pi_\lambda(\alpha) = L^{-1}(\lambda^{-e}L(\alpha))$. Pour sa transformée de Walsh, nous avons :

$$\mathcal{W}_F^{\mathcal{S}'}(\alpha,\beta\lambda) = \sum_{x\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(\beta\lambda x^s + \alpha x)} = \sum_{x\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(\beta y^s + \alpha\lambda^{-e}y)} = \mathcal{W}_F^{\mathcal{S}'}(\lambda^{-e}\alpha,\beta) .$$

Nous obtenons donc

$$\mathcal{W}_F^{\mathcal{S}}(\alpha,\beta\lambda)^2 = \mathcal{W}_F^{\mathcal{S}'}(\lambda^{-e}(L^{-1})^*(\alpha),\beta)^2 = \mathcal{W}_F^{\mathcal{S}}(\psi_\lambda(\alpha),\beta)^2$$

avec $\psi_\lambda(\alpha) = L^*(\lambda^{-e}(L^{-1})^*(\alpha))$, car $(L^{-1})^* = (L^*)^{-1}$. De manière évidente, π_λ et ψ_λ sont des permutations pour tout $\lambda \neq 0$. □

Remarquons que le fait qu'une permutation soit à dérivées (resp. transformée de Walsh) invariantes par multiplication n'implique pas qu'il en est de même pour son inverse : la proposition 4.2 n'est pas vraie lorsque la permutation puissance est composée à gauche par une permutation affine. Par ailleurs, la proposition 4.4 ci-dessous montre que les permutations dont les dérivées (resp. la transformée de Walsh) sont invariantes par multiplication ne sont pas toutes affinement équivalentes à une permutation puissance. Elle concerne les permutations dites *crooked*, qui est une classe de permutations contenant les permutations APN quadratiques (voir section 1.5.2).

Définition 4.3. *[BFDF98] Une fonction $\mathcal{S}$ de $\mathbf{F}_{2^m}$ dans $\mathbf{F}_{2^m}$ est dite* crooked *si, pour tout élément non nul α de $\mathbf{F}_{2^m}$, $\mathsf{Im}(D_\alpha\mathcal{S})$ est un sous-espace affine de codimension* 1.

Toutes les permutations crooked sont APN (*i.e.* telles que $\Delta(\mathcal{S}) = 2$) et presque-courbes (*i.e.* telles que $\mathcal{L}(\mathcal{S}) = 2^{(m+1)/2}$) [BFDF98], et existent pour m impair uniquement. Il est évident que toute permutation APN quadratique est crooked, et il est conjecturé que les fonctions crooked sont exactement les fonctions APN quadratiques. Cela a été prouvé dans [Kyu07] dans le cas des fonctions monomiales et dans [BK08] dans le cas des fonctions binomiales. De plus, ces fonctions ont des propriétés d'invariance multiplicative.

Proposition 4.4. *Soit $\mathcal{S}$ une permutation crooked. Alors les dérivées de $\mathcal{S}$ sont invariantes par multiplication et la transformée de Walsh de $\mathcal{S}^{-1}$ est invariante par multiplication.*

Pour démontrer cette proposition, rappelons qu'un hyperplan affine de $\mathbf{F}_{2^m}$ est un sous-espace affine de codimension 1 et est de la forme $\{x \in \mathbf{F}_{2^m} | \mathrm{Tr}(ax) = 1\}$, $a \in \mathbf{F}_{2^m}^*$. Donc une fonction crooked $\mathcal{S}$ est telle que les ensembles $\mathsf{Im}(D_\alpha \mathcal{S})$, pour tout élément non nul α de $\mathbf{F}_{2^m}$, sont des hyperplans affines. De plus, si $\mathcal{S}$ est une permutation, ces hyperplans sont tous distincts.

Lemme 4.5. *[BFDF98, CC03] Soit $\mathcal{S}$ une permutation crooked de $\mathbf{F}_{2^m}$. Alors les ensembles $\mathsf{Im}(D_\alpha \mathcal{S})$, $\alpha \in \mathbf{F}_{2^m}^*$, sont distincts et correspondent à tous les hyperplans affines de $\mathbf{F}_{2^m}$.*

Démonstration. Puisque $\mathcal{S}$ est une permutation, pour tout α non nul, la dérivée $D_\alpha \mathcal{S}$, définie par $D_\alpha \mathcal{S} : x \mapsto \mathcal{S}(x+\alpha) + \mathcal{S}(x)$, ne s'annule pas, ce qui implique que $\mathsf{Im}(D_\alpha \mathcal{S})$ est un hyperplan affine (qui ne contient pas 0). Si ces hyperplans affines ne sont pas distincts, il existe α et α' distincts tels que $\mathsf{Im}(D_\alpha \mathcal{S}) = \mathsf{Im}(D_{\alpha'} \mathcal{S})$. Pour tout $x \in \mathbf{F}_{2^m}$, nous pouvons écrire :

$$D_{\alpha+\alpha'}\mathcal{S}(x) = (\mathcal{S}(x+\alpha+\alpha') + \mathcal{S}(x+\alpha)) + (\mathcal{S}(x+\alpha) + \mathcal{S}(x)) = D_{\alpha'}\mathcal{S}(x+\alpha) + D_\alpha \mathcal{S}(x),$$

i.e. tout élément de $\mathsf{Im}(D_{\alpha+\alpha'}\mathcal{S})$ s'écrit comme la somme de deux éléments de l'hyperplan affine $\mathsf{Im}(D_\alpha \mathcal{S})$ (puisque $\mathsf{Im}(D_\alpha \mathcal{S}) = \mathsf{Im}(D_{\alpha'} \mathcal{S})$). Donc $\mathsf{Im}(D_{\alpha+\alpha'}\mathcal{S})$ est inclus dans un hyperplan vectoriel. Or ces deux espaces sont de même dimension, donc ils sont égaux. L'ensemble $\mathsf{Im}(D_{\alpha+\alpha'}\mathcal{S})$ serait alors un hyperplan vectoriel, contenant 0, ce qui est impossible. □

Avec ce lemme, nous pouvons démontrer la proposition 4.4.

Démonstration. Notons φ la fonction qui associe à $\alpha \in \mathbf{F}_{2^m}^*$ l'élément $\varphi(\alpha) \in \mathbf{F}_{2^m}^*$ tel que $\mathsf{Im}(D_\alpha \mathcal{S}) = \{x \in \mathbf{F}_{2^m} \mid \mathrm{Tr}(\varphi(\alpha)x) = 1\}$ et telle que $\varphi(0) = 0$. D'après le lemme 4.5, les $(2^m - 1)$ hyperplans affines correspondant à $\mathsf{Im}(D_a \mathcal{S})$ pour tous $a \neq 0$ sont distincts. Donc φ est une permutation de $\mathbf{F}_{2^m}$. Soit $\lambda \in \mathbf{F}_{2^m}^*$. Alors $\delta_F(\alpha, \lambda x) = 0$ si $\mathrm{Tr}(\varphi(\alpha)\,\lambda x) = 0$ et $\delta_F(\alpha, \lambda x) = 2$ si $\mathrm{Tr}(\varphi(\alpha)\,\lambda x) = 1$. Posons $\pi_\lambda(\alpha) = \varphi^{-1}(\varphi(\alpha)\lambda)$. Alors, pour tout $x \in \mathbf{F}_{2^m}$, nous avons :

$$\begin{aligned} \delta(\pi_\lambda(\alpha), x) = 0 &\Leftrightarrow \mathrm{Tr}(\varphi(\pi_\lambda(\alpha))\, x) = 0 \\ &\Leftrightarrow \mathrm{Tr}(\varphi(\alpha)\, \lambda x) = 0 \\ &\Leftrightarrow \delta(\alpha, \lambda x) = 0. \end{aligned}$$

Donc les dérivées de $\mathcal{S}$ sont invariantes par multiplication. De plus, d'après la proposition 1.52, nous avons que, pour tous les éléments non nuls u, v de $\mathbf{F}_{2^m}$,

$$\mathcal{W}_F(u,v)^2 = \sum_{a,b \in \mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(au+bv)} \delta_F(a,b) = 2^m + \sum_{a,b \in \mathbf{F}_{2^m}, a \neq 0} (-1)^{\mathrm{Tr}(au+bv)} \delta_F(a,b).$$

Or nous avons vu que le spectre différentiel de $\mathcal{S}$ est déterminé par φ : pour tout $a \neq 0$, $\delta_F(a,b) = 1 - (-1)^{\mathrm{Tr}(\varphi(a)b)}$. Alors pour tout $v \neq 0$, nous avons

$$\begin{aligned}\mathcal{W}_F(u,v)^2 &= 2^m + \sum_{a,b\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au+bv)} - \sum_{a,b\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au+bv+\varphi(a)b)} \\ &= 2^m - \sum_{a\in\mathbf{F}_{2^m},a\neq 0} (-1)^{\mathrm{Tr}(au)} \left(\sum_{b\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(b(v+\varphi(a)))} \right).\end{aligned}$$

Or la somme $\sum_{b\in\mathbf{F}_{2^m}} (-1)^{\mathrm{Tr}(b(v+\varphi(a)))}$ est égale à 2^m si $v + \varphi(a) = 0$, c'est-à-dire si $a = \varphi^{-1}(v)$, et à 0 sinon. Comme $\varphi^{-1}(v) \neq 0$ lorsque $v \neq 0$, nous obtenons

$$\mathcal{W}_F(u,v)^2 = 2^m - 2^m(-1)^{\mathrm{Tr}(u\varphi^{-1}(v))}.$$

Il s'ensuit que

$$\mathcal{W}_F(xy,v)^2 = 2^m - 2^m(-1)^{\mathrm{Tr}(xy\varphi^{-1}(v))} = 2^m - 2^m(-1)^{\mathrm{Tr}(y\varphi^{-1}(\psi_x(v)))} = \mathcal{W}_F(y,\psi_x(v))^2$$

où $\psi_x(v) = \varphi(x\varphi^{-1}(v))$. De plus, pour tout x non nul, ψ_x est une permutation. □

La proposition 4.4 s'applique par exemple à la famille infinie de permutations APN de degré 2 suivante :

$$x \longmapsto x^{2^i+1} + ux^{2^{j\frac{m}{3}}+2^{(3-j)\frac{m}{3}+i}} \text{ avec } \mathrm{pgcd}(i,m) = 1 \text{ et } j = im/3 \bmod 3$$

sur $\mathbf{F}_{2^m}$, m impair, divisible par 3 et pas par 9, qui n'est pas affinement équivalente à une fonction puissance [BCL08].

4.2 Nouvelles bornes pour les boîtes-S à invariance multiplicative

Nous allons montrer que pour les boîtes-S à dérivées (resp. à transformée de Walsh) invariantes par multiplication, les bornes établies dans le théorème 3.6 se simplifient. Ces bornes simplifiées dépendent des quantités suivantes, qui correspondent à une version simplifiée des quantités $\mathcal{B}_u(\mu)$ et $\mathcal{B}_u^{\perp}(\mu)$ définies dans le chapitre précédent (notation 3.5) qui ne font pas intervenir le paramètre λ. Autrement dit, nous considérons uniquement les translations des éléments des colonnes de la table des différences (resp. de la table de Walsh) au lieu des transformations affines.

Notation 4.6. Soient d et $d^{\perp}$ deux entiers strictement positifs. Pour tout $\mu \in \mathbf{F}_{2^m}$, définissons

$$\begin{aligned}\mathcal{B}'_u(\mu) &= \max_{\alpha,\beta\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta_F(\alpha,\gamma)^u\, \delta_F(\gamma+\mu,\beta)^{(d-u)}, \text{ avec } 1 \leqslant u < d, \\ \mathcal{B}'^{\perp}_u(\mu) &= \max_{\alpha,\beta\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha,\gamma)^{2u}\, \mathcal{W}_F(\gamma+\mu,\beta)^{2(d^{\perp}-u)}, \text{ avec } 1 \leqslant u < d^{\perp}.\end{aligned}$$

Pour des permutations dont les dérivées (resp. la transformée de Walsh) sont invariantes par multiplication, les quantités définies ci-dessus et celle de la notation 3.5 sont égales.

Proposition 4.7. *Soient d et $d^\perp$ deux entiers strictement positifs. Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$ telle que soit $\mathcal{S}$, soit $\mathcal{S}^{-1}$ est à dérivées (resp. à transformée de Walsh) invariantes par multiplication. Alors, nous avons*

$$\mathcal{B}_u(0) = \mathcal{B}'_u(0), \quad \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}'_u(\mu)$$

pour tout $1 \leqslant u < d$, et

$$\mathcal{B}^\perp_u(0) = \mathcal{B}'^\perp_u(0), \quad \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}^\perp_u(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}'^\perp_u(\mu)$$

pour tout $1 \leqslant u < d^\perp$, où $\mathcal{B}_u(\mu)$ et $\mathcal{B}^\perp_u(\mu)$ sont les valeurs définies à la notation 3.5.

De nouveau, nous allons donner une preuve dans le cas générique. La proposition 4.7 peut s'écrire de la manière suivante.

Proposition 4.8. *Soit m et d deux entiers positifs et Λ une matrice de taille $2^m \times 2^m$ dont les coefficients $\Lambda(\alpha, \beta)$, $(\alpha, \beta) \in (\mathbf{F}_{2^m})^2$, vérifient les conditions (2.4). Soit*

$$\mathcal{B}'_u(\mu) = \max_{\alpha, \beta \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma + \mu, \beta)^{(d-u)}, \text{ avec } 1 \leqslant u < d\,.$$

Supposons que l'une des deux conditions suivantes est vraie :

(i) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation π_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\alpha, \lambda\beta) = \Lambda(\pi_\lambda(\alpha), \beta)$ $\forall \beta \in \mathbf{F}_{2^m}^*$;*

(ii) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation ψ_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\lambda\beta, \alpha) = \Lambda(\beta, \psi_\lambda(\alpha))$ $\forall \beta \in \mathbf{F}_{2^m}^*$.*

Alors les quantités $\mathcal{B}_u(\mu)$, $1 \leqslant u < d$, définies dans le théorème 3.7 vérifient

$$\mathcal{B}_u(0) = \mathcal{B}'_u(0) \text{ et } \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}_u(\mu) = \max_{\mu \in \mathbf{F}_{2^m}^*} \mathcal{B}'_u(\mu)\,.$$

Démonstration. En posant $\gamma' = \gamma\lambda + \mu$, nous avons pour tous $\alpha, \beta, \lambda \in \mathbf{F}_{2^m}^*$ et tout $\mu \in \mathbf{F}_{2^m}$:

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u} = \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \lambda^{-1}(\gamma' + \mu))^u \Lambda(\gamma', \beta)^{d-u}.$$

De plus, si la condition (i) est vérifiée, alors nous obtenons

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u} = \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\pi_{\lambda^{-1}}(\alpha), \gamma' + \mu)^u \Lambda(\gamma', \beta)^{d-u}.$$

Le résultat s'ensuit. D'autre part, pour tous $\alpha, \beta, \lambda \in \mathbf{F}_{2^m}^*$ et tout $\mu \in \mathbf{F}_{2^m}$, nous avons :

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma\lambda + \mu, \beta)^{d-u} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\lambda(\gamma + \lambda^{-1}\mu), \beta)^{d-u}.$$

De plus, si la condition (ii) est vérifiée, alors nous obtenons

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*}\Lambda(\alpha,\gamma)^u\Lambda(\gamma\lambda+\mu,\beta)^{d-u}=\sum_{\gamma\in\mathbf{F}_{2^m}^*}\Lambda(\alpha,\gamma)^u\Lambda(\gamma+\lambda^{-1}\mu,\psi_\lambda(\beta))^{d-u}.$$

Le résultat s'ensuit. □

Ce résultat implique que les bornes supérieures définies dans le théorème 3.6 se simplifient.

Corollaire 4.9. *Soit E un chiffrement de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$, où $\mathcal{M}$ est linéaire sur $\mathbf{F}_{2^m}$, de branch number différentiel d et de branch number linéaire $d^\perp$. Si soit $\mathcal{S}$, soit $\mathcal{S}^{-1}$ est à dérivées (resp. à transformée de Walsh) invariantes par multiplication, alors nous avons :*

$$\mathrm{MEDP}_2^E\leqslant 2^{-md}\max_{1\leqslant u<d}\max_{\mu\in\mathbf{F}_{2^m}}\mathcal{B}'_u(\mu),$$

$$\mathrm{MELP}_2^E\leqslant 2^{-2md^\perp}\max_{1\leqslant u<d^\perp}\max_{\mu\in\mathbf{F}_{2^m}}\mathcal{B}'^{\perp}_u(\mu).$$

Nous observons ici que, pour ces boîtes-S particulières, la nouvelle borne supérieure sur MEDP_2 et MELP_2 ne fait plus intervenir de multiplication dans $\mathbf{F}_{2^m}$ (ni la forme particulière de la fonction π qui varie avec la representation du corps). Elle est donc indépendante du choix de la base de $\mathbf{F}_{2^m}$, contrairement à la borne générale comme nous l'avons vu à l'exemple 3.17.

4.3 Une borne inférieure universelle pour les boîtes-S à invariance multiplicative

En plus de diminuer le temps de calcul de la nouvelle borne, grâce à cette propriété, nous obtenons des bornes inférieures universelles sur le MEDP_2 et le MELP_2, *i.e.*, des bornes inférieures sont valides pour *toute* fonction de diffusion *de branch number maximal*.

Théorème 4.10. *Soit $\mathcal{S}$ une permutation de $\mathbf{F}_{2^m}$. Alors,* pour toute *fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^m}$-linéaire sur $(\mathbf{F}_{2^m})^t$ de branch number maximal $d=t+1$, les valeurs* MEDP_2 *et* MELP_2 *de tout chiffrement E de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ vérifient les conditions suivantes.*

- *Si $\mathcal{S}$ et $\mathcal{S}^{-1}$ sont à dérivées invariantes par multiplication, alors*

$$\mathrm{MEDP}_2^E\geqslant 2^{-m(t+1)}\max_{1\leqslant u<d}\mathcal{B}'_u(0);$$

- *si $\mathcal{S}$ et $\mathcal{S}^{-1}$ sont à transformée de Walsh invariante par multiplication, alors*

$$\mathrm{MELP}_2^E\geqslant 2^{-2m(t+1)}\max_{1\leqslant u<d}\mathcal{B}'^{\perp}_u(0)\ ;$$

- *si $\mathcal{S}$ est à dérivées (respectivement transformée de Walsh) invariantes par multiplication, alors*

$$\mathrm{MEDP}_2^E\geqslant 2^{-m(t+1)}\mathcal{B}'_t(0),\quad \mathrm{MELP}_2^E\geqslant 2^{-2m(t+1)}\mathcal{B}'^{\perp}_t(0).$$

– *si $\mathcal{S}^{-1}$ est à dérivées (respectivement transformée de Walsh) invariantes par multiplication, alors*

$$\mathrm{MEDP}_2^E \geqslant 2^{-m(t+1)}\mathcal{B}'_1(0), \quad \mathrm{MELP}_2^E \geqslant 2^{-2m(t+1)}\mathcal{B}'^{\perp}_1(0).$$

Le théorème 4.10 s'écrit dans le cas générique de la façon suivante.

