Security of balanced and unbalanced Feistel Schemes with Linear Non Equalities

Jacques Patarin

Université de Versailles
45 avenue des Etats-Unis, 78035 Versailles Cedex, France
jacques.patarin@prism.uvsq.fr

Abstract. In this paper we will study 2 security results "above the birthday bound" related to secret key cryptographic problems.
1. The classical problem of the security of 4, 5, 6 rounds balanced Random Feistel Schemes.
2. The problem of the security of unbalanced Feistel Schemes with contracting functions from $2n$ bits to n bits. This problem was studied by Naor and Reingold [14] and by [32] with a proof of security up to the birthday bound.
These two problems are included here in the same paper since their analysis is closely related, as we will see. In problem 1 we will obtain security result very near the information bound (in $O(\frac{2^n}{n})$) with improved proofs and stronger explicit security bounds than previously known. In problem 2 we will cross the birthday bound of Naor and Reingold. For some of our proofs we will use [2] submitted to Crypto 2010.

Key words: Luby-Rackoff constructions, Balanced random Feistel schemes, Unbalanced random Feistel schemes, Security Proofs, linear equalities and linear non equalities.

1 Introduction

A balanced "random Feistel scheme", also called a "Luby-Rackoff construction" or "generic balanced Feistel scheme" is a balanced Feistel scheme where all the internal round functions are random. We will denote by ψ^d the balanced random Feistel scheme with d rounds. From d random functions on n bits, ψ^d generates a random permutation on $2n$ bits. In their famous paper [8], M. Luby and C.Rackoff proved in 1988 the CPA-2 security (Adaptive Chosen Plaintext Attacks) of ψ^3 and the CPCA-2 security (Adaptive Chosen Plaintext and Chosen Cipher text Attacks) of ψ^4 when the number of queries q satisfies $q \ll \sqrt{2^n}$. Instead of using truly random internal functions, it is also possible to use pseudorandom functions. As pointed out in [8] the security is then obtained by a triangular argument: the security will be good with pseudorandom functions if the security is good with random functions and if the pseudorandomness of the functions is good. Luby-Rackoff constructions, and more generally Luby-Rackoff style analysis has inspired a considerable amount of research. One direction of research was to improve the security bound $q \ll \sqrt{2^n}$ in these schemes, or in other generic cryptographic constructions. This will be also the thema of this paper. This bound $q \ll \sqrt{2^n}$ is called the birthday bound because since we use random functions on n bits, we can avoid collisions on the outputs of these functions as long as $q \ll \sqrt{2^n}$. Another bound, called the information bound is of interest. It is the bound on the number of queries such that an adversary with infinite computer power would be able to find all the secret functions used. With secret functions on n bits, the information bound is $q \ll 2^n$. There is at present very little hope to obtain proofs of security in the number of computations instead of in the number of queries beyond the information bound since this will give a proof of $P \neq NP$ (or a slightly weaker version) the most famous problem in theoretical computer science. However the number of computations to be performed by an attacker is always at least the number of queries that he needs. Moreover in some schemes it is possible do deal with the multiple collisions and to obtain proofs of security beyond the birthday bound and smaller or equal to the information bound. It was proved by Patarin [17], and independently by Aiello and Verkatesan [1], that for ψ^3 and ψ^4 the birthday bound $q \ll 2^n$ is the best possible bound. However, when the number of rounds increases, and in other cryptographic constructions better security bounds can be proved. These proofs "beyond the birthday bound" are generally not easy. Four very different strategies have been used to obtain these proofs of security. We quickly present here some of the papers based on these strategies.

Strategy 1

Aiello and Verkatesan ([1],[24]) have obtained a proof of security in $q \ll 2^n$ (the information bound) for the Benes schemes, a non invertible analog of the Feistel schemes. The main idea of the proof relies on the analysis of "circles" of equalities. This is clever but unfortunately it seems to be relatively easy to use this strategy only for some specific schemes such as Benes schemes.

Strategy 2

Maurer and Prietrzak [12] increased the number d of rounds in ψ^d in order to prove a security that tends to the information bound when d tends to infinity. Similarly Lucks [9] has proved that the Xor of d independent pseudorandom permutations is indistinguishable from a random function on n bits with security $q \ll 2^{\frac{d}{d+1}n}$. This bound tends to the information bound $q \ll 2^n$ when d tends to infinity. From a practical point of view $d =$ about 10 for Lucks and $d =$ about 100 for Maurer and Prietrzak will give security bounds very near the information bound.

Strategy 3

Bellare and Impagliazzo [3] have proved that for the problem of Lucks, even for $d = 2$, the security is already in $q \ll 2^n$, the information bound. Their proof strategy is based on some probability theory theorems such as Azuma inequality and Chernoff bounds. Unfortunately they did not present all the parts of their proof in detail, and some o functions are introduced that are not given explicitly.

Strategy 4

Patarin [17, 22, 23] kept the number of rounds d relatively small, and near the minimum for security above the birthday bound: $d = 5, 6$, or 7. He has successively imposed for these schemes the proved security bound. The last results [22, 23] gave the CPCA-2 security for ψ^5 when $q \ll 2^n$, i.e. the information bound for the minimum number of rounds (5). Patarin strategy is based on transforming the cryptographic security problem in the analysis of the number of solutions of systems of linear equalities and linear non equalities. However, its last papers on ψ^d schemes are often considered difficult to read, and they introduce some O functions that are not given explicitly. More precisely, he has published an explicit bound for $q \ll 2^{\frac{2}{3}n}$, $q \ll 2^{\frac{3}{4}n}$, for $q \ll 2^{\frac{4}{5}n}$ and he has explained how to get one in $q \ll 2^{\frac{k-1}{k}n}$ for all integer k. Now since the coefficients in these bounds do not grow faster than exponentially in k, he could claim $q \ll 2^n$, but it was difficult to obtain a precise general bound without O functions.

In this paper, we will also follow Patarin general strategy. However, we have found some improvements in the proofs, some simplifications, and we will be able to obtain explicit security bounds, with no non specified O functions. We will also apply these techniques to random unbalanced Feistel schemes with contracting functions (that we will denote G_k^d). Since it is a bit long to speak of "the analysis of the number of solutions of linear equalities and linear non equalities in finite groups", this theory has been called in [2] by a nickname: "mirror theory". The term "mirror" refers to the multiple induction properties that we have in this theory, sometimes these inductions are obvious, sometimes they are not. We will explain why we believe that this almost impossible to avoid Mirror Theory in the analysis of random Feistel schemes or similar schemes near the information bounds when the number of rounds is fixed. In some of his papers, Patarin calls "coefficient H technique" his technique to transform a cryptographic security problem in a Mirror theory problem. However, when this transformation is done, we have to evaluate the Mirror theory problem.

