Generic Attacks on Unbalanced Feistel Schemes with Expanding Functions

Jacques Patarin[1], Valérie Nachef[2], and Côme Berbain[3]

[1] Université de Versailles
45 avenue des Etats-Unis, 78035 Versailles Cedex, France
[2] Department of Mathematics
University of Cergy-Pontoise
CNRS UMR 8088
2 avenue Adolphe Chauvin, 95011 Cergy-Pontoise Cedex, France
[3] France Telecom Research and Development
38-40 rue du Général Leclerc, 92794 Issy-les-Moulineaux, France
jacques.patarin@prism.uvsq.fr
valerie.nachef@u-cergy.fr
come.berbain@orange-ftgroup.com

Abstract. Unbalanced Feistel schemes with expanding functions are used to construct pseudo-random permutations from kn bits to kn bits by using random functions from n bits to $(k-1)n$ bits. At each round, all the bits except n bits are changed by using a function that depends only on these n bits. C.S.Jutla [6] investigated such schemes, which he denotes by F_k^d, where d is the number of rounds. In this paper, we describe novel Known Plaintext Attacks (KPA) and Non Adaptive Chosen Plaintext Attacks (CPA-1) against these schemes. With these attacks we will often be able to improve the result of C.S.Jutla. We also give precise formulas for the complexity of our attacks in d, k and n.

Key words: Unbalanced Feistel permutations, pseudo-random permutations, generic attacks on encryption schemes, Block ciphers.

1 Introduction

A Feistel scheme from $\{0,1\}^l$ to $\{0,1\}^l$ with d rounds is a permutation built from rounds functions $f_1, \dots, f_d$. When these round functions are randomly chosen, we obtain what is called a "Random Feistel Scheme". The attacks on these "random Feistel schemes" are called "generic attacks" since these attacks are valid for most of the round functions $f_1, \dots f_d$.

• When $l = 2n$ and when the f_i functions are from $\{0,1\}^n$ to $\{0,1\}^n$ we obtain the most classical Feistel schemes, also called "balanced" Feistel schemes. Since the famous paper of M.Luby and C.Rackoff [11], many results have been obtained on the security of such classical Feistel schemes (see [12] for an overview of these results). When the number of rounds is lower than 5, we know attacks with less than $2^l (= 2^{2n})$ operations: for 5 rounds, an attack in $O(2^n)$ operations is given in [15] and for 3 or 4 rounds an attack in $\sqrt{2^n}$ is given in [1],[13]. When the functions are permutations, similar attacks for 5 rounds are given in [7] and [9]. Therefore, for security, at least 6 rounds are recommended, i.e. each bit will be changed at least 3 times.

• When $N = kn$ and when the round functions are from $(k-1)n$ bits to n bits, we obtain what is called an "Unbalanced Feistel Scheme with contracting functions". In [12] some security proofs are given for such schemes when for the first and the last rounds pairwise independent functions are used instead of random contracting functions. At Asiacrypt 2006 ([16]) generic attacks on such schemes have been studied.

• When $N = kn$ and when the rounds functions are from n bits to $(k-1)n$ bits, we obtain what is called an "Unbalanced Feistel Scheme with expanding functions", also called "complete target heavy unbalanced Feistel networks"(see [17]). Generic attacks on Unbalanced Feistel Schemes with expanding functions is the theme of this paper. One advantage of these schemes is that it requires much less memory to store a random function of n bits to $(k-1)n$ bits than a random function of $(k-1)n$ bits to n bits. BEAR and LION [2] are two block ciphers which employ both expanding and contracting unbalanced Feistel networks. The AES-candidate MARS is also using a similar structure.

Attacks on Unbalanced Feistel Schemes with expanding functions have been previously studied by C.S.Jutla ([6]). We will often be able to improve his attacks by attacking more rounds, or by using a smaller complexity. Moreover we will generalize these attacks by analyzing KPA (Known Plaintext Attacks), not only CPA-1 (non adaptive plaintext attacks) and by giving explicit formulas for the complexities. We will not introduce adaptive attacks, or chosen plaintext and chosen ciphertext attacks, since we have not found anything significantly better than CPA-1.

The paper is organized as follows. First, we give some notation. Then, we describe our different attacks when $k = 3$. For the clarity of the exposition we have decided to present in the main part of this paper only the case $k = 3$, only the best known attacks, and only when the complexity is less then $O(2^N) = O(2^{kn}) = O(2^{3n})$ when $k = 3$. The attacks with a complexity greater than or equal to $O(2^{3n})$ and $k = 3$ will be presented in Appendices A, B,C and D. Finally, the attacks for any k, $k \geq 3$ will be presented in Appendices E,F,G,H, I. Our attacks for any $k \geq 3$ are in fact a generalisation of our attacks for $k = 3$. We will have essentially two families of attacks called "2 point attacks" (TWO) and "rectangle attacks" (SQUARE, R1, R2, R3, R4). It can be noticed that $k = 2$ is very different from $k = 3$ (and $k \geq 3$), since we do not have the analog of the "rectangle" attacks.

2 Notations

We first describe Unbalanced Feistel Scheme with Expanding Functions F_k^d and introduce some useful notations. F_k^d is a Feistel scheme of d rounds. At each round j, we denote by f_j the round function from n bits to $(k-1)n$ bits. f_j is defined as $f_j = (f_j^{(1)}, f_j^{(2)}, \ldots, f_j^{(k-1)})$, where each function $f_j^{(l)}$ is defined from $\{0,1\}^n$ to $\{0,1\}^n$. On some input $[I^1, I^2, \ldots, I^k]$ F_k^d produces an output denoted by $[S^1, S^2, \ldots, S^k]$ by going through d rounds. At round j, the first n bits of the round entry are used as an input to the round function f_j, which produces $(k-1)n$ bits. Those bits are xored to the $(k-1)n$ last bits of the round entry and the result is rotated by n bits.

The first round is represented on Figure 1 below:

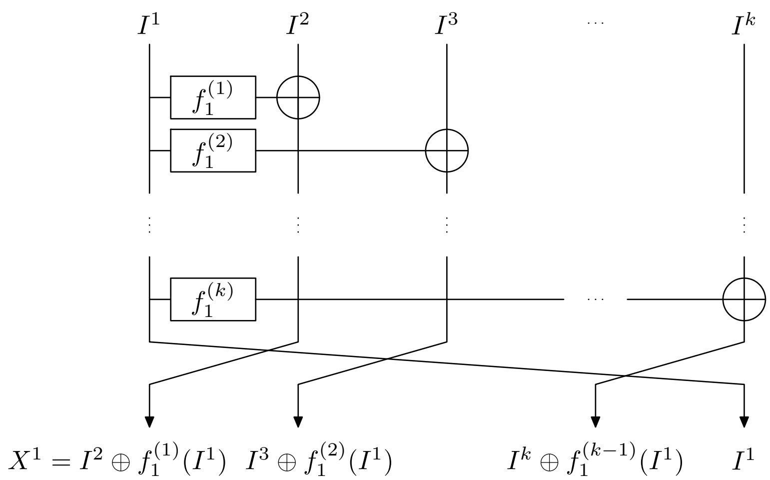

Fig. 1. First Round of F_k^d

We introduce notation X^j: we denote by X^j the n-bit value produced by round j, which will be the input of next round function f_{j+1}. We have

$$X^1 = I^2 \oplus f_1^{(1)}(I^1)$$
$$X^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(X^1)$$

$$X^3 = I^4 \oplus f_1^{(3)}(I^1) \oplus f_2^{(2)}(X^1) \oplus f_3^{(1)}(X^2)$$
$$\dots$$

More generally, we can express the X^j recursively:

$$\forall \xi < k,\ X^\xi = I^{\xi+1} \oplus f_1^{(\xi)}(I^1) \oplus_{i=2}^{\xi} f_i^{(\xi-i+1)}(X^{i-1})$$
$$\forall \xi \geq 1,\ X^{k+\xi} = X^\xi \oplus_{i=2}^{k} f_{\xi+i}^{(k-i+1)}(X^{\xi+i-1})$$

After d rounds ($d \geq k+1$), the output $[S^1, S^2, \dots, S^k]$ can be expressed by using the introduced values X^j:

$$S^k = X^{d-1}$$
$$S^{k-1} = X^{d-2} \oplus f_d^{(k-1)}(X^{d-1})$$
$$S^{k-2} = X^{d-3} \oplus f_{d-1}^{(k-1)}(X^{d-2}) \oplus f_d^{(k-2)}(X^{d-1})$$
$$\dots$$

More generally, we can express the S^j recursively:

$$\forall \xi,\ 1 \leq \xi \leq k-1 \quad S^\xi = X^{d-1-k+\xi} \oplus_{i=d-k+\xi}^{d-1} f_{i+1}^{(\xi+d-i-1)}(X^i)$$

Inversion of F_k^d

For all $A_1, \dots, A_k \in I_n$, let

$$\sigma[A_1, A_2, \dots, A_k] = [A_k, A_{k-1}, \dots, A_2, A_1]$$

We have $\sigma \circ \sigma$ =Identity and $(F_k^1(f_1^{(1)}, \dots, f_1^{(k-1)}))^{-1} = \sigma \circ F_k^1(f_1^{(k-1)}, \dots, f_1^{(1)})) \circ \sigma$. Therefore by composition, we see that the inverse of an F_k^d is another F_k^d if we take the k inputs, the k outputs and the $d(k-1)$ functions in the inverse order:

$$F_k^d(f_1^{(1)}, \dots, f_1^{(k-1)}, \dots, f_d^{(1)}, \dots, f_d^{(k-1)})^{-1} = \sigma \circ F_k^d(f_d^{(k-1)}, \dots, f_d^{(1)}, \dots, f_1^{(k-1)}, \dots, f_1^{(1)}) \circ \sigma$$

3 Attacks "TWO" with $k = 3$ and $d \leq 5$

In this section, we will describe a family of attacks called "TWO". These attacks will use correlations on pairs of cleartext/ciphertexts. Therefore, they can be called "2 points" attacks. When $k = 2$ (i.e. on classical balanced Feistel Schemes) these attacks give the best known generic attacks (cf [15]). However these attacks where have not been studied in [6]. As we will see, TWO attacks are more efficient than the attacks of [6] when the number of rounds is very small, or very large but, surprisingly, not when the number of rounds is intermediate.

Remark. We present here TWO only for $k = 3$ and $d \leq 5$. TWO for $k = 3$ and $d \geq 6$ will be presented in Appendix A and TWO for any $k \geq 3$ will be presented in Appendix E.

3.1 Attack TWO against F_3^1

We just test if $S^3 = I^1$. We need one message and about one computation in KPA and CPA-1.

3.2 Attack TWO against F_3^2

We will concentrate the attack on the equation: $X^1 = I^2 \oplus f_1^{(1)}(I^1)$, i.e. here $S^3 = I^2 \oplus f_1^{(1)}(I^1)$.

- For the CPA-1 attack, we choose two messages such that I^1 is constant. Then we test if $S^3 \oplus I^2$ is constant. Thus in CPA-1, we need only 2 messages (and about 2 computations).
- For the KPA attack, we can transform this CPA-1 attack. If we have two indices $i < j$ such that $I^1(i) = I^1(j)$, then we test if $S^3(i) \oplus S^3(j) = I^2(i) \oplus I^2(j)$. Here, from the birthday paradox, this KPA attack is in $O(\sqrt{2^n})$ messages and $O(\sqrt{2^n})$ computations.

3.3 Attack TWO against F_3^3

We will concentrate the attack on the equation: $S^3 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(I^2 \oplus f_1^{(1)}(I^1))$ (since here we have $S^3 = X^2$).

- For the CPA-1 attack, we choose two messages such that I^1 and I^2 are constant. Then we test if $S^3 \oplus I^3$ is constant. Thus in CPA-1, we need only 2 messages (and about 2 computations).
- For the KPA attack, we can transform this CPA-1 attack. If we have two indices $i < j$ such that $I^1(i) = I^1(j)$ and $I^2(i) = I^2(j)$, then we test if $S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j)$. Here, from the birthday paradox, this KPA requires $O(2^n)$ messages and $O(2^n)$ computations.

3.4 Attack TWO against F_3^4

CPA-1 Attack

We will concentrate the attack on the equation: $S^2 = X^2 \oplus f_4^{(2)}(S^3)$ (since here $X^3 = S^3$) with

$$X^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(I^2 \oplus f_1^{(1)}(I^1)).$$

For the CPA-1 attack, we will choose m messages ($m \simeq \sqrt{2^n}$) such that I^1 and I^2 are constant. Therefore for all i, j we will have: $X^2(i) \oplus X^2(j) = I^3(i) \oplus I^3(j)$. Now when $m \geq O(\sqrt{2^n})$, from the birthday paradox, we know that we will have with a good probability at least one (i, j), $i < j$ such that $S^3(i) = S^3(j)$. If this occurs, we will test if $S^2(i) \oplus S^2(j) = I^3(i) \oplus I^3(j)$. This appears with probability about $\frac{1}{2^n}$ for a random permutation and with probability 1 on F_3^4 when $S^3(i) = S^3(j)$, $I^1(i) = I^1(j)$ and $I^2(i) = I^2(j)$. Thus we have obtained a CPA-1 attack with $O(\sqrt{2^n})$ messages and $O(\sqrt{2^n})$ complexity.

KPA Attack

We can transform this CPA-1 attack in a KPA attack in the usual way: we wait for collisions on I^1, I^2, and S^3, and we test if $S^2(i) \oplus S^2(j) = I^3(i) \oplus I^3(j)$. From the birthday paradox, we will get with a good probability at least one collision on I^1, I^2, S^3 when $m^2 \geq O(2^{3n})$. Therefore the number of messages and the complexity are here in $O(2^{\frac{3n}{2}})$.

Remark: There is also another KPA attack on F_3^4: we just have to count the number of $i < j$ such that $S^3(i) \oplus S^3(j) = I^1(i) \oplus I^1(j)$. It is possible to prove that there is a small deviation of this value and that this attack also has a complexity in $O(2^{\frac{3n}{2}})$ messages and computations. (We do not give the details since it gives the same complexity).