Théorème 4.11. *Soient m et t deux entiers positifs et Λ une matrice de taille $2^m \times 2^m$ dont les coefficients $\Lambda(\alpha,\beta)$, $(\alpha,\beta) \in (\mathbf{F}_{2^m})^2$, vérifient les conditions (2.4). Supposons que l'une des conditions suivantes est vraie :*

(i) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation π_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\alpha,\lambda\beta) = \Lambda(\pi_\lambda(\alpha),\beta)\ \forall\beta \in \mathbf{F}_{2^m}^*$;*

(ii) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation ψ_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\lambda\beta,\alpha) = \Lambda(\beta,\psi_\lambda(\alpha))\ \forall\beta \in \mathbf{F}_{2^m}^*$.*

Définissons $M\Lambda$ par

$$M\Lambda = \max_{a,b\neq 0} \sum_{c\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i,c_i)\right)\left(\prod_{j=1}^{t} \Lambda(c_{t+j},b_j)\right) ,$$

où $\mathcal{C}$ est un code $\mathbf{F}_{2^m}$-linéaire de longueur $2t$, de dimension t et de distance minimale $t+1$. Alors

– *Si (i) et (ii) sont vérifiées en même temps, alors $M\Lambda \geqslant \max_{1\leqslant u<d} \mathcal{B}'_u(0)$.*
– *Si (i) est vérifiée, alors $M\Lambda \geqslant \mathcal{B}'_t(0)$.*
– *Si (ii) est vérifiée, alors $M\Lambda \geqslant \mathcal{B}'_1(0)$.*

Démonstration. Pour tout u tel que $1 \leqslant u \leqslant t$, considérons des valeurs $\widehat{\alpha},\widehat{\beta} \in \mathbf{F}_{2^m}^*$ telles que

$$\sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha},\gamma)^u \Lambda(\gamma,\widehat{\beta})^{(d-u)} = \mathcal{B}'_u(0) .$$

Soit $a \in (\mathbf{F}_{2^m})^t$ un mot de support $\{1,\ldots,u\}$ et $b \in (\mathbf{F}_{2^m})^t$ un mot de support $\{1,\ldots,t+1-u\}$. Puisque $\mathcal{C}$ est MDS, tout ensemble de $(t+1)$ positions est le support d'un mot de poids minimum [MS77, Page 319]. Soit $c \in \mathcal{C}$ de support $I = \{1,\ldots,u\} \cup \{t+1,\ldots,2t+1-u\}$. D'après le lemme 3.4, les mots de support I sont les éléments $\gamma c, \gamma \in \mathbf{F}_{2^m}^*$. Examinons les trois cas.

– Si (i) et (ii) sont vérifiées, alors nous avons

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{x\in\mathcal{C}} \left(\prod_{i=1}^{t} \Lambda(a_i,x_i)\right)\left(\prod_{j=1}^{t} \Lambda(x_{t+j},b_j)\right) \\
&= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i,\gamma c_i)\right)\left(\prod_{j=1}^{t} \Lambda(\gamma c_{t+j},b_j)\right) \\
&= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(\pi_{c_i}(a_i),\gamma)\right)\left(\prod_{j=1}^{t} \Lambda(\gamma,\psi_{c_{t+j}}(b_j))\right) .
\end{aligned}$$

Choisissons a et b tels que $a_i = \pi_{c_i}^{-1}(\widehat{\alpha})$ pour $1 \leqslant i \leqslant u$ et $a_i = 0$ sinon, et $b_j = \psi_{c_{t+j}}^{-1}(\widehat{\beta})$ pour $1 \leqslant j \leqslant t+1-u$, $b_j = 0$ sinon. Alors pour ces valeurs :

$$\Lambda_{a,b} = \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{u} \Lambda(\widehat{\alpha}, \gamma) \right) \left(\prod_{j=1}^{(t+1-u)} \Lambda(\gamma, \widehat{\beta}) \right) = \mathcal{B}'_u(0) \ .$$

Puisqu'une telle paire (a, b) peut être définie pour tout $1 \leqslant u < d$, nous en déduisons que

$$M\Lambda \geqslant \max_{1 \leqslant u < d} \mathcal{B}'_u(0).$$

– Si (i) est vérifiée, posons $u = t$ et définissons a et b par $a_i = \pi_{c_i c_{t+1}^{-1}}^{-1}(\widehat{\alpha})$ pour $1 \leqslant i \leqslant t$, $b_1 = \widehat{\beta}$ et $b_j = 0$ pour $j > 1$. Alors nous obtenons :

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i, \gamma c_i) \right) \Lambda(\gamma c_{t+1}, b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(a_i, \gamma' c_i c_{t+1}^{-1}) \right) \Lambda(\gamma', b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{t} \Lambda(\pi_{c_i c_{t+1}^{-1}}(a_i), \gamma') \right) \Lambda(\gamma', b_1) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha}, \gamma')^t \Lambda(\gamma', \widehat{\beta}) = \mathcal{B}'_t(0) \ .
\end{aligned}$$

– Si (ii) est vérifiée, posons $u = 1$ et définissons a et b par $a_1 = \widehat{\alpha}$ et $a_i = 0$ pour $i > 1$, et $b_j = \varphi_{c_{t+j} c_1^{-1}}^{-1}(\widehat{\beta})$ pour $1 \leqslant j \leqslant t$. Alors nous obtenons :

$$\begin{aligned}
\Lambda_{a,b} &= \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(a_1, \gamma c_1) \left(\prod_{j=1}^{t} \Lambda(\gamma c_{t+j}, b_j) \right) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(a_1, \gamma') \left(\prod_{j=1}^{t} \Lambda(\gamma' c_{t+j} c_1^{-1}, b_j) \right) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(a_1, \gamma') \left(\prod_{j=1}^{t} \Lambda(\gamma', \psi_{c_{t+j} c_1^{-1}} b_j) \right) \\
&= \sum_{\gamma' \in \mathbf{F}_{2^m}^*} \Lambda(\widehat{\alpha}, \gamma') \Lambda(\gamma', \widehat{\beta})^t = \mathcal{B}'_1(0) \ .
\end{aligned}$$

□

4.4 Étude de la superboîte-S de l'AES pour différentes boîtes-S affinement équivalentes

Nous nous intéressons ici à la valeur du MEDP_2 et du MELP_2 de l'AES, c'est-à-dire aux propriétés différentielles et linéaires de la superboîte-S représentée à la figure 1.10, lorsque la boîte-S de l'AES est remplacée par une boîte-S affinement équivalente. La boîte-S de l'AES correspond à la fonction inverse de $\mathbf{F}_{2^8}$ composée avec une permutation $\mathbf{F}_2$-affine de $\mathbf{F}_{2^8}$: $\mathcal{S}(x) = A(x^{254})$. Lorsque la boîte-S de l'AES a été conçue, la fonction inverse a été choisie car elle a de bonnes propriétés de résistance aux crytanalyses différentielle et linéaire. Cependant, cette fonction, ainsi que les autres composantes du chiffrement, ont une description très simple dans $\mathbf{F}_{2^8}$, ce qui implique que des attaques algébriques, utilisant les équations décrivant le chiffrement, peuvent être utilisées pour obtenir des informations sur la clé. La composition par la permutation affine A rend les équations décrivant un tour du chiffrement plus complexes et permet d'éviter ce type d'attaque, mais elle n'avait pas pour but d'augmenter la résistance du chiffrement aux attaques différentielles et linéaires. Toutefois, nous allons voir que la permutation affine A a un impact sur les propriétés cryptographiques de deux tours d'une superboîte-S (*i.e.* sur deux tours d'un réseau de substitution-permutation).

Comme la boîte-S de l'AES correspond à la fonction inverse composée à gauche par une permutation affine A, l'inverse $\mathcal{S}^{-1}$ de la boîte-S de l'AES est la fonction inverse composée à droite par A^{-1} : $\mathcal{S}^{-1}(x) = (A^{-1}(x))^{254}$. D'après la proposition 4.2, $\mathcal{S}^{-1}$ est à dérivées et transformée de Walsh invariantes par multiplication. Nous pouvons donc appliquer les théorèmes 3.6 et 4.10, ainsi que la proposition 4.7, avec $t = 4$. Pour toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^8}$-linéaire et de branch number maximal, nous obtenons :

$$2^{-40}\mathcal{B}'_1(0) \leqslant \mathrm{MEDP}_2 \leqslant 2^{-40} \max_{1\leqslant u\leqslant 4} \max_{\mu\in\mathbf{F}^*_{2^8}} \mathcal{B}'_u(\mu),$$

$$2^{-80}\mathcal{B}'^{\perp}_1(0) \leqslant \mathrm{MELP}_2 \leqslant 2^{-80} \max_{1\leqslant u\leqslant 4} \max_{\mu\in\mathbf{F}^*_{2^8}} \mathcal{B}'^{\perp}_u(\mu)\ .$$

Ces bornes ne dépendent pas de l'isomorphisme entre $\mathbf{F}_2^8$ et $\mathbf{F}_{2^8}$ puisque leur expression ne contient pas de multiplication dans $\mathbf{F}_{2^8}$. Alors, pour différents choix de permutation affine A, nous avons les résultats suivants.

– Pour la fonction affine A utilisée dans l'AES, tout chiffrement $\mathrm{SPN}_F(8, 4, \mathcal{S}, \mathcal{M})$ vérifie

$$53 \times 2^{-34} \leqslant \mathrm{MEDP}_2 \leqslant 55,5 \times 2^{-34}$$

et

$$1,638 \times 2^{-28} \leqslant \mathrm{MELP}_2 \leqslant 1,86 \times 2^{-28}$$

pour toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^8}$-linéaire et MDS et tout isomorphisme entre $\mathbf{F}_{2^8}$ et $\mathbf{F}_2^8$. D'après la proposition 2.23, les valeurs exactes pour la fonction de diffusion utilisée dans l'AES correspondent aux bornes inférieures dans les deux cas. Cependant, nous avons présenté dans la proposition 3.12 une fonction de diffusion MDS pour laquelle $\mathrm{MELP}_2 \geqslant 1,66 \times 2^{-28}$. Donc le choix de la fonction de diffusion linéaire de branch number maximal est susceptible de modifier la valeur de MELP_2 dans l'intervalle défini ci-dessus.

– Pour la fonction affine A' utilisée dans les chiffrements SHARK [RDP+96] et SQUARE [DKR97] (deux prédécesseurs de l'AES), tout chiffrement de la forme $\mathrm{SPN}_F(8,4,\mathcal{S},\mathcal{M})$ vérifie

$$53\times 2^{-34}\leqslant \mathrm{MEDP}_2\leqslant 56\times 2^{-34} \text{ et } 1,7169\times 2^{-28}\leqslant \mathrm{MELP}_2\leqslant 1,9847\times 2^{-28}$$

pour toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^8}$-linéaire et MDS. Donc la fonction affine choisie dans la boîte-S de l'AES augmente légèrement la sécurité du chiffrement par rapport à celle choisie dans SQUARE. En effet, il est impossible avec la fonction affine de SQUARE d'obtenir une valeur pour le MELP sur deux tours aussi petite que celle de l'AES. Rappelons que l'isomorphisme entre $\mathbf{F}_2^8$ et $\mathbf{F}_{2^8}$ est différent dans SQUARE et dans l'AES [Bar] : pour l'AES, les opérations sur le corps sont effectuées modulo le polynôme $X^8+X^4+X^3+X+1$ tandis qu'elles sont effectuées modulo le polynôme $X^8+X^7+X^6+X^5+X^4+X^2+1$ pour SQUARE.

– Nous avons trouvé une permutation affine A'' de $\mathbf{F}_2^8$ pour laquelle la boîte-S est telle que tout chiffrement de la forme $\mathrm{SPN}_F(8,4,\mathcal{S},\mathcal{M})$ vérifie

$$\mathrm{MEDP}_2 = 57\times 2^{-34} \text{ et } 1,8713\times 2^{-28}\leqslant \mathrm{MELP}_2\leqslant 1,957\times 2^{-28}$$

pour toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^8}$-linéaire et MDS. La permutation A'' est définie par :

$$\begin{bmatrix} y_0\\ y_1\\ y_2\\ y_3\\ y_4\\ y_5\\ y_6\\ y_7 \end{bmatrix} = \begin{bmatrix} 0\,0\,0\,1\,1\,0\,0\,0\\ 0\,1\,0\,1\,0\,0\,1\,1\\ 1\,1\,0\,1\,1\,0\,1\,1\\ 0\,0\,1\,0\,1\,0\,0\,0\\ 1\,0\,1\,0\,1\,0\,1\,0\\ 1\,1\,1\,0\,0\,0\,1\,0\\ 0\,1\,0\,1\,0\,0\,1\,0\\ 1\,0\,0\,0\,1\,1\,0\,1 \end{bmatrix} \begin{bmatrix} x_0\\ x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{bmatrix} + \begin{bmatrix} 1\\ 1\\ 0\\ 0\\ 0\\ 0\\ 1\\ 1 \end{bmatrix}.$$

Donc le MEDP sur deux tours d'un chiffrement doté de cette boîte-S est strictement plus élevé que lorsque la boîte-S de l'AES est utilisée, quelle que soit la fonction de diffusion $\mathbf{F}_{2^8}$-linéaire de branch number maximal choisie.

Boîte-S	Bornes sur MEDP_2	Bornes sur MELP_2
AES	$53\times 2^{-34}\leqslant \mathrm{MEDP}_2\leqslant 55,5\times 2^{-34}$	$1,638\times 2^{-28}\leqslant \mathrm{MELP}_2\leqslant 1,86\times 2^{-28}$
SHARK	$53\times 2^{-34}\leqslant \mathrm{MEDP}_2\leqslant 56\times 2^{-34}$	$1,7169\times 2^{-28}\leqslant \mathrm{MELP}_2\leqslant 1,9847\times 2^{-28}$
$A''(x^{254})$	$\mathrm{MEDP}_2 = 57\times 2^{-34}$	$1,8713\times 2^{-28}\leqslant \mathrm{MELP}_2\leqslant 1,957\times 2^{-28}$

FIGURE 4.2 – Bornes sur les valeurs de MEDP_2 et MELP_2 pour l'AES avec différentes boîtes-S.

Même si nous ne sommes pas capables de construire explicitement une permutation affine A qui minimise les valeurs de MEDP_2 et MELP_2, nos résultats simplifient le travail des concepteurs. En effet, la permutation affine A et la fonction de diffusion $\mathcal{M}$ peuvent être choisies presque séparément puisqu'une bonne estimation du MEDP_2 et du MELP_2 est obtenue indépendamment de la fonction de diffusion. Cette méthode est naturellement beaucoup plus rapide que de calculer ces valeurs avec l'algorithme de Keliher et Sui [KS07] pour un grand nombre de paires $(A, \mathcal{M})$.

La section suivante s'intéresse au cas particulier où $A = \mathrm{Id}$, *i.e.* au cas de la boîte-S naïve.

4.5 Involutions avec invariance multiplicative

Un cas intéressant consiste à considérer une boîte-S involutive à dérivées (ou transformée de Walsh) invariantes par multiplication. Dans ce cas, la borne inférieure de la section précédente correspond à la borne supérieure de théorème 3.6, et ces deux valeurs sont égales à la borne de FSE 2003.

Corollaire 4.12. *Soit $\mathcal{S}$ une involution de $\mathbf{F}_{2^m}$ à dérivées (respectivement transformée de Walsh) invariantes par multiplication. Alors, pour tout t et toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^m}$-linéaire de $\mathbf{F}_{2^m}^t$ de branch number $t+1$, tout chiffrement de la forme $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ vérifie*

$$\begin{aligned} \mathrm{MEDP}_2^E &= 2^{-m(t+1)} \max_{\alpha \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\alpha, \gamma)^{t+1}, \\ \mathrm{MELP}_2^E &= 2^{-2m(t+1)} \max_{\alpha \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha, \gamma)^{2(t+1)}. \end{aligned}$$

La boîte-S naïve, *i.e.* la fonction inverse dans $\mathbf{F}_{2^m}$, vérifie les hypothèses du corollaire précédent. Les valeurs exactes du MEDP_2 et du MELP_2 pour un réseau de substitution-permutation composé de la boîte-S naïve et de toute fonction de diffusion $\mathbf{F}_{2^m}$-linéaire de branch number maximal sont alors toujours égales à la borne de FSE 2003. Par exemple, pour deux tours de l'AES avec la boîte-S naïve, nous avons $\mathrm{MEDP}_2 = 79 \times 2^{-34}$ et $\mathrm{MELP}_2 = 48,193,409 \times 2^{-52}$, et ces résultats sont indépendants de la fonction de diffusion $\mathbf{F}_{2^8}$-linéaire et MDS choisie. En particulier, les valeurs exactes du MEDP_2 et du MELP_2 ne dépendent pas de la représentation du corps puisque le corollaire 4.12 fournit la même valeur pour toute base. Ceci explique pourquoi, parmi toutes les boîtes-S de la même classe d'équivalence affine, la boîte-S naïve est celle qui amène à la plus grande valeur pour le MEDP_2 et à la plus grande valeur pour le MELP_2 pour toute fonction de diffusion $\mathbf{F}_{2^m}$-linéaire de branch number maximal. Et cette situation est indépendante de la taille de la boîte-S et du choix de la fonction de diffusion $\mathbf{F}_{2^m}$-linéaire et MDS.

4.6 Utilisation de ces bornes pour un nombre de tours plus élevés

À partir de la valeur MEDP_2 (respectivement MELP_2) et du théorème 2.11, nous pouvons obtenir une borne sur la valeur MEDP (respectivement MELP) pour quatre tours d'un réseau de substitution-permutation. Cela est possible pour tout chiffrement pouvant se représenter avec des superboîtes-S (*cf* page 10), mais nous allons décrire cette méthode dans le cas particulier de l'AES [PSLL03].

Avec cette représentation, quatre tours de l'AES correspondent à deux tours du chiffrement de la forme $\mathrm{SPN}_F(32, 4, \mathrm{SBS}, \mathrm{SR} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR})$. Or l'espérance maximale de la probabilité d'une différentielle (respectivement le maximum du potentiel linéaire moyen) sur une superboîte-S est égale à la valeur MEDP_2 (respectivement MELP_2) pour le chiffrement. De plus, la fonction de diffusion $\mathrm{SR} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR}$ a un branch number égal à 5 relativement à $\mathbf{F}_2^{32}$. En appliquant le théorème 2.11 avec la valeur exacte du MEDP_2 (respectivement du MELP_2) calculée grâce à l'algorithme de Keliher et Sui (proposition 2.23), nous obtenons

$$\mathrm{MEDP}_4 \leqslant (53 \times 2^{-34})^4 = 1,881 \times 2^{-114}$$

et

$$\mathrm{MELP}_4 \leqslant (1,638 \times 2^{-28})^4 = 1,802 \times 2^{-110}.$$

Cependant, il n'est possible d'utiliser ni la borne de FSE 2003 ni la nouvelle borne pour améliorer ce résultat. En effet, ces bornes nécessitent la connaissance de la distribution de l'espérance $\mathrm{EDP}_2(a, b)$ de la probabilité d'une différentielle (a, b) lorsque a et b varient dans $(\mathbf{F}_{2^8})^4 \backslash \{0\}$ (respectivement la distribution du potentiel linéaire moyen $\mathrm{ELP}_2(u, v)$ d'un masque (u, v) lorsque u et v varient dans $(\mathbf{F}_{2^8})^4 \backslash \{0\}$), qui est l'équivalent de la table des différences (respectivement la table de Walsh) de la superboîte-S (en moyenne sur les clés). Or ces deux distributions n'ont pas encore été entièrement déterminées pour l'AES.