Unbalanced Feistel Schemes
Let G_k^d be the unbalanced Feistel scheme that use random functions from kn bits to n bits in order to obtain a pseudorandom permutation from $(k+1)n$ bits to $(k+1)n$ bits. (For example: $\psi^d = G_1^d$). Here since we use random functions from kn bits to n bits, the birthday bound is $q \ll \sqrt{2^{kn}}$, and the information bound is $q \ll 2^{kn}$. It is not easy to evaluate the security of G_k^d schemes, $k \geq 2$, even beyond the birthday bound. It is however a very interesting problem from a practical point of view since, as pointed out by Naor and Reingold [14] it would allow to design secure permutations with relatively small length. In [14] a very beautiful idea was used to prove the birthday bound security of schemes similar to G_k^d: they have introduced 2-wise independent permutations. 2-wise permutations are very simple, such as $y = ax + b$, a and b secret in a finite field. In [14], they proved CPA-2 security up to the birthday bound for $G_k^d \circ g$ where g is a secret 2-wise permutation when $d \geq k+1$, and CPCA-1 security up to the birthday bound for $h^{-1} \circ G_k^d \circ g$ where g and h are two secret 2-wise permutations when $d \geq k+1$. Thanks to this clever idea to introduce these simple 2-wise permutations, the proof of security became magically simple, but

is however limited to the birthday bound. Recently Yun,Park and Lee [32] have proved the birthday bound tor G_k^d schemes when $d \geq 2k+1$ in CPA-2 and when $d \geq 3k+1$ for CPCA-2. At the end of this paper we will improve these bounds and obtain proof of security for G_k^d schemes above the birthday bound. For G_2^6 for example, the birthday bound is in $q \ll 2^n$ while we will prove in this paper security for $q \ll 2^{1.5n}$. Here the information bound is $q \ll 2^{2n}$.

2 Balanced Feistel Schemes

Definition of Ψ^k

We recall the definition of the balanced Feistel Schemes, i.e. the classical Feistel schemes. Let n be an integer. $I_n = \{0,1\}^n$. Let F_n be the set of all applications from I_n to I_n. Let B_n be the set of all permutations from I_n to I_n. Let f_1 be a function of F_n. Let L, R, S and T be four n-bit strings in I_n. Let $\Psi(f_1)$ denotes the permutation of B_{2n} such that:

$$\Psi(f_1)[L,R] = [S,T] \overset{\text{def}}{\Leftrightarrow} \begin{cases} S = R \\ T = L \oplus f_1(R) \end{cases}$$

More generally if $f_1, f_2, \ldots, f_k$ are k are functions of F_n, let $\Psi^k(f_1, \ldots, f_k)$ denotes the permutation of B_{2n} such that:

$$\Psi^k(f_1, \ldots, f_k) = \Psi(f_k) \circ \cdots \circ \Psi(f_2) \circ \Psi(f_1).$$

The permutation $\Psi^k(f_1, \ldots, f_k)$ is called a 'balanced Feistel scheme with k rounds' or shortly Ψ^k. When $f_1, \ldots, f_k$ are randomly and independently chosen in F_n, then $\Psi^k(f_1, \ldots, f_k)$ is called a 'random Feistel scheme with k rounds' or a 'Luby-Rackoff construction with k rounds'.

Notations for 4 rounds

- We will denote by $[L_i, R_i]$, $1 \leq i \leq q$, the q cleartexts. These cleartexts can be assumed to be pairwise distinct, i.e. $i \neq j \Rightarrow L_i \neq L_j$ or $R_i \neq R_j$.
- $[R_i, X_i]$ is the output after one round, i.e.

$$\forall i, 1 \leq i \leq q, X_i = L_i \oplus f_1(R_i).$$

- $[X_i, Y_i]$ is the output after two rounds, i.e.

$$\forall i, 1 \leq i \leq q, Y_i = R_i \oplus f_2(X_i) = R_i \oplus f_2(L_i \oplus f_1(R_i)).$$

- $[Y_i, S_i]$ is the output after three rounds, i.e.

$$\forall i, 1 \leq i \leq q, S_i = X_i \oplus f_3(Y_i) = L_i \oplus f_1(R_i) \oplus f_3(Y_i).$$

- $[S_i, T_i]$ is the output after 4 rounds, i.e.

$$\forall i, 1 \leq i \leq q, T_i = Y_i \oplus f_4(S_i).$$

When we will study Ψ^5 we will keep a central Ψ^4 but add one more round at the beginning. When we will study Ψ^6 we will keep a central Ψ^4 but add one more round at the beginning, and one more round at the end. Our aim is to present proofs of security in KPA (Known Plaintext Attacks) for Ψ^4, in CPA-2 for Ψ^5, and in CPCA-2 for Ψ^6 when the number of queries satisfies $q \ll \frac{2^n}{n}$, with a simple explicit bound for q (not just the undefined $q \ll \frac{2^n}{n}$). It is possible to improve the coefficient $\frac{2^n}{n}$ and to have $q \ll 2^n$ instead of $q \ll \frac{2^n}{n}$, and it is also possible to prove CPCA-2 for Ψ^5 instead of Ψ^6 when $q \ll \frac{2^n}{n}$ but we will not consider these refinements here, we will concentrate on getting simple explicit bounds.

Hints about these refinements:

• $q \ll \frac{2^n}{\sqrt{n}}$ instead of $q \ll \frac{2^n}{n}$ can be obtained by noticing that the value of α of the Theorem $P_i \oplus P_j$ of [2] is $\alpha = O(\frac{q^2}{2^n})$.

• $q \ll 2^n$ instead of $q \ll \frac{2^n}{n}$ can be obtained by noticing that most of the frameworks $\mathcal{F}$ with $\xi_{max} \geq n$ also have many very small blocks with 2 or 3 variables for example. However in this paper we look for bounds in $q \ll \frac{2^n}{n}$.

• CPCA-2 security for Ψ^5 instead of Ψ^6 can be proved for $q \ll \frac{2^n}{n}$ by using Theorem 14 of Appendix A. and by distinguishing direct and inverse queries. However in this paper we will analyze CPCA-1 security for Ψ^6 and not Ψ^5.

Definition of H for Ψ^k

When $[L_i, R_i], [S_i, T_i], 1 \leq i \leq q$, is a given sequence of $2q$ values of I_{2n}, we will denote by $H_k(L, R, S, T)$ or in short by H_k, or simply by H, the number of k-tuples of functions $(f_1, \ldots f_k)$ of F_n^k such that:

$$\forall i,\ 1 \leq i \leq q,\ \Psi^k(f_1, \ldots, f_p)[L_i, R_i] = [S_i, T_i].$$

We will analyze the properties of these H values in order to obtain our security results.

Frameworks for Ψ^k

Definition.

For 4 rounds, Ψ^4, let us define a "framework" as a set of equations $X_i = X_j$ or $Y_i = Y_j$. We will say that two frameworks are equal if they imply exactly the same set of equalities in X and Y.

Definition

Let $\mathcal{F}$ be a framework. We will denote by Weight $(\mathcal{F})$ the number of variables (X_i, Y_i), $1 \leq i \leq q$, $X_i \in I_n$, $Y_i \in I_n$, that satisfy $\mathcal{F}$, i.e. such that the equalities $X_i = X_j$ or $Y_i = Y_i$ are exactly those in $\mathcal{F}$.

Theorem 1 *If we denote by x the number of independent equalities $X_i = X_j$ of a framework $\mathcal{F}$ and by y the number of independent equalities of $\mathcal{F}$, we have: Weight $(\mathcal{F}) = J_{q-x} \cdot J_{q-y}$ where J_i denotes $J_i = 2^n(2^n - 1) \ldots (2^n - i + 1)$.*

Proof. We have fixed x values X_i from the other values, and all the other variables X_j must be pairwise distinct. Therefore for $(X_1, X_2, \ldots, X_q)$ we have exactly J_{q-x} possibilities. Similarly for $(Y_1, Y_2, \ldots, Y_q)$ we have exactly J_{q-y} possibilities.