3.5 Attack TWO against F_3^5

CPA-1 Attack

We will concentrate the attack on the equation: $S^1 = X^2 \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4)$, with

$$X^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(I^2 \oplus f_1^{(1)}(I^1)), \quad X^4 = S^3, \quad X^3 = S^2 \oplus f_5^{(2)}(S^3)$$

For the CPA-1 attack, we will choose m messages ($m \simeq 2^n$) such that I^1 and I^2 are constant. Therefore for all i, j, we will have: $X^2(i) \oplus X^2(j) = I^3(i) \oplus I^3(j)$. Now when $m \geq O(2^n)$, from the birthday paradox, we know that we will have with a good probability at last one (i, j), $i < j$, such that $S^2(i) = S^2(j)$ and $S^3(i) = S^3(j)$. This means here (since $S^2 = X^3 \oplus f_5^{(2)}(X^4)$ and $S^3 = X^4$) that $X^4(i) = X^4(j)$ and $X^3(i) = X^3(j)$. If we get such an (i, j), we will test if: $S^1(i) \oplus S^1(j) = I^3(i) \oplus I^3(j)$. This appears with probability about $\frac{1}{2^n}$ for a random permutation and with probability 1 on F_3^5, when $I^1(i) = I^1(j)$, $I^2(i) = I^2(j)$, $S^2(i) = S^2(j)$, $S^3(i) = S^3(j)$. Thus we have obtained a CPA-1 attack with $O(2^n)$ messages and $O(2^n)$ complexity.

Remark: If we get no such (i, j) for some value I^1 and I^2 (we then have at most 2^n possibilities for I^3), we can try again with some other fixed values I^1, I^2. With a high probability, we will get a solution after only a few tries.

KPA Attack

We can trasform this CPA-1 attack in a KPA attack in the usual way: we wait for collisions on I^1, I^2, S^2, S^3, and then we test if $S^1(i) \oplus S^1(j) = I^3(i) \oplus I^3(j)$. From the birthday paradox, we will get with a good probability at least one collision on I^1, I^2, S^2, S^3 when $m^2 \geq O(2^{4n})$. Therefore the number of messages and the complexity are here in $O(2^{2n})$.

Remark: There is also another possible KPA attack on F_3^5: we just have to count the number of (i,j), $i<i$ such that $I^1(i)=I^1(j)$ and $X^4(i)\oplus I^2(i)=X^4(j)\oplus I^2(j)$. It is possible to prove that there is a small deviation of this value and that this attack also has complexity in $O(2^{2n})$ messages and computations. (We do not give the details here since it gives the same complexity).

4 "SQUARE" Attack on F_3^6

We will present here our best attack on F_3^6. This attack belongs to a family of attacks that we have called "SQUARE" ('SQUARE" attacks will be a special case of "R1" attacks when we use only a square of 4 points in the attack. More general description of the SQUARE and R1 attacks will be given below and in Appendix G). We have $F_3^6[I^1,I^2,I^3]=[S^1,S^2,S^3]$ with

$$\begin{cases} S^1 = X^3 \oplus f_5^{(2)}(X^4) \oplus f_6^{(1)}(X^5) \\ S^2 = X^4 \oplus f_6^{(2)}(X^5) \\ S^3 = X^5 \end{cases}$$

with

$$\begin{cases} X^1 = I^2 \oplus f_1^{(1)}(I^1) \\ X^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(X^1) \\ X^3 = I^1 \oplus f_2^{(2)}(X^1) \oplus f_3^{(1)}(X^2) \\ X^4 = X^1 \oplus f_3^{(2)}(X^2) \oplus f_4^{(1)}(X^3) \\ X^5 = X^2 \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4) \end{cases}$$

Let i_1, i_2, i_3, i_4 be four indices of messages (so these values are between 1 and m). We will denote by $[I^1(\alpha),I^2(\alpha),I^3(\alpha)]$ the plaintext of message i_α, and by $[S^1(\alpha),S^2(\alpha),S^3(\alpha)]$ the ciphertext of message i_α. (i.e. for simplicity we use the notation $I^1(\alpha)$ and $S^1(\alpha)$ instead of $I^1(i_\alpha)$ and $S^1(i_\alpha)$, $1\le\alpha\le 4$). The idea of the attack is to count the number N of indices (i_1,i_2,i_3,i_4) such that:

$$\begin{cases} I^1(1) = I^1(2) \\ I^1(3) = I^1(4) \\ I^2(1) \oplus I^2(2) = I^2(3) \oplus I^2(4) \\ I^3(1) \oplus I^3(2) = I^3(3) \oplus I^3(4) \\ S^3(1) = S^3(2) \\ S^3(3) = S^3(4) \\ S^2(1) = S^2(2) \\ S^2(3) = S^2(4) \\ S^1(1) \oplus S^1(2) = S^1(3) \oplus S^1(4) \end{cases}$$

We will call the 4 first equations the "input equations", and we will call the 5 last equations the "output equations".

KPA.

If the messages are randomly chosen we will have $E(N)\simeq \frac{m^4}{2^{9n}}$. (The standard deviation $\sigma(N)$ can also be computed, cf Appendix D, however the standard deviation is not needed here since $E(N)$ will be about the double for F_3^6). For a F_3^6 permutation we will have about 2 times more solutions since the 5 output equations can occur at random, or due to these 5 internal equations:

$$\begin{cases} X^1(1) = X^1(3) \\ X^2(1) = X^2(2) \\ X^3(1) = X^3(3) \\ X^4(1) = X^4(2) \\ X^5(1) = X^5(2) \end{cases}$$

Therefore here we have: $\varphi = 4$, $a = 2$, $n_I = 4$, $n_S = 5$, $n_X = 5$, where φ denotes the number of points linked with the equalitites, a denotes the number of equations in X between the indices 1 and 3, n_I denotes the number of input equations, n_S the number of output equations and n_X the number of needed equations in X.

These equations are summarized in figure 2 below.

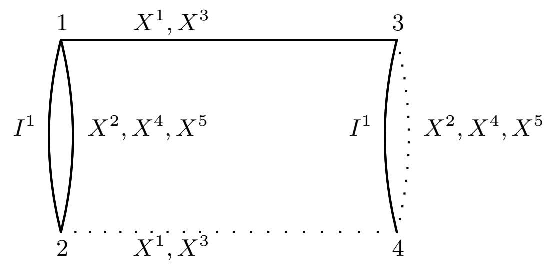

Fig. 2. SQUARE Attack on F_3^6

In this figure 2 two poinst are joined by an edge if the values are equal (for example $I^1(1) = I^1(2)$). We draw a solid edge if the probability appears with probability $\frac{1}{2^n}$ and a dotted line if the equality follows conditionally with probability 1 from other imposed equalities. For example here, from $X^1(1) = X^1(3)$ we get $X^1(2) = X^1(4)$ (since $X^1(1) \oplus X^1(2) \oplus X^1(3) \oplus X^1(4) = I^2(1) \oplus I^2(2) \oplus I^2(3) \oplus I^2(4) = 0$). Similarly

$$\begin{cases} X^2(1) = X^2(2) \,\text{gives}\, X^2(3) = X^2(4) \\ X^3(1) = X^3(3) \,\text{gives}\, X^3(2) = X^3(4) \\ X^4(1) = X^4(2) \,\text{gives}\, X^4(3) = X^4(4) \\ X^5(1) = X^5(2) \,\text{gives}\, X^5(3) = X^5(4) \end{cases}$$

Now since $S^3 = X^5$, $S^2 = X^4 \oplus f_6^{(2)}(X^5)$ and $S^1 = X^3 \oplus f_5^{(2)}(X^4) \oplus f_6^{(1)}(X^5)$, we get the 5 output equations written above. Therefore, in KPA, for a F_3^6 permutation, the expectancy of N is larger than for a random permutation by a value about $\frac{m^4}{2^{9n}}$ (since we have 5 equations in X and 4 in I), i.e. we expect to have about 2 times more solutions for N: $E(N) \simeq \frac{2m^4}{2^{9n}}$ for F_3^6. So we will be able to distinguish with a high probability F_3^6 from a random permutation by counting N when $N \neq 0$, with high probability i.e. when $m^4 \geq 2^{9n}$, or $m \geq 2^{\frac{9n}{4}}$. We have found a KPA with $O(2^{\frac{9n}{4}})$ complexity and $O(2^{\frac{9n}{4}})$ messages. (This is better than the $O(2^{\frac{5n}{2}})$ complexity found in section 3).

CPA-1.

We can transform this KPA in CPA-1. We will choose only two fixed different values a and b, $a \neq b$ for I^1: $\frac{m}{2}$ plaintexts will have $I^1 = a$ and $\frac{m}{2}$ plaintexts will have $I^1 = b$. Let α be a fixed integer between 0 and n (the best value for α will be chosen below). We will generate all the possible messages $[I^1, I^2, I^3]$ such that I^1 has the value a or b, the first α bits of I^2 are 0, and the first α bits of I^3 are 0. Therefore we have $m = 2 \cdot 2^{n-\alpha} \cdot 2^{n-\alpha}$. How many solutions (i_1, i_2, i_3, i_4) will satisfy our 4 input equations? For i_1 we have m possibilities. Then, when i_1 is fixed, for i_2 such that $I^1(2) = I^1(1)$ we have $\frac{m}{2}$ possibilities, and for i_3 such that $I^1(3) \neq I^1(1)$ we have $\frac{m}{2}$ possibilitites. Then, when i_1, i_2, i_3 are fixed, we have one and exactly one possibility for i_4, since $I^1(4)$, $I^2(4)$ and $I^3(4)$ are now fixed from the input equations. Therefore, for (i_1, i_2, i_3, i_4) that satisfy our 4 input equations, we have exactly $\frac{m^3}{4}$ solutions. For a random permutation we will have $E(N) \simeq \frac{m^4}{4 \cdot 2^{5n}}$ (since we have 5 output equations). For a permutation F_3^6 we will have $E(N) \simeq \frac{m^3}{2 \cdot 2^{5n}}$, i.e. about 2 times more solutions, since these 5 output equations can occur at random, or due to 5 internal equations in X, as we have seen. So this CPA-1 will succed with a high probability when $N \neq 0$ with a high probability i.e. when $m \geq O(2^{\frac{5n}{3}})$. (Therefore we will choose $\alpha \simeq \frac{n}{6}$ for $m \simeq 2^{\frac{5n}{3}}$. We have found here a CPA-1 with $O(2^{\frac{5n}{3}})$ complexity and $O(2^{\frac{5n}{3}})$ messages. (This is better than the $O(2^{2n})$ complexity found in section 3).

Complexity

Here the complexity is in $O(m)$ because we can compute N in $O(m)$. For this we can proceed in 3 steps.

Step 1: we compute all the solutions (i, j) such that $S^3(i) = S^3(j)$, $S^2(i) = S^2(j)$ and $I^1(i) = I^1(j)$. We need here $O(m)$ computations and we will find about $\frac{m^2}{2^{2n}} \simeq 2^{\frac{4n}{3}}$ solutions. We store these solutions in two sets A and B: A with $I^1 = a$, B with $I^1 = b$.

Step 2: we compute $A' = \{S^1(i) \oplus S^1(j),\ (i, j) \in A\}$ and $B' = \{S^1(i) \oplus S^1(j),\ (i, j) \in B\}$. We have about $2^{\frac{4n}{3}}$ solutions in A', and $2^{\frac{4n}{3}}$ solutions in B'.

Step 3: now we look for a common value in A' and B'. This can be done with $2^{\frac{4n}{3}}$ computations (and memory), and we will find about $(2^{\frac{4n}{3}})^2/2^n$ solutions, i.e. $O(2^{\frac{5n}{3}})$ solutions. The number of these solutions gives N.

Remark. This attack on F_3^6, unlike our attacks on F_3^7, F_3^8, F_3^9, and unlike the TWO attacks of the previous sections can be seen as using only ideas already present in Jutla's paper [6] (except the fact that we have also designed a KPA, not only a CPA-1).

5 Attack "R1" on F_3^7

We will now describe our "R1" attack on F_3^7. As we will see, we will obtain here a complexity in $O(2^{2n})$ in CPA-1 and in $O(2^{\frac{5n}{2}})$ in KPA. This is better than the $O(2^{3n})$ of the TWO attacks. In [6], Jutla shows that he can obtain on F_k^d attacks with complexity less than $O(2^{kn})$ when $d \leq 3k - 3$. For $d = 3$, this gives attacks up to only 6 rounds, unlike here where we will reach 7 rounds with a complexity less than 2^{3n}. We have $F_3^7[I^1, I^2, I^3] = [S^1, S^2, S^3]$ with

$$\begin{cases} S^1 = X^4 \oplus f_6^{(2)}(X^5) \oplus f_7^{(1)}(X^6) \\ S^2 = X^5 \oplus f_7^{(2)}(X^6) \\ S^3 = X^6 \end{cases}$$

with

$$\begin{cases} X^1 = I^2 \oplus f_1^{(1)}(I^1) \\ X^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(X^1) \\ X^3 = I^1 \oplus f_2^{(2)}(X^1) \oplus f_3^{(1)}(X^2) \\ X^4 = X^1 \oplus f_3^{(2)}(X^2) \oplus f_4^{(1)}(X^3) \\ X^5 = X^2 \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4) \\ X^6 = X^3 \oplus f_5^{(2)}(X^4) \oplus f_6^{(1)}(X^5) \end{cases}$$

Let $i_1, i_2, i_3, i_4, i_5, i_6$ be six indices of messages (so these values are between 1 and m). We will denote by $[I^1(\alpha), I^2(\alpha), I^3(\alpha)]$ the plaintext of message i_α, and by $[S^1(\alpha), S^2(\alpha), S^3(\alpha)]$ the ciphertext of message i_α. (i.e. for simplicity we use the notation $I^1(\alpha)$ and $S^1(\alpha)$ instead of $I^1(i_\alpha)$ and $S^1(i_\alpha)$, $1 \leq \alpha \leq 6$). The idea of the attack is to count the number N of indices $(i_1, i_2, i_3, i_4, i_5, i_6)$ such that:

$$\begin{cases} I^1(1) = I^1(2) \text{ and } I^1(3) = I^1(4) \text{ and } I^1(5) = I^1(6) \\ I^2(1) \oplus I^2(2) = I^2(3) \oplus I^2(4) = I^2(5) \oplus I^2(6) \\ I^3(1) \oplus I^3(2) = I^3(3) \oplus I^3(4) = I^3(5) \oplus I^3(6) \\ \quad\text{and} \\ S^3(1) = S^3(2) \text{ and } S^3(3) = S^3(4) \text{ and } S^3(5) = S^3(6) \\ S^2(1) = S^2(2) \text{ and } S^2(3) = S^2(4) \text{ and } S^2(5) = S^2(6) \\ S^1(1) \oplus S^1(2) = S^1(3) \oplus S^1(4) = S^1(5) \oplus S^1(6) \end{cases}$$

We will call the 7 first equations the "input equations" and we will call the 8 last equations the "output equations".