Dans [DLP+09], les auteurs ont déterminé dans le cas différentiel une partie de cette distribution lorsque la boîte-S de l'AES est remplacée par la boîte-S naïve, *i.e.* la fonction inverse dans $\mathbf{F}_{2^8}$. Plus exactement, ils ont déterminé la valeur $\mathrm{EDP}_2(a, b)$ pour des différentielles (a, b) sur une superboîte-S avec a de poids 1, ainsi qu'une borne sur la valeur $\mathrm{EDP}_2(a, b)$ sur les différentielles de (a, b) avec a de poids 2 et b de poids 3 ou 4. Avec cette distribution partielle, les auteurs ont pu utiliser la borne de FSE 2003 sur le chiffrement de la forme $\mathrm{SPN}_F(32, 4, \mathrm{SBS'}, \mathrm{SR} \circ \mathrm{AK} \circ \mathrm{MC} \circ \mathrm{SR})$ (où SBS' est la superboîte-S de l'AES où la boîte-S est remplacée par la fonction inverse) pour obtenir une borne sur la valeur $\mathrm{EDP}_4(x, y)$ pour des différentielles $(x, y) \in (\mathbf{F}_{2^8}^{16})^2$ telles que cinq superboîtes-S sont actives, chacune contenant cinq ou six boîtes-S actives. Si (x, y) est une différentielle vérifiant ces conditions, alors $\mathrm{EDP}_4(x, y) \leqslant 5 \times 2^{-128}$.

5

MEDP_2 atteint par une différentielle de poids non minimal

Pour améliorer la résistance à la cryptanalyse différentielle d'un chiffrement de type SPN_F, il est conseillé de choisir une fonction de diffusion avec un branch number élevé. En effet, nous avons vu au chapitre 2 que plus le branch number est élevé, plus la probabilité maximale d'une caractéristique différentielle est petite. Dans les analyses de sécurité, une borne sur la probabilité maximale d'une caractéristique différentielle est considérée comme suffisante pour que le chiffrement soit supposé sûr. En effet, il est communément admis que la probabilité d'une différentielle se comporte de la même façon que celle d'une caractéristique : les différentielles de grandes probabilités sont celles qui contiennent des caractéristiques ayant le moins de boîtes-S actives. Ceci est aussi utilisé pour les attaques sur les chiffrements : lorsqu'un attaquant cherche une différentielle de grande probabilité, il la cherche généralement à partir de caractéristiques minimisant le nombre de boîtes-S actives. Cependant, la probabilité d'une différentielle est égale à la somme des probabilités des caractéristiques qui la composent. Or, dans un chiffrement de type SPN_F, lorsque le nombre de boîtes-S actives augmente, le nombre de caractéristiques dans cette différentielle peut augmenter. Dans ce cas, s'il existe un grand nombre de caractéristiques de probabilité petite mais non nulle, en ajoutant ces probabilités, il est possible d'obtenir une différentielle de grande probabilité (par rapport à la probabilité d'une caractéristique unique).

5.1 Quelques situations où le MEDP_2 est atteint par une différentielle de poids minimal

Dans ce travail, nous nous sommes intéressés aux différentielles sur deux tours d'un réseau de substitution-permutation. Dans ce cas, le nombre de boîtes-S actives correspond au poids de la différentielle et il n'y a aucune raison pour que la probabilité maximale MEDP de deux tours d'un chiffrement de type SPN_F soit atteinte par une différentielle de poids minimal. Pourtant, c'est le cas pour la plupart des réseaux de substitution-permutation. En effet, les bornes inférieures sur le MEDP_2 (resp. le

MELP$_2$) que nous avons obtenues au théorème 4.10 dans le cas où les dérivées (resp. la transformée de Walsh) de la boîte-S et de son inverse sont invariantes par multiplication sont atteintes par les différentielles (resp. les masques linéaires) de poids minimal.

Proposition 5.1. *Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$ où $\mathcal{S}$ est une permutation telle que les dérivées (respectivement la transformée de Walsh) de $\mathcal{S}$ et celles de $\mathcal{S}^{-1}$ sont invariantes par multiplication et où $\mathcal{M}$ est une fonction $\mathbf{F}_{2^m}$-linéaire de $(\mathbf{F}_{2^m})^t$ de branch number maximal $d = t+1$. Alors le* MEDP$_2$ *(respectivement le* MELP$_2$*) est atteint par une différentielle (respectivement un masque linéaire) de poids minimal si et seulement si* $\mathrm{MEDP}_2^E = 2^{-m(t+1)} \max_{1\leqslant u\leqslant t} \mathcal{B}'_u(0)$ *(respectivement* $\mathrm{MELP}_2^E = 2^{-2m(t+1)} \max_{1\leqslant u\leqslant t} \mathcal{B}'^{\perp}_u(0)$*).*

Nous allons démontrer cette proposition dans le cas générique, c'est-à-dire que nous allons démontrer le résultat suivant.

Proposition 5.2. *Soient m et t deux entiers positifs. Soit Λ une matrice de taille $2^m \times 2^m$ dont les coefficients vérifient les conditions (2.4). Supposons que les deux conditions suivantes sont vérifiées :*

(i) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation π_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\alpha,\lambda\beta) = \Lambda(\pi_\lambda(\alpha),\beta)\ \forall \beta \in \mathbf{F}_{2^m}^*$;*

(ii) *pour tout $\lambda \in \mathbf{F}_{2^m}^*$, il existe une permutation ψ_λ de $\mathbf{F}_{2^m}^*$ telle que $\Lambda(\lambda\beta,\alpha) = \Lambda(\beta,\psi_\lambda(\alpha))\ \forall \beta \in \mathbf{F}_{2^m}^*$.*

Alors pour tout code $\mathcal{C}$ $[2t,t,t+1]$ MDS $\mathbf{F}_{2^m}$-linéaire, la valeur $\max_{a,b\in\mathbf{F}_{2^m}^*} \Lambda_{a,b}$ est atteinte par un couple de poids minimal $t+1$ si et seulement si $\max_{a,b\in\mathbf{F}_{2^m}^*} \Lambda_{a,b} = \mathcal{B}(0)$, où $\mathcal{B}(0) = \max_{1\leqslant u<d} \max_{\alpha,\beta\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \Lambda(\alpha,\gamma)^u \Lambda(\gamma,\beta)^{t+1-u}$.

Démonstration. Soit $\mathcal{C}$ un code $[2t,t,t+1]$ MDS $\mathbf{F}_{2^m}$-linéaire. Nous allons montrer que

$$\max_{\substack{\widehat{a},\widehat{b}\neq 0 \\ wt(\widehat{a},\widehat{b})=t+1}} \Lambda_{\widehat{a},\widehat{b}} = \max_{1\leqslant u\leqslant t} \mathcal{B}'_u(0).$$

Soit $c \in \mathcal{C}$ un mot de même support que $(\widehat{a},\widehat{b})$. Si $wt(\widehat{a},\widehat{b}) = t+1$, alors l'ensemble des mots de $\mathcal{C}$ de même support que $(\widehat{a},\widehat{b})$ est $\{\gamma c, \gamma \in \mathbf{F}_{2^m}^*\}$. Donc, en notant $u = wt(\widehat{a})$, nous avons :

$$\begin{aligned}
\Lambda_{\widehat{a},\widehat{b}} &= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{u} \Lambda(\widehat{a}_i,\gamma c_i)\right)\left(\prod_{j=1}^{t+1-u} \Lambda(\gamma c_{t+j},\widehat{b}_j)\right) \\
&= \sum_{\gamma\in\mathbf{F}_{2^m}^*} \left(\prod_{i=1}^{u} \Lambda(\pi_{c_i}(\widehat{a}_i),\gamma)\right)\left(\prod_{j=1}^{t+1-u} \Lambda(\gamma,\psi_{c_{t+j}}(\widehat{b}_j))\right),
\end{aligned}$$

où la dernière égalité est obtenue grâce aux conditions (i) et (ii). Appliquons de nouveau

le lemme 2.16 deux fois de suite :

$$\begin{aligned}
\Lambda_{\widehat{a},\widehat{b}} &\leqslant \prod_{i=1}^{u} \left[\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\pi_{c_i}(\widehat{a}_i), \gamma)^u \left(\prod_{j=1}^{t+1-u} \Lambda(\gamma, \psi_{c_{t+j}}(\widehat{b}_j)) \right) \right]^{\frac{1}{u}} \\
&\leqslant \prod_{i=1}^{u} \left[\sum_{\gamma \in \mathbf{F}_{2^m}^*} \left(\prod_{j=1}^{t+1-u} \Lambda(\pi_{c_i}(\widehat{a}_i), \gamma)^{\frac{u}{t+1-u}} \Lambda(\gamma, \psi_{c_{t+j}}(\widehat{b}_j)) \right) \right]^{\frac{1}{u}} \\
&\leqslant \prod_{i=1}^{u} \left[\prod_{j=1}^{t+1-u} \left(\sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\pi_{c_i}(\widehat{a}_i), \gamma)^u \Lambda(\gamma, \psi_{c_{t+j}}(\widehat{b}_j))^{t+1-u} \right)^{\frac{1}{t+1-u}} \right]^{\frac{1}{u}}.
\end{aligned}$$

En prenant le maximum sur les valeurs $\pi_{c_i}(\widehat{a}_i)$, $\psi_{c_{t+j}}(\widehat{b}_j)$ et u, nous obtenons

$$\Lambda_{\widehat{a},\widehat{b}} \leqslant \max_{1 \leqslant u \leqslant t} \max_{\alpha,\beta \in \mathbf{F}_{2^m}^*} \sum_{\gamma \in \mathbf{F}_{2^m}^*} \Lambda(\alpha, \gamma)^u \Lambda(\gamma, \beta)^{t+1-u},$$

c'est-à-dire $\Lambda_{\widehat{a},\widehat{b}} \leqslant \max_{1 \leqslant u \leqslant t} \mathcal{B}'_u(0)$.

Réciproquement, nous avons montré dans la preuve du théorème 4.11 qu'il était toujours possible de construire un couple $(\widehat{a}, \widehat{b})$ tel que $\Lambda_{\widehat{a},\widehat{b}} = \max_{1 \leqslant u \leqslant t} \mathcal{B}'_u(0)$. □

De la même façon, si seulement une des fonctions parmi $\mathcal{S}$ et $\mathcal{S}^{-1}$ est à dérivées (respectivement transformée de Walsh) invariantes par multiplication, nous avons démontré dans la preuve du théorème 4.11 que la borne inférieure du théorème 4.10 était atteinte par une différentielle (respectivement un masque linéaire) de poids minimal.

Ainsi, puisque le MEDP_2 de l'AES correspond à la borne inférieure du théorème 4.10 d'après les résultats de Keliher et Sui [KS07] (*cf* proposition 2.23), nous déduisons que le MEDP_2 de l'AES est atteint pour une différentielle de poids minimal. Plus précisément, Daemen et Rijmen ont démontré que le MEDP_2 de l'AES était atteint pour exactement 12 différentielles, toutes de poids 5 [DR06, section 9]. Ces différentielles optimales sont composées de 75 chemins : un de probabilité 2^{-30} et 74 de probabilité 2^{-35}.

De plus, lorsque la boîte-S est une involution à dérivées (respectivement transformée de Walsh) invariantes par multiplication, le MEDP_2 (respectivement MELP_2) est toujours atteint par une différentielle (respectivement un masque linéaire) de poids minimal.

Corollaire 5.3. *Soit $\mathcal{S}$ une involution de $\mathbf{F}_{2^m}$ à dérivées (respectivement transformée de Walsh) invariantes par multiplication. Alors, pour tout t et toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_{2^m}$-linéaire de $\mathbf{F}_{2^m}^t$ de branch number maximal $t+1$, le MEDP_2 (respectivement MELP_2) pour tout chiffrement de la forme $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$ est atteint par des différentielles (respectivement des masques linéaires) de poids minimal.*

Démonstration. Nous avons démontré (corollaire 4.12) que, sous ces hypothèses, le

MEDP$_2$ et le MELP$_2$ étaient toujours égaux à la borne inférieure du théorème 4.10 :

$$\begin{aligned}\mathrm{MEDP}_2 &= 2^{-m(t+1)} \max_{\alpha\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \delta_F(\alpha,\gamma)^{t+1} = 2^{-m(t+1)} \max_{1\leqslant u\leqslant t} \mathcal{B}'_u(0),\\ \mathrm{MELP}_2 &= 2^{-2m(t+1)} \max_{\alpha\in\mathbf{F}_{2^m}^*} \sum_{\gamma\in\mathbf{F}_{2^m}^*} \mathcal{W}_F(\alpha,\gamma)^{2(t+1)} = 2^{-2m(t+1)} \max_{1\leqslant u\leqslant t} \mathcal{B}'^{\perp}_u(0).\end{aligned}$$

Nous déduisons de la proposition 5.1 que, pour toute permutation $\mathcal{M}$ $\mathbf{F}_{2^m}$-linéaire MDS, le MEDP$_2$ (resp. MELP$_2$) est toujours atteint par des différentielles (resp. des masques linéaires) de poids minimal. □

Ces résultats nous ont donc conduit à nous demander si cette situation était plus générale ou s'il existait des cas pour lesquels la probabilité MEDP sur deux tours du chiffrement de type SPN$_F$ est atteinte par une différentielle de poids non minimal. Dans ce chapitre, nous donnerons des exemples de tels chiffrements avec différents paramètres. Plus précisément, nous allons chercher des conditions sur la boîte-S et la fonction de diffusion linéaire qui peuvent conduire à une telle situation. Ces travaux ont donnés lieu à des présentations dans les conférences *Codes, Cryptology and Information Security 2015* [CR15a] and *12th International Conference on Finite Fields and their applications Fq12*.

Définir un chiffrement de type SPN$_F(m,t,\mathcal{S},\mathcal{M})$ consiste à choisir une permutation de $\mathbf{F}_{2^m}$ (la boîte-S) et une permutation $\mathcal{M}$ $\mathbf{F}_{2^m}$-linéaire de $\mathbf{F}_{2^m}^t$, ou, de façon équivalente, un code $\mathcal{C}_{\mathcal{M}}$ linéaire sur $\mathbf{F}_{2^m}$ de longueur $2t$, de dimension t et de distance minimale d.

Rappelons que la probabilité de la différentielle (a,b) est égale à la somme des probabilités des caractéristiques qui la composent, c'est-à-dire des caractéristiques Q ayant a comme différence en entrée du premier tour et b en sortie du deuxième étage de boîtes-S (*cf* figure 5.1). Comme dans les chapitres précédents, nous considérons dans tout ce chapitre deux tours de chiffrement amputés du dernier étage linéaire dans la mesure où celui-ce ne joue pas de rôle sur le MEDP et le MELP.

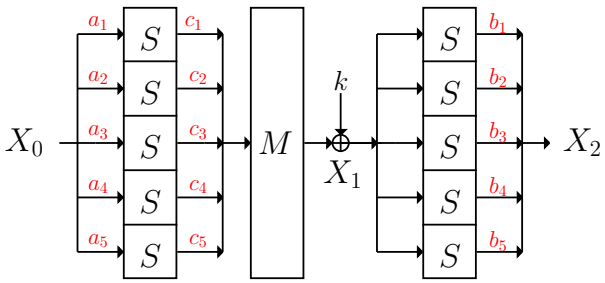


FIGURE 5.1 – Une caractéristique sur deux tours d'un chiffrement de la forme SPN$(m,5,S,M)$.

Pour un chiffrement de type $\mathrm{SPN}_F(m, t, \mathcal{S}, \mathcal{M})$, rappelons que :

$$\mathrm{EDP}_2(a,b) = 2^{-mwt(a,b)} \sum_{\substack{c \in \mathcal{C}_\mathcal{M}: \\ \mathsf{Supp}(c)=\mathsf{Supp}(a,b)}} \left(\prod_{i \in \mathsf{Supp}(a)} \delta_F(a_i, c_i) \right) \left(\prod_{j \in \mathsf{Supp}(b)} \delta_F(c_j, b_j) \right). \quad (5.1)$$

Pour calculer la probabilité d'une différentielle, il faut donc compter le nombre de caractéristiques de probabilité non nulle (aussi appelés chemins différentiels) qui la composent.

Notation 5.4. Nous noterons $\mathcal{A}_w(a, b)$ le nombre de chemins composant une différentielle (a, b) de poids w.

Chaque caractéristique correspond à un mot c de $\mathcal{C}_\mathcal{M}$ de même support que (a, b). Donc $\mathcal{A}_w(a, b)$ correspond aussi au nombre de mots c du code $\mathcal{C}_\mathcal{M}$ ayant le même support que (a, b) et tels que

$$\left(\prod_{i \in \mathsf{Supp}(a)} \delta_F(a_i, c_i) \right) \left(\prod_{j \in \mathsf{Supp}(b)} \delta_F(c_j, b_j) \right) \neq 0.$$

Pour trouver un chiffrement pour lequel le MEDP_2 ne soit pas atteint pour une différentielle de poids minimal, nous nous focalisons sur deux critères. Le premier critère est le nombre de mots de même support dans le code $\mathcal{C}_\mathcal{M}$, puisqu'il conditionne le nombre de chemins dans une différentielle. Comme la répartition des mots selon leur poids est connue pour les codes MDS, les chiffrements de type SPN_F où le branch number de la fonction de diffusion est maximal sont les plus simples à étudier. De plus, ces chiffrements sont ceux utilisés en pratique car ils ont une meilleure résistance à différentes cryptanalyses. Sauf mention contraire, les fonctions de diffusion (resp. les codes) considérées dans ce chapitre seront de branch number maximal (resp. MDS). Dans la partie suivante, les résultats sur la répartition des mots de codes MDS sont rappelés, ils permettent d'obtenir une première borne sur le nombre de chemins dans une différentielle. Un deuxième critère est l'uniformité différentielle de la boîte-S : en fonction de sa valeur, la probabilité MEDP_2 peut beaucoup varier. Nous présentons dans la troisième partie du chapitre les résultats pour les chiffrements dont la boîte-S est une fonction APN. En effet, le calcul de la probabilité EDP_2 (formule (5.1)) est simplifié dans ce cas, puisque les coefficients $\delta_F(\alpha, \beta)$ de la table des différences de la boîte-S ne prennent que deux valeurs : 0 ou 2. Dans la quatrième partie, nous nous intéresserons au cas où l'uniformité différentielle est plus grande, cas dans lequel nous avons trouvé des exemples de chiffrements tels que le MEDP_2 est atteint par des différentielles de poids non minimal.