An exact formula for H and Ψ^4 with "frameworks"

Let $[L_i, R_i], [S_i, T_i]$, $1 \leq i \leq q$ be a given sequence of $2q$ values of I_{2n}. Let r be the number of independent equalities $R_i = R_j$, $i \neq j$, and let s be the number of independent equalities $S_i = S_j$, $i \neq j$. As above, let x be the number of independent equalities $X_i = X_j$ for a framework $\mathcal{F}$, and by y be the number of independent equalities $Y_i = Y_j$ in $\mathcal{F}$.

Theorem 2 *The exact formula for H_4 (i.e. for Ψ^4) is:*

$$H_4 = \frac{|F_n|^4 \cdot 2^{n(r+s)}}{2^{4nq}} \sum_{\textit{all frameworks } \mathcal{F}} 2^{n(x+y)} [\textit{ Number of } X_i \textit{ satisfying } (C1)]$$

$$\cdot [\textit{ Number of } Y_i \textit{ satisfying } (C2)]$$

where

$$(C_1) : \begin{cases} R_i = R_j \Rightarrow X_i \oplus X_j = L_i \oplus L_j \\ Y_i = Y_j \textit{ is in } \mathcal{F} \Rightarrow X_i \oplus X_j = S_i \oplus S_j \\ \textit{The only equations } X_i = X_j, i < j, \textit{ are exactly those implied by } \mathcal{F}. \end{cases}$$

$$(C_2) : \begin{cases} S_i = S_j \Rightarrow Y_i \oplus Y_j = T_i \oplus T_j \\ X_i = X_j \textit{ is in } \mathcal{F} \Rightarrow Y_i \oplus Y_j = R_i \oplus R_j \\ \textit{The only equations } Y_i = Y_j, i < j, \textit{ are exactly those implied by } \mathcal{F}. \end{cases}$$

Proof. Let (D) be these conditions:
1. $\forall i,\ 1 \leq i \leq q,\ R_i = S_i$.
2. $\forall i, j,\ 1 \leq i \leq q,\ 1 \leq j \leq q,\ R_i = R_j \Rightarrow T_i \oplus L_i = T_j \oplus L_j$
For one round Feistel, Ψ^1, it is easy to see that if (D) is not satisfied then $H_1 = 0$, and if (D) is satisfied, then $H_1 = \frac{|F_n| \cdot 2^{nr}}{2^{nq}}$. This is an exact formula for H_1. By composing two Ψ^1 we get like this an exact formula for H_2, then H_3 and finally (by looking for all possible intermediate variables X_i, Y_i) for H_4 the formula of Theorem 2. We do not give all the details here because this proof is not difficult, and also because Theorem 2 is not new: it was already mentioned in [22].

3 KPA security for Ψ^4

Our results on Ψ^4 for proving KPA security will be based on this Theorem.

Theorem 3 *For random values $[L_i, R_i], [S_i, T_i]$, $1 \leq i \leq q$ such that $[L_i, R_i]$, $1 \leq i \leq q$ are pairwise distinct, with probability $\geq 1 - \beta$ we have:*
1. The number H of $(f_1, f_2, f_3, f_4) \in F_n^4$ such that $\Psi^4(f_1, f_2, f_3, f_4)[L_i, R_i] = [S_i, T_i]$ satisfies $H \geq \frac{|F_n^4|}{2^{nq}}(1 - \alpha)$
2. α and β can be chosen $\ll 1$ when $q \ll \frac{2^n}{n}$. (Moreover we will obtain explicit values for α and β: $\alpha = \frac{2q}{2^n}$, $\beta = \frac{2q}{2^n}$ when $q \leq \frac{2^n}{67}$).

To prove Theorem 3, we will evaluate H from the formula given by Theorem 2, i.e. with the sum on all frameworks $\mathcal{F}$. When a framework $\mathcal{F}$ is fixed, 5 cases can occur:

Case 1. A contradiction appears by linearity in the equations generated by $(C1)$. In this case we will say that we have a "circle" of equalities in $R, Y_{\mathcal{F}}$ that generates a "circle" of equalities in X by $(C1)$.

Case 2. A contradiction appears by linearity in the equations generated by $(C2)$. In this case we will say that we have a "circle" of equalities in $S, X_{\mathcal{F}}$ that generates a "circle" of equations in Y by $(C2)$.

Case 3. No contradiction occurs by linearity but $\xi_X q$ is not $\ll \frac{2^n}{n}$, where ξ_X denotes the maximum number of X_j variables fixed in the equations $(C1)$ when one variable X_i is fixed. In this case we will say that we have a big "line" of equalities $R, Y_{\mathcal{F}}$ that generates a big "line" of equalities in X by $(C1)$.

Case 4. No contradiction occurs by linearity but $\xi_Y q$ is not $\ll \frac{2^n}{n}$, where ξ_Y denotes the maximum number of Y_j variables fixed in the equations $(C1)$ when one variable Y_i is fixed. In this case we will say that we have a big "line" of equalities $S, X_{\mathcal{F}}$ that generates a big "line" of equalities in Y by $(C2)$.

Case 5. No contradiction occurs by linearity and $\xi_X q \ll \frac{2^n}{n}$ and $\xi_Y q \ll \frac{2^n}{n}$.

In order to prove Theorem 3 we will first prove that Cases 1, 2, 3, 4 appear with a negligible probability when the values $[L_i, R_i], [S_i, T_i]$ are randomly chosen, and when $\mathcal{F}$ is randomly chosen with a distribution of probability proportional to $Weight(\mathcal{F})$. That is what is done in Appendix B

Theorem 4 *If $q \leq \frac{2^n}{67n}$, then for every KPA with q (random) known plaintexts we have: $Adv^{PRF} \leq \frac{4q}{2^n}$ where Adv denotes the advantage to distinguish Ψ^4 from a random function $f \in_R F_{2n}$, and $Adv^{PRP} \leq \frac{4q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}}$ where Adv^{PRP} denotes the advantage to distinguish Ψ^4 from a random permutation $f \in_R B_{2n}$.*

Proof. This comes immediately from Theorem 10 of Appendix A, Theorem 3 and the classical pseudorandom function/ pseudorandom permutation switching lemma (i.e. First property in the proof of Lemma 3).

4 CPA-2 security for Ψ^5

Theorem 5 *If $q \leq \frac{2^n}{67n}$, then for every CPA-2 with q adaptive chosen plaintexts, we have: $Adv^{PRF} \leq \frac{5q}{2^n}$ and $Adv^{PRR} \leq \frac{5q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}}$. Where Avd^{PRF} denotes the advantage to distinguish Ψ^5 from $f \in_R F_{2n}$, and Adv^{PRP} denotes the advantage to distinguish Ψ^5 from $f \in_R B_{2n}$.*

Proof. See appendix C.