KPA. If the messages are randomly chosen, for a random permutation we will have $E(N) \simeq \frac{m^6}{2^{15n}}$. (The standard deviation σ can also be computed, cf Appendix D, however the standard deviation is not needed here since $E(N)$ will be about the double for F_3^7). For a F_3^7 permutation we will have about 2 times more solutions since the 8 output equations can occur at random, or due to these 8 internal equations:

$$\begin{cases} X^1(1) = X^1(3) = X^1(5) \\ X^2(1) = X^2(2) \\ X^3(1) = X^3(2) \\ X^4(1) = X^4(3) = X^4(5) \\ X^5(1) = X^5(2) \\ X^6(1) = X^6(2) \end{cases}$$

Therefore here we have: $\varphi = 6$, $a = 2$, $n_I = 7$, $n_S = 8$, $n_X = 8$, where φ denotes the number of points linked with the equalitites, a denotes the number of equations in X between the indices 1 and 3, n_I denotes the number of input equations, n_S the number of output equations and n_X the number of needed equations in X.

These equations are summarized in figure 3 below.

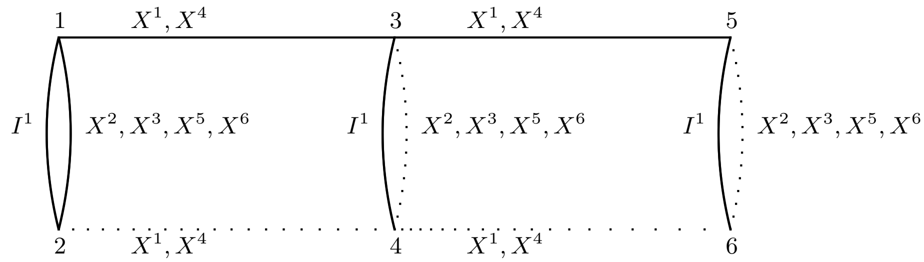

Fig. 3. R1 Attack on F_3^7

In this figure 3 (as in figure 2), two points are joined by an edge if the values are equal (for example $I^1(1) = I^1(2)$). We draw a solid edge if the probability appears with probability $\frac{1}{2^n}$ and a dotted line if the equality follows conditionally with probability 1 from other imposed equalities. For example here, from $X^1(1) = X^1(3) = X^1(5)$ we get $X^1(2) = X^1(4) = X^1(6)$ (since $X^1(1) \oplus X^1(2) \oplus X^1(3) \oplus X^1(4) = I^2(1) \oplus I^2(2) \oplus I^2(3) \oplus I^2(4) = 0$ and in the same way $X^1(1) \oplus X^1(2) \oplus X^1(5) \oplus X^1(6) = 0$). Similarly

$$\begin{cases} X^2(1) = X^2(2) \text{ gives } X^2(3) = X^2(4) \text{ and } X^2(5) = X^2(6) \\ X^3(1) = X^3(2) \text{ gives } X^3(3) = X^2(4) \text{ and } X^3(5) = X^3(6) \\ X^4(1) = \ X^4(3) = X^4(5) \text{ gives } X^4(2) = X^4(4) = X^4(6) \\ X^5(1) = X^5(2) \text{ gives } X^5(3) = X^5(4) \text{ and } X^5(5) = X^5(6) \\ X^6(1) = X^6(2) \text{ gives } X^6(3) = X^6(4) \text{ and } X^6(5) = X^6(6) \end{cases}$$

Now since $S^3 = X^6$, $S^2 = X^5 \oplus f_7^{(2)}(X^6)$ and $S^1 = X^4 \oplus f_6^{(2)}(X^5) \oplus f_7^{(1)}(X^6)$, we get the 8 output equations written above. Therefore, in KPA, for a F_3^7 permutation, the expectancy of N is larger than for a random permutation by a value of about $\frac{m^6}{2^{15n}}$ (since we have 8 equations in X and 7 in I), i.e. we expect to have about 2 times more solutions for N: $E(N) \simeq \frac{2m^6}{2^{15n}}$ for F_3^7. So we will be able to distinguish with a high probability F_3^7 from a random permutation by counting N when $N \neq 0$ with a high probability, i.e. when $m^6 \geq O(2^{15n})$, or $m \geq O(2^{\frac{5n}{2}})$. We have found here a KPA with $O(2^{\frac{5n}{2}})$ complexity and $O(2^{\frac{5n}{2}})$ messages. This is better than the $O(2^{3n)}$ complexity of the attack TWO, and it shows that we can attack 7 rounds, not only 6 with a complexity less than 2^{3n}.

CPA-1

We can transform this KPA in CPA-1. We will choose only 3 fixed different values a, b, c for I^1: $\frac{m}{3}$ plaintexts will have $I^1 = a$, $\frac{m}{3}$ plaintexts will have $I^1 = b$, and $\frac{m}{3}$ plaintexts will have $I^1 = c$. We will generate all (or almost all) possible messages $[I^1, I^2, I^3]$ with such I^1. Therefore, $m = 3 \cdot 2^{2n}$. How many solutions $(i_1, i_2, i_3, i_4, i_5, i_6)$ will satisfy our 7 input equations? For i_1, we have m possibilities. Then, when i_1 is fixed, for i_2 such that $I^1(2) = I^1(1)$, we have $\frac{m}{3}$ possibilities.

Then for i_3, such that $I^1(3) \neq I^1(1)$, we have $\frac{2m}{3}$ possibilities. Then for i_5 such that $I^1(5) \neq I^1(1)$ and $I^1(5) \neq I^1(3)$, we have $\frac{m}{3}$ possibilities. Now for i_4 and i_6, we have one and only one possibility since their values I^1, I^2, I^3 are fixed from the input equations when i_1, i_2, i_3, i_5 are fixed. Therefore for $(i_1, i_2, i_3, i_4, i_5, i_6)$ that satisfy the 7 input equations, we have $\frac{2m^4}{27}$ solutions. For a random permutation we will have $E(N) \simeq \frac{2m^4}{27 \cdot 2^{8n}}$ (since we have 8 output equations). For a permutation F_3^7, we will have $E(N) \simeq \frac{4m^4}{27 \cdot 2^{8n}}$, i.e. about 2 times more solutions, since the 8 output equations can occur at random, or due to 8 internal equations in X as we have seen. So this CPA-1 will succeed when $N \neq 0$ with a high probability, i.e. when $m^4 \geq O(2^{8n})$, or $m \geq O(2^{2n})$. Here we have $m \simeq 3 \cdot 2^{2n}$, the probability of success is not negligible. Moreover if it fails for some values (a, b, c) for I^1, we can start again with another (a, b, c). Therefore this CPA-1 is in $O(2^{2n})$ complexity and $O(2^{2n})$ messages. (This is better than the $O(2^{3n})$ attack TWO found in Section 3).

6 Attack "R2" on F_3^8

We will present here our best attack on F_3^8. These attacks belong to a family of attacks that we have called "R2". In fact, R2 attacks are very similar to M1 attacks: the main difference is the position of the equations in I. (A more general description and analysis of the R2 attacks will be given in Appendix H). Therefore we present here only the main ideas (our notations and conventions for R2 are similar to those for M1). The ideas of the attack R2 on F_3^8 is to count the number N of indices $(i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8)$ such that:

$$\begin{cases} I^1(1) = I^1(3) \;=\; I^1(5) = I^1(7) \\ I^1(2) = I^1(4) \;=\; I^1(6) = I^1(8) \\ I^2(1) \oplus I^2(2) \;=\; I^2(3) \oplus I^2(4) \;=\; I^2(5) \oplus I^2(6) \;=\; I^2(7) \oplus I^2(8) \\ I^3(1) \oplus I^3(2) \;=\; I^3(3) \oplus I^3(4) \;=\; I^3(5) \oplus I^3(6) \;=\; I^3(7) \oplus I^3(8) \\ \text{ and} \\ S^3(1) = S^3(2) \text{ and } S^3(3) = S^3(4) \text{ and } S^3(5) = S^3(6) \text{ and } S^3(7) = S^3(8) \\ S^2(1) = S^2(2) \text{ and } S^2(3) = S^2(4) \text{ and } S^2(5) = S^2(6) \text{ and } S^2(7) = S^2(8) \\ S^1(1) \oplus S^1(2) \;=\; S^1(3) \oplus S^1(4) \;=\; S^1(5) \oplus S^1(6) \;=\; S^1(7) \oplus S^1(8) \end{cases}$$

We will call the 12 first equations the "input equations" and we will call the last 11 equations the "output equations. In the same way as we did for M1 on F_3^6 and F_3^7, we can easily prove that the expectancy for N is about double in F_3^8 compared with a random permutation, since in F_3^8 the 11 output equations can occur at random or due to these 11 equations in X:

$$\begin{cases} X^1(1) = X^1(2) \\ X^2(1) = X^2(3) = X^2(5) = X^2(7) \\ X^3(1) = X^3(2) \\ X^4(1) = X^4(2) \\ X^5(1) = X^5(3) = X^5(5) = X^5(7) \\ X^6(1) = X^6(2) \\ X^7(1) = X^7(2) \end{cases}$$

(Remember that here $S^3 = X^7$, $S^2 = X^6 \oplus f_8^{(2)}(S^3)$ and $S^1 = X^5 \oplus f_7^{(2)}(X^6) \oplus f_8^{(1)}(S^3)$).

Therefore here we have: $\varphi = 8$, $a = 2$, $n_I = 12$, $n_S = 11$, $n_X = 11$, with the usual notations for φ, a, n_I, n_S, n_X.

These equations are summarized in figure 4 below.

KPA

If the messages are randomly chosen we will have $E(N) \simeq \frac{m^8}{2^{23n}}$ for a random permutation, and $E(N) \simeq \frac{2m^8}{2^{23n}}$ for F_3^8 permutations. Therefore with a good probability $N \neq 0$ (and the attack will succeed) when $m \geq O(2^{\frac{23n}{8}})$. (This is less than 2^{3n}).

CPA-1

We can transform this KPA in a CPA-1 in the usual way. Here we will choose all the possible (or almost all) I^1, I^2, I^3 such that the $\frac{n}{2}$ first bits of I^1 are 0. Therefore we have here $m = 2^{\frac{n}{2}} \cdot 2^n \cdot 2^n = 2^{\frac{5n}{2}}$ possible inputs. Here $E(N) \simeq \frac{m^8}{2^{20}}$ (each collision in I^1 has probability about $\frac{1}{\sqrt{2^n}}$) for a random permutation and $E(N) \simeq \frac{2m^8}{2^{20n}}$ for a F_3^8 permutation. Here

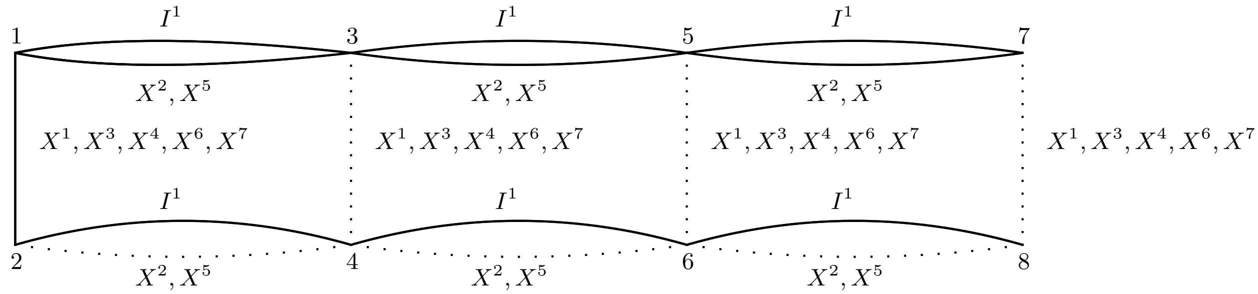

Fig. 4. R2 Attack on F_3^8

we have $m = O(2^{\frac{5n}{2}})$ so the probability of success is not negligible. (Moreover if we find no solution we can try again by fixing α bits of I^1 at 0 different from the first α bits).

Remark. On points 1 et 2 we have 5 equations (in X^1, X^3, X^4, X^6, X^7). Therefore a necessary condition for the attack to succeed is $m^2 \geq 2^{5n}$. This condition is satisfied here. Similarly, on points 1, 2, 3 we need $m^3 \geq 2^{8n}$ in KPA (and $m^3 \geq 2^{7.5n}$ in CPA-1). These conditions are also satisfied here.

7 Experimental results

We have implemented the CPA-1 attacks SQUARE and R1 against F_3^6 and F_3^7 described in section 4 and 5.

The attack against F_3^6 is using 4 points (i.e. $\varphi = 4$) and $2^{\frac{5n}{3}}$ plaintexts, while the attack against F_3^7 is using 6 points (i.e. $\varphi = 6$) and 2^{2n} plaintexts. Our experiments confirm our ability to distinguish between a F_3^6 or F_3^7 scheme and a random permutation. Our experiments were done as follows:

- choose randomly an instance of F_3^6 or F_3^7
- choose randomly a permutation: for this we use classical balanced Feistel scheme with a large number of rounds (more than 20)
- launch the attack in CPA-1 with $2^{\frac{5n}{3}}$ or 2^{2n} plaintexts
- count the number of structures satisfying the input and output relations for the F_3^6 or F_3^7 function and for the permutation
- if this number is higher than a fixed threshold, declare the function to be a F_3^6 or F_3^7 function and otherwise a random permutation

All this procedure is iterated a large number of time (at least 1000 times) to evaluate the efficiency of our distinguisher. We give the percentage of success, i.e. the number of F_3^6 or F_3^7 scheme that have been correctly distinguished and the percentage of false alarm, i.e. the number of random permutation that have incorrectly been declared as F_3^6 or F_3^7 scheme.

Table 1. Experimental results for CPA-1 attacks

scheme	n	threshold	Percentage of success of the attack	Percentage of false alarm
F_3^6	8	2	54%	4%
F_3^7	6	1	33%	8%

Our experiments show that the distinguisher on F_3^6 is more efficient than the one on F_3^7. But in both case they confirm our theoretical analysis.