5.2 Répartition des mots dans un code MDS

Le nombre de chemins différentiels composant une différentielle (a, b) donnée influence fortement la probabilité de cette différentielle. Afin qu'une caractéristique différentielle dans (a, b) ait une probabilité non nulle, il faut que le mot de code qui lui

correspond ait le même support que (a, b). Donc, nous allons commencer par un résultat simple sur la répartition des mots de poids w ayant un support donné dans un code MDS. Ce résultat montre que le nombre de mots ayant un support donné ne dépend que de la taille de ce support.

Paquets de mots

Les codes que nous considérons sont linéaires sur $\mathbf{F}_{2^m}$ de paramètres $[2t, t, d = t+1]$. Rappelons que les mots d'un code linéaire sur $\mathbf{F}_{2^m}$ peuvent être regroupés par paquets : si c est un mot de $\mathcal{C}_{\mathcal{M}}$, le paquet de c est $\mathcal{P}(c) = \{\gamma c, \gamma \in \mathbf{F}_{2^m}^*\}$. Les mots d'un même paquet ont le même support, donc le nombre de mots ayant le même support dans le code $\mathcal{C}_{\mathcal{M}}$ est un multiple de $(2^m - 1)$.

D'autre part, pour tout couple (α, β) de $(\mathbf{F}_{2^m}^*)^2$, les coefficients $\delta_F(\alpha, \gamma\beta)$, $\gamma \in \mathbf{F}_{2^m}^*$, correspondent à une ligne de la table des différences de la boîte-S. Pour toute permutation de $\mathbf{F}_{2^m}$, au moins $(2^{m-1} - 1)$ de ces coefficients sont nuls (environ la moitié), avec égalité si et seulement si la permutation est APN. En effet, on a :

$$\sum_{\gamma \in \mathbf{F}_{2^m}^*} \delta_F(\alpha, \gamma\beta) = 2^m.$$

Si l'uniformité différentielle est supérieure à 2, il existe des lignes ayant plus de $(2^{m-1}-1)$ coefficients nuls. Donc dans un paquet de mots, il y a au plus 2^{m-1} mots correspondant à une caractéristique différentielle de probabilité non nulle.

Mots de poids minimum

D'après le théorème de répartition des mots d'un code MDS linéaire de paramètres $[2t, t, d = t+1]$ (théorème 1.14), il y a $\binom{2t}{t+1}(2^m - 1)$ mots de poids $(t+1)$. Sachant qu'il y a $\binom{2t}{t+1}$ supports possibles pour un mot de poids minimum, il y a $(2^m - 1)$ mots de poids $(t+1)$ ayant le même support dans un code MDS, c'est-à-dire que les mots ayant le même support de poids minimum appartiennent au même paquet.

Nous avons vu précédemment que dans un paquet, il y a au plus 2^{m-1} mots correspondant à une caractéristique différentielle de probabilité non nulle. Nous obtenons donc le lemme suivant.

Lemme 5.5. *Le nombre de chemins dans une différentielle (a, b) de poids minimal $(t+1)$ satisfait :*

$$\mathcal{A}_{t+1}(a, b) \leqslant 2^{m-1}.$$

Mots de poids $w > t + 1$

D'après le théorème de répartition des mots d'un code MDS, il y a

$$\binom{2t}{w}(2^m - 1)\sum_{j=0}^{w-t-1}(-1)^j\binom{w-1}{j}2^{m(w-t-1-j)}$$

mots de poids w et il y a $\binom{2t}{w}$ supports possibles pour un mot de poids w. La proposition suivante indique que les mots de poids w sont uniformément répartis selon les supports. Dans un code MDS, il n'existe pas de support meilleur qu'un autre, au sens où il y aurait plus de mots ayant ce support.

Proposition 5.6. *Soit $\mathcal{C}$ un code MDS $\mathbf{F}_{2^m}$-linéaire de paramètres $[2t, t, d = t+1]$. Pour tout support de taille w, il y a exactement*

$$(2^m - 1) \sum_{j=0}^{w-t-1} (-1)^j \binom{w-1}{j} 2^{m(w-t-1-j)}$$

mots ayant ce support.

Démonstration. Nous allons démontrer que le nombre de mots de code ayant un support donné ne dépend que de la taille de ce support. Procédons par récurrence sur la taille w du support. Notons A'_w le nombre de mots de code ayant un support de taille w. Pour $w = t+1$, nous avons vu que ce résultat est vrai et que $A'_{t+1} = 2^m - 1$.

Soit k un entier compris entre 1 et t. Supposons que le résultat est vrai pour tout $w = t+i$, $i < k$. Soit I un support de taille $(t+k)$ et J l'ensemble formé des $(t-k)$ coordonnées qui n'appartiennent pas à I. Les mots de code dont le support est inclus dans I correspondent alors aux mots de code qui s'annulent sur J. Comme tout ensemble de t coordonnées de $\mathcal{C}_{\mathcal{M}}$ est un ensemble d'information [MS77, Page 321], nous déduisons qu'il y a exactement $(2^{mk} - 1)$ mots de code non nuls dont le support est inclus dans I. Puisque nous voulons connaître le nombre de mots de code dont le support est égal à I, il faut enlever les mots de poids strictement inférieurs à $(t+k)$ contenus dans l'ensemble précédent, c'est-à-dire les mots de poids $(t+i)$ avec $1 \leqslant i < k$. Nous obtenons alors

$$A'_{t+k} = 2^{mk} - 1 - \sum_{i=1}^{k-1} \binom{t+k}{i} A'_{t+i}.$$

Le résultat est démontré, puisque ce nombre ne dépend que de k, et non du support I choisi.

Puisqu'il y a A'_w mots de code ayant un support de taille w et $\binom{2t}{w}$ supports de taille w, d'après le théorème de répartition des mots d'un code MDS (théorème 1.14), nous avons :

$$\binom{2t}{w} A'_w = \binom{2t}{w} (2^m - 1) \sum_{j=0}^{w-t-1} (-1)^j \binom{w-1}{j} 2^{m(w-t-1-j)},$$

c'est-à-dire

$$A'_w = (2^m - 1) \sum_{j=0}^{w-t-1} (-1)^j \binom{w-1}{j} 2^{m(w-t-1-j)}.$$

□

Il y a donc $\sum_{j=0}^{w-t-1} (-1)^j \binom{w-1}{j} 2^{m(w-t-1-j)}$ paquets de mots ayant le même support de taille w et au maximum 2^{m-1} mots dans un paquet correspondant à une

caractéristique différentielle de probabilité non nulle. Nous obtenons alors le résultat suivant.

Corollaire 5.7. *Le nombre de chemins dans une différentielle (a, b) de poids w satisfait :*

$$\mathcal{A}_w(a,b) \leqslant 2^{m-1} \sum_{j=0}^{w-t-1} (-1)^j \binom{w-1}{j} 2^{m(w-t-1-j)}.$$

Remarquons que plus t est grand, moins il y a de chemins différentiels de poids $t+2$ dans un code MDS. Pour trouver un chiffrement dont la valeur MEDP sur deux tours est atteint par une différentielle de poids supérieur au poids minimum, nous nous sommes donc intéressés aux codes avec une petite dimension t.

5.3 Quand la boîte-S est une permutation APN

Nous nous intéressons dans un premier temps aux chiffrements SPN$_F$ dont la boîte-S est une permutation APN. Dans ce cas, les conditions pour que les meilleures différentielles soient de poids minimal sont simples. En effet, lorsque la boîte-S est une permutation APN, la probabilité d'une différentielle sur cette boîte-S est nulle ou égale à $2^{-(m-1)}$. Donc pour deux tours d'un réseau de substitution-permutation, la probabilité de toute caractéristique dans une différentielle de poids w ($w \geqslant t+1$) est soit nulle, soit égale à $2^{-w(m-1)}$. D'après l'égalité (5.1), nous obtenons donc que la probabilité d'une différentielle (a, b) de poids w sur deux tours du chiffrement est

$$\mathrm{EDP}_2(a,b) = 2^{-w(m-1)} \mathcal{A}_w(a,b).$$

Dans le cas où $t = 2^{m-1}$, qui correspond à la plus grande valeur possible d'après la conjecture MDS, nous pouvons montrer le lemme suivant.

Lemme 5.8. *Soit $(E_k)_k$ un chiffrement* SPN$_F(m, t, \mathcal{S}, \mathcal{M})$ *avec $t = 2^{m-1}$, où $\mathcal{S}$ est une permutation APN de $\mathbf{F}_{2^m}$ et $\mathcal{M}$ une fonction $\mathbf{F}_{2^m}$-linéaire de $(\mathbf{F}_{2^m})^t$ de branch number maximal. S'il existe une différentielle (α, β) de poids $(t+2)$ telle que*

$$\mathrm{EDP}_2(\alpha,\beta) > \max_{\substack{a\neq 0, b\in\mathbf{F}_{2^m} \\ wt(a,b)=t+1}} \mathrm{EDP}_2(a,b)$$

alors toute différentielle (a, b) de poids minimal vérifie

$$\mathcal{A}_{t+1}(a,b) < 2^{m-1}.$$

Démonstration. La probabilité de la différentielle (α, β) est plus grande que celle de (a, b) si les nombres de chemins différentiels $\mathcal{A}_{t+1}(a, b)$ et $\mathcal{A}_{t+2}(\alpha, \beta)$ vérifient :

$$\begin{aligned} \mathrm{EDP}_2(\alpha,\beta) \geqslant \mathrm{EDP}_2(a,b) &\Leftrightarrow 2^{(t+2)(1-m)} \mathcal{A}_{t+2}(\alpha,\beta) \geqslant 2^{(t+1)(1-m)} \mathcal{A}_{t+1}(a,b) \\ &\Leftrightarrow \mathcal{A}_{t+2}(\alpha,\beta) \geqslant 2^{m-1} \mathcal{A}_{t+1}(a,b). \end{aligned} \tag{5.2}$$

Le nombre de chemins différentiels dans la différentielle (α, β) doit être au moins 2^{m-1} fois plus grand que le nombre maximal de chemins différentiels composant une différentielle de poids $(t+1)$.

D'après le corollaire 5.7, nous savons que le nombre $\mathcal{A}_{t+2}$ de chemins différentiels composant une différentielle de poids $(t+2)$ est majoré par $2^{m-1}(2^m-(t+1))$. Pour que le code permette d'avoir une différentielle (α,β) de poids $(t+2)$ de plus grande probabilité que les différentielles de poids $(t+1)$, il faut que le nombre de chemins différentiels de poids $(t+2)$ imposé par l'inégalité (5.2) reste inférieur au nombre maximal de chemins différentiels de poids $(t+2)$, c'est-à-dire :

$$(2^m-(t+1))\times 2^{m-1} \geqslant 2^{m-1}\mathcal{A}_{t+1}(a,b).$$

Nous en déduisons que toute différentielle (a,b) de poids $(t+1)$ doit vérifier

$$\mathcal{A}_{t+1}(a,b) \leqslant (2^m-(t+1)) = 2^{m-1}-1.$$

□

5.3.1 Boîtes-S APN sur le corps à 8 éléments

Les fonctions APN étant des permutations si m est impair, nous allons travailler dans un premier temps avec $m=3$. Nous allons montrer que pour deux tours d'un chiffrement SPN_F dont la boîte-S est une permutation APN de $\mathbf{F}_8$ et dont la fonction de diffusion a un branch number maximal, la valeur MEDP_2 est toujours atteinte par une différentielle de poids minimal $(t+1)$. Ce résultat est dû aux propriétés particulières des permutations APN de ce corps et au fait que le corps est suffisamment petit pour que des calculs de probabilité puissent être faits en quelques heures.

Les permutations APN du corps à 8 éléments

Les résultats sur les différentielles de poids minimal qui seront présentés dans la suite sont obtenus grâce à la particularité des permutations APN de $\mathbf{F}_8$: elles correspondent à des permutations quadratiques de ce corps et leurs inverses sont aussi quadratiques.

Puisque ces permutations sont quadratiques, elles sont crooked, *i.e.*, chaque ligne de leur table des différences décrit un hyperplan affine (*cf* lemme 4.5). Plus exactement, pour tout hyperplan affine H de $\mathbf{F}_8$, il existe $a\in\mathbf{F}_8^*$ tel que $H=\{b\in\mathbf{F}_8 \mid \delta_F(a,b)=2\}$.

Les fonctions inverses des permutations APN de $\mathbf{F}_8$ étant aussi des permutations APN quadratiques, elles ont les mêmes propriétés. Autrement dit, les supports des colonnes de la table des différences d'une telle permutation correspondent à l'ensemble de tous les hyperplans affines de $\mathbf{F}_8$.

Une étude sur la table des différences de toutes les permutations APN de $\mathbf{F}_8$ peut donc être réduite à l'étude sur la table des différences d'une seule de ces permutations, les résultats pour les autres étant déductibles par permutation des lignes et des colonnes de la table des différences. Il est aussi possible de travailler uniquement sur les lignes puisque pour toute colonne considérée, il existe une ligne qui lui est égale, *i.e.*, telle que $\delta_F(a,\gamma)=\delta_F(\gamma,b)$ pour tout élément non nul γ de $\mathbf{F}_8$.

Différentielles de poids minimum

Les propriétés précédentes nous permettent de démontrer que lorsque la boîte-S d'un chiffrement SPN_F est une permutation APN de $\mathbf{F}_8$ et que le branch number de la

fonction de diffusion est maximal, la probabilité maximale d'une différentielle de poids minimal est égale à 2^{-2t}. Cette valeur est atteinte pour toute fonction de diffusion $\mathbf{F}_8$-linéaire de $(\mathbf{F}_8)^t$, de branch number maximal.

Proposition 5.9. *Soit $\mathcal{S}$ une permutation APN de $\mathbf{F}_8$. Pour tout entier t et pour toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_8$-linéaire de $(\mathbf{F}_8)^t$ de branch number maximal, tout chiffrement de type* $\mathrm{SPN}_F(3, t, \mathcal{S}, \mathcal{M})$ *vérifie :*

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=t+1}} \mathrm{EDP}_2(a,b) = 2^{-2t}.$$

Démonstration. Soit $I = \{i_1, \ldots, i_{t+1}\}$ un sous-ensemble de $\{1, \ldots, 2t\}$ de taille $(t+1)$. Nous allons mettre en évidence une différentielle (a, b) de support I telle que le nombre $\mathcal{A}_{t+1}(a, b)$ de chemins différentiels la composant est égal à 4. Nous en déduirons la valeur maximale de la probabilité d'une différentielle de poids minimal car nous savons d'après le lemme 5.5 que toute différentielle de poids minimal satisfait $\mathcal{A}_{t+1}(a, b) \leqslant 2^{m-1} = 4$.

Soit c un mot de $\mathcal{C}_{\mathcal{M}}$ tel que $\mathsf{Supp}(c) = I$ (nous avons vu qu'il en existe toujours exactement $(2^m - 1) = 7$). Choisissons un élément non nul $a_{i_1} \in \mathbf{F}_8$ et considérons l'ensemble $H = \{\beta \in \mathbf{F}_8 \,|\, \delta_F(a_{i_1}, \beta) = 2\}$. Alors H est un hyperplan affine. Définissons

$$\Gamma = \{c_{i_1}^{-1}\lambda, \lambda \in H\} .$$

Cet ensemble Γ est aussi un hyperplan affine. Donc les quatre mots du paquet de c qui s'écrivent $c' = \gamma c$ avec $\gamma \in \Gamma$ vérifient :

$$\delta_F(a_{i_1}, c'_{i_1}) = \delta_F(a_{i_1}, \lambda c_{i_1}^{-1} c_{i_1}) = 2.$$

De plus, pour toute position i_j de I telle que $i_j \leqslant t$, les coordonnées de ces quatre mots à la position i_j prennent leurs valeurs dans l'ensemble $c_{i_j}\Gamma$ qui est un hyperplan affine. Donc il existe a_{i_j} tel que cet ensemble correspond à $\{\beta \in \mathbf{F}_8 \,|\, \delta_F(a_{i_j}, \beta) = 2\}$. De même, pour toute position $i_j \in I$ telle que $i_j > t$, il existe b_{i_j} tel que l'hyperplan affine $c_{i_j}\Gamma$ correspond à $\{\alpha \in \mathbf{F}_8 \,|\, \delta_F(\alpha, b_{i_j}) = 2\}$. Pour ce choix de valeurs pour les coordonnées de (a, b), nous obtenons

$$\mathcal{A}_{t+1}(a, b) = 4 ,$$

ce qui implique que

$$\mathrm{EDP}_2(a, b) = 4 \times 2^{-2(t+1)} = 2^{-2t} .$$

□

Remarque 5.10. La démonstration ci-dessus montre plus généralement que, pour tout paquet $\mathcal{P}(c)$ de poids quelconque, on peut trouver une différentielle contenant quatre chemins différentiels correspondant à des mots de $\mathcal{P}(c)$. Cependant, pour un mot de code de poids strictement supérieur à $(t+1)$, les mots de même support se répartissent dans plusieurs paquets. Lorsque nous choisissons un paquet, nous savons qu'il existe une différentielle telle que quatre caractéristiques correspondant à des mots de ce paquet sont de probabilité non nulle, mais nous n'avons aucune information sur le nombre de chemins différentiels dans les autres paquets de mots de même support. Nous ne pouvons donc pas calculer la probabilité de cette différentielle.

Différentielles de poids supérieur

Puisque $\mathcal{C}_{\mathcal{M}}$ est un code MDS sur $\mathbf{F}_8$, la valeur de t est au plus 4 (conjecture des codes MDS, voir page 15). De plus, nous pouvons déduire du lemme 5.8 que le maximum de EDP_2 ne peut pas être atteint par une différentielle de poids $(t+2)$ lorsque $t = 4$, puisque cela impliquerait que toutes les différentielles de poids minimal vérifient $\mathcal{A}_{t+1}(a,b) \leqslant 3$ alors que nous avons prouvé que $A_{t+1}(a,b)$ est égal à 4.

Nous devons donc considérer tous les codes MDS de longueur $2t$ et de dimension t sur $\mathbf{F}_8$ pour $t \in \{2,3\}$. Pour chacun de ces codes, nous avons calculé la plus grande valeur possible de $\mathcal{A}_{t+2}(a,b)$ lorsque (a,b) est une différentielle de poids $(t+2)$. Comme les tables des différences des boîtes-S crooked sur $\mathbf{F}_8$ ont la même structure, la valeur maximale de $\mathcal{A}_{t+2}(a,b)$ sur toutes les différentielles (a,b) de support I correspond au plus grand ensemble Γ de mots de code c de support I tel que, pour chaque coordonnée $i \in I$, les c_i quand c varie dans Γ appartiennent au même hyperplan affine.