5 CPCA-2 security for Ψ^6

Of course, from our CPA-2 security for Ψ^5 by using [12] we obtain a CPCA-2 proof of security for Ψ^{10}, with a precise bound for Ψ^{10} since we have a precise bound for Ψ^5. I.e. the composition of two CPA-1 schemes (and therefore of two CPA-2 schemes also) with independent keys gives a CPCA-2 scheme: $Adv^{CPCA-2}_{F\circ G^{-1}} \leq Adv^{CPA-1}_{F} + Adv^{CPA-1}_{G}$. It is also possible to write $\Psi^6 = \Psi^1 \circ \Psi^4 \circ \Psi^1$, and to proceed with the $[S_i, T_i]$ values of Ψ^6 exactly as we did in the previous sections with the $[L_i, R_i]$ values of Ψ^5. Then for Ψ^6 we obtain:

Theorem 6 *For all pairwise distinct $[L_i, R_i]$, $1 \leq i \leq q$ and for all pairwise distinct $[S_i, T_i]$, $1 \leq i \leq q$ the number H of $(f_1, f_2, f_3, f_4, f_5, f_6) \in F_n^6$ such that $\forall i,\ 1 \leq i \leq q$,*

$$\Psi^6(f_1, f_2, f_3, f_4, f_5, f_6)[L_i, R_i] = [S_i, T_i]$$

satisfies $H \geq \frac{|F_n|^6}{2^{2nq}}(1-\alpha)$ where α can be chosen $\ll 1$ when $q \ll \frac{2^n}{n}$. More precisely, we can choose $\alpha = \frac{8q}{2^n}$ if $q \leq \frac{2^n}{67n}$.

Proof. We write $\Psi^6 = \Psi^1 \circ \Psi^4 \circ \Psi^1$. We use these notation:
Cleartexts: $[L_i, R_i]$, $1 \leq i \leq q$
1 round: $[R_i, X'_i]$ with $X'_i = L_i \oplus f_1(R_i)$
2 rounds: $[X'_i, X_i]$
3 rounds: $[X_i, Y_i]$
4 rounds: $[Y_i, Y'_i]$ with $Y'_i = T_i \oplus f_6(S_i)$
5 rounds: $[Y'_i, S_i]$
6 rounds: $[S_i, T_i]$
The analysis of the circle and lines in $Y', X_{\mathcal{F}}$ ca now be done exactly as we did the analysis of the circles end lines in $X', Y_{\mathcal{F}}$ for Ψ^5 (symmetry of the hypothesis). Therefore, if $q \leq \frac{2^n}{64}$, for all pairwise distinct values $[L_i, R_i]$, $1 \leq i \leq q$ and for all pairwise distinct values $[S_i, T_i]$, $1 \leq i \leq q$:

$$\sum_{f_1\in F_n}\sum_{f_6\in F_n}\sum_{\text{all frameworks } \mathcal{F} \text{ of Case 5}} Weight(\mathcal{F}) \geq |F_n|^2 \cdot 2^{2qn}\left(1-\frac{8q}{2^n}\right) \tag{1}$$

(same proof as for Lemma 12). Thus, if $q \leq \frac{2^n}{67n}$, we can use the theorem $P_i \oplus P_j$ with $\xi_{max} \leq n$ i.e. Theorem 6 of [2], and (1) gives Theorem 6 with $\alpha = \frac{8q}{2^n}$.

Theorem 7 *If $q \leq \frac{2^n}{128n}$, then for every CPCA-2 with q adaptive chosen plaintexts or chosen ciphertexts, we have: $Adv^{PRP} \leq \frac{8q}{2^n} + \frac{q^2}{2\cdot 2^{2n}}$ where Adv^{PRP} denote the advantage to distinguish Ψ^6 from $f \in_R B_{2n}$.*

Proof. This comes immediately from Theorem 14 (or the variant given) of Appendix A, and Theorem 6.

6 Unbalanced Feistel schemes with Contracting Functions

The balanced Feistel schemes Ψ^k that we have seen were permutation from $2n$ bits to $2n$ bits generated by k functions from n bits to n bits. They were obtained by composition of the scheme Ψ^1 such that

$$\Psi(f_1)[L,R] = [S,T] \overset{\text{def}}{\Leftrightarrow} \begin{cases} S = R \\ T = L \oplus f_1(R) \end{cases}$$

Similarly, let consider the scheme G_3^1 defined by

$$\forall [I^1, I^2, I^3] \in I_n^3,\ \forall [S^1, S^2, S^3] \in I_n^3,$$

$$G_3^1(f_1)[I^1, I^2, I^3] = [S^1, S^2, S^3] \overset{\text{def}}{\Leftrightarrow} \begin{cases} S^1 = I^2 \\ S^2 = I^3 \\ S^3 = I^1 \oplus f_1([I^2, I^3]) \end{cases}$$

where f_1 is a function from $2n$ bits to n bits. And let $G_3^d(f_1, \ldots, f_d) = G_3^1(f_d) \circ \ldots \circ G_3^1(f_2) \circ G_3^1(f_1)$. G_3^d schemes are permutations from $3n$ bits to $3n$ bits generated from d functions from $2n$ bits to n bits. More general G_k^d schemes can be defined, with any k integer, $k \geq 2$, as permutation from kn bits to kn bits generated from functions from $(k-1)n$ bits to n bits, and many different design of unbalanced Feistel schemes exist. In this paper, however, we will only study G_3^d schemes as unbalanced Feistel schemes. Similar security results above the birthday bound can be obtained on more general unbalanced Feistel schemes, but we will not study these generalizations. We know from [28] that CPA-1 attacks on G_3^5 exist with complexity about 2^n. On these G_3^d schemes, since we use internal functions from $2n$ bits to n bits, the birthday bound is 2^n, and the information bound is 2^{2n}. Recently, Yun, Park and Lee [32] have proved the CPA-2 security of G_k^d schemes when $d \geq 2k+1$ and the CPCA-2 security when $d \geq 3k+1$ up to the birthday bound. For $k = 3$, this means CPA-2 security when $q \ll 2^n$ for G_3^7 and CPCA-2 security when $q \ll 2^n$ for G_3^{10}. Noar and Reingold, as explained in the introduction of this paper, have also obtained the birthday bound security on small transformations of the G_k^d schemes. Our aim in the paper is to obtain a proof of security for $q \ll 2^{\frac{3n}{2}}$, i.e. beyond the birthday bound, but not yet the information bound. We know that in CPA-1 we will need at least 6 rounds. In fact G_3^6 will play a central role in our proof on G_k^d securities, in a similar way as Ψ^4 has played a central role in our proof on Ψ^d securities, and we will first study the KPA security of G_3^6. Our first step is to obtain an exact value for H for G_3^d. We will obtain first such formulas for G_3^1, then G_3^2 etc... until G_3^6. Due to the lack of space, we present here only the formulas for G_3^1 and G_3^6 (it is however not difficult to find them). We denote by $F_{n,2}$ the set of all applications from I_{2n} to I_n. Therefore $|F_{n,2}| = 2^{n.2^n}$.

1 Round

Theorem 8 *Let H_1 be the number of $f_1 \in F_{n,2}$ such that*

$$\forall i,\ 1 \leq i \leq q,\ [S_i^1, S_i^2, S_i^3] = G_3^1[I_i^1, I_i^2, I_i^3]$$

$$\textit{i.e. such that } [S_i^1, S_i^2, S_i^3] = [I_i^2, I_i^3, I_i^1 \oplus f_1([I_i^2, I_i^3])]$$

Let (D) be these conditions:

$$(D) : \begin{cases} \forall i,\ 1 \leq i \leq q,\ (I_i^2 = S_i^1) \,\text{and}\, (I_i^3 = S_i^2) \\ \forall i,\ 1 \leq i \leq q,\ \forall j,\ 1 \leq j \leq q,\ [I_i^2, I_i^3] = [I_j^2, I_j^3] \Rightarrow S_i^3 = S_j^3 \end{cases}$$

If (D) is not satisfied, then $H_1 = 0$. If (D) is satisfied, then $H_1 = \frac{|F_{n,2}|}{2^{nq}} \cdot 2^{nr}$ where r is the number of independent equalities $[I_i^2, I_i^3] = [I_j^2, I_j^3],\ i \neq j$.