8 Conclusion for $k = 3$ and more results

For F_3^d, $d \geq 9$, our best attacks require more than $O(2^{3n})$ computations. More generally for any $k \geq 3$ for F_k^d, when $d \geq 3k$ our best attacks require more than $O(2^{kn})$ computations.

Attack by the signature

A classical theorem proves that all the permutations F_k^d have an even signature. Therefore, by computing the signature of F_k^d we are able to distinguish F_k^d from a random permutation with $O(2^{kn})$ computations when all the 2^{kn} cleartext/ciphertext are known with a non-negligible probability. However if we do not have access to the complex codebook of size 2^{kn}, or if we want to distinguish F_k^d from a random permutation with an even signature, this "attack" obviously fails.

Brute force attack

A possible attack is the exhaustive search on the d round functions $f_1, \ldots, f_d$ from $\{0,1\}^n$ to $\{0,1\}^{(k-1)n}$ that have been used in the unbalanced Feistel construction. This attack always exists, but since we have $2^{d(k-1)n \cdot 2^n}$ possibilities for $f_1, \ldots, f_d$, this attack requires about $2^{d(k-1)n \cdot 2^n}$ computations and about $\frac{d(k-1) \cdot 2^n}{k}$ queries but only one permutation of the generator. This attack means that an adversary with infinite computing power will be able to distinguish F_k^d from a random permutation (or from a truly random permutation with an even signature) when $m \geq \frac{d(k-1) \cdot 2^n}{k}$. Brute force attack requires a small value for m but a huge computing power.

TWO and Rectangle attacks

In Appendix we explain how to extend the attack TWO and the rectangle attacks when $d \leq 3k$, i.e. when we want to attack a generator of F_3^d, not only a single F_3^d. When $3k \leq d \leq k^2$, rectangle attacks are the best known attacks, in term of complexity, when we want to distinguish a generator of F_3^d from a random permutation with an even signature. When $d \geq k^2 + 1$, TWO becomes the best known attack for this problem.

Finally, the results that we have obtained for $k = 3$ are summarized in table 2 below, and for any $k \geq 3$ in table 4 of Appendix I.

Table 2. Results on F_3^d. For example for F_3^7, this table means that the best attack that we have found in KPA is the attack M1 and this attack needs $m \simeq 2^{\frac{5}{2}n}$ and has a complexity of $\simeq 2^{\frac{5}{2}n}$ computations. For $d \geq 9$ more than one permutation is needed or $\geq 2^{3n}$ computations are needed in the best known attacks.

	KPA	CPA-1
F_3^1	1	1
F_3^2	$2^{\frac{n}{2}}$, TWO	2
F_3^3	2^n, TWO	2
F_3^4	$2^{\frac{3}{2}n}$, TWO	$2^{\frac{n}{2}}$, TWO
F_3^5	2^{2n}, TWO	2^n, TWO
F_3^6	$2^{\frac{9}{4}n}$, SQUARE	$2^{\frac{5}{3}n}$, SQUARE
F_3^7	$2^{\frac{5}{2}n}$, R1, $\varphi = 6$	2^{2n}, R1, $\varphi = 6$
F_3^8	$2^{\frac{23}{8}n}$, R2, $\varphi = 8$	$2^{\frac{5}{2}n}$, R2, $\varphi = 8$
F_3^9	2^{3n}, R2, $\varphi \geq 10$	2^{3n}, R2, $\varphi \geq 10$
F_3^{10}	2^{7n}, TWO	2^{7n}, TWO
F_3^{11}	2^{8n}, TWO	2^{8n}, TWO
F_3^d, $d \geq 10$	$2^{(d-6+\lfloor \frac{d}{3} \rfloor)n}$, TWO	$2^{(d-6+\lfloor \frac{d}{3} \rfloor)n}$, TWO

9 Open problems

There are still many open problems on Unbalanced Feistel Schemes with Expanding Functions.

• One of them is to get proofs of security, not only design of attacks. Classical proofs "a la Luby-Rackoff" will give security within the "birthday bound" (i.e. in $m << \sqrt{2^n}$). A better proof is given in [6] with security in $m \leq 2^{(1-\frac{1}{k})n}$. It is probably possible to improve this result (for example by using generalisation of [15]) in order to get security in $m << 2^n$ ("information theory bound"). However here 2^n is very small compared with a security in, say, 2^{nk} that we would like to get. At present, proving a security in $2^{\alpha n}$, for $\alpha > 1$ looks a very difficult problem, not mentioning $\alpha = k$ or $\alpha > k$.

• Another problem is to design better attacks than the attacks of this paper. For example, instead of 2 points attacks (TWO) or rectangle attacks (M1, R2), we have tried attacks with different geometries of the equations (hexagons instead of rectangles, 3-dimension cubes instead of 2-dimension rectangles, etc...). So far our most promising new geometries are "Multi-Rectangles attacks" (see Appendix J). These new attacks are very promising (at least from a theoretical point of view) but still under investigation.

10 Conclusion

The attacks of this paper improve C.S.Jutla's results [6]. We follow many C.S.Jutla's ideas: we employ generalizations of the birthday paradox, and we use in our attacks SQUARE, R1, R2, R3, R4 a"rectangle framework" of equalities. Usual birthday attacks (see [1], [10], [3]) are based on requiring two variables to be the same. Generalizations to more than one coincidence have been studied in [5], [4], [8].

To improve the attacks of C.S.Jutla, we have first made a systematic analysis of the different ways to optimize the parameters. For example, we have optimized the position of the internal equalities and of the equalities in the input and the ouput variables in the rectangle framework and we have computed the optimal number of points of this rectangle framework. In CPA-1, we have also introduced a fixed number of 0 at the beginning of $I^2, I^3, \ldots, I^k$. We have described 5 general attacks TWO, M1, R2, R3, R4 and the best of these 5 attacks is sometimes TWO, sometimes M1, sometimes R2, or R3, or R4 depending on the number of rounds (cf Appendix I).

One of our main result is that we can attack with KPA with a complexity strictly lower than 2^{kn} when $d \leq 3k - 1$ (unlike $d \leq 3k - 3$ with CPA-1 for C.S.Jutla). Therefore we have obtained "generic attacks" (with a complexity less than 2^{kn}) on two more rounds by using rectangle attacks.

Another of our result is that when k and d are fixed the complexity of our attacks are generally smaller than [6]. Another of our main result is the fact that we have shown that the "TWO" attacks are the best known attacks for very small, or very large values of d (but not for intermediate values). For very large values of d we assume that we want to attack a generator of F_k^d (not only one F_k^d).

Moreover, we have shown that there exists also another very promising family of attacks, that we have called "Multi-Rectangle" attacks. We think that with "Multi-Rectangle" attacks we will be able to attack more rounds (when $k \geq 4$) and decrease the complexity, but the precise results are not exactly known since these attacks are still under investigation.

In conclusion, there are much more possibilities for generic attacks on unbalanced Feistel schemes with expanding functions than with other Feistel schemes (classical or with contracting functions). So these constructions must be designed with great care and with sufficiently many rounds. However, if sufficiently many rounds are used, these schemes are very interesting since the memory needed to store the functions is much smaller compared with other generic Feistel schemes.

References

1. William Aiello and Ramarathnam Venkatesan. Foiling Birthday Attacks in Length-Doubling Transformations - Benes: A Non-Reversible Alternative to Feistel. In Ueli M. Maurer, editor, *Advances in Cryptology – EUROCRYPT '96*, volume 1070 of *Lecture Notes in Computer Science*, pages 307–320. Springer-Verlag, 1996.
2. Ross J. Anderson and Eli Biham. Two Practical and Provably Secure Block Ciphers: BEARS and LION. In Dieter Gollman, editor, *Fast Software Encryption*, volume 1039 of *Lecture Notes in Computer Science*, pages 113–120. Springer-Verlag, 1996.
3. Don Coppersmith. Another Birthday Attack. In Hugh C. Williams, editor, *Advances in Cryptology – CRYPTO '85*, volume 218 of *Lecture Notes in Computer Science*, pages 14–17. Springer-Verlag, 1985.
4. Don Coppersmith. Luby-Rackoff: Four rounds is not enough. Technical Report RC20674, IBM Research Report, December 1996.

5. Marc Girault, Robert Cohen, and Mireille Campana. A Generalized Birthday Attack. In C. G. Gnther, editor, *Advances in Cryptology – EUROCRYPT '88*, volume 330 of *Lecture Notes in Computer Science*, pages 129–156. Springer-Verlag, 1988.
6. Charanjit S. Jutla. Generalized Birthday Attacks on Unbalanced Feistel Networks. In Hugo Krawczyk, editor, *Advances in Cryptology – CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*, pages 186–199. Springer-Verlag, 1998.
7. Lars R. Knudsen. DEAL - A 128-bit Block Cipher. Technical Report 151, University of Bergen, Department of Informatics, Norway, february 1998.
8. Lars R. Knudsen, Xuejia Lai, and Bart Preneel. Attacks on Fast Double Block Length Hash Functions. *J. Cryptology*, 11(1):59–72, 1998.
9. Lars R. Knudsen and Vincent Rijmen. On the Decorrelated Fast Cipher (DFC) and Its Theory. In Lars R. Knudsen, editor, *Fast Software Encrytion – FSE '99*, volume 1636 of *Lecture Notes in Computer Science*, pages 81–94. Springer-Verlag, 1999.
10. Michael Luby. *Pseudorandomness and cryptographic applications*. Princeton University Press, 1996.
11. Michael Luby and Charles Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM J. Comput.*, 17(2):373–386, 1988.
12. Moni Naor and Omer Reingold. On the Construction of Pseudorandom Permutations: Luby-Rackoff Revisited. *J. Cryptology*, 12(1):29–66, 1999.
13. Jacques Patarin. New Results on Pseudorandom Permutation Generators Based on the DES Scheme. In Joan Feigenbaum, editor, *Advances in Cryptology – CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 301–312. Springer-Verlag, 1991.
14. Jacques Patarin. Generic Attacks on Feistel Schemes. In Colin Boyd, editor, *Advances in Cryptology – ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 222–238. Springer-Verlag, 2001.
15. Jacques Patarin. Security of Random Feistel Schemes with 5 or More Rounds. In Matthew K. Franklin, editor, *Advances in Cryptology – CRYPTO 2004*, volume 3152 of *Lecture Notes in Computer Science*, pages 106–122. Springer-Verlag, 2004.
16. Jacques Patarin, Valérie Nachef, and Côme Berbain. Generic Attacks on Unbalanced Feistel Schemes with Contracting Functions. In Xuejia Lai and Kefei Chen, editors, *Advances in Cryptology – ASIACRYPT 2006*, volume 4284 of *Lecture Notes in Computer Science*, pages 396–411. Springer-Verlag, 2006.
17. Bruce Schneier and John Kelsey. Unbalanced Feistel Networks and Block Cipher Design. In Dieter Gollmann, editor, *Fast Software Encrytion – FSE '96*, volume 1039 of *Lecture Notes in Computer Science*, pages 121–144. Springer-Verlag, 1996.

Appendices A,B,C,D for $k = 3$

A Attacks TWO with $k = 3$ and $d \geq 6$

A.1 Attack TWO against F_3^6

In this sub-section we will describe the "MO" attack on F_3^6. Unlike for 1,2,3,4,5 rounds, this attack is not the best attack that we have found against F_3^6. However, it is interesting to describe it in order to compare it with the other attacks.

KPA Attack

We will concentrate the attack on the equation: $S^3 = X^2 \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4)$ i.e. on

$$S^3 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(I^2 \oplus f_1^{(1)}(I^1)) \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4)$$

In this attack, we will count the number N of (i, j), $i < j$ such that

$$\begin{cases} S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j) \\ I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \end{cases}$$

For a random permutation, the expectancy of N is $E(N) \simeq \frac{m(m-1)}{2 \cdot 2^{3n}}$ with a standard deviation $\sigma(N) \simeq \sqrt{E(N)} \simeq \frac{m}{2^{\frac{3n}{2}}}$. (The way to compute the standard deviation is explained in Appendix D).

For F_3^6, we can notice that when $I^1(i) = I^1(j)$ and $I^2(i) = I^2(j)$, we have

$$S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j) \Leftrightarrow f_4^{(2)}(X^3(i)) \oplus f_4^{(2)}(X^3(j)) = f_5^{(1)}(X^4(i)) \oplus f_5^{(1)}(X^4(j))$$

and this can occur if

$$\begin{cases} X^3(i) = X^3(j) \\ X^4(i) = X^4(j) \end{cases}$$

(with a probability about $\frac{1}{2^{2n}}$) or due to the functions $f_4^{(2)}$ and $f_5^{(1)}$ (with a probability about $\frac{1}{2^n}$). So the expectancy of N slightly larger for F_3^6 than for a random permutation. More precisely, $|E(N)_{F_3^6} - E(N)_{perm}| \simeq \frac{m(m-1)}{2 \cdot 2^{4n}}$. Moreover, this value is larger than the standard deviation of N when $\frac{m^2}{2^{4n}} \geq \frac{m}{2^{\frac{3n}{2}}}$, i.e. when $m \geq 2^{\frac{5n}{2}}$. Therefore, we have a KPA on F_3^6 with $O(2^{\frac{5n}{2}})$ messages and complexity $O(2^{\frac{5n}{2}})$.

CPA-1 Attack

We can transform this attack in a CPA-1 attack with a better complexity. Let μ be an integer (μ will be chosen below about 2^n). We will choose μ possible values for (I^1, I^2) and we will ask for the $2^n \cdot \mu$ ciphertexts of (I^1, I^2, I^3) for all possible I^3.

We will count the number N of (i, j) such that:

$$\begin{cases} I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \\ S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j) \end{cases}$$

For a random permutation the expectancy of N is $E(N) \simeq \mu \cdot \frac{2^{2n}}{2^n} = \mu \cdot 2^n$, with a standard deviation $\sigma(N) \simeq \sqrt{\mu \cdot 2^n}$. (The way to compute the standard deviation is explained in Appendix D). For F_3^6, $E(N)$ is slightly larger, as we have seen above: we expect to have about $\mu \cdot \frac{2^{2n}}{2^{2n}}$ more solutions, i.e. about μ more solutions. This is larger than $\sigma(N)$ when $\mu \geq \sqrt{\mu \cdot 2^n}$, i.e. when $\mu \geq 2^n$. Thus we have obtained a CPA-1 on F_3^6 with a complexity $O(\mu \cdot 2^n) = O(2^{2n})$ and $O(2^{2n})$ messages.