Pour $t = 2$, la plus grande valeur possible de $\mathcal{A}_4(a,b)$ lorsque (a,b) est une différentielle de poids $(t+2)$ a été calculée pour tous les codes $[4,2,3]$-linéaires MDS sur $\mathbf{F}_8$ (2058 codes). Pour tous ces codes, la valeur maximale pour $\mathcal{A}_4(a,b)$ est égale à 8. Nous en déduisons que

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=3}} \mathrm{EDP}_2(a,b) = 2^{-2\times 2} = 2^{-4} \text{ et } \max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a,b) = 2^{-4\times 2} \times 8 = 2^{-5}.$$

Dans ce cas, la valeur MEDP pour deux tours est atteinte par une différentielle de poids minimum.

Pour $t = 3$, la plus grande valeur possible de $\mathcal{A}_5(a,b)$ a été calculée pour tous les codes $[6,3,4]$-linéaires MDS sur $\mathbf{F}_8$ (environ 6,5 millions de codes), et nous obtenons que pour tous ces codes, la valeur maximale de $\mathcal{A}_5(a,b)$ est 4, ce qui implique que

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a,b) = 2^{-6} \text{ et } \max_{\substack{a\neq 0,\, b \\ wt(a,b)=5}} \mathrm{EDP}_2(a,b) = 2^{-5\times 2} \times 4 = 2^{-8}.$$

De plus, nous avons vérifié pour tous ces codes que la valeur maximale de $\mathcal{A}_6(a,b)$ est 32, donc

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=6}} \mathrm{EDP}_2(a,b) = 2^{-12} \times 32 = 2^{-7}.$$

Nous en déduisons le résultat suivant.

Proposition 5.11. *Soit $\mathcal{S}$ une permutation APN de $\mathbf{F}_8$. Pour tout entier t et toute fonction de diffusion $\mathcal{M}$ $\mathbf{F}_8$-linéaire sur $(\mathbf{F}_8)^t$ de branch number maximal, les chiffrements de la forme $\mathrm{SPN}_F(3,t,\mathcal{S},\mathcal{M})$ vérifient*

$$\mathrm{MEDP}_2 = 2^{-2t},$$

et cette valeur est uniquement atteinte pour des différentielles de poids minimal.

5.3.2 Les permutations APN du corps à 32 éléments

Les permutations APN du corps $\mathbf{F}_{32}$ ont été classifiées (à équivalence affine près) dans [BL08]. Comme ces permutations n'ont pas la même structure algébrique que celles de $\mathbf{F}_8$, il faut étudier chacune des fonctions de cette classification. De plus, le nombre de codes MDS linéaires sur $\mathbf{F}_{32}$ est beaucoup plus grand que dans le cas précédent.

Nous avons donc calculé la valeur maximale de $\mathcal{A}_{t+1}(a,b)$ pour certaines permutations APN (environ 180) et certaines fonctions de diffusion de branch number maximal pour $t \in \{2,3\}$.

Pour $t=2$, nous avons observé pour tous les codes essayés (environ 200 par permutation) que la valeur maximale de $\mathcal{A}_{t+1}(a,b)$ est au moins 10. Il faudrait donc trouver une différentielle de poids 4 contenant $\mathcal{A}_{t+2}(a,b) = \mathcal{A}_4(a,b) = 10 \times 2^{5-1} = 160$ caractéristiques pour atteindre le même EDP_2 que pour la meilleure différentielle de poids minimum. Cependant, les plus grandes valeurs que nous observons pour $\mathcal{A}_4(a,b)$ pour les codes testés sont comprises entre 83 et 92, c'est-à-dire que la valeur maximale de EDP_2 pour une différentielle de poids 4 est légèrement supérieure à la moitié de la valeur maximale de EDP_2 pour des différentielles de poids minimal.

Lorsque $t=3$, nous avons testé environ 30 permutations et 200 codes par permutation. Avec ces composants, nous observons que la valeur maximale de $\mathcal{A}_{t+1}(a,b)$ est au moins 9. Il faudrait donc trouver une différentielle de poids 5 avec $\mathcal{A}_{t+2}(a,b) = \mathcal{A}_5(a,b) = 9 \times 2^{5-1} = 144$, alors que les plus grandes valeurs que nous observons pour $\mathcal{A}_5(a,b)$ sont comprises entre 54 et 60, ce qui n'est pas suffisant.

5.4 Cas où le MEDP_2 n'est pas atteint par une différentielle de poids minimum

Il semble que le nombre $\mathcal{A}_w(a,b)$ de chemins dans une différentielle de poids $w > t+1$ n'est pas assez grand pour obtenir une probabilité plus grande que celle qui est obtenue pour les différentielles de poids minimal. Cependant, dans les cas précédemment étudiés, la plus grande probabilité pour une caractéristique est toujours égale à la valeur maximale $(\Delta(\mathcal{S})/2^m)^{t+1}$. Si la valeur EDP_2 était minimisée pour toute différentielle de poids minimal, c'est-à-dire si le nombre $\mathcal{A}_{t+1}$ était petit et que les probabilités des caractéristiques composant ces différentielles étaient différentes de $(\Delta(\mathcal{S})/2^m)^{t+1}$, il serait peut-être possible d'avoir une différentielle de poids $w > t+1$ avec une probabilité plus grande que celles des différentielles de poids minimal. C'est cette piste que j'ai explorée.

5.4.1 Deux exemples où le MEDP_2 est atteint par une différentielle de poids $(t+2)$

Les boîtes-S telles que peu de coefficients de la table des différences sont égaux à $\Delta(\mathcal{S})$ sont un bon choix pour éviter l'existence de caractéristiques de probabilité $(\Delta(\mathcal{S})/2^m)^{t+1}$ dans une différentielle de poids minimal. Mais pour les différentielles de poids $(t+2)$, la probabilité des caractéristiques doit aussi être assez élevée. Une boîte-S avec 4 à 6 coefficients de la table des différences égaux à $\Delta(\mathcal{S})$ semblent être

un bon compromis, comme nous allons le voir dans les exemples suivants. Les boîtes-S présentées ci-dessous sont définies sur l'espace vectoriel $\mathbf{F}_2^m$ alors que les fonctions de diffusion sont définies sur le corps $\mathbf{F}_{2^m}$, comme souvent dans les spécifications d'un chiffrement.

Exemple sur le corps à 8 éléments

Soit S la permutation de $\mathbf{F}_8$ définie par

x	0	1	2	3	4	5	6	7
$S(x)$	0	1	2	3	4	6	7	5

.

L'uniformité différentielle de cette fonction est $\Delta(S) = 4$ et il y a 6 coefficients égaux à 4 dans sa table des différences. Alors, pour certaines fonctions de diffusion $\mathbf{F}_8$-linéaires de $(\mathbf{F}_8)^t$ de branch number maximal, il existe des différentielles de poids $(t+2)$ dont la probabilité est supérieure à celle des différentielles de poids minimal. Par exemple, pour $t = 2$, considérons

$$\mathcal{M} = \begin{pmatrix} \alpha & \alpha+1 \\ \alpha^2 & \alpha^2+1 \end{pmatrix}$$

où α est une racine de $X^3 + X + 1$. Nous avons calculé la valeur exacte du maximum de EDP_2 pour les différentielles de poids minimal d'un côté, et pour les différentielles de poids $t + 2 = 4$ de l'autre. Nous avons obtenu :

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=3}} \mathrm{EDP}_2(a,b) = 2^{-4}$$

puisque la meilleure différentielle de poids minimal ne contient qu'une unique caractéristique de probabilité 2^{-4}, et

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a,b) = 2^{-3}$$

puisque plusieurs différentielles de poids 4 sont composées de deux caractéristiques de probabilité 2^{-4}. Ces différentielles sont :

$$\begin{aligned} a = (1,1), b = (2,2), \\ a = (2,2), b = (3,3), \\ a = (3,3), b = (1,1). \end{aligned}$$

Exemple sur le corps à 16 éléments

Un deuxième exemple est la permutation S de $\mathbf{F}_2^4$ définie par

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	2	1	5	4	9	15	8	12	11	6	7	3	14	10	13

.

Son uniformité différentielle est $\Delta(S) = 6$ et il y a 4 coefficients égaux à 6 dans sa table des différences. Alors, pour certaines fonctions de diffusions $\mathbf{F}_{16}$-linéaire de $(\mathbf{F}_{16})^t$ de

branch number maximal, il existe des différentielles de poids $(t+2)$ dont la probabilité est supérieure à celle des différentielles de poids minimal. Par exemple, pour $t = 4$, considérons

$$\mathcal{M} = \begin{pmatrix} 1 & 1 & \alpha^3 & \alpha^3 \\ \alpha^2+\alpha+1 & 1 & 1 & \alpha^2+\alpha \\ \alpha^2 & \alpha^3+1 & 1 & \alpha^3+\alpha^2+1 \\ \alpha^2+1 & \alpha^3+\alpha^2+\alpha & \alpha^3+\alpha & 1 \end{pmatrix}$$

où α est une racine de $X^4 + X + 1$.

Nous avons calculé la valeur exacte du maximum de EDP_2 pour tous les poids possibles pour une différentielle. Nous obtenons

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=5}} \mathrm{EDP}_2(a,b) = 1,2656 \times 2^{-8}, \qquad \max_{\substack{a\neq 0,\, b \\ wt(a,b)=6}} \mathrm{EDP}_2(a,b) = 1,4238 \times 2^{-8},$$

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=7}} \mathrm{EDP}_2(a,b) = 1,0942 \times 2^{-10} \text{ et } \max_{\substack{a\neq 0,\, b \\ wt(a,b)=8}} \mathrm{EDP}_2(a,b) = 1,292 \times 2^{-12}.$$

Le MEDP_2 est donc atteint par une différentielle de poids supérieur au poids minimal. Cette différentielle est

$$a = (7,7,7,7), b = (0,7,7,0).$$

5.4.2 Un exemple où le MEDP_2 est atteint par une différentielle de poids $(t+3)$

De même, nous pouvons présenter des réseaux de substitution-permutation pour lesquels le MEDP_2 est uniquement atteint par des différentielles de poids $(t+3)$.

Soit S la permutation de $\mathbf{F}_2^4$ définie par

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$S(x)$	0	4	3	7	9	14	11	12	10	13	15	8	6	5	2	1

.

Son uniformité différentielle est $\Delta(S) = 8$ et il y a 4 coefficients égaux à 8 dans sa table des différences. Notons $\mathcal{M}$ la permutation $\mathbf{F}_{16}$-linéaire de $(\mathbf{F}_{16})^3$ $(t = 3)$ de branch number maximal définie par

$$\mathcal{M} = \begin{pmatrix} 1 & \alpha & \alpha^3+\alpha^2+\alpha \\ \alpha^2 & \alpha+1 & \alpha^3+\alpha^2+\alpha+1 \\ \alpha^2+1 & \alpha^2+1 & \alpha^2+1 \end{pmatrix}$$

où α est une racine $X^4 + X + 1$.

Pour un réseau de substitution-permutation ayant ces composants, nous avons calculé que

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=4}} \mathrm{EDP}_2(a,b) = \max_{\substack{a\neq 0,\, b \\ wt(a,b)=5}} \mathrm{EDP}_2(a,b) = 2^{-6}$$

et

$$\max_{\substack{a\neq 0,\, b \\ wt(a,b)=6}} \mathrm{EDP}_2(a,b) = 524288 \times 2^{-24} = 2^{-5}\ .$$

La différentielle qui atteint le MEDP_2 est

$$a = (15, 15, 15), b = (7, 7, 7).$$

Dans les exemples de cette section, les boîtes-S sont choisies telles que peu de coefficients de leur table des différences soient égaux à la valeur maximale $\Delta(S)$. Au contraire, pour les permutations APN, où tous les coefficients non nuls de la table des différences sont égaux à la valeur maximale, le MEDP_2 est atteint par une différentielle de poids minimal. Une question se pose alors : existerait-il une borne telle que, si le nombre de coefficients de la table des différences égaux à $\Delta(S)$ est supérieur à cette borne, nous pouvons prouver que le MEDP_2 est atteint par une différentielle de poids minimal pour toute fonction de diffusion de branch number maximal ?

6

Étude de la probabilité d'une différentielle à clé fixée

Dans les chapitres précédents, nous nous sommes intéressés au maximum de l'espérance de la probabilité d'une différentielle et au potentiel linéaire moyen d'un masque sur deux tours d'un réseau de substitution-permutation. Ces deux valeurs sont considérées comme de bonnes approximations de la résistance de ces chiffrements aux cryptanalyses linéaires et différentielles. Cependant, cette affirmation repose sur l'hypothèse d'équivalence stochastique, qui suppose que le comportement des fonctions de chiffrement E_k est similaire pour toutes les clés. Autrement dit, sur l'hypothèse que, pour la plupart des clés, la probabilité d'une différentielle à clé fixée $\mathrm{DP}^{E_k}(a, b)$ est proche de l'espérance sur les clés de la probabilité de cette différentielle $\mathrm{EDP}(a, b)$. Or, dans [DR07a], les auteurs ont montré que pour l'AES, il existe une différentielle sur deux tours dont la probabilité pour certaines clés est au moins deux fois plus grande que la valeur du MEDP sur deux tours. Ainsi, il peut y avoir une différence assez grande entre la probabilité pour certaines clés d'une différentielle et l'espérance de la probabilité de cette différentielle. C'est pourquoi nous nous intéressons dans ce chapitre aux probabilités à clé fixée des différentielles sur deux tours d'un réseau de substitution-permutation. Il s'agit d'un travail en cours dont nous détaillons les premiers résultats.

Pour calculer la probabilité des différentielles à clé fixée, il faut en général commencer par calculer la probabilité d'une caractéristique différentielle. Dans ce cas, lorsque la clé est fixée, cette probabilité n'est pas égale au produit des probabilités des différentielles sur chacun des tours puisque la probabilité d'une différentielle dépend des textes clairs : lorsque la clé est fixée, la propriété de Markov n'existe pas. Donc la valeur de la probabilité d'une caractéristique sur deux tours d'un réseau de substitution-permutation peut varier en fonction de la clé.

Pour certains réseaux de substitution-permutation, en particulier ceux dont l'uniformité différentielle de la boîte-S est inférieure ou égale à 4, le calcul de la probabilité des caractéristiques a été fait dans [DR07a, DR09]. Les auteurs ont montré que l'ensemble des textes clairs qui suivent une caractéristique sur deux tours d'un tel chiffrement ont

une structure qui permet de déterminer la probabilité de ces caractéristiques de manière plus simple que dans le cas général. En particulier, pour l'AES et la variante de l'AES où la boîte-S est remplacée par la fonction inverse de $\mathbf{F}_{2^8}$, les différentes valeurs possibles pour la probabilité de la caractéristique (a, c, b) ont été déterminées pour toutes les valeurs a, c et b.

Une fois la probabilité à clé fixée des caractéristiques connue, il faut utiliser ces résultats pour calculer la probabilité d'une différentielle à clé fixée. La probabilité d'une différentielle sur deux tours d'un réseau de substitution-permutation (a, b) est égale à la somme des probabilités des caractéristiques qui ont pour différence en entrée a et pour différence en sortie b. Pour faire ce calcul, il faut déterminer selon la clé les valeurs prises par les probabilités des caractéristiques dans la différentielle parmi l'ensemble des valeurs possibles. Dans ce chapitre, nous cherchons donc à comprendre la structure de l'ensemble des clés telles que la probabilité à clé fixée d'une caractéristique sur deux tours est non nulle.

6.1 Caractéristiques plateau sur deux tours

Nous rappelons ici les travaux menés dans [DR07a] pour calculer la probabilité à clé fixée d'une caractéristique sur deux tours d'un réseau de substitution-permutation $(E_k)_k$ de la forme $\mathrm{SPN}(m, t, S, M)$. Pour calculer la probabilité d'une caractéristique sur deux tours, nous allons dans un premier temps compter le nombre de textes clairs qui suivent une différentielle $(\alpha, \beta) \in (\mathbf{F}_2^m)^2$ sur la boîte-S du chiffrement, c'est-à-dire le nombre de textes clairs x vérifiant $S(x + \alpha) + S(x) = \beta$.

Notation 6.1. Pour toute fonction S de $\mathbf{F}_2^m$ dans $\mathbf{F}_2^m$ et pour toute différentielle (α, β) de $\mathbf{F}_2^m \times \mathbf{F}_2^m$, notons $\mathcal{U}_{\alpha,\beta}(S)$ l'ensemble des solutions $x \in \mathbf{F}_2^m$ de

$$S(x + \alpha) + S(x) = \beta \ ,$$

et $\mathcal{V}_{\alpha,\beta}(S)$ l'ensemble des éléments $S(x)$ où $x \in \mathcal{U}_{\alpha,\beta}(S)$.

L'ensemble $\mathcal{U}_{\alpha,\beta}(S)$ est dit ensemble des entrées valides de la différentielle (α, β), l'ensemble $\mathcal{V}_{\alpha,\beta}(S)$ correspond aux images par S des entrées valides de la différentielle.

De même, pour une fonction Sub de $\mathbf{F}_2^{mt}$ dans $\mathbf{F}_2^{mt}$, notons $\mathcal{U}_{a,b}(\mathsf{Sub})$ l'ensemble des entrées valides de la différentielle (a, b), c'est-à-dire l'ensemble des solutions $x \in \mathbf{F}_2^{mt}$ de

$$\mathsf{Sub}(x + a) + \mathsf{Sub}(x) = b \ ,$$

et $\mathcal{V}_{a,b}(\mathsf{Sub})$ l'ensemble des éléments $\mathsf{Sub}(x)$ où $x \in \mathcal{U}_{a,b}(\mathsf{Sub})$.

Il a été remarqué que, sous certaines conditions, les ensembles des entrées valides d'une boîte-S (et d'un étage de boîtes-S) ont une structure d'espace affine. Cela induit une structure pour les caractéristiques, qui nous permet de calculer leur probabilité à clé fixée assez simplement.

Proposition 6.2. *[DR07a] Soit S une fonction de $\mathbf{F}_2^m$ dans $\mathbf{F}_2^m$ d'uniformité différentielle $\Delta(S) \leqslant 4$. Alors pour toute différentielle (α, β) de $\mathbf{F}_2^m \times \mathbf{F}_2^m$, les ensembles $\mathcal{U}_{\alpha,\beta}(S)$ et $\mathcal{V}_{\alpha,\beta}(S)$ sont des sous-espaces affines de $\mathbf{F}_2^m$. De plus, si $\#\mathcal{U}_{\alpha,\beta}(S) = \#\mathcal{V}_{\alpha,\beta}(S) \neq 4$,*

alors l'espace vectoriel $V_{\alpha,\beta}(S)$ sous-jacent à $\mathcal{V}_{\alpha,\beta}(S)$ est entièrement déterminé par sa dimension et par la différence β et l'espace vectoriel $U_{\alpha,\beta}(S)$ sous-jacent à $\mathcal{U}_{\alpha,\beta}(S)$ est entièrement déterminé par sa dimension et par la différence α.