6 rounds
We will denote by $X_1, \dots, X_q$, $Y_1, \dots, Y_q$, $Z_1, \dots, Z_q$ the intermediate variables that appear when we compute G_3^6 from 6 compositions of G_3^1.
0 round: $[I^1, I^2, I^3]$.
1 round: $[I^2, I^3, X]$.
2 rounds: $[I^3, X, Y]$.
3 rounds: $[X, Y, Z]$.
4 rounds: $[Y, Z, S^1]$.
5 rounds: $[Z, S^1, S^2]$.
6 rounds: $[S^1, S^2, S^3]$.
The inputs (i.e. the cleartext) for message i will be denoted $[I_i^1, I_i^2, I_i^3]$, and similarly the output (i.e. the ciphertext) $[S_i^1, S_i^2, S_i^3]$ and after 3 rounds $[X_i, Y_i, Z_i]$.
A framework $\mathcal{F}$, for G_3^6, is a set of equalities of these types:
$[I_i^3, X_i] = [I_j^3, X_j]$, $i \neq j$, or $[X_i, Y_i] = [X_j, Y_j]$, $i \neq j$, or $[Y_i, Z_i] = [Y_j, Z_j]$, $i \neq j$, or $[Z_i, S_i^1] = [Z_j, S_j^1]$, $i \neq j$.
The number of such independent equalities in $\mathcal{F}$ will be denoted by f. We denote by X the sequence $(X_1, \dots, X_q)$, by Y the sequence $(Y_1, \dots, Y_q)$ and by Z the sequence $(Z_1, \dots, Z_q)$. With these notations we have:

Theorem 9 *Let H_6 be the number of $(f_1, \dots, f_6) \in F_{n,2}^6$ such that*

$$\forall i,\ 1 \leq i \leq q,\ [S_i^1, S_i^2, S_i^3] = G_3^6[I_i^1, I_i^2, I_i^3]$$

$$Then H_6 = \frac{|F_{n,2}| \cdot 2^{n(r+s)}}{2^{6nq}} \sum_{all\ frameworks\ \mathcal{F}} 2^{nf}[\text{ Number of } X, Y, Z \text{ satisfying } (C)]$$

- *where r denotes the number of independent equations $[I_i^2, I_i^3] = [I_j^2, I_j^3]$, $i \neq j$.*
- *s denotes the number of independent equations $[S_i^1, S_i^2] = [S_j^1, S_j^2]$, $i \neq j$.*
- *f denotes, as seen above, the number of equalities in $\mathcal{F}$.*

and (C) denotes these relations: $\forall i,\ 1 \leq i \leq q,\ \ \forall j,\ 1 \leq j \leq q,\ \ i \neq j$:

$$(C): \left\{ \begin{array}{l} [I_i^2, I_i^3] = [I_j^2, I_j^3] \Rightarrow X_i \oplus X_j = I_i^1 \oplus I_j^1 \\ [I_i^3, X_i] = [I_j^3, X_j] \Rightarrow Y_i \oplus Y_j = I_i^2 \oplus I_j^2 \\ [X_i, Y_i] = [X_j, Y_j] \Rightarrow Z_i \oplus Z_j = I_i^3 \oplus I_j^3 \\ [Y_i, Z_i] = [Y_j, Z_j] \Rightarrow S_i^1 \oplus S_j^1 = X_i \oplus X_j \\ [Z_i, S_i^1] = [Z_j, S_j^1] \Rightarrow S_i^2 \oplus S_j^2 = Y_i \oplus Y_j \\ [S_i^1, S_i^2] = [S_j^1, S_j^2] \Rightarrow S_i^3 \oplus S_j^3 = Z_i \oplus Z_j \end{array} \right.$$

Proof. As explained above, Theorem 9, despite its length, can easily be proved from Theorem 8.

7 Proof of security for G_3^6, G_3^7 and G_3^8

We present here only the main ideas, since the proofs are very similar to what we have done for Ψ^4, Ψ^5, Ψ^6. We start from the exact value H for G_3^6 given in Theorem 9, i.e. with the sum on all frameworks $\mathcal{F}$. When a framework is fixed,

3 cases can occur:
Case 1. A contradiction appears by linearity in the equations generated by (C). In this case we have in (C) a circle of the equations in X, or a circle in Y, or a circle in Z. If we have a circle in X, it comes from a circle in $[I^2, I^3]$, and $[Y, Z]_{\mathcal{F}}$. If we have a circle in Y, it comes from a circle in $[I^3, X]_{\mathcal{F}}$, and $[Z, S^1]_{\mathcal{F}}$. If we have a circle in Z, it comes from a circle in $[X, Y]_{\mathcal{F}}$, and $[S^1, S^2]$.
Case 2. No contradiction occurs by linearity but $\xi_X > n$, or $\xi_Y > n$, or $\xi_Z > n$, where ξ_X is the maximum number of X_j variables fixed in (C) when one X_i variable is fixed (similar definition for ξ_Y and $\xi_X = Z$).
Case 3. No contradiction occurs by linearity and $\xi_X \leq n$, $\xi_Y \leq n$ and $\xi_Z \leq n$.
For a framework $\mathcal{F}$, we define:$Weight(\mathcal{F})$ = Number of sequences $[X, Y, Z]$ that satisfy exactly the equalities of $\mathcal{F}$. (This means that the only equalities $[I_i^3, X_i] = [I_j^3, X_j]$ are exactly those in $\mathcal{F}$, and similarly for $[X_i, Y_i]$, $[Y_i, Z_i]$ and $[Z_i, S_i^1]$ collisions). Therefore, $\sum_{\text{all frameworks } \mathcal{F}} Weight(\mathcal{F}) = 2^{3nq}$. We will prove that if $\mathcal{F}$ is randomly chosen (i.e. with a distribution probability proportional to $Weight(\mathcal{F})$), then Case 3 is dominant when $\alpha^2 \ll 2^{3n}$, in KPA for G_3^6, in CPA-2 for G_3^7 and in CPCA-2 for G_3^8 (1). In order to prove (1), we analyze the probability of circles and lines of length $\geq n$ in a similar way as we did for Ψ^d. More precisely, we will use this lemma (which is the analog of Lemma 6 for Ψ^d schemes):

Lemma 1 *For all sequences of pairwise distinct* $[I_i^1, I_i^2, I_i^3]$, $1 \leq i \leq q$, *the number* N *of* (f_1, i, j) *such that* $X'_i = X'_j$, $i \neq j$, *satisfies* $N \leq |F_{n,2}| \cdot \frac{q(q-1)}{2^n}$ *where* $X'_i = I_i^1 \oplus f_1([I_i^2, I_i^3])$ *and* $X'_j = I_j^1 \oplus f_1([I_j^2, I_j^3])$

Proof of lemma 1. The proof is analog to the proof of Lemma 7: $X'_i = X'_j$ implies $[I_i^2, I_i^3] \neq [I_j^2, I_j^3]$ and then we have at most $\frac{|F_{n,2}|}{2^n}$ functions f_1 solutions when i, j are fixed. We see that the analysis for G_3^6 (in KPA), G_3^7 (in CPA-2) and G_3^8 (in CPCA-2) can be done in a very similar way as we did for Ψ^4, Ψ^5, Ψ^6, and we will obtain proofs of security when $q \ll \frac{2^{1.5n}}{n}$ (more precisely when $q \leq \frac{2^{1.5n}}{67n}$), since we will then have $H \geq \frac{|F_{n,2}|^d}{2^{3nq}}(1-\epsilon)$ with ϵ small in KPA for $d = 6$, in CPA-2 for $d = 7$, and in CPCA-2 for $d = 8$.
Remark. A natural problem is to try to achieve security for $\alpha \ll 2^{2n}$ instead of $\alpha \ll 2^{1.5n}$. This is not easy since now we need to prove that $H \geq \frac{|F_{n,2}|^d}{2^{3n}(2^{3n}-1)\ldots(2^{3n}-q+1)}(1-\epsilon)$ instead of $H \geq \frac{|F_{n,2}|^d}{2^{3nq}}(1-\epsilon)$ with small ϵ.