Remark: On F_3^6 if we start from equation S^1 instead of S^3, we will obtain a similar KPA but we will obtain a chosen ciphertext attack in 2^{2n} instead of a chosen plaintext attack. This is why we have presented here the attacks from the equation S^3.

A.2 Attack TWO against F_3^7

We present here only the main ideas, since the attack is similar as before. We can concentrate the attack on the equation of S^2 (or with S^1, since with S^1, we have a similar result).

$$S^2 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(X^1) \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4) \oplus f_7^{(1)}(S^3)$$

We will count the number N of (i, j), $i < j$, such that:

$$\begin{cases} I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \\ S^3(i) = S^3(j) \\ S^2(i) \oplus S^2(j) = I^3(i) \oplus I^3(j) \end{cases}$$

For a random permutation, we have $E(N) \simeq \frac{m^2}{2 \cdot 2^{4n}}$, with a standard deviation $\sigma(N) \simeq O(\frac{m}{2^{2n}})$. For F_3^7, we will have about $\frac{m^2}{2 \cdot 2^{5n}}$ more solutions (they came from $X^3(i) = X^3(j)$ and $X^4(i) = X^4(j)$). Therefore this attack will succeed when $\frac{m^2}{2^{5n}} \geq O(\frac{m}{2^{2n}})$, i.e. $m \geq O(2^{3n})$. This gives a KPA with complexity about 2^{3n} and about 2^{3n} messages. (We have here nothing better in CPA-1).

A.3 Attack TWO against F_3^d when $d \geq 8$ and $d = 2 \bmod 3$

Here we will assume that we want to attack not only one F_3^d but a generator of F_3^d permutations, i.e. we have access to α such permutations with μ messages per permutation (μ will be about 2^{3n}). In this TWO attack, (here $d = 2 \bmod 3$), we will count the number N of (i, j), $i < j$ such that:

$$\begin{cases} I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \\ S^3(i) = S^3(j) \\ S^2(i) = S^2(j) \\ S^1(i) \oplus S^1(j) = I^3(i) \oplus I^3(j) \end{cases}$$

For α random permutations we have $E(N) \simeq \frac{\alpha\mu(\mu-1)}{2\cdot 2^{5n}}$ with a standard deviation $\sigma(N) = O(\sqrt{\frac{\alpha\mu^2}{2^{5n}}})$. (The way to compute the standard deviation is explained in Appendix D).

- F_3^8: For F_3^8 we have

$$S^1 = I^3 \oplus f_1^{(2)}(I^1) \oplus f_2^{(1)}(I^2 \oplus f_1^{(1)}(I^1)) \oplus f_4^{(2)}(X^3) \oplus f_5^{(1)}(X^4) \oplus f_7^{(2)}(S^2 \oplus f_8^{(2)}(S^3)) \oplus f_8^{(1)}(S^3)$$

Therefore for F_3^8 we will have about $\frac{\alpha\mu^2}{2^{6n}}$ more solutions in N (they come from $X^3(i) = X^3(j)$ and $X^4(i) = X^4(j)$). This is larger than $\sigma(N)$ (and the attack will succeed) if $\frac{\alpha\mu^2}{2^{6n}} \geq 0(\sqrt{\frac{\alpha\mu^2}{2^{5n}}})$ i.e. when $\alpha\mu^2 \geq O(2^{7n})$. With $\mu \simeq 2^{3n}$, this gives $\alpha \geq 0(2^n)$. Therefore we have obtained a KPA against a generator of F_3^8 with complexity $\alpha \cdot \mu = O(2^{4n})$ and $O(2^{4n})$ messages.

- **F_3^d, $d \geq 8$, $d = 2 \bmod 3$**

More generally for F_3^d, $d \geq 8$, $d = 2 \bmod 3$, we will have about $\dfrac{\alpha\mu^2}{2^{(\frac{2d+2}{3})n}}$ more solutions in N (since each time we increase d by 3 we have 2 more variables in S^1). This is larger than $\sigma(N)$ if $\frac{\alpha\mu^2}{2^{(\frac{2d+2}{3})n}} \geq O(\sqrt{\frac{\alpha\mu^2}{2^{5n}}})$. With $\mu \simeq 2^{3n}$ this gives $\alpha \cdot 2^{6n} \geq O(2^{(\frac{4d-11}{3})n})$, $\alpha \geq O(2^{(\frac{4d-29}{3})n})$. Therefore we have obtained a KPA against a generator of F_3^d, $d = 2 \bmod 3$ with complexity $\alpha \cdot \mu = O(2^{(\frac{4d-20}{3})n})$. Since $d = 2 \bmod 3$, this is also $O(2^{(d-6+\lfloor \frac{d}{3} \rfloor)n})$.

A.4 Attack TWO against F_3^d when $d \geq 9$ and $d = 0 \bmod 3$

Here we will again assume that we have access to α permutations with μ messages per permutation, $\mu \simeq 2^{3n}$. When $d = 0 \bmod 3$, we will count the number N of (i,j), $i < j$ such that:

$$\begin{cases} I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \\ S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j) \end{cases}$$

For α random permutations, we have $E(N) \simeq \frac{\alpha\mu(\mu-1)}{2\cdot 2^{3n}}$ with a standard deviation $\sigma(N) = O(\sqrt{\frac{\alpha\mu^2}{2^{3n}}})$. For F_3^d, $d = 0 \bmod 3$, we will have about $\frac{\alpha\mu^2}{2^{(\frac{2d}{3})n}}$ more solutions in N (by writing the expression of S^3 similarly as before). This is larger than $\sigma(N)$ if $\frac{\alpha\mu^2}{2^{(\frac{2d}{3})n}} \geq O(\sqrt{\frac{\alpha\mu^2}{2^{3n}}})$. With $\mu \simeq 2^{3n}$, this gives $\alpha \cdot 2^{6n} \geq O(2^{(\frac{4d}{3}-3)n})$, $\alpha \geq 2^{(\frac{4d}{3}-9)n}$. Therefore we have obtained a KPA against a generator of F_3^d, $d = 0 \bmod 3$, with complexity $\alpha \cdot \mu = O(2^{(\frac{4d}{3}-6)n})$. Since $d = 0 \bmod 3$, this is also $O(2^{(d-6+\lfloor \frac{d}{3} \rfloor)n})$.

A.5 Attack TWO against F_3^d when $d \geq 10$ and $d = 1 \bmod 3$

Here we will again assume that we have access to α permutations with μ messages per permutations, $\mu \simeq 2^{3n}$. When $d = 1 \bmod 3$, we will count the number N of (i,j), $i < j$ such that:

$$\begin{cases} I^1(i) = I^1(j) \\ I^2(i) = I^2(j) \\ S^3(i) = S^3(j) \\ S^2(i) \oplus S^2(j) = I^3(i) \oplus I^3(j) \end{cases}$$

(Remark: another possible attack with the same complexity will be to count the number N of (i, j), $i < j$ such that: $I^1(i) = I^1(j)$, $S^3(i) = S^3(j)$, $S^2(i) = S^2(j)$ and $S^1(i) \oplus S^1(j) = I^2(i) \oplus I^2(j)$).

For α random permutations we have $E(N) \simeq \frac{\alpha\mu(\mu-1)}{2\cdot 2^{4n}}$ with a standard deviation $\sigma(N) = O(\sqrt{\frac{\alpha\mu^2}{2^{4n}}})$. For F_3^d, $d = 1 \bmod 3$, we will have about $\frac{\alpha\mu^2}{2^{(\frac{2d+1}{3})n}}$ more solutions in N (by writing the expression of S^2 similarly as before). This is larger than $\sigma(N)$ if $\frac{\alpha\mu^2}{2^{(\frac{2d+1}{3})n}} \geq O(\sqrt{\frac{\alpha\mu^2}{2^{4n}}})$. With $\mu \simeq 2^{3n}$, this gives $\alpha \cdot 2^{6n} \geq O(2^{(\frac{4d-10}{3})n})$, $\alpha \geq O(2^{(\frac{4d-28}{3})n})$. Therefore we have obtained a KPA against a generator of F_3^d, $d = 1 \bmod 3$, with complexity $\alpha \cdot \mu = O(2^{(\frac{4d-19}{3})n})$. Since $d = 1 \bmod 3$, this is also $O(2^{(d-6+\lfloor\frac{d}{3}\rfloor)n})$.

A.6 Conclusion for the attacks TWO on F_3^d

We summarize the results obtained in this section on the TWO attacks in the table 1 below. These are the best attacks that we have found by using correlation on only two indices i and j. In the next section, we will study attacks by correlation on more than two indices.

Table 3. Summary of the complexity of the Attacks TWO with $k = 3$. For $d = 6, 7, 8, 9$, the attacks SQUARE, R1 or R2 will be better.

d	KPA	CPA-1
1	1	1
2	$2^{\frac{n}{2}}$	2
3	2^n	2
4	$2^{\frac{3}{2}n}$	$2^{\frac{n}{2}}$
5	2^{2n}	2^n
6	$2^{\frac{5}{2}n}$	2^{2n}
7	2^{3n}	2^{3n}
8	2^{4n}	2^{4n}
9	2^{6n}	2^{6n}
10	2^{7n}	2^{7n}
11	2^{8n}	2^{8n}
F_3^d, $d \geq 7$	$2^{(d-6+\lfloor\frac{d}{3}\rfloor)n}$	$2^{(d-6+\lfloor\frac{d}{3}\rfloor)n}$

The horizontal line shows when the complexity reaches 2^{3n}, i.e. when we need a generator.

B Attack "R2" on F_3^9

We will present here our best attack on F_3^9. Here, the complexity of the attack and the number of messages m needed are in $O(2^{3n})$. Therefore, when we have only one F_3^9 we can distinguish F_3^9 from a random permutation with a non-negligible probability p when $m = O(2^{3n})$ with $O(2^{3n})$ computations. (However, if we want p to be arbitrary near 1, we will need more than one F_3^9; i.e. a generator of F_3^9). Since 2^{3n} is the total number of possible inputs for one F_3^9, we see that 9 rounds for F_3^9 plays the same limit role as 6 rounds for the classical Feistel schemes F_2^d: for this number of rounds the complexity of the best known attack is about the number of all possible inputs (for F_2^6 the best known attacks are in $O(2^{2n})$, see [14], [15]).

The general properties of R2 on F_k^d are presented in Appendix H. We present here only the main ideas. More details about R2 are given in Appendix H. For $k = 3$ and $d = 9$ we have with $a = 2$:

$$\begin{cases} n_I = 2\varphi - 4 \\ n_S = \frac{3\varphi}{2} - 1 \\ n_X = \varphi + 4 \end{cases}$$

(Same notations as in Appendix H). In order to have $n_X \leq n_S$, we will choose $\varphi \geq 10$. We can prove that the attack will succeed if these 3 conditions are satisfied:

1. (On all the points): $m^\varphi \geq 2^{(n_I+n_X)n}$, i.e. $m^\varphi \geq 2^{3\varphi}$.
2. (On points 1 and 2): $m^2 \geq 2^{6n}$.
3. (On points 1, 2, 3): $m^3 \geq 2^{9n}$.

We see that all these conditions mean $m \geq O(2^{3n})$. Therefore, R2 (with $\varphi \geq 10$) gives a KPA in $0(2^{3n})$ (and the same in CPA-1).

C Why SQUARE, R1 and R2 are worse than TWO on F_3^d, $d \geq 10$

TWO is at present our best known attack on F_3^d, $d \geq 10$. This may look surprising, since SQUARE, R1 and R2 are better for $d = 6, 7, 8$ and 9. We will quickly present the main reason why we were not able to find anything better than TWO on F_3^d, $d \geq 10$. Essentially, the problem comes from the fact that with attacks like SQUARE, R1 or R2, we cannot have $n_X \leq n_S$, when $d \geq 10$. Therefore, N is still slightly larger for F_3^d than for a random permutation, but not by a factor of 2 (or more) anymore. We have computed the advantage obtained (by computing the standard deviation σ as explained in Appendix D) but we do not give the details since this gives an attack with a complexity larger than TWO. We have also tried different geometries for the equalities, (with $n_X \leq n_S$), but it has given a larger complexity than TWO.

Remark. Alternatively, we can see the problem like this: since $a \geq \lfloor \frac{d-1}{k} \rfloor$ we need at least 3 equations in X between indices 1 and 2. Therefore, when φ is changed in $\varphi + 2$, n_S becomes $n_S + 3$, but n_X becomes at least $n_X + 3$. When $d \geq 10$, we have to start on a rectangle with $n_X > n_S$. Therefore we will have $n_X > n_S$ for any φ. In fact, when $d \geq 10$, when φ increases, the probability of existence of the set of equations now decreases fast (instead of being about the same). Therefore small φ become better, and $\varphi = 2$ becomes better than $\varphi \geq 4$: TWO becomes again better than SQUARE, R1 and R2. Exactly the same problem occurs on F_k^d when $d \geq k^2 + 1$, for any $k \geq 3$. This is why TWO is the best known attack on generators F_k^d when $d \geq k^2 + 1$.

D Computation of the Standard deviations

In the attacks TWO, we have sometimes to compute the standard deviation $\sigma(N)$ of a variable N. We will explain here how these values $\sigma(N)$ can be computed. In TWO we will have $\sigma(N) \simeq \sqrt{E(N)}$ but this is not always true in M1, R2, R3, R4. $\sigma(N)$ can be computed in the same way for SQUARE, R1, R2, R3, R4, but we do not need it, as explained above. We will compute $\sigma(N)$ as explained in [16]. The starting point of the computation is to use this classical formula on the covariances:

If x_i are variables (independent or not), we have:

$$V(\sum_{i=1}^{\alpha} x_i) = \sum_{i=1}^{\alpha} V(x_i) + 2\sum_{i=1}^{\alpha} \sum_{j=i+1}^{\alpha} cov(x_i, x_j)$$

Where $cov(x_i, x_j)$ is the covariance of x_i and x_j:

$$cov(x_i, x_j) = E(x_j\, x_j) - E(x_i)\, E(x_j)$$

So

$$V(\sum_{i=1}^{\alpha} x_i) = \sum_{i=1}^{\alpha} V(x_i) + \sum_{i \neq j} [E(x_j\, x_j) - E(x_i)\, E(x_j)] \quad (1)$$

We will present here just one example of explicit computations of $\sigma(N)$ from this formula (1). All the other cases lead to similar computations.