Dans toute la suite, les espaces affines seront notés par des lettres calligraphiques et les lettres droites seront réservées aux espaces vectoriels associés.

Démonstration. Soit (α, β) une différentielle pour la fonction S. Si $(\alpha, \beta) = (0, 0)$, alors

$$\mathcal{U}_{\alpha,\beta}(S) = \mathbf{F}_2^m \text{ et } \mathcal{V}_{\alpha,\beta}(S) = \mathbf{F}_2^m \ .$$

Donc ces deux ensembles sont des sous-espaces affines. Si l'uniformité de S est au plus quatre, alors les ensembles $\mathcal{U}_{\alpha,\beta}(S)$ et $\mathcal{V}_{\alpha,\beta}(S)$ contiennent 0, 2 ou 4 éléments lorsque $(\alpha, \beta) \neq (0, 0)$.

- si $\mathcal{U}_{\alpha,\beta}(S)$ (resp. $\mathcal{V}_{\alpha,\beta}(S)$) est vide, alors c'est un espace affine ;
- si $\mathcal{U}_{\alpha,\beta}(S)$ contient deux éléments, alors il existe $x \in \mathbf{F}_2^m$ tel que

  $$\mathcal{U}_{\alpha,\beta}(S) = \{x, x + \alpha\} = x + \langle\alpha\rangle$$

  qui est un espace affine (de même, $\mathcal{V}_{\alpha,\beta}(S) = S(x) + \langle\beta\rangle$) ;
- si $\mathcal{U}_{\alpha,\beta}(S)$ contient quatre éléments, alors il existe x et $y \in \mathbf{F}_2^m$ tels que

  $$\mathcal{U}_{\alpha,\beta}(S) = \{x, x + \alpha, y, y + \alpha\} = x + \langle\alpha, x + y\rangle$$

  qui est un espace affine (de même, $\mathcal{V}_{\alpha,\beta}(S) = S(x) + \langle\beta, S(x) + S(y)\rangle$).

Lorsque l'uniformité différentielle de S est égale à 2, le dernier cas ne se présente pas : l'ensemble $V_{\alpha,\beta}(S)$ ne dépend que de β et l'ensemble $U_{\alpha,\beta}(S)$ ne dépend que de α. □

Définition 6.3. *[DR07a] Soit S une fonction de $\mathbf{F}_2^m$ dans $\mathbf{F}_2^m$ d'uniformité différentielle $\Delta(S) \leqslant 4$. Une différentielle $(\alpha, \beta) \in \mathbf{F}_2^m \times \mathbf{F}_2^m$ est dite* double différentielle *si $\#\mathcal{U}_{\alpha,\beta}(S) = \#\mathcal{V}_{\alpha,\beta}(S) = 4$.*

Exemple 6.4. Considérons la boîte-S naïve, c'est-à-dire la fonction inverse sur $\mathbf{F}_{2^m}$. Cette fonction étant définie sur le corps $\mathbf{F}_{2^m}$, nous la noterons $\mathcal{S}$ dans cet exemple (nous obtenons une boîte-S sur $\mathbf{F}_2^m$ à partir de cette fonction en identifiant le corps $\mathbf{F}_{2^m}$ et l'espace vectoriel $\mathbf{F}_2^m$). Soit (α, β) une différentielle de $(\mathbf{F}_{2^m})^2$.

- Si $\alpha = \beta = 0$,

  $$\mathcal{U}_{\alpha,\beta}(\mathcal{S}) = \mathbf{F}_{2^m} \text{ et } \mathcal{V}_{\alpha,\beta}(\mathcal{S}) = \mathbf{F}_{2^m} \ ;$$

- Si $\alpha \neq 0$ et $\beta = \alpha^{-1}$,

  $$\mathcal{U}_{\alpha,\beta}(\mathcal{S}) = \langle\alpha, \alpha\tau\rangle \text{ et } \mathcal{V}_{\alpha,\beta}(\mathcal{S}) = \langle\beta, \beta\tau\rangle \ ,$$

  où τ est un élément de $\mathbf{F}_{2^m} \setminus \mathbf{F}_2$ tel que $\tau^3 = 1$;
- Si $\alpha \neq 0$, $\beta \neq \alpha^{-1}$ et $\mathrm{Tr}((\alpha\beta)^{-1}) = 0$, il existe (x, x') tels que

  $$\mathcal{U}_{\alpha,\beta}(\mathcal{S}) = \{x, x + \alpha\} \text{ et } \mathcal{V}_{\alpha,\beta}(\mathcal{S}) = \{x', x' + \beta\} \ .$$

- Dans les autres cas, $\mathcal{U}_{\alpha,\beta}(\mathcal{S}) = \mathcal{V}_{\alpha,\beta}(\mathcal{S}) = \emptyset$.

Lorsque $S' = A \circ S$ où S est la permutation de $\mathbf{F}_2^m$ correspondant à la boîte-S naïve et A est une permutation affine $\mathbf{F}_2^m$, par exemple lorsque S' est la boîte-S de l'AES, nous avons pour tout $(\alpha, \beta) \in (\mathbf{F}_2^m)^2$:

$$S'(x+\alpha) + S'(x) = \beta \Leftrightarrow S(x+\alpha) + S(x) = A^{-1}(\beta) \ .$$

Donc l'ensemble des entrées valides $\mathcal{U}_{\alpha,\beta}(S')$ et l'ensemble des images $\mathcal{V}_{\alpha,\beta}(S')$ pour une boîte-S affinement équivalente à la boîte-S naïve présentent des cas similaires à ceux de la boîte-S naïve (exemple 6.4).

Comme un produit cartésien d'espaces vectoriels est un espace vectoriel, pour toute fonction Sub constituée de t applications en parallèle d'une fonction d'uniformité différentielle inférieure ou égale à 4, les ensembles $\mathcal{U}_{a,b}(\mathsf{Sub})$ et $\mathcal{V}_{a,b}(\mathsf{Sub})$ sont des sous-espaces affines de $\mathbf{F}_2^{mt}$ pour toute différentielle (a, b) de $\mathbf{F}_2^{mt} \times \mathbf{F}_2^{mt}$.

Pour calculer la probabilité de la caractéristique sur deux tours, il faut compter le nombre de textes clairs qui suivent la caractéristique.

Notation 6.5. Soit (a, c, b) une caractéristique différentielle sur deux tours d'un réseau de substitution-permutation de la forme $\mathrm{SPN}(m, t, S, M)$, a, b et c étant des éléments de $\mathbf{F}_2^{mt}$. Notons $\mathcal{X}_k(a, c, b)$ l'ensemble des entrées valides de la caractéristique (a, c, b) pour la clé $k \in \mathbf{F}_2^{\kappa}$, c'est-à-dire l'ensemble des éléments x de $\mathbf{F}_2^{mt}$ tels que les textes clairs x et $x+a$ conduisent à deux textes ayant une différence égale à c après le premier étage de boîtes-S et à des textes chiffrés ayant une différence égale à b après le deuxième étage de boîtes-S :

$$\begin{aligned}\mathcal{X}_k(a, c, b) = \{x \in \mathbf{F}_2^{mt} \,|\, &\mathsf{Sub}(x+a) + \mathsf{Sub}(x) = c \\ &\text{et } \mathsf{Sub}(M(\mathsf{Sub}(x+a) + k)) + \mathsf{Sub}(M(\mathsf{Sub}(x) + k)) = b\} \ ,\end{aligned}$$

où Sub est un étage de boîtes-S.

La définition suivante présente un type de caractéristiques dont la probabilité ne prend qu'une valeur non nulle lorsque la clé varie.

Définition 6.6. *[DR07a] Une caractéristique Q est dite plateau de hauteur $h(Q)$ si pour toute clé $k \in \mathbf{F}_2^{\kappa}$, la probabilité de la caractéristique Q est nulle ou égale à $2^{h(Q)-mt}$.*

Ainsi, calculer la distribution sur les clés de la probabilité d'une telle caractéristique est simplifiée : il n'y a qu'une probabilité non nulle à déterminer. Or les caractéristiques sur deux tours de l'AES, entre autres, sont plateau.

Théorème 6.7. *[DR07a, Théorème 21] Soit $Q = (a, c, b)$ une caractéristique sur deux tours d'un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$ tel que l'uniformité différentielle de S est inférieure ou égale à 4. Alors Q est une caractéristique plateau de hauteur*

$$h(Q) = \dim(U_{M(c),b}(\mathsf{Sub}) \cap M(V_{a,c}(\mathsf{Sub}))).$$

Démonstration. Pour une clé k, il existe un élément $x \in \mathcal{X}_k(a,c,b)$ si et seulement si il existe un élément x' de $\mathcal{V}_{a,c}(\mathsf{Sub})$ et que l'élément $z = M(x') + k$ appartient à $\mathcal{U}_{M(c),b}(\mathsf{Sub})$. Donc le nombre de textes clairs qui suivent la caractéristique $Q = (a,c,b)$ est égal au cardinal de l'ensemble

$$I_k = \mathcal{U}_{M(c),b}(\mathsf{Sub}) \cap (k + M(\mathcal{V}_{a,c}(\mathsf{Sub}))).$$

Pour les clés k telles que l'ensemble I_k est vide, la probabilité de la caractéristique Q est nulle. Supposons maintenant que I_k soit non vide pour une certaine clé k. Montrons que l'ensemble I_k est un sous-espace affine, c'est-à-dire que pour tout élément $w \in I_k$, $w + I_k$ est un sous-espace vectoriel. Soit w un élément de I_k, alors w est un élément de $\mathcal{U}_{M(c),b}(\mathsf{Sub})$, donc

$$\mathcal{U}_{M(c),b}(\mathsf{Sub}) = w + U_{M(c),b}(\mathsf{Sub}).$$

Et w est aussi un élément de $k + M(\mathcal{V}_{a,c}(\mathsf{Sub}))$, donc $w + k \in M(\mathcal{V}_{a,c}(\mathsf{Sub}))$ et

$$M(\mathcal{V}_{a,c}(\mathsf{Sub})) = w + k + M(V_{a,c}(\mathsf{Sub})).$$

Nous en déduisons que $I_k = (w + U_{M(c),b}(\mathsf{Sub})) \cap (w + M(V_{a,c}(\mathsf{Sub})))$, et par translation,

$$w + I_k = U_{M(c),b}(\mathsf{Sub}) \cap M(V_{a,c}(\mathsf{Sub})).$$

Or $U_{M(c),b}(\mathsf{Sub})$ et $M(V_{a,c}(\mathsf{Sub}))$ sont des espaces vectoriels car $\Delta(S) \leqslant 4$ et M est une fonction linéaire sur $\mathbf{F}_2$.

Donc lorsque la probabilité de la caractéristique Q est non nulle, elle est égale à :

$$\begin{aligned} \mathrm{DCP}_2^{E_k}(Q) &= 2^{-mt} \#\mathcal{X}_k(a,c,b) \\ &= 2^{-mt} \#(U_{M(c),b}(\mathsf{Sub}) \cap M(V_{a,c}(\mathsf{Sub}))) \\ &= 2^{-mt} \times 2^{\dim(U_{M(c),b}(\mathsf{Sub}) \cap M(V_{a,c}(\mathsf{Sub})))}. \end{aligned}$$

□

Ce théorème permet de déterminer les probabilités de toutes les caractéristiques sur deux tours d'un tel réseau de substitution-permutation en fonction de a, c et b pour toutes les clés.

En particulier, il s'applique aux probabilités des caractéristiques sur deux tours de l'AES. La plus grande probabilité possible pour une caractéristique est $2^{-27} = 128 \times 2^{-34}$ pour une partie des clés [DR07a]. Donc la différentielle qui contient cette caractéristique a pour probabilité au minimum 128×2^{-34} sur une partie des clés. Or le MEDP_2 de l'AES est égal à 53×2^{-34}. Donc il existe une différentielle dont la probabilité pour certaines clés est presque trois fois supérieure au MEDP_2. Cette différentielle contient peut-être d'autres caractéristiques, donc sa probabilité pourrait être encore plus grande. Dans ce cas, la différence entre la probabilité à clé fixée de certaines différentielles sur deux tours de l'AES et la valeur du MEDP_2 pourrait être relativement grande. Nous nous intéressons donc à la distribution de la probabilité à clé fixée d'une différentielle sur deux tours d'un réseau de substitution-permutation afin de déterminer à quel point l'espérance de la probabilité d'une différentielle est une bonne estimation. Pour cela, nous allons caractériser l'ensemble des clés pour lesquelles une caractéristique a une probabilité non nulle.

6.2 Ensemble des clés définissant un chemin différentiel

La probabilité d'une différentielle (a, b) est égale à la somme des probabilités des caractéristiques $Q = (a, c, b)$ dans cette différentielle :

$$\mathrm{DP}_2^{E_k}(a, b) = \sum_{\substack{c \in \mathcal{C}_M : \\ \mathsf{Supp}(c) = \mathsf{Supp}(a,b)}} \mathrm{DCP}_2^{E_k}(a, c, b),$$

où $\mathcal{C}_M$ est le code associé à la fonction de diffusion M. Pour une clé k fixée, la probabilité d'une différentielle (a, b) dépend fortement du nombre de caractéristiques dans la différentielle ayant une probabilité non nulle.

Notation 6.8. Pour une caractéristique différentielle (a, c, b) sur deux tours d'un réseau de substitution-permutation de la forme $\mathrm{SPN}(m, t, S, M)$, notons $\mathcal{K}(a, c, b)$ l'ensemble des clés k pour lesquelles la caractéristique (a, c, b) a une probabilité non nulle, c'est-à-dire pour lesquelles l'ensemble $\mathcal{X}_k(a, c, b)$ est non vide.

Si deux caractéristiques (a, c_1, b) et (a, c_2, b) sont de probabilité non nulle pour une clé k, alors k appartient à l'intersection des ensembles $\mathcal{K}(a, c_1, b)$ et $\mathcal{K}(a, c_2, b)$. Dans cette section, nous étudions les ensembles de clés pour lesquelles une caractéristique est de probabilité non nulle pour déterminer à quelle condition deux caractéristiques dans une même différentielle sont de probabilité non nulle pour une même clé. Lorsque l'uniformité différentielle de la boîte-S est inférieure ou égale à 4, l'ensemble des entrées valides et l'ensemble des images par S de ces entrées valides sont des espaces affines. Nous déduisons que les ensembles de clés ont aussi une structure d'espace affine.

Proposition 6.9. *Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(m, t, S, M)$ dont la boîte-S a une uniformité différentielle inférieure ou égale à 4. Soit (a, c, b) une caractéristique différentielle sur deux tours de $(E_k)_k$. Alors $\mathcal{K}(a, c, b)$ est l'espace affine sur $\mathbf{F}_2$ défini par*

$$M(\mathcal{V}_{a,c}(\mathsf{Sub})) + \mathcal{U}_{M(c),b}(\mathsf{Sub}) \ .$$

Démonstration. Pour une clé k, il existe un élément $x \in \mathcal{X}_k(a, c, b)$ si et seulement si il existe un élément x' de $\mathcal{V}_{a,c}(\mathsf{Sub})$ et que l'élément z défini par $z = M(x') + k$ appartient à $\mathcal{U}_{M(c),b}(\mathsf{Sub})$. Nous en déduisons que les clés k pour lesquelles $\mathcal{X}_k(a, c, b)$ n'est pas vide sont exactement les clés qui s'écrivent

$$k = M(x') + z$$

avec $x' \in \mathcal{V}_{a,c}(\mathsf{Sub})$ et $z \in \mathcal{U}_{M(c),b}(\mathsf{Sub})$. Comme $M(V_{a,c}(\mathsf{Sub}))$ et $U_{M(c),b}(\mathsf{Sub})$ sont des espaces vectoriels, l'ensemble $\mathcal{K}(a, c, b)$ est un espace affine dont l'espace vectoriel sous-jacent est la somme directe $M(V_{a,c}(\mathsf{Sub})) + U_{M(c),b}(\mathsf{Sub})$. □

De plus, l'espace vectoriel sous-jacent de l'espace affine $\mathcal{K}(a, c, b)$ peut se décrire en fonction des espaces vectoriels correspondant aux espaces affines $\mathcal{U}_{M(c),b}(\mathsf{Sub})$ et $\mathcal{V}_{a,c}(\mathsf{Sub})$. Nous utilisons la notation suivante pour définir les matrices génératrices de ces espaces vectoriels.

Notation 6.10. Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}(m,t,S,M)$, où S est une fonction de $\mathbf{F}_2^m$ à valeurs dans $\mathbf{F}_2^m$ d'uniformité différentielle $\Delta(S) \leqslant 4$. Soit $(a,b) = (a_1,\ldots,a_t,b_1,\ldots,b_t) \in \mathbf{F}_2^{mt} \times \mathbf{F}_2^{mt}$ une différentielle sur deux tours de $(E_k)_k$. Pour tout $c = (c_1,\ldots,c_t) \in \mathbf{F}_2^{mt}$, les ensembles $V_{a_1,c_1}(S) \times \cdots \times V_{a_t,c_t}(S)$ et $U_{M(c)_1,b_1}(S) \times \cdots \times U_{M(c)_t,b_t}(S)$ sont les espaces vectoriels sous-jacents de $\mathcal{V}_{a,c}(\mathsf{Sub})$ et $\mathcal{U}_{M(c),b}(\mathsf{Sub})$. Pour tout i compris entre 1 et t, notons A_i une matrice génératrice de l'espace vectoriel $V_{a_i,c_i}(S)$ et B_i une matrice génératrice de l'espace vectoriel $U_{M(c)_i,b_i}(S)$.

La proposition suivante donne la forme d'une matrice génératrice de l'espace vectoriel sous-jacent de l'espace affine $\mathcal{K}(a,c,b)$.

Proposition 6.11. *Soit (a,c,b) une caractéristique différentielle sur deux tours d'un chiffrement de la forme $\mathrm{SPN}(m,t,S,M)$ dont la boîte-S a une uniformité différentielle inférieure ou égale à 4. Alors l'espace affine $\mathcal{K}(a,c,b)$ a pour direction l'espace vectoriel $K(a,c,b)$ généré par les vecteurs binaires de taille mt correspondant aux lignes de la matrice*

$$G_{a,c,b} = \mathcal{D} \times \begin{pmatrix} M \\ \mathrm{Id}_t \end{pmatrix} ,$$

où M et Id_t sont des matrices dont les colonnes sont des vecteurs de $\mathbf{F}_2^m$ et $\mathcal{D}$ est la matrice bloc avec $2mt$ colonnes définie par

$$\mathcal{D} = \begin{bmatrix} A_1 & 0 & & & & 0 \\ 0 & \ddots & & & & \\ & & A_t & & & \\ & & & B_1 & & \\ & & & & \ddots & 0 \\ 0 & & & & & B_t \end{bmatrix}$$

avec A_i et B_i définis selon la notation 6.10. En particulier, si la caractéristique ne comporte pas de double différentielle, alors $K(a,c,b)$ dépend uniquement de c.