8 Conclusion

In this paper we have proved the security "beyond the birthday bound" for Ψ^4, Ψ^5, Ψ^6 (respectively in KPA, CPA-2, CPCA-2) when $q \ll \frac{2^n}{n}$ and for G_3^6, G_3^7, G_3^8 (respectively in KPA, CPA-2, CPCA-2) when $q \ll \frac{2^{3n/2}}{n}$, where q is the number of queries, with explicit and relatively simple security bounds. The general proof strategy used follows the general proof strategy of [23], but we have found some improvements, simplifications, and extensions. For example the interactions between the "frameworks of equalities" and the values called H has been significantly simplified. This proof strategy, i.e. essentially to derive the cryptographic security from "mirror theory" might look esoteric but we really believe that it is in fact rather natural. This is because if the values called H were, with a non negligible probability, significantly very different from their average value H_0, then some attacks could be created, i.e. this condition is not only a sufficient condition for security, but also a necessary condition for security. Moreover, Theorem 2 for Ψ^d, or Theorem 9 for G_3^d show that the values H have an exact formulation (not just an evaluation, but an exact formulation) in term of systems of linear equalities and non equalities. In fact what may be surprising is why the analysis of these systems ("mirror theory") was not done by mathematicians before (since, as far as we know it was not done). From a cryptographic point of view, in a way, after Theorem 2 (section 2) and Theorem 9 (section 6), the problems become purely mathematical, and are not cryptographic anymore. In fact, many proof strategy to obtain security "beyond the birthday bound" try to avoid this "mirror theory" by various clever ways (for example by increasing the number of rounds). What we show in this paper is that we can deal with these systems of equalities and non equalities in order to obtain good security bounds with a small number of rounds. We also believe that these proof strategies will be used in the future on many other cryptographic problems.

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. Anonymous. Introduction to Mirror Theory: Analysis of Systems of Linear Equalities and Linear Non equalities for Cryptography. In *Paper submitted to Crypto 2010*, 2010.
3. Mihir Bellare and Russell Impagliazzo. A Tool for Obtaining Tighter Security Analyses of Pseudorandom Function Based Constructions, with Applications to PRP to PRF Conversion. ePrint Archive 1999/024: Listing for 1999.
4. Mihir Bellare, Ted Krovetz, and Phillip Rogaway. Luby-Rackoff Backwards: Increasing Security by Making Block Ciphers Non-invertible. In Kaisa Nyberg, editor, *Advances in cryptology – EUROCRYPT 1998*, volume 1403 of *Lecture Notes in Computer Science*, pages 266–280. Springer-Verlag, 1998.
5. Chris Hall, David Wagner, John Kelsey, and Bruce Schneier. Building PRFs from PRPs. In Hugo Krawczyk, editor, *Advances in Cryptology – CRYPTO 1998*, volume 1462 of *Lecture Notes in Computer Science*, pages 370–389. Springer-Verlag, 1998.
6. Marshall Hall Jr. A Combinatorial Problem on Abelian Groups. *Proceedings of the Americal Mathematical Society*, 3(4):584–587, 1952.
7. Lars R. Knudsen. DEAL - A 128-bit Block Cipher. Technical Report 151, University of Bergen, Department of Informatics, Norway, february 1998.
8. Michael Luby and Charles Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM J. Comput.*, 17(2):373–386, 1988.

9. Stefan Lucks. The Sum of PRPs Is a Secure PRF. In Bart Preneel, editor, *Advances in Cryptology – EUROCRYPT 2000*, volume 1807 of *Lecture Notes in Computer Science*, pages 470–487. Springer-Verlag, 2000.
10. U. Maurer. A simplified and generalized treatment of Luby-Rackoff pseudorandom permutation generators. In *Advances in Cryptology – EUROCRYPT '92*, Lecture Notes in Computer Science, pages 239–255. Springer-Verlag, 1992.
11. U. Maurer. Indistinguishability of Random Systems. In *Advances in Cryptology – EUROCRYPT '02*, volume 2332 of *Lecture Notes in Computer Science*, pages 100–132. Springer-Verlag, 2002.
12. Ueli Maurer and Krzysztof Pietrzak. The Security of Many-Round Luby-Rackoff Pseudo-Random Permutations. In Eli Biham, editor, *Advances in Cryptology – EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 544–561. Springer-Verlag, 2003.
13. Valérie Nachef. Random Feistel Schemes for $m = 3$. *available from the author at: valerie.nachef@u-cergy.fr.*
14. Moni Naor and Omer Reingold. On the Construction of Pseudorandom Permutations: Luby-Rackoff Revisited. *J. Cryptology*, 12(1):29–66, 1999.
15. Jacques Patarin. Pseudorandom Permutations based on the DES Scheme. In *Eurocode '91*, volume 514 of *Lecture Notes in Computer Science*, pages 193–204. Springer-Verlag, 1990.
16. Jacques Patarin. Etude de Générateurs de Permutations Basés sur les Schémas du DES. In *Ph. Thesis.* Inria, Domaine de Voluceau, France, 1991.
17. Jacques Patarin. New results on pseudorandom permutation generators based on the DES Scheme. In *Advances in Cryptology – CRYPTO 1991*, Lecture Notes in Computer Science, pages 301–312. Springer-Verlag, 1991.
18. Jacques Patarin. Improved Security Bounds for Pseudorandom Permutations. In *4th ACM Conference on Computer and Communication Security*, pages 142 – 150, 1997.
19. Jacques Patarin. About Feistel Schemes with 6 (or more) Rounds). In Serge Vaudenay, editor, *FSE 1998*, volume 1372 of *Lecture Notes in Computer Science*, pages 103 – 121. Springer-Verlag, 1998.
20. 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.
21. Jacques Patarin. Luby-Rackoff: 7 Rounds are Enough for $2^{n(1-\epsilon)}$ Security. In Dan Boneh, editor, *Advances in Cryptology – CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 513–529. Springer-Verlag, 2003.
22. Jacques Patarin. Security of Random Feistel Schemes with 5 or more rounds. In Matthew K. Franklin, editor, *Advances in Cryptology – CRYPTO' 04*, volume 3152 of *Lecture Notes in Computer Science*, pages 106–122. Springer-Verlag, 2004.
23. Jacques Patarin. On linear systems of equations with distinct variables and Small block size. In Dongho Wan and Seungjoo Kim, editors, *ICISC 2005*, volume 3935 of *Lecture Notes in Computer Science*, pages 299–321. Springer-Verlag, 2006.
24. Jacques Patarin. A proof of security in $O(2^n)$ for the Benes schemes. In Serge Vaudenay, editor, *Progress in Cryptology – AFRICACRYPT' 08*, volume 5023 of *Lecture Notes in Computer Science*, pages 209–220. Springer-Verlag, 2008.
25. Jacques Patarin. A Proof of Security in $O(2^n)$ for the Xor of Two Random Permutations - Extended Version. *Cryptology ePrint archive: 2008/010: Listing for 2008*, 2008.
26. Jacques Patarin. Generic Attacks for the Xor of k Random Permutations. *Cryptology ePrint archive: 2008/009: Listing for 2008*, 2008.