Example: Computation of $\sigma(N)$ for F_3^6

Here we choose μ values for (I^1, I^2), for example we can assume that I^1 is constant, and that we have μ distinct values for I^2. Since I^1 is constant, we want to count the number N of (i,j) such that

$$I^2(i) = I^2(j) \text{ and } S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j)$$

The way to compute $E(N)$ and $\sigma(N)$ in such cases was explained in [16] p. 410-411. We give here only some details for F_3^6. Let E be the set of all possible (i,j), $i \neq j$, such that $I^2(i) = I^2(j)$. We have $|E| = \mu \cdot 2^n(2^n-1) \simeq \mu \cdot 2^{2n}$ (since I^1 is constant and we have 2^n possibilities for I^3). For $(i,j) \in E$, let $\delta_{ij} = 1 \Leftrightarrow S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j)$. We have:

$$N = \sum_{(i,j)\in E} \delta_{ij}$$

$$E(N) = \sum_{(i,j)\in E} E(\delta_{ij})$$

$$V(N) = \sum_{(i,j)\in E} V(\delta_{ij}) + \sum_{\substack{(i,j)\in E\ (k,l)\in E \\ (i,j)\neq(k,l)}} E(\delta_{ij}\cdot\delta_{kl}) - E(\delta_{ij})E(\delta_{kl}) \quad (*)$$

$$E(\delta_{ij}) = Pr_{f\in_R B_{3n}}(S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j))$$

where B_{3n} is the set of all permutations from $3n$ bits to $3n$ bits. For a random function, we have $E(\delta_{ij}) = \frac{1}{2^n}$. For a random permutation $E(\delta_{ij})$ is just slightly different: $E(\delta_{ij}) \simeq \frac{1}{2^n}$. More precisely, since $I^1(i) = I^1(j)$ and $I^2(i) = I^2(j)$ and $I^3(i) \neq I^3(j)$ we can prove that the exact value here is $E(\delta_{ij}) = \frac{2^{2n}}{2^{3n}-1} \simeq \frac{1}{2^n} + \frac{1}{2^{4n}}$.

$$V(\delta_{ij}) = E(\delta_{ij}^2) - (E(\delta_{ij}))^2 \simeq \frac{1}{2^n} - \frac{1}{2^{2n}}$$

$$E(\delta_{ij}\cdot\delta_{kl}) = Pr_{f\in_R B_{3n}}[S^3(i) \oplus S^3(j) = I^3(i) \oplus I^3(j) \text{ and } S^3(k) \oplus S^3(l) = I^3(k) \oplus I^3(l)]$$

Case 1: i, j, k, l are 4 distinct values. Then the computation shows that

$$E(\delta_{ij}\cdot\delta_{kl}) - E(\delta_{ij})E(\delta_{kl}) \leq \frac{4}{2^{5n}} + O(\frac{1}{2^{6n}})$$

Case 2: In $\{i, j, k, l\}$ we have 3 values. Then the computation shows that

$$E(\delta_{ij}\cdot\delta_{kl}) - E(\delta_{ij})E(\delta_{kl}) \leq \frac{3}{2^{5n}} + O(\frac{1}{2^{6n}})$$

Therefore, from $(*)$ we have:

$$V(N) = \frac{|E|}{2^n} + \frac{4|E|^2}{2^{5n}} + \text{negigeable terms}$$

$$V(N) = \frac{\mu\cdot 2^{2n}}{2^n} + \frac{4\mu^2\cdot 2^{4n}}{2^{5n}} + \text{negigeable terms}$$

$$V(N) \simeq \mu \cdot 2^n$$

Therefore, $\sigma(N) \simeq \sqrt{\mu \cdot 2^n}$, as claimed.

Appendices E,F,G,H,I for $k \geq 3$

E Attacks "TWO" for any $k \geq 3$

In this section, we explain the attack TWO. This attack does not use a rectangle but multiple collisions on 2 points (except for F_k^1) and is interesting for a small number of rounds or when we are attacking generators.

E.1 Attack TWO against F_k^1

We need one message in KPA and CPA-1. We just test if $S^k = I^1$.

E.2 Attack TWO against F_k^2

For the CPA-1 attack, we have $m = 2$. We choose two messages such that I^1 is constant. Then we test if $S^k \oplus I^2$ is constant. With a random permutation, the probability is $\frac{1}{2^n}$ and with F_k^2 the probability is 1.

We transform this attack into a KPA attack. We count the number of (i, j) such that $I^1(i) = I^1(j)$ and then we test if $S^k(i) \oplus I^2(i) = S^k(j) \oplus I^2(j)$. If $m \geq 2^{\frac{n}{2}}$, we can get such collisions and then the attack succeeds.

E.3 Attack TWO against F_k^d, $3 \leq d \leq k$

For the CPA-1 attack, we have $m = 2$ messages again. We choose $I^1, I^2, \ldots, I^{d-1}$ constant. then $X^1, X^2, \ldots, X^{d-2}$ will be constant but the X^{d-1} values will be pairwise distinct and $\forall i, j,\ X^{d-1}(i) \oplus X^{d-1}(j) = I^d(i) \oplus I^d(j)$.

Then we test if $S^k \oplus I^d$ is constant. As before with a random permutation, the probability is $\frac{1}{2^n}$ and one with F_k^d.

We transform this attack into a KPA attack. We look for $i < j$ such that:

$$I^1(i) = I^1(j), I^2(i) = I^2(j), \ldots, I^{d-1}(i) = I^{d-1}(j)$$

and then we test if $S^k(i) \oplus I^d(i) = S^k(j) \oplus I^d(j)$. When $m^2 \geq 2^{(d-1)n}$, we can get such collisions and the attack succeeds. Thus we have $m \geq 2^{\frac{d-1}{2}n}$ and the same complexity.

E.4 Attack TWO against F_k^{k+1}

We will concentrate the attack on the equation:

$$S^{k-1} = X^{k-1} \oplus f_{k+1}^{(k-1)}(X^k) \quad \text{with} \quad X^{k-1} = I^k \oplus f_1^{(k-1)}(I^1) \oplus_{i=2}^{k-1} f_i^{(k-i)}(X^{i-1})$$

The attack proceeds as follows:

1. We choose $I^1, I^2, \ldots, I^{k-1}$ constant. Then we have that $I^1, X^1, \ldots, X^{k-2}$ are constant and that

   $$\forall i, j,\ X^{k-1}(i) \oplus X^{k-1}(j) = I^k(i) \oplus I^k(j)$$

   and this implies that $i \neq j \Rightarrow X^{k-1}(i) \neq X^{k-1}(j)$.
2. Then, we look for indexes $i, j,\ i \neq j$ such that $S^k(i) = S^k(j)$. (Here we notice that $S^k = X^k$ since $d = k+1$). Then we test if $S^{k-1}(i) \oplus S^{k-1}(j) = I^k(i) \oplus I^k(j)$. (We have here $S^{k-1} = X^{k-1} \oplus f_{k+1}^{(k-1)}(S^k)$). When $m \simeq \sqrt{2^n}$, we can find such collisions and distinguish a random permutation from F_k^{k+1} and the complexity is about $\sqrt{2^n}$.
   As previously, we transform this attack into a KPA attack. We need to have $k-1$ equalities on the variables I^i, $1 \leq i \leq k-1$ and one equality on S^{k-1}. So, this attack is possible if $m \geq 2^{\frac{kn}{2}}$ with the same complexity.

E.5 Attack TWO against F_k^{k+2}

We will concentrate the attack on the equation:

$$S^{k-2} = X^{k-1} \oplus f_{k+1}^{(k-1)}(X^k) \oplus f_{k+2}^{(k-2)}(X^{k+1}) \quad \text{with} \quad X^{k-1} = I^k \oplus f_1^{(k-1)}(I^1) \oplus_{i=2}^{k-1} f_i^{(k-i)}(X^{i-1})$$

The attack proceeds as follows:

1. We choose $I^1, I^2, \ldots, I^{k-1}$ constant. Then we have that $I^1, X^1, \ldots, X^{k-2}$ are constant and that

$$\forall i, j,\ X^{k-1}(i) \oplus X^{k-1}(j) = I^k(i) \oplus I^k(j)$$

and this implies that $i \neq j \Rightarrow X^{k-1}(i) \neq X^{k-1}(j)$.

2. Then we look for indexes i, j such that $S^k(i) = S^k(j)$ and $S^{k-1}(i) = S^{k-1}(j)$. Here, we have the following relations:

$$\begin{aligned} &S^k = X^{k+1}; \text{since } d = k+2 \\ &S^{k-1} = X^k \oplus f_{k+2}^{(k-1)}(X^{k+1}) \\ &S^{k-2} = X^{k-1} \oplus f_{k+1}^{(k-1)}(X^k) \oplus f_{k+2}^{(k-2)}(X^{k+1}) \end{aligned}$$

So the X^{k-1} are pairwise distinct but we can get a collision (i, j) for the X^{k+1} variables and the X^k variables. Then we test if

$$S^{k-2}(i) \oplus S^{k-2}(j) = I^k(i) \oplus I^k(j)$$

So when $m^2 \geq 2^{2n}$ i.e. $m \geq 2^n$, we can get such collisions and the attack follows. We notice that this attack is possible since we have the condition $m \leq 2^n$ (only the variables I^k can take all the possible values).

This attack leads to a KPA attack with $m^2 \geq 2^{(k+1)n}$. This gives $m \geq 2^{\frac{k+1}{2}n}$ and this attack is valid since $2^{\frac{k+1}{2}n} \leq 2^{kn}$ (here $k \geq 3$).

E.6 Attack TWO against F_k^{k+u}, $2 \leq u \leq k-1$

We will concentrate the attack on the equation;

$$S^{k-u} = X^{k-1} \oplus_{i=k}^{d-1} f_{i+1}^{(2k-i-1)}(X^i) \quad (\text{here} \quad d = k+u)$$

We will count the number N of (i, j) such that $I^1(i) = I^1(j), I^2(i) = I^2(j), \ldots, I^{k-1}(i) = I^{k-1}(j), S^k(i) = S^k(j), S^{k-1}(i) = S^{k-1}(j), \ldots, S^{k-u+1}(i) = S^{k-u+1}(j)$ and $S^{k-u}(i) \oplus S^{k-u}(j) = I^k(i) \oplus I^k(j)$. For F_k^{k+u}, this last equation is a consequence of the other equations, i.e. of these $k-1$ equations in I and u equations in S. Therefore, the attack will succeed in KPA when $m^2 \geq 2^{(k+u-1)n}$, i.e. when $m \geq 2^{\frac{k+u-1}{2}n}$. In CPA-1, we will fix $I^1, I^2, \ldots, I^k$ to some values, and we will do this α times. The attack will succeed with $\alpha = 2^{(u-2)n}$ and the complexity in CPA-1 is here $\alpha \cdot 2^n = 2^{(u-1)n}$.

F SQUARE Attacks

We have already seen examples of SQUARE attacks with $k = 3$. Here we will present SQUARE for any value of k. (Remark: SQUARE attacks are a special case of R1 attacks when $\varphi = 4$, i.e. when we have a square of only 4 points in the rectangle of equations). Since we have seen some examples of SQUARE attacks, and since we will present in more detail R1 attacks, we will present here only the main ideas and results. We use the same notations as before. In these attacks, we have $n_X = d-1$ equations in X and $n_S = 2k-1$ equations in S. Therefore, we will have: $n_S \geq n_X \Leftrightarrow d \leq 2k$. When $d \geq 2k+1$, SQUARE attacks fail (but more general attacks like R1 attacls may still be valid). When $d \leq 2k$, the SQUARE attack will succeed in KPA if

$$\begin{cases} m^2 \geq 2^{\lceil \frac{d}{2} \rceil n} & \text{(condition between points 1 and 2 or 2 and 3 in figure 5)} \\ m^3 \geq 2^{dn} & \text{(condition between points 1, 2, and 3)} \\ m^4 \geq 2^{(d+k)n} & \text{(condition between points 1, 2, 3, 4)} \end{cases}$$

The last condition is dominant, therefore when $k+2 \le d \le 2k$, we have a complexity of SQUARE in KPA of: $2^{\frac{d+k}{4}n}$.

In CPA-1, the conditions become:

$$\begin{cases} m^2 \ge 2^{\lceil \frac{d-1}{2} \rceil n} & \text{(condition between points 1 and 2 or 2 and 3)} \\ m^3 \ge 2^{(d-1)n} & \text{(condition between points 1, 2, and 3)} \end{cases}$$

The last condition is dominant, therefore when $k+2 \le d \le 2k$, we have a complexity of SQUARE in CPA-1 of: $2^{\frac{d-1}{3}n}$.

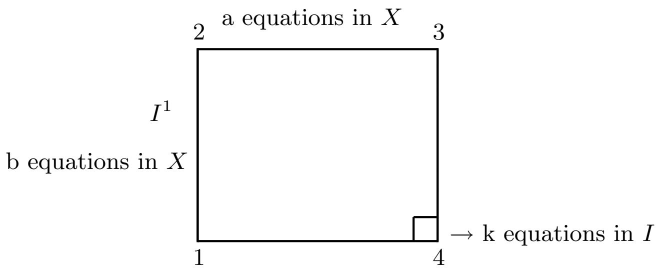

Fig. 5. SQUARE attack with $a+b=d-1$. Generally we will choose $a \simeq b \simeq \frac{d-1}{2}$

G Attacks "R1" for any $k \ge 3$ with $d \ge k+1$

We have already seen examples of M1 with $k=3$. Here we will present R1 for any value of k. When k is fixed, for very small values of d, TWO will be the best known attack. Then, when d increases, SQUARE and after that R1 will become the best known attack. Then, when d increases again, R2, R3 or R4 that we will see in Appendix H will become the best known attack. Finally, for very large d, TWO will become again the best known attack (see Appendix I).

Remark. The idea of R1 is to minimize the total number $n_I + n_X$ of needed equations in I and X. When this criteria is dominant, R1 will be the best attack.