Démonstration. D'après la proposition 6.11, le sous-espace affine $\mathcal{K}(a,c,b)$ correspond à l'espace $M(\mathcal{V}_{a,c}(\mathsf{Sub})) + \mathcal{U}_{M(c),b}(\mathsf{Sub})$. Le sous-espace linéaire sous-jacent est donc la somme directe de $M_c(V_{a,c}(\mathsf{Sub}))$ et $U_{M(c),b}(\mathsf{Sub})$.

Si la caractéristique ne comporte pas de double différentielle, d'après la démonstration de la proposition 6.2, l'espace vectoriel $V_{a_i,c_i}(S)$ est engendré par c_i pour tout $i \in \{1,\ldots,t\}$ et l'espace vectoriel $U_{(M(c))_j,b_j}(S)$ est engendré par la j-ème coordonnée de $M(c)$ pour tout $j \in \{1,\ldots,t\}$. □

Exemple 6.12. Considérons deux tours du chiffrement de la forme $\mathrm{SPN}(8,4,S,M)$ où M est la fonction de diffusion de l'AES et S est la fonction de $\mathbf{F}_2^8$ qui correspond à la fonction inverse $\mathcal{S}$ de F_{2^8}, *i.e.* $S = \varphi^{-1} \circ \mathcal{S} \circ \varphi$, φ étant l'isomorphisme de l'AES qui identifie l'espace vectoriel $\mathbf{F}_2^8$ au corps $\mathbf{F}_{2^8}$:

$$\begin{array}{rccc} \varphi : & \mathbf{F}_2^8 & \longrightarrow & \mathbf{F}_{2^8} \\ & (x_0,\ldots,x_7) & \longmapsto & \sum_{i=0}^{7} x_i X^i \end{array}$$

où les opérations sont effectuées modulo le polynôme irréductible $X^8+X^4+X^3+X+1$. Ce chiffrement correspond à la superboîte-S de l'AES où la boîte-S a été remplacée par la boîte-S naïve et est considérée sur l'espace vectoriel $\mathbf{F}_2^8$.

Choisissons $c = (\texttt{0e}, \texttt{09}, \texttt{0d}, \texttt{0b})$, où chaque coordonnée en notation hexadécimale correspond à un mot de 8 bits. Alors $M(c) = (\texttt{01}, \texttt{0}, \texttt{0}, \texttt{0})$. Soit $(a, b) \in (\mathbf{F}_2^8)^2$ une différentielle sur deux tours telle que $b_1 \neq 1$ et, pour tout $1 \leqslant i \leqslant 4$, $\varphi(a_i) \neq \varphi(c_i)^{-1}$, c'est-à-dire qu'il n'y a pas de doubles différentielles (définition 6.3) dans la caractéristique (a, c, b). D'après la proposition 6.2, les ensembles $V_{a_i,c_i}(S)$, $1 \leqslant i \leqslant 4$, sont les espaces vectoriels de dimension 1 générés par c_i, l'ensemble $U_{(M(c))_1,b_1}(S)$ est engendré par le vecteur $\texttt{01} = (0,0,0,0,0,0,0,1)$ et les ensembles $U_{(M(c))_i,b_i}(S)$, $2 \leqslant i \leqslant 4$, sont égaux à $\mathbf{F}_2^8$. Donc les matrices A_i, $1 \leqslant i \leqslant 4$, sont composées du vecteur c_i, $B_1 = (10000000)$ et les matrices B_i, $2 \leqslant i \leqslant 4$, sont égales à la matrice identité Id_8.

D'après la proposition 6.11, l'espace vectoriel sous-jacent de $\mathcal{K}(a, c, b)$ est l'espace généré par la matrice

$$G_{a,c,b} = \begin{bmatrix} G_1 & G_2 \\ 0 & \mathrm{Id}_{24} \end{bmatrix},$$

avec

$$[G_1 G_2] = \begin{bmatrix} 00111000011100000111000001001000 \\ 11011000010010001001000010010000 \\ 10110000111010000101100010110000 \\ 11010000110100001011100001101000 \\ 10000000000000000000000000000000 \end{bmatrix}$$

Nous pouvons écrire $G_{a,c,b}$ sous forme systématique, c'est-à-dire en appliquant une permutation π aux colonnes de la matrice $G_{a,c,b}$:

$$G'_{a,c,b} = \begin{bmatrix} \mathrm{Id}_{28} & Z \\ 0 & 0 \end{bmatrix}.$$

Le sous-espace vectoriel généré par les lignes de $G_{a,c,b}$ est donc de dimension 28. La matrice de parité correspondante est

$$H_{a,c,b} = \begin{bmatrix} {}^tZ & \mathrm{Id}_4 \end{bmatrix},$$

c'est-à-dire qu'un élément $k = (k_0, k_1, k_2, k_3, \ldots, k_{30}, k_{31}) \in \mathbf{F}_2^{32}$ appartient à $K(a, c, b)$ si et seulement si

$${}^t(k_{\pi(0)}, k_{\pi(1)}, \ldots, k_{\pi(31)})\, H = 0,$$

où π est la permutation qui a été appliquée aux colonnes de $G_{a,c,b}$.

Ici, nous avons $k_\pi = (k_0, k_1, k_2, k_{31}, k_4, k_{30}, k_{29}, k_{28}, k_8, \ldots, k_{26}, k_{27}, k_7, k_6, k_5)$ et

$${}^tZ = \begin{bmatrix} 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 1 & 0 & \ldots & 0 \end{bmatrix}.$$

Nous obtenons ainsi quatre relations linéaires caractérisant les clés de $\mathcal{K}(a, c, b)$: l'espace vectoriel sous-jacent de $\mathcal{K}(a, c, b)$ est défini par

$$K(a, c, b) = \{(k_0\ k_1\ k_2\ (k_1 + k_2)\ k_4\ 0\ 0\ 0 || K'), k_0 k_1 k_2 k_4 \in \mathbf{F}_2^4, \text{ et } K' \in \mathbf{F}_2^{24}\}.$$

En utilisant l'identification de $\mathbf{F}_2^8$ au corps $\mathbf{F}_{2^8}$ définie par l'isomorphisme φ, l'espace vectoriel $K(a,c,b)$ peut aussi être décrit par

$$K(a,c,b) = \left\{ x \in \mathbf{F}_{2^8}^4 \;\middle|\; \sum_{i=0}^{3} \mathrm{Tr}(\lambda_i x_i) = 0 \right\}$$

avec

$$\lambda \in \langle (\mathtt{ed},\mathtt{0},\mathtt{0},\mathtt{0}), (\mathtt{53},\mathtt{0},\mathtt{0},\mathtt{0}), (\mathtt{a4},\mathtt{0},\mathtt{0},\mathtt{0}), (\mathtt{52},\mathtt{0},\mathtt{0},\mathtt{0}) \rangle \ .$$

Exemple 6.13. Dans le même contexte que précédemment, choisissons $c = (\mathtt{0},\mathtt{01},\mathtt{04},\mathtt{07})$. Nous avons $M(c) = (\mathtt{0},\mathtt{09},\mathtt{0},\mathtt{0b})$. Soit $(a,b) \in (\mathbf{F}_2^8)^2$ une différentielle sur deux tours telle qu'il n'y a pas de doubles différentielles dans la caractéristique (a,c,b). Alors l'espace vectoriel généré par les 29 lignes de la matrice $G_{a,c,b}$ est de dimension 27. Nous obtenons que l'espace vectoriel sous-jacent à $\mathcal{K}(a,c,b)$ contient les éléments $k = (k_0,k_1,k_2,k_3,\ldots,k_{30},k_{31}) \in \mathbf{F}_2^{32}$ vérifiant :

$$\begin{aligned} k_{25} &= k_9 + k_{24}, \\ k_{27} &= k_8 + k_9 + k_{11} + k_{24} + k_{26}, \\ k_{29} &= k_{12} + k_{13}, \\ k_{30} &= k_{13} + k_{14}, \\ k_{31} &= k_{14} + k_{15} \ . \end{aligned}$$

Le corollaire suivant donne la dimension de l'espace $\mathcal{K}(a,c,b)$ en fonction de la caractéristique (a,c,b).

Corollaire 6.14. *Pour toute clé $k \in \mathcal{K}(a,c,b)$, le nombre d'entrées valides de la caractéristique (a,c,b) est égal à 2^v, où v est la dimension du noyau de $G_{a,c,b}$. En particulier, nous avons*

$$v \leq \min\left(\dim V_{a,c}(\mathsf{Sub}), \dim U_{M(c),b}(\mathsf{Sub})\right) \ .$$

De plus, nous avons

$$\mathrm{codim}\mathcal{K}(a,c,b) - v = (m-1)s + mt - s_2$$

où $s = wt(a) + wt(b)$ est le nombre de boîtes-S actives et s_2 le nombre de doubles différentielles.

Démonstration. Nous avons déjà vu dans le théorème 6.7 que $v = \dim(U_{M(c),b}(\mathsf{Sub}) \cap M(V_{a,c}(\mathsf{Sub})))$. Donc $v \leq \min\left(\dim V_{a,c}(\mathsf{Sub}), \dim U_{M(c),b}(\mathsf{Sub})\right)$.

De plus, le cardinal de l'intersection $\mathcal{U}_{M(c),b}(\mathsf{Sub}) \cap (k + M(\mathcal{V}_{a,c}(\mathsf{Sub})))$ correspond au nombre de combinaisons linéaires non nulles des lignes de la matrice $G_{a,c,b}$. Puisque le nombre de lignes de $G_{a,c,b}$ est donné par

$$m(2n-s) + s + s_2 \ ,$$

la dimension de $\mathcal{K}(a,c,b)$ est donc égale à

$$m(2n-s) + s + s_2 - v$$

ce qui implique que

$$\mathrm{codim}\mathcal{K}(a,c,b) = mn - m(2n-s) - s - s_2 + v = (m-1)s - mn - s_2 + v \ .$$

□

Le tableau suivant contient une description de tous les sous-espaces vectoriels $K(a,c,b)$ pour un chiffrement $\mathrm{SPN}(8,t,S,M)$ où la boîte-S est celle présentée dans l'exemple 6.12 (la boîte-S naïve) et la fonction de diffusion est celle de l'AES et lorsque la caractéristique (a,c,b) contient le nombre minimum de boîtes-S actives (5) et ne contient pas de double différentielle. Ce tableau utilise la description de $K(a,c,b)$ sur le corps $\mathbf{F}_{2^8}$ décrite dans l'exemple 6.12 :

$$K(a,c,b) = \left\{ x \in \mathbf{F}_{2^8}^4 \ \middle| \ \sum_{i=0}^{3} \mathrm{Tr}(\lambda_i x_i) = 0 \right\}$$

avec

$$\lambda \in \Lambda_{a,c,b} \ .$$

$(c, M(c))$	codim$\mathcal{K}$	générateurs de $\Lambda_{a,c,b}$				
(1,0,0,0,2,1,1,3)	4	(71,5,5d,c4)	(51,50,a7,f4)	(d3,a5,8,37)	(8c,fe,a2,ca)	
(2,1,0,0,7,0,3,7)	4	(57,0,9c,74)	(a5,0,f6,52)	(51,0,f7,a4)	(a2,0,f5,53)	
(0,1,1,0,2,1,3,0)	4	(6,9,5,0)	(a6,a2,f5,0)	(f5,f3,2,0)	(51,56,f4,0)	
(0,0,1,3,2,0,7,7)	4	(74,0,a6,3a)	(a4,0,a5,52)	(53,0,51,a4)	(a6,0,a2,53)	
(2,0,0,3,7,1,7,0)	5	(5,5,5,0)	(a6,a6,a6,0)	(53,53,53,0)	(a4,a4,a4,0)	(52,52,52,0)
(2,0,1,0,5,1,0,7)	5	(74,b5,0,ed)	(1,a5,0,52)	(2,51,0,a4)	(4,a2,0,53)	(a6,13,0,d7)
(0,1,0,3,0,1,4,7)	5	(0,e2,52,99)	(0,a5,53,52)	(0,51,a6,a4)	(0,a2,57,53)	(0,57,cb,74)
(0,1,4,7,0,9,0,b)	5	(0,b0,0,99)	(0,f6,0,52)	(0,f7,0,a4)	(0,f5,0,53)	(0,9c,0,74)
(5,1,0,7,e,0,d,0)	5	(c1,0,ed,0)	(b5,0,d7,0)	(a6,0,53,0)	(53,0,a4,0)	(a4,0,52,0)
(7,1,3,0,e,0,0,b)	4	(e8,0,0,74)	(a4,0,0,52)	(53,0,0,a4)	(a6,0,0,53)	
(2,0,3,7,0,0,d,b)	4	(0,0,a6,a6)	(0,0,52,52)	(0,0,a4,a4)	(0,0,53,53)	
(2,1,7,0,0,9,d,0)	4	(0,4e,3a,0)	(0,f5,53,0)	(0,f7,a4,0)	(0,f6,52,0)	
(7,0,7,7,e,9,0,0)	5	(e8,9c,0,0)	(74,4e,0,0)	(d2,bb,0,0)	(81,4c,0,0)	(25,ba,0,0)
(e,9,d,b,1,0,0,0)	4	(ed,0,0,0)	(53,0,0,0)	(a4,0,0,0)	(52,0,0,0)	

Un premier résultat reliant les espaces des clés de deux caractéristiques est le suivant. Il est valide lorsque la fonction de diffusion est linéaire sur $\mathbf{F}_{2^m}$ et concerne les caractéristiques qui ne diffèrent que d'une constante multiplicative. Nous utilisons donc à nouveau la notation calligraphique pour les fonctions définies sur le corps $\mathbf{F}_{2^m}$ et des indices pour indiquer les quantités qui sont considérées sur ce corps.

Corollaire 6.15. *Soit $(E_k)_k$ un chiffrement de la forme $\mathrm{SPN}_F(m,t,\mathcal{S},\mathcal{M})$, où $\mathcal{S}$ est une fonction de $\mathbf{F}_{2^m}$ dans $\mathbf{F}_{2^m}$ d'uniformité différentielle $\Delta(\mathcal{S}) \leqslant 4$. Soit γ un élément non nul de $\mathbf{F}_{2^m}$. Soient $Q = (a,c,b)$ et $\gamma Q = (\gamma a, \gamma c, \gamma b)$ deux caractéristiques qui ne contiennent pas de double différentielle. Alors, $K(Q)$ et $K(\gamma Q)$ ont la même dimension. De plus, en notant*

$$\Lambda_Q = K(Q)^{\perp},$$

i.e.,

$$\Lambda_Q = \left\{ \lambda \in \mathbf{F}_{2^m}^t \;\middle|\; \sum_{i=0}^{t-1} \mathrm{Tr}(\lambda_i x_i) = 0, \forall x \in K(Q) \right\},$$

nous avons

$$\Lambda_{\gamma Q} = \{\gamma^{-1}\lambda, \lambda \in \Lambda_Q\} .$$

Remarquons que, dans ce résultat, l'inverse γ^{-1} ne dépend pas du choix de la boîte-S, en particulier de l'utilisation de la fonction inverse comme boîte-S.

Démonstration. Si $Q = (a, c, b)$ ne contient pas de double différentielle, les blocs de la matrice $\mathcal{D}$ définie dans la proposition 6.11 correspondent soit à la matrice génératrice de $\mathbf{F}_{2^m}$, soit à un unique mot qui est une coordonnée du mot de code $(c, \mathcal{M}(c))$. Donc, lorsque Q est multiplié par γ, tous les blocs de $\mathcal{D}$ sont multipliés par γ. De plus, les mots $x \in K(\gamma Q)$ sont donnés par

$$\gamma x' \text{ avec } x' \in K(Q) .$$

Alors,

$$\sum_{i=0}^{n-1} \mathrm{Tr}(\lambda_i x'_i) = 0 \quad \text{si et seulement si} \quad \sum_{i=0}^{n-1} \mathrm{Tr}((\lambda_i \gamma^{-1})(\gamma x'_i)) = 0 .$$

Nous en déduisons que

$$\Lambda_{\gamma Q} = \{\gamma^{-1}\lambda, \lambda \in \Lambda_Q\} .$$

□

Exemple 6.16. Reprenons l'exemple 6.12. Si nous multiplions $c = (\mathtt{0e}, \mathtt{09}, \mathtt{0d}, \mathtt{0b})$ par un élément non nul $\gamma \in \mathbf{F}_{2^8}$, nous obtenons (pour une différentielle (a, b) telle qu'il n'y a pas de double différentielle dans les caractéristiques (a, c, b) et $(a, \gamma c, b)$) que $\mathcal{K}(a, \gamma c, b)$ est un sous-espace vectoriel de codimension 4, mais les relations définissant les éléments k qui appartiennent à $K(a, \gamma c, b)$ sont différentes de celles définissant les éléments de $K(a, c, b)$. Par exemple, prenons $\gamma = \mathtt{02}$. Alors l'espace vectoriel sous-jacent de $\mathcal{K}(a, \gamma c, b)$ est défini par

$$K(a, \gamma c, b) = \{(0\ k_1\ k_2\ k_3\ (k_2 + k_3)\ k_5\ 0\ 0||K'), k_1k_2k_3k_5 \in \mathbf{F}_2^4,\ \text{et}\ K' \in \mathbf{F}_2^{24}\} .$$

Les éléments de $K(a, \gamma c, b)$ sont donc ceux de $K(a, c, b)$ multipliés par $\gamma = \mathtt{02}$. De même, tous les coefficients des éléments λ de $\mathbf{F}_{2^8}$ correspondant sont multipliés par γ^{-1}. Les relations définissant $K(a, \gamma c, b)$ correspondent dans le corps $\mathbf{F}_{2^8}$ à

$$\lambda \in \langle(\mathtt{fb}, \mathtt{0}, \mathtt{0}, \mathtt{0}), (\mathtt{a4}, \mathtt{0}, \mathtt{0}, \mathtt{0}), (\mathtt{52}, \mathtt{0}, \mathtt{0}, \mathtt{0}), (\mathtt{29}, \mathtt{0}, \mathtt{0}, \mathtt{0})\rangle .$$

Remarque 6.17. Lorsque que les caractéristiques (a, c, b) et $(\gamma a, c, \gamma b)$, $\gamma \in \mathbf{F}_{2^m}^*$, n'ont pas de doubles différentielles, alors la matrice $\mathcal{D}$ est identique pour ces deux caractéristiques. Donc les ensembles $\mathcal{K}(a, c, b)$ et $\mathcal{K}(\gamma a, c, \gamma b)$ (respectivement $K(a, c, b)$ et $K(\gamma a, c, \gamma b)$) sont égaux. C'est pourquoi, dans l'exemple précédent, nous avons considéré les caractéristiques (a, c, b) et $(a, \gamma c, b)$.

Dans la suite de ces travaux, il nous faut étudier les intersections des espaces vectoriels $K(a, \gamma c, b)$, $\gamma \in \mathbf{F}_{2^m}^*$, afin de déterminer à quelle condition la différentielle (a, b) contient plusieurs chemins différentiels pour une clé k fixée. Il serait aussi intéressant de déterminer une caractérisation des ensembles $\mathcal{K}(a, \gamma c, b)$, en plus de celle que nous avons des espaces vectoriels sous-jacents de ces ensembles.