27. Jacques Patarin. The coefficient H technique . In Roberto Avanzi, Liam Keliher, and Francesco Sica, editors, *SAC 2008*, volume 5381 of *Lecture Notes in Computer Science*, pages 66–81. Springer-Verlag, 2009.
28. 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.
29. F. Salzborn and G. Szekeres. A Problem in Combinatorial Group Theory. *Ars Combinatoria*, 7:3–5, 1979.
30. 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.
31. S. Vaudenay. Provable security for block ciphers by decorrelation. In *STACS '98*, volume 1373 of *Lecture Notes in Computer Science*, pages 249–275. Springer-Verlag, 1998.
32. Aaram Yun, Je Hong Park, and Jooyoung Lee. Lai-Massey Scheme and Quasi-Feistel Networks. *Cryptology ePrint archive: 2007/347: Listing for 2007.*

A The “Coefficient H technique”

We present here the general theorems on KPA, CPA-2 and CPCA-2 that are often used when we want to prove security of generic schemes from the number H of secret keys that send some cleartexts on some ciphertexts. These Theorems were proved in 1991 and named by J.Patarin the “coefficient H theorems”. The general idea is not surprising: when H is very near the average value H_0 for H except for some “bad” and “rare” events, then we will get the wanted security. The proofs are done by considering the set of all the transcripts of a distinguisher D that gives the output 1 and by showing that the probability that D gives 1 will be larger or very near when D deals a generic construction compared to the cases when D deals with a random permutation. Then by symmetry a similar result is obtained when D gives the output 0. We do not give here more details about the proofs since they are standard proofs now, used and proved again by various authors, and since details about their proofs can be found for example in [27]. Notice that we do not need $H \simeq H_0$ except on rare events, by that $[(H \geq H_0)$ or $(H \simeq H_0)]$ except on rare events is enough.

Notation for this sections

- N is an integer. I_N is $\{0,1\}^N$.
- F_N is the set of all functions from I_N to I_N.
- B_N is the set of all permutations from I_N to I_N.
- t and n are two integers.
- $F_{t,n}$ is the set of all functions from t bits to n bits.
- K is a set of k-uples of functions of $F_{t,n}$. The elements of K will be called “keys”.
- PRF means “Pseudorandom Functions”.
- PRP means “Pseudorandom Permutations”.

– G is an application of $K \to F_N$. Therefore G is a way to design a function from F_N from k-uples $(f_1, \dots f_k)$ of functions of $F_{t,n}$ of K.
– Let q be an integer (q will be the number of queries). Let $(a_i)_{1 \le i \le q}$ be a sequence of pairwise distinct elements of I_N. Let $(b_i)_{1 \le i \le q}$ be a sequence of elements of I_N. By definition, we will denote by $H(a, b)$ or simply by H if the context of the a_i and $b_i$is clear, the number of $(f_1, \dots f_k) \in K$ such that $\forall i$, $1 \le i \le q$, $G(f_1, \dots, f_k)(a_i) = b_i$. Therefore H is the number of "keys" (i.e. elements of K) that send all the a_i inputs to the exact values b_i.

Theorem 10 *(KPA,PRF)*
Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$.
If
(1) For random values a_i, b_i, $1 \le i \le q$ of I_N such that the a_i are pairwise distinct , with probability $\ge 1 - \beta$, we have: $H \ge \frac{|K|}{2^{Nq}}(1 - \alpha)$.
Then
(2) For every KPA with q (random) known plaintexts we have: $Adv^{PRF} \le \alpha + \beta$ where Adv^{PRF} denotes the advantage to distinguish $G(f_1, \dots, f_k)$ when $(f_1, \dots, f_k) \in_R K$ from a function $f \in_R F_N$.
(By "advantage", we mean as usual, for a distinguisher the absolute value of the difference of the two probabilities to output 1).

Remark. We did not ask here for the b_i values to be distinct.

Theorem 11 *(KPA,PRP)*
Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$.
If
(1) For random values a_i, b_i, $1 \le i \le q$ of I_N such that the a_i are pairwise distinct and the b_i are pairwise distinct, with probability $\ge 1 - \beta$, we have: $H \ge \frac{|K|}{2^N(2^N-1)\dots(2^N-q+1)}(1 - \alpha)$.
Then
(2) For every KPA with q (random) known plaintexts we have: $Adv^{PRP} \le \alpha + \beta$ where Adv^{PRP} denotes the advantage to distinguish $G(f_1, \dots, f_k)$ when $(f_1, \dots, f_k) \in_R K$ from a permutation $f \in_R B_N$.

Remark. We have $2^{Nq}(1 - \frac{q(q-1)}{2 \cdot 2^N}) \le 2^N(2^N-1)\dots(2^N-q+1) \le 2^{Nq}$, therefore if $q \ll \sqrt{2^N}$ we can use Theorem 10 or Theorem 11 and we will get similar security results (i.e. we can "switch" between random permutations and random functions as long as $q \ll \sqrt{2^N}$).

Theorem 12 *(CPA-2,PRF)*
Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$. Let E be a subset of I_N such that $|E| \ge (1 - \beta)2^{Nq}$.
If
(1) For all sequences a_i, $1 \le i \le q$ of pairwise distinct elements of I_N and for all sequences b_i, $1 \le i \le q$ of E we have $H \ge \frac{|K|}{2^{Nq}}(1 - \alpha)$.

Then
(2) For every CPA-2 with q chosen plaintexts we have: $Adv^{PRF} \leq \alpha + \beta$ *where* Adv^{PRF} *denotes the advantage to distinguish* $G(f_1, \ldots, f_k)$ *when* $(f_1, \ldots, f_k) \in_R K$ *from a function* $f \in_R F_N$.

Theorem 13 *(CPA-2,PRP)*
Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$. Let E be a subset of I_N such that $|E| \geq (1-\beta)2^N(2^N-1)\ldots(2^N-q+1)$
If
(1) For all sequences a_i, $1 \leq i \leq q$ of pairwise distinct elements of I_N and for all sequences b_i, $1 \leq i \leq q$ of E we have $H \geq \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha)$.
Then
(2) For every CPA-2 with q chosen plaintexts we have: $Adv^{PRP} \leq \alpha + \beta$ *where* Adv^{PRP} *denotes the advantage to distinguish* $G(f_1, \ldots, f_k)$ *when* $(f_1, \ldots, f_k) \in_R K$ *from a permutation* $f \in_R B_N$.

Theorem 14 *(CPCA-2,PRP)*
Let α be a real number, $\alpha > 0$.
If
(1) For all sequences of pairwise distinct elements a_i, $1 \leq i \leq q$, and for all sequences of pairwise distinct elements b_i, $1 \leq i \leq q$, we have:
$H \geq \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha)$. *Then*
(2) For every CPCA-2 with q chosen plaintexts we have: $Adv^{PRP} \leq \alpha$ *where* Adv^{PRP} *denotes the advantage to distinguish* $G(f_1, \ldots, f_k)$ *when* $(f_1, \ldots, f_k) \in_R K$ *from a permutation* $f \in_R B_N$.