In R1 we will count the number N of sets of plaintext/ciphertext pairs satisfying some conditions (I) and (S). We use a "rectangle" set of equalitites between the coordinates of the input variables $[I^1(i), \ldots, I^k(i)]$ and between the coordinates of the internal variables $X^i(j)$. We call φ the number of points of the rectangle, so φ is always greater than or equal to 4 in order to have a rectangle. (Attacks with equalities on only two points are the attacks "TWO").

G.1 Definition of R1

Let us consider φ plaintext/ciphertext pairs. The i-th pair is denoted by $[I^1(i), I^2(i), \ldots I^k(i)]$ for the plaintext and by $[S^1(i), S^2(i), \ldots, S^k(i)]$ for the ciphertext. We will fix some conditions on the inputs of the φ pairs. On the case of F_k^d, those conditions will turn into conditions on the internal state variables X^j due to the structure of the Feistel scheme. This conditions will imply equations on the outputs. On the case of a random permutation, equations on the outputs will only appear at random. By counting the sets of φ pairs satisfying the conditions on inputs and outputs, we can distinguish between F_k^d and a random permutation, since in the case of F_k^d the equations on the outputs appear not only at random, but a part of them is due to the conditions we set.

We first set the following conditions on the input variables:

$$(I) = \begin{cases} I^1(1) = I^1(2) \text{ and } I^1(3) = I^1(4) \text{ and } I^1(5) = I^1(6) \ldots \text{and } I^1(\varphi-1) = I^1(\varphi) \\ \forall i,\ 2 \le i \le k,\ I^i(1) \oplus I^i(2) = I^i(3) \oplus I^i(4) = \ldots = I^i(\varphi-1) \oplus I^i(\varphi) \end{cases}$$

Conditions on the first block I^1 are here to cancel the impact of function f_1, while conditions on other blocks are used to obtain differential equations on the internal state variables. These equations will then propagate to other rounds with some probability until they turn to equations on the outputs, which then can be detected.

In order for the previous conditions to propagate with high probability, we need some extra conditions on the internal state variables. We have $d-2$ internal state variables $X^1, X^2, \dots, X^{d-2}$ and $X^{d-1} = S^k$ is an output variable.

Let a be an integer, $1 \le a \le d-1$. We will choose a values of $\{1, 2, \dots, d-k\}$. (Therefore in R1 we have 2 parameters: φ and a. These values will be optimized depending on k and d). Let E be the set of these a values, and let F be the set of all integers i, $1 \le i \le d-1$ such that $i \notin E$. We have $|E| = a$ and $|F| = d-a-1$. Let (X) be the set of these equalities:

$$(X) = \begin{cases} \forall i \in E,\ X^i(1) = X^i(3) = \ldots = X^i(\varphi-1) \\ \forall i \in F,\ X^i(1) = X^i(2) \end{cases}$$

Between two different plaintext/ciphertext pairs i and j, $i \ne j$, we can have at most $k-1$ successive equalities on the variables $I^1, X^1, X^2, \dots, X^{d-1}$. Otherwise from k successive equalities we would get $I_i^1 = I_j^1, I_i^2 = I_j^2, \dots, I_i^k = I_j^k$, so the two messages would be the same. Therefore we must have: $\lfloor \frac{d}{k} \rfloor \le a \le d-1-\lfloor \frac{d-1}{k} \rfloor$. For the same reason we must have $\{d-k\} \in E$ since $d-1$, $d-2$, …, $d-k+1$ are in F.

From the conditions (I) and (X) and considering the equalities that we can derive from them with probability one, we will have:

$$(S) = \begin{cases} \forall i,\ 2 \le i \le k,\ S^i(1) = S^i(2) \text{ and } S^i(3) = S^i(4) \ldots \text{and } S^i(\varphi-1) = S^i(\varphi) \\ S^1(1) \oplus S^1(2) = S^1(3) \oplus S^1(4) = \ldots = S^1(\varphi-1) \oplus S^1(\varphi) \end{cases}$$

Consequently the conditions (S) can appear by chance, or due to the conditions (X).

Our KPA attack consists in counting the number N of rectangle sets of plaintext/ciphertext pairs satisfying the conditions (I) and (S). The obtained number can be divided into two parts: either the conditions (I) and (S) appear completely at random, or conditions (I) appear and conditions (S) are satisfied because (X) happened.

Figure 5 illustrates one rectangle set of our attack. Plaintext/ciphertext pairs are denoted by $1, 2, \dots, \varphi$. Two points are joined by an edge if the values are equal (for example $I^1(1) = I^1(2)$). We draw a solid edge if the equality appears with probability $\frac{1}{2^n}$ and a dotted line if the equality follows conditionally with probability 1 from other imposed equalities.

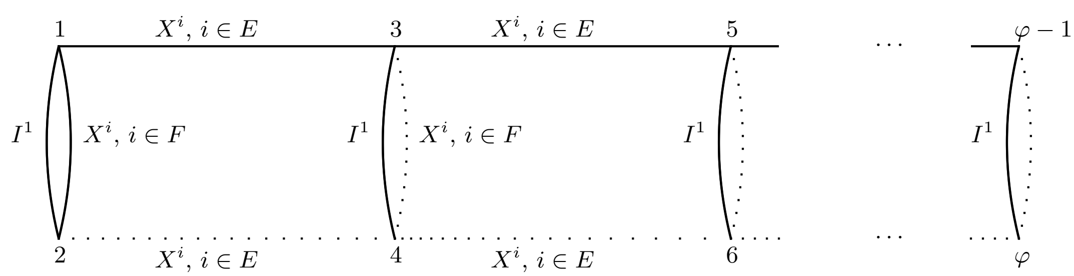

Fig. 6. Attack R1 on F_k^d

G.2 Properties of R1

We will denote by n_I the number of equalities in (I), and by n_S the number of equalities in (S). Similarly, we will denote by n_X the number of equalities in (X). Therefore n_X is the number of independant equalities in the X^i variables needed in order to get (S) from (I). In this attack R1 we have:

$$\begin{cases} \cdot n_I = \frac{k\varphi}{2} - k + 1 \\ \cdot n_S = \frac{k\varphi}{2} - 1 \\ \cdot n_X = a(\frac{\varphi}{2} - 2) + d - 1 \end{cases}$$

The value N is expected to be larger for a F_k^d than for a random permutation due to the fact that (S) can come from random reasons or from (X) in F_k^d. Therefore, it is natural, in order to get necessary and sufficient condition of success for R1, to evaluate the expectancy and the standard deviation of N in the case of F_k^d and in the case of random permutations. This can be done (by using the covariance formula as in [16] or by using approximation as in [6]), and in fact we did it. We have found that each time that R1 was better than TWO, we had $n_X \le n_S$. However, when $n_X \le n_S$ we can easily obtain sufficient condition of success for R1 without computing the standard deviations, since when $n_X \le n_S$ we will have for most permutations about 2 times more (or more) solutions with F_k^d than this random permutation. Therefore, a sufficient condition of success for R1 when $n_X \le n_S$ is to have that (X) and (I) can be satisfied with a non negligible probability. A sufficient condition for this is to have:

In KPA

Condition 1: $n_X \le n_S$.

Condition 2: $m^\varphi \ge 2^{n(n_I+n_X)}$.

Condition 3: $m^2 \ge 2^{(d-a)n}$.

Condition 4: $m^3 \ge 2^{dn}$ and more generally $\forall i$, $0 \le i \le \frac{\varphi}{2} - 1$, $m^{3+i} \ge 2^{(d+ia)n}$.

Condition 5: $m^4 \ge 2^{(d+k)n}$.

(Conditions 2, 3, 4, 5 are necessary. Conditions 1, 2, 3, 4, 5 are sufficient for success. Condition 1 is not necessary, but the computation of $\sigma(N)$ shows that R1 is not better than TWO when $n_X > n_S$.)

Condition 2 comes from the fact that we have about m^φ rectangles with φ points, and the probability that (I) and (X) are satisfied on one rectangle is $\frac{1}{2^{n(n_I+n_X)}}$.

Condition 3 comes from the fact that between points 1 and 2 we have $|F|$ equations in X^i, and one equation in I^1. Therefore in KPA we must have $m^2 \ge 2^{(|F|+1)n} = 2^{(d-a)n}$.

Condition 4 comes from the fact that between points 1, 2 and 3 we have $d-1$ equations in X^i, and one equation in I^1. Therefore we must have $m^3 \ge 2^{dn}$. Similarly between the points 1, 2, 3, 5, we must have: $m^4 \ge 2^{(d+a)n}$. And similarly between the points 1, 2, 3, 5, 7, ..., $(\varphi - 1)$, we must have: $m^{\frac{\varphi}{2}+1} \ge 2^{(d+a(\frac{\varphi}{2}-2))n}$.

Condition 5 comes from the fact that between points 1, 2, 3, 4, we have $d-1$ equations in X^i, 2 equations in I^1 and $(k-1)$ in I^2, I^3, ..., I^{k-1}.

It is easy to see that the conditions on any points are consequences of these 5 conditions. Moreover, if $m \ge 2^{an}$ (we will often, but not always, choose a like this), condition 4 can be changed with only $m^3 \ge 2^{dn}$.

CPA-1

In CPA-1 the sufficient conditions when $m \le 2^{(k-1)n}$ are:

Condition 1: $n_X \le n_S$.

Condition 2: $m^{(\frac{\varphi}{2}+1)} \ge 2^{n \cdot n_X}$.

Condition 3: $m^2 \ge 2^{(d-a-1)n}$.

Condition 4 and Condition 5: $m^3 \ge 2^{(d-1)n}$.

From these conditions we can compute the best parameters a and φ for any d and k, when d and k are fixed.

Remark. If we choose $n_X < n_S$ (instead of $n_X \le n_S$), the attacks are slightly less efficient but more spectacular since with a non negligible probability (I) and (S) are satisfied with F_k^d and not with random permutations. Moreover with $n_X < n_S$ it is still possible (with R2) to attack $3k-1$ rounds with less than 2^{kn} complexity.

Example 1

For F_3^7 in KPA we see from condition 5 that $m \ge 2^{\frac{5}{2}n}$, and for F_3^7 in CPA-1 we see from condition 4 that $m \ge 2^{2n}$. Since we have seen that with $a = 2$ and $\varphi = 6$ these bounds are obtained, it shows that $a = 2$ and $\varphi = 6$ give the optimal R1 attack on F_3^7.

Example 2

When d is small, in R1, condition 2 is the dominant condition. By definition, we denote by A and B the integers such that when $A \le d \le B$, condition 2 dominates in R1, and R1 is better than TWO. In order to have $n_I + n_X$ minimum, i.e.

$\frac{k\varphi}{2} - k + a(\frac{\varphi}{2} - 2) + d$ minimum, we will choose a minimum and φ minimum. Therefore, we will choose $a = \lfloor \frac{d}{k} \rfloor$ and from condition 1 we get that the minimum value for φ is $\varphi = \frac{2d-4a}{k-a}$, and then we have $n_X = n_S$.

Then the complexity in KPA given by condition 2 gives: $m \geq 2^{(k-\frac{k}{\varphi})n}$, with $\varphi = \frac{2d-4a}{k-a}$, with $a = \lfloor \frac{d}{k} \rfloor$.

In CPA-1 a similar computation gives a complexity in $2^{(k-1)(1-\alpha)n}$ with $\alpha = \frac{2k-\varphi}{k\varphi-\varphi+2k-2}$ and $\varphi = \frac{2d-4a}{k-a}$.

These are the best parameters and complexities for R1, when d is not too large (i.e. when condition 2 is dominant).

H Attacks “R2”, “R3”, “R4” for any $k \geq 3$ with $d \geq k$

In R1 we had two kinds of equations that we will call “horizontal” and “vertical” equations. By definition, the equations between indices 1 and 2, 3 and 4, ..., $\varphi - 1$ and φ will be call “vertical”, and the equations between indices i and j will be called “horizontal” when i and j have the same parity (i and j even, or i and j odd). For example, in R1, we will say that I^1 is vertical (since $I^1(1) = I^1(2)$, $I^1(3) = I^1(4)$, ..., $I^1(\varphi-1) = I^1(\varphi)$). In R1, S^i, $2 \leq i \leq k$ are also vertical, and therefore in R1 we have chosen X^i, $d-k+1 \leq i \leq d-1$ to be also vertical. When the dominant term in the complexity is given by $n_I + n_X$, this is the best choice. However, when d increases, sometimes the dominant term in the complexity is not given by $n_I + n_X$ (as we have seen for F_3^8). We have defined 4 families of attacks, R1, R2, R3, R4 from these 4 possibilities:

R1. I^1 is vertical and S^i, $2 \leq i \leq k$ are vertical. Therefore we will choose X^i, $d-k+1 \leq i \leq d-1$ to be vertical.

R2. I^1 is horizontal and S^i, $2 \leq i \leq k$ are vertical. Therefore we will choose X^i, $d-k+1 \leq i \leq d-1$ to be vertical.

R3. I^1 is vertical and S^i, $3 \leq i \leq k$ are horizontal. Therefore we will choose X^i, $d-k+2 \leq i \leq d-1$ to be horizontal.

R4. I^1 is horizontal and S^i, $4 \leq i \leq k$ are horizontal. Therefore we will choose X^i, $d-k+3 \leq i \leq d-1$ to be horizontal.

We will always denote by a the number of equations in X between the indices 1 and 3, by φ the total number of indices of the rectangle, by n_I the number of equations in I, by n_S the number of equations in S and by n_X the number of independant equations in X.

The R2 attacks

For example for R2 we will choose a values of $\{1, 2, \ldots, d-k\}$. Let E be the set of these a values, and let F be the set of all integers i, $1 \leq i \leq d-1$ such that $i \notin E$. We have $|E| = a$, $|F| = d-a-1$, and F contains all the $k-1$ values i, $d-k+1 \leq i \leq d-1$. For R2 we have:

$$(I) = \begin{cases} I^1(1) = I^1(3) = I^1(5) = \ldots = I^1(\varphi-1) \\ I^1(2) = I^1(4) = I^1(6) = \ldots = I^1(\varphi) \\ \forall i,\ 2 \leq i \leq k,\ I^i(1) \oplus I^i(2) = I^i(3) \oplus I^i(4) = \ldots = I^i(\varphi-1) \oplus I^i(\varphi) \end{cases}$$

$$(X) = \begin{cases} \forall i \in E,\ X^i(1) = X^i(3) = \ldots = X^i(\varphi-1) \\ \forall i \in F,\ X^i(1) = X^i(2) \end{cases}$$

$$(S) = \begin{cases} \forall i,\ 2 \leq i \leq k,\ S^i(1) = S^i(2) \text{ and } S^i(3) = S^i(4) \ldots \text{ and } S^i(\varphi-1) = S^i(\varphi) \\ S^1(1) \oplus S^1(2) = S^1(3) \oplus S^1(4) = \ldots = S^1(\varphi-1) \oplus S^1(\varphi) \end{cases}$$

The equations (X) have been chosen such that (S) is just a consequence of (I) and (X). Our attacks consist in counting the number N of rectangle sets of plaintext/ciphertext pairs satisfying the conditions (I) and (S). Figure 6 illustrates the equations for R2.