Conclusion

Dans une première partie de cette thèse, nous avons raffiné les critères classiques de résistance des réseaux de substitution-permutation aux attaques linéaires et différentielles. Nous avons présenté une nouvelle borne sur le MEDP (respectivement MELP) sur deux tours d'un chiffrement tel que la fonction de diffusion est linéaire sur $\mathbf{F}_{2^m}$, comme dans l'AES. Cette nouvelle borne ne dépend que de la boîte-S et du branch number de la fonction de diffusion, et n'est plus invariante par composition avec des permutations affines contrairement aux résultats précédents. De plus, pour toute boîte-S et toute valeur t, nous avons montré qu'il existe toujours au moins une permutation linéaire de $(\mathbf{F}_{2^m})^t$ de branch number maximal pour laquelle le MEDP_2 (respectivement MELP_2) dépasse une certaine quantité.

Nous avons aussi introduit une nouvelle propriété des boîtes-S qui simplifie le calcul de la nouvelle borne, vérifiée par exemple par les fonctions puissance. Nous avons montré que, lorsque S et S^{-1} vérifient cette propriété, notre borne inférieure est satisfaite pour toute fonction M de branch number maximal. En conséquence, si S est l'inversion dans $\mathbf{F}_{2^m}$ par exemple, alors la valeur exacte de MEDP_2 (respectivement MELP_2) est toujours la plus grande possible parmi toutes les fonctions de la même classe d'équivalence.

Il serait intéressant de pouvoir utiliser cette nouvelle borne pour améliorer la borne supérieure sur le MEDP (respectivement le MELP) sur quatre tours d'un réseau de substitution-permutation. Cependant, il faudrait pour cela déterminer la distribution de l'espérance $\mathrm{EDP}_2(a, b)$ de la probabilité d'une différentielle (a, b) sur deux tours (respectivement la distribution du potentiel linéaire $\mathrm{ELP}_2(u, v)$ d'un masque (u, v) sur deux tours). Or, seule une partie de cette distribution a été déterminée dans le cas différentiel pour la variante de l'AES où la boîte-S est remplacée par la boîte-S naïve, *i.e.* la fonction inverse dans $\mathbf{F}_{2^8}$.

D'autre part, nous avons démontré que, contrairement à l'idée communément admise, le MEDP_2 n'est pas toujours atteint par une différentielle ayant le plus petit nombre possible de boîtes-S actives. En effet, nous avons présenté les premiers exemples de réseaux de substitution-permutation pour lesquels le MEDP_2 est atteint par une différentielle dont le nombre de boîtes-S actives est supérieur au branch number de M.

Cependant, en pratique, le MEDP_2 est atteint par une différentielle de plus petit poids possible pour un grand nombre de chiffrements. Nous avons démontré cette observation pour certaines familles de boîtes-S lorsque la fonction de diffusion M a un branch number maximal. Existe-t-il d'autres familles de boîtes-S pour lesquelles nous pouvons

démontrer que, lorsque la fonction de diffusion M a un branch number maximal, le MEDP_2 est atteint par une différentielle de poids minimal ?

Un autre prolongement intéressant serait de démontrer qu'à partir d'une certaine valeur, toute différentielle sur deux tours de poids supérieur à cette valeur ne peut pas atteindre le MEDP_2. Il suffirait alors de calculer la probabilité des différentielles de poids inférieur à cette valeur pour déterminer le MEDP_2. En particulier, cela pourrait diminuer la complexité de l'algorithme de Keliher et Sui [KS07] pour calculer la valeur exacte du MEDP_2.

Enfin, dans le dernier chapitre, nous avons présenté des résultats de travaux en cours sur le calcul de la probabilité d'une différentielle sur deux tours d'un réseau de substitution-permutation lorsque la clé est fixée. En particulier, nous avons montré que sous certaines conditions (vérifiées par l'AES par exemple), l'ensemble des clés telles qu'une caractéristique a une probabilité non nulle est un espace affine et nous avons déterminé l'espace vectoriel sous-jacent. Il nous faut maintenant caractériser ces espaces affines et utiliser ces résultats pour déterminer la probabilité d'une différentielle à clé fixée.

Bibliographie

[AÅBL12] Mohamed Ahmed Abdelraheem, Martin Ågren, Peter Beelen, and Gregor Leander. On the Distribution of Linear Biases : Three Instructive Examples. In *Advances in Cryptology - CRYPTO 2012*, volume 7417 of *LNCS*, pages 50–67. Springer, 2012.

[ADK+14] Martin R. Albrecht, Benedikt Driessen, Elif Bilge Kavun, Gregor Leander, Christof Paar, and Tolga Yalçin. Block Ciphers - Focus on the Linear Layer (feat. PRIDE). In *Advances in Cryptology - CRYPTO 2014*, volume 8616 of *LNCS*, pages 57–76. Springer, 2014.

[Bar] Paulo S.L.M. Barreto. Implementation of the SQUARE block cipher. http://www.larc.usp.br/~pbarreto/sqjava21.zip.

[BB02] Elad Barkan and Eli Biham. In How Many Ways Can You Write Rijndael ? In *Advances in Cryptology - ASIACRYPT 2002*, volume 2501 of *LNCS*, pages 160–175. Springer, 2002.

[BBL13] Céline Blondeau, Andrey Bogdanov, and Gregor Leander. Bounds in Shallows and in Miseries. In *Advances in Cryptology - CRYPTO 2013, part I*, volume 8042 of *LNCS*, pages 204–221. Springer, 2013.

[BCG+12a] Julia Borghoff, Anne Canteaut, Tim Güneysu, Elif Bilge Kavun, Miroslav Knezevic, Lars R. Knudsen, Gregor Leander, Ventzislav Nikov, Christof Paar, Christian Rechberger, Peter Rombouts, Søren S. Thomsen, and Tolga Yalçin. PRINCE - A Low-latency Block Cipher for Pervasive Computing Applications (Full version). IACR Cryptology ePrint Archive 529, 2012.

[BCG+12b] Julia Borghoff, Anne Canteaut, Tim Güneysu, Elif Bilge Kavun, Miroslav Knezevic, Lars R. Knudsen, Gregor Leander, Ventzislav Nikov, Christof Paar, Christian Rechberger, Peter Rombouts, Søren S. Thomsen, and Tolga Yalçin. PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. In *Advances in Cryptology - ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 208–225. Springer, 2012.

[BCL08] Lilya Budaghyan, Claude Carlet, and Gregor Leander. Two Classes of Quadratic APN Binomials Inequivalent to Power Functions. *IEEE Transactions on Information Theory*, 54(9) : pages 4218–4229, 2008.

[BDB12] Simeon Ball and Jan De Beule. On sets of vectors of a finite vector space in which every subset of basis size is a basis II. *Designs, Codes and Cryptography*, 65(1-2) : pages 5–14, 2012.

[BDMW10] K.A. Browning, J.F. Dillon, M.T. McQuistan, and A.J. Wolfe. An APN permutation in dimension six. In *Finite Fields : Theory and Applications*, volume 518 of *Contemporary Mathematics*, pages 33–42. AMS, 2010.

[BFDF98] T. D. Bending and Dmitry Fon-Der-Flaass. Crooked Functions, Bent Functions, and Distance Regular Graphs. *Electronic Journal of Combinatorics*, 5, 14 pages, 1998.

[BGT11] Céline Blondeau, Benoît Gérard, and Jean-Pierre Tillich. Accurate estimates of the data complexity and success probability for various cryptanalyses. *Designs, Codes and Cryptography*, 59(1-3) : pages 3–34, 2011.

[BK08] Jürgen Bierbrauer and Gohar M.M. Kyureghyan. Crooked binomials. *Designs, Codes and Cryptography*, 46(3) : pages 269–301, 2008.

[BKL+07] Andrey Bogdanov, Lars Ramkilde Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and Charlotte Vikkelsoe. PRESENT : An Ultra-Lightweight Block Cipher. In *Cryptographic Hardware and Embedded Systems - CHES 2007*, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.

[BL08] Marcus Brinkmann and Gregor Leander. On the classification of APN functions up to dimension five. *Designs, Codes and Cryptography*, 49(1-3) : pages 273–288, 2008.

[BN13] Céline Blondeau and Kaisa Nyberg. New Links between Differential and Linear Cryptanalysis. In *Advances in Cryptology - EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 388–404. Springer, 2013.

[BS90] Eli Biham and Adi Shamir. Differential Cryptanalysis of DES-like Cryptosystems. In *Advances in Cryptology - CRYPTO'90*, volume 537 of *LNCS*, pages 2–21. Springer, 1990.

[BS91] Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, pages 3–72, 1991.

[Car10] Claude Carlet. *Boolean Functions for Cryptography and Error Correcting Codes*, chapter 8 of *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, pages 257–397. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010.

[CC03] Anne Canteaut and Pascale Charpin. Decomposing Bent Functions. *IEEE Transactions on Information Theory*, 49(8) :2004–19, 2003.

[Cha98] Pascale Charpin. Théorie Algébrique des codes-correcteurs d'erreurs, 1998. Cours de Master Algorithmique.

[CKL+03] Kilsoo Chun, Seungjoo Kim, Sangjin Lee, Soo Hak Sung, and Seonhee Yoon. Differential and linear cryptanalysis for 2-round SPNs. *Information Processing Letters*, 87(5) : pages 277–282, 2003.

[CR14] Anne Canteaut and Joëlle Roué. Extended differential properties of cryptographic functions. In *Topics in Finite Fields*, volume 632 of *Contemporary Mathematics*, pages 43–70. AMS, 2014.

[CR15a] Anne Canteaut and Joëlle Roué. Differential Attacks against SPN : a Thourough Analysis. In *C2SI 2015*, volume 9084 of *LNCS*, pages 45–62. Springer, 2015.

[CR15b] Anne Canteaut and Joëlle Roué. On the behaviors of affine equivalent Sboxes regarding differential and linear attacks. In *Advances in Cryptology - EUROCRYPT 2015, part I*, volume 9056 of *LNCS*, pages 45–74. Springer, 2015.

[CV94] Florent Chabaud and Serge Vaudenay. Links Between Differential and Linear Cryptoanalysis. In *Advances in Cryptology - EUROCRYPT '94*, volume 950 of *LNCS*, pages 356–365. Springer, 1994.

[Dae95] Joan Daemen. *Cipher and hash function design strategies based on linear and differential cryptanalysis.* PhD thesis, K.U. Leuven, 1995.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6) : pages 644–654, 1976.

[DKR97] Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The Block Cipher Square. In *Fast Software Encryption - FSE'97*, volume 1267 of *LNCS*, pages 149–165. Springer, 1997.

[DLP+09] Joan Daemen, Mario Lamberger, Norbert Pramstaller, Vincent Rijmen, and Frederik Vercauteren. Computational aspects of the expected differential probability of 4-round AES and AES-like ciphers. *Computing*, 85(1-2) : pages 85–104, 2009.

[DR00] Joan Daemen and Vincent Rijmen. Rijndael for AES. In *AES Candidate Conference*, pages 343–348, 2000.

[DR01] Joan Daemen and Vincent Rijmen. The Wide Trail Design Strategy. In *IMA International Conference - Coding and Cryptography 2001*, volume 2260 of *LNCS*, pages 222–238. Springer, 2001.

[DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael : AES - The Advanced Encryption Standard.* Information Security and Cryptography. Springer, 2002.

[DR05] Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. IACR Cryptology ePrint Archive 212, 2005.

[DR06] Joan Daemen and Vincent Rijmen. Understanding Two-Round Differentials in AES. In *Security and Cryptography for Networks - SCN 2006*, volume 4116 of *LNCS*, pages 78–94. Springer, 2006.

[DR07a] Joan Daemen and Vincent Rijmen. Plateau characteristics. *IET Information Security*, 1(1) : pages 11–17, 2007.

[DR07b] Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. *Journal of Mathematical Cryptology*, 1(3) : pages 221–242, 2007.

[DR09] Joan Daemen and Vincent Rijmen. New criteria for linear maps in AES-like ciphers. *Cryptography and Communications*, 1(1) : pages 47–69, 2009.

[DR11] Joan Daemen and Vincent Rijmen. *Advanced Linear Cryptanalysis of Block and Stream Ciphers*, chapter Correlation Analysis in $GF(2^n)$, pages 115–131. Cryptology and information security. IOS Press, 2011.

[Fei74] Horst Feistel. Block Cipher Cryptographic System. *U.S. Patent 3798359*, 1974.

[FIP77] FIPS 46. Data Encryption Standard. Federal Information Processing Standards Publication 46, 1977. U.S. Department of Commerce/N.I.S.T.

[FIP01] FIPS 197. Advanced Encryption Standard. Federal Information Processing Standards Publication 197, 2001. U.S. Department of Commerce/N.I.S.T.

[GNL12] Zheng Gong, Svetla Nikova, and Yee Wei Law. KLEIN : A New Family of Lightweight Block Ciphers. In *RFID, Security and Privacy - RFIDSec 2011*, volume 7055 of *LNCS*, pages 1–18. Springer, 2012.

[GPPR11] Jian Guo, Thomas Peyrin, Axel Poschmann, and Matthew Robshaw. The LED Block Cipher. In *Cryptographic Hardware and Embedded Systems - CHES 2011*, volume 6917 of *LNCS*, pages 326–341. Springer, 2011.

[Har96] Carlo Harpes. *Cryptanalysis of iterated block ciphers*, volume 7 of *ETH Series in Information Processing*. Hartung-Gorre Verlag, Konstanz, 1996.

[HKM95] Carlo Harpes, Gerhard G. Kramer, and James L. Massey. A Generalization of Linear Cryptanalysis and the Applicability of Matsui's Piling-Up Lemma. In *Advances in Cryptology - EUROCRYPT'95*, volume 921 of *LNCS*, pages 24–38. Springer, 1995.

[HLL+00] Seokhie Hong, Sangjin Lee, Jongin Lim, Jaechul Sung, Dong Hyeon Cheon, and Inho Cho. Provable Security against Differential and Linear Cryptanalysis for the SPN Structure. In *Fast Software Encryption - FSE 2000*, volume 1978 of *LNCS*, pages 273–283. Springer, 2000.

[HLP52] Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities*. Cambridge University Press, 1952.

[KLL+14] Elif Bilge Kavun, Martin M. Lauridsen, Gregor Leander, Christian Rechberger, Peter Schwabe, and Tolga Yalçın. Prøst v1.1. Submission to the CAESAR competition, 2014. http://proest.compute.dtu.dk/proestv11.pdf.

[KS07] Liam Keliher and Jiayuan Sui. Exact maximum expected differential and linear probability for two-round Advanced Encryption Standard. *IET Information Security*, 1(2) : pages 53–57, 2007.

[Kyu07] Gohar M.M. Kyureghyan. Crooked maps in $\mathbf{F}_{2^n}$. *Finite Fields and Their Applications*, 13(3) : pages 713–726, 2007.

[LK06] Chae Hoon Lim and Tymur Korkishko. mCrypton - A Lightweight Block Cipher for Security of Low-Cost RFID Tags and Sensors. In *Information Security Applications - WISA 2005*, volume 3786 of *LNCS*, pages 243–258. Springer, 2006.

[LMM91] Xuejia Lai, James L. Massey, and Sean Murphy. Markov ciphers and differential cryptanalysis. In *Advances in Cryptology - EUROCRYPT'91*, volume 547 of *LNCS*, pages 17–38. Springer-Verlag, 1991.

[LN83] Rudolf Lidl and Harald Niederreiter. *Finite Fields*. Cambridge University Press, 1983.

[Mac63] Florence Jessie MacWilliams. A theorem on the distribution of weights in a systematic code. *Bell System Technical Journal*, 42 : pages 79–94, 1963.

[Mat94] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology - EUROCRYPT'93*, volume 765 of *LNCS*, pages 386–397. Springer-Verlag, 1994.

[Mat95] Mitsuru Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In *Advances in Cryptology - CRYPTO'94*, volume 839 of *LNCS*, pages 1–11. Springer-Verlag, 1995.

[McE87] Robert J. McEliece. *Finite Fields for Computer Scientists and Engineers*. Kluwer Academic Publishers, 1987.

[MS77] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The Theory of Error-Correcting Codes*. North-Holland, 1977.

[MvOV97] Alfred John Menezes, Paul C. van Oorshot, and Scott Alexander Vanstone. *Handbook of applied cryptography*. CRC Press, 1997. http://www.cacr.math.uwaterloo.ca/hac/.

[MWGP12] Nicky Mouha, Qingju Wang, Dawu Gu, and Bart Preneel. Differential and Linear Cryptanalysis Using Mixed-Integer Linear Programming. In *Information Security and Cryptology - Inscrypt 2011*, volume 7537 of *LNCS*, pages 57–76. Springer, 2012.

[NK93] Kaisa Nyberg and Lars Ramkilde Knudsen. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92*, volume 740 of *LNCS*, pages 566–574. Springer-Verlag, 1993.

[Nyb94] Kaisa Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT '93*, volume 765 of *LNCS*, pages 53–64. Springer-Verlag, 1994.

[Nyb95] Kaisa Nyberg. Linear Approximation of Block Ciphers. In *Advances in Cryptology - EUROCRYPT'94*, volume 950 of *LNCS*, pages 439–444. Springer-Verlag, 1995.

[PSLL03] Sangwoo Park, Soo Hak Sung, Sangjin Lee, and Jongin Lim. Improving the Upper Bound on the Maximum Differential and the Maximum Linear Hull Probability for SPN Structures and AES. In *Fast Software Encryption - FSE 2003*, volume 2887 of *LNCS*, pages 247–260. Springer, 2003.

[RDP+96] Vincent Rijmen, Joan Daemen, Bart Preneel, Antoon Bosselaers, and Erik De Win. The Cipher SHARK. In *Fast Software Encryption - FSE'96*, volume 1039 of *LNCS*, pages 99–111. Springer, 1996.

[Rot76] Oscar S. Rothaus. On bent functions. *Journal of Combinatorial Theory, Series A*, 20 : pages 300–305, 1976.

[Seg55] Beniamino Segre. Curve razionali normali ek-archi negli spazi finiti. *Annali di Matematica Pura ed Applicata*, 39(1) : pages 357–379, 1955.

[Sha49] Claude E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, pages 656–715, 1949.

[Sin64] Richard C. Singleton. Maximum Distance q-nary codes. *IEEE Transactions on Information Theory*, 10 : pages 116–118, 1964.

[TCG91] Anne Tardy-Corfdir and Henri Gilbert. A known plaintext attack of FEAL-4 and FEAL-6. In *Advances in Cryptology - CRYPTO'91*, volume 576 of *LNCS*, pages 172–182. Springer-Verlag, 1991.

[ZZ99] Yuliang Zheng and Xian-Mo Zhang. Plateaued functions. In *Information and Communication Security - ICICS'99*, volume 1726 of *LNCS*, pages 224–300. Springer-Verlag, 1999.