Between two different plaintext/ciphertext pairs i and j, $i \neq j$, we can have at most $k-1$ successive equalities on the variables I^1, X^1, ..., X^{d-1}. Therefore we have for R2: $\lfloor \frac{d-1}{k} \rfloor \leq a \leq d-1-\lfloor \frac{d}{k} \rfloor$. For R2 we have:

$$\begin{cases} \bullet n_I = \frac{k\varphi}{2} + \frac{\varphi}{2} - k - 1 \\ \bullet n_S = \frac{k\varphi}{2} - 1 \\ \bullet n_X = a(\frac{\varphi}{2} - 2) + d - 1 \end{cases}$$

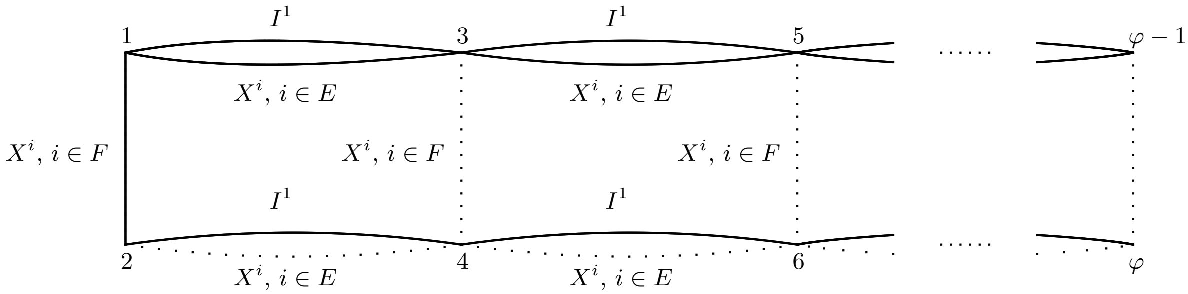

Fig. 7. Attack R2 on F_k^d

As we have explained for R1, sufficient conditions of success for R2 in KPA are these 5 conditions:
Condition 1: $n_X \leq n_S$.
Condition 2: $m^\varphi \geq 2^{n(n_I+n_X)}$, or for a generator with α functions: $\alpha \cdot 2^{kn\varphi} \geq 2^{n(n_I+n_X)}$.
Condition 3: $m^3 \geq 2^{dn}$, or for a generator with α functions: $\alpha \geq 2^{(d-3k)n}$.
Condition 4: $m^2 \geq 2^{(d-a-1)n}$, or for a generator with α functions: $\alpha \geq 2^{(d-1-a-2k)n}$.
Condition 5: $m^4 \geq 2^{(d+k)n}$, or for a generator with α functions: $\alpha \geq 2^{(d-3k)n}$.

Example for R2

In this example we want to attack a generator F_k^d, $3k \leq d \leq k^2$ with R2. Condition 1 gives $\varphi \geq \frac{2d-4a}{k-a}$ and $a \leq k-1$. From condition 4, we see that it will be interesting to choose the maximum value for a. Therefore we will choose $a = k-1$, since this is compatible with the other conditions. Now for $a = k-1$ and $\varphi \geq \frac{2d-4a}{k-a}$, we have $\alpha = O(2^{(d-3k)n})$ and the complexity is $O(\alpha \cdot 2^{kn}) = O(2^{(d-2k)n})$. This is better than TWO.

Remark. The condition $d \leq k^2$ comes from $a \leq k-1$ and $a \geq \lfloor \frac{d-1}{k} \rfloor$.

Example for R4

We will now present how to attack F_k^{3k-1} when $k \geq 5$ with a complexity less than 2^{kn}. This example is interesting since $3k-1$ is the maximum number of rounds that we can attack with a complexity lee than 2^{kn} (for $d = 3k$ the complexity of the best known attacks become $O(2^{kn})$ and for $d \geq 3k+1$ we need more than $O(2^{kn})$ computations). It is also interesting since in [6] Jutla was able to attack only $3k-3$ rounds with a complexity less than 2^{kn}. We will present only the main ideas. We will use the attack R4 with $a = k-1$, i.e. between 1 and 3 we have these $k-1$ equations: X^{d-1}, X^{d-2}, ..., X^{d-k+3}, plus X^k and X^{2k}.

Remark. With R2 (but not with R1) we can also attack F_k^{3k-1} (with $\varphi = 2k+2$ and $a = k-1$) with a complexity less than 2^{kn}, but the complexity of R4 will be slightly better.

In R4 with $a = k-1$, we have:

$$\begin{cases} \bullet n_I = \frac{k\varphi}{2} + \frac{\varphi}{2} - k - 1 \\ \bullet n_S = k\varphi - \frac{3\varphi}{2} - 2k + 3 \\ \bullet n_X = \frac{k\varphi}{2} + d - 2k - \frac{\varphi}{2} + 1 \end{cases}$$

Therefore when $d = 3k-1$, we have $n_X = \frac{k\varphi}{2} + k - \frac{\varphi}{2}$. $n_X \leq n_S$ gives $\varphi \geq 6 + \frac{6}{k-2}$. For $k \geq 5$, this means $\varphi \geq 8$ (φ is always even). Now if wee look at all the 5 conditions for the complexity, these conditions give: $m \geq 2^{(k-\frac{1}{8})n}$ in KPA, and $m \geq 2^{(k-\frac{1}{2})n}$ in CPA-1. These complexities are less than 2^{kn} as claimed.

I Summary of the results on F_k^d, $k \geq 3$, on TWO, SQUARE and Rectangle attacks

Table 4. Results on F_k^d for $k \geq 3$, on TWO, SQUARE and Rectangle attacks (i.e. without Multi-rectangle attacks. Multi-Rectangle attacks may have sometimes better complexities

	KPA	CPA-1
F_k^1	1	1
F_k^2	$2^{\frac{n}{2}}$, TWO	2
F_k^3	2^n, TWO	2
F_k^d, $2 \leq d \leq k$, TWO	$2^{\frac{d-1}{2}n}$,TWO	2
F_k^{k+1}	$2^{\frac{k}{2}n}$, TWO	$2^{\frac{n}{2}}$, TWO
F_k^{k+2}	$2^{\frac{k+1}{2}n}$, TWO and SQUARE	2^n, TWO
F_k^{k+3}	$2^{\frac{2k+3}{4}n}$, SQUARE	2^{2n},(TWO) or $2^{\frac{k+2}{3}n}$, SQUARE
F_k^d, $k+2 \leq d \leq 2k$	$2^{\frac{d+k}{4}n}$, SQUARE	$2^{(d-k-1)n}$,TWO or $2^{\frac{d-1}{3}n}$, SQUARE
F_k^{2k}	$2^{\frac{3k}{4}n}$, SQUARE	$2^{\frac{2k-1}{3}n}$,SQUARE
$\vdots$	$\vdots$	$\vdots$
F_k^{3k-1}	$2^{(k-\frac{1}{8})n}$, R2 $k=3$, R3 $k=4$, R4 $k \geq 5$	$2^{(k-\frac{1}{2})n}$, R2 $k=3$ or $k=4$, R4 $k \geq 5$
F_k^{3k}	2^{kn}, R2	2^{kn}, R2

When $d \geq k+4$, rectangle attacks R1, R2, R3, R4 and multi-rectangle attacks become more efficient than TWO attacks. The general formulas are difficult to get since there are many possible attacks (moreover we have not yet completely analysed all the possibilities for multi-rectangle attacks). It is possible to attack more than $3k$ rounds (with a complexity less than 2^{kn}) when $k \geq 6$. The exact number of rounds that we can get with complexity less than 2^{kn} with Multi-rectangle attacks is still on investigation.

J Multi-Rectangle attacks

An interesting problem is to design better attacks than 2 points attacks, or rectangle attacks. We have tried attacks with different geometries of equations (hexagons instead of rectangles, multi-dimensional cubes instead of 2-dimension rectangles, etc...). So far the best new attacks that we have found are "Multi-Rectangles attacks", i.e; attacks where some "rectangles" in I equations are linked with S equations. We will present here only two examples. These new attacks are very promising asymptotically (i.e. when n becomes large) but their efficiency from a practical point of view and the best design are still under investigation.

J.1 Example 1: Attack on F_6^{18}

With a 2 rectangles attack as in figure 8, it seems that we can attack F_6^{18} with a complexity less than 2^{6n}. Such an attack is much better than rectangle attacks. However if we use 2 rectangles of, say, 2×20 points, we will have a constant factor of $2 \times 20!$ in the complexity (since by symmetry each times we have a solution, we have $2 \times 20!$ solutions) and therefore such a theoretical attack might be of no practical interest.

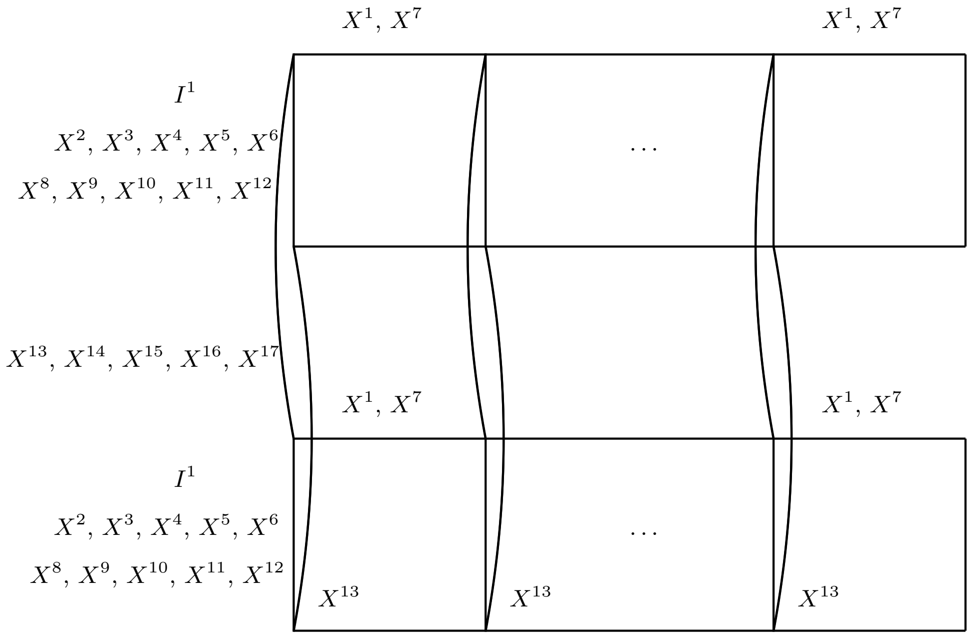

Fig. 8. Multi-rectangle attack on on F_6^{18}

J.2 Example 2: Attack on F_3^{12}

With hypercubes in dimension 4 as in figure 9, it may be possible to attack F_3^{12} with a complexity less than 2^{3n} (between 2 points of this structure we also have at most 3 equations in X). More generally it may be possible to attack $F_k^{k^2+k}$ with a complexity less than 2^{kn} by using hypercubes in dimension $k+1$ with at most k equations in X between 2 points. However these attacks are still under analysis and moreover there is certainly a huge constant factor in the complexity due to the fact that we use many points with many symmetries.

Remark: Multi-Rectangles attacks are also of interest for less rounds, for example in order to attack F_k^{2k} with a smaller complexity than rectangle attacks.

We will give more details in the extended version of this paper.

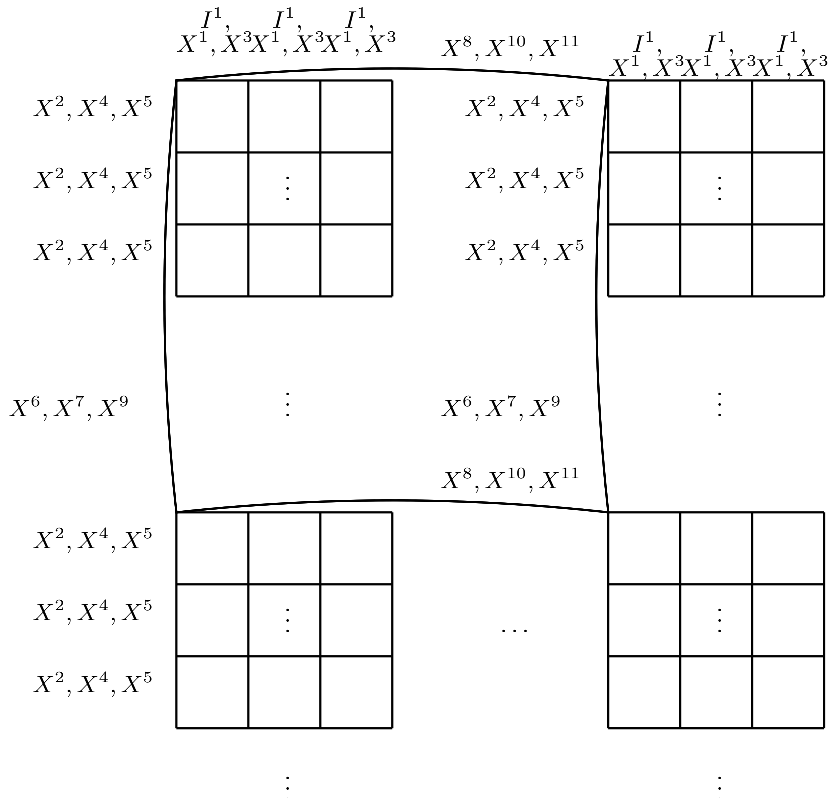

Fig. 9. Multi-rectangle attack on F_3^{12}