Variant of Theorem 14
We have $2^{Nq}(1-\frac{q(q-1)}{2\cdot 2^N}) \leq 2^N(2^N-1)\ldots(2^N-q+1) \leq 2^{Nq}$. Therefore a variant of Theorem 14 is to have $H \geq \frac{|K|}{2^{Nq}}(1-\alpha)$ in (1), and $Adv^{PRP} \leq \alpha + \frac{q(q-1)}{2\cdot 2^N}$ in (2). When $q \ll \sqrt{2^N}$ this variant is sufficient.

Theorem 15 *(CPCA-2,PRP with more general conditions)*
Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$. Let P be the set of all sequences (a_i, b_i), $1 \leq i \leq q$ such that all the a_i are pairwise distinct and all the b_i are pairwise distinct. (Therefore $|P| = (2^N(2^N-1)\ldots(2^N-q+1))^2$*).*
If
(1) There exists a subset E of P such that
(1a) For all $(a,b) \in E$, we have $H \geq \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha)$.
(1b) For all CPCA-2 acting on a random permutation f of B_N, the probability that $(a,b) \in E$ is $\geq 1-\beta$ where (a,b) denotes here the successive $b_i = f(a_i)$ or $a_i = f^{-1}(b_i)$, $1 \leq i \leq q$ that will appear.
Then

(2) For every CPCA-2 with q chosen plaintexts we have: $Adv^{PRP} \leq \alpha + \beta$ where Adv^{PRP} denotes the advantage to distinguish $G(f_1, \ldots, f_k)$ when $(f_1, \ldots, f_k) \in_R K$ from a permutation $f \in_R B_N$.

Variant of Theorem 15
As above, a variant of theorem 15 is to have $H \geq \frac{|K|}{2^{Nq}}(1-\alpha)$ in (1), and $Adv^{PRP} \leq \alpha + \beta + \frac{q(q-1)}{2 \cdot 2^N}$ in (2). When $q \ll \sqrt{2^N}$ this variant is sufficient.

B Proof of KPA security for $\boldsymbol{\Psi}^4$

Circles in $R, Y_{\mathcal{F}}$
Definition
We will say that we have a 'circle in $R, Y_{\mathcal{F}}$' if there are k indices $i_1, \ldots, i_k$ with $k \geq 3$ and such that:
1. $i_1, i_2, \ldots, i_{k-1}$ are pairwise distinct and $i_k = i_1$.
2. $\forall \lambda,\ 1 \leq \lambda \leq k-2$ we have at least one of these conditions: $R_{i_\lambda} = R_{i_{\lambda+1}}$ or $Y_{i_\lambda} = Y_{i_{\lambda+1}}$ is one of the equalities of $\mathcal{F}$.
Example. If $R_1 = R_2$ and $Y_1 = Y_2$ is one of the equalities of $\mathcal{F}$, we have a circle in $R, Y_{\mathcal{F}}$.
Clearly, if we have a circle in $R, Y_{\mathcal{F}}$, from it we can generate a "minimum circle" in $R, Y_{\mathcal{F}}$, i.e. keeping only one equation $R_{i_1} = R_{i_l}$ per line of equations $R_{i_1} = R_{i_2} = R_{i_3} = \ldots = R_{i_l}$, and keeping only one equation $Y_{i_1} = Y_{i_l}$ per line of equations $Y_{i_1} = Y_{i_2} = Y_{i_3} = \ldots = Y_{i_l}$. Therefore, we have a circle in $R, Y_{\mathcal{F}}$ if and only if there is an even integer μ and there are μ pairwise distinct indices $i_1, \ldots, i_\mu$ with $(R_{i_1} = R_{i_2})$, ($Y_{i_2} = Y_{i_3}$ is in $\mathcal{F}$), $(R_{i_3} = R_{i_4}) \ldots$ ($Y_{i_\mu} = Y_{i_1}$ is in $\mathcal{F}$). μ will be called the length of the circle.

Lemma 2 *When the values $[L_i, R_i]$ are randomly chosen, pairwise distinct, and when the framework $\mathcal{F}$ is randomly chosen, with a distribution of probability proportional to $Weight(\mathcal{F})$, then the probability to have a circle in $R, Y_{\mathcal{F}}$ is $\leq \frac{q^2}{2^{2n}(1-\frac{q^2}{2 \cdot 2^{2n}})}$.*

Proof of Lemma 2 When we will say that $\mathcal{F}$ is "randomly chosen" it will always mean with a distribution of probability proportional to $Weight(\mathcal{F})$
First Property.
First, we can notice that with probability $\geq 1 - \frac{q^2}{2 \cdot 2^{2n}}$ we can assume that the $[L_i, R_i]$ values, for $1 \leq i \leq q$, are random, without considering the fact that they are pairwise distinct, since the probability to distinguish these two distributions is $\leq \frac{q^2}{2 \cdot 2^{2n}}$. (This is the Functions/Permutations switching lemma on permutations on $2n$ bits.).
Circle of length 2
To have a circle of length 2, we must find two indices i and j, $i < j$ such that $(R_i = R_j)$ and ($Y_i = Y_j$ is in $\mathcal{F}$). For $(i, j), i < j$, we have $\frac{q(q-1)}{2}$ possibilities. Now when i and j are fixed, the probability to have $R_i = R_j$ is $\frac{1}{2^n}$ if the R_i

values are random, and the probability to have $Y_i = Y_j$ is $\frac{1}{2^n}$ since $\mathcal{F}$ is randomly chosen (i.e. generated with a distribution proportional to $Weight(\mathcal{F})$). Therefore the probability to have $(R_i = R_j)$ and $(Y_i = Y_j$ is in $\mathcal{F}$ is $\leq \frac{q(q-1)}{2\cdot 2^n}$ if the R_i values are randomly chosen.

Circles of length μ, μ even

To have a circle of length μ, we must find μ pairwise distinct indices $i_1, i_2, \ldots, i_\mu$ such that $(R_{i_1} = R_{i_l})$, $(Y_{i_1} = Y_{i_2}$ is in $\mathcal{F})$, $(R_{i_3} = R_{i_4})\ldots(Y_{i_\mu} = Y_{i_1}$ is in $\mathcal{F})$. If the R_i values are randomly chosen, and if $\mathcal{F}$,is randomly chosen (i.e. with $Weight(\mathcal{F})$ distribution), then this probability is $\leq \frac{q^\mu}{\mu! 2^{\mu n}}$. Therefore, by using the First property above, we see that the probability to have a circle in $R, Y_\mathcal{F}$ is $\leq \frac{q^2}{2\cdot 2^{2n}} + \sum_{i=1}^{+\infty} \frac{q^{2i}}{(2i)! 2^{2in}} \leq \frac{q^2}{2^{2n}(1-\frac{q^2}{2\cdot 2^{2n}})}$ as claimed.

Lines in $R, Y_\mathcal{F}$

Definition

We will say that we have a “line in $R, Y_\mathcal{F}$” of length θ if there are $\theta + 1$ pairwise distinct indices $i_1, i_2, \ldots, i_{\theta+1}$ such that $\forall\lambda$, $1 \leq \lambda \leq \theta$, we have at least one of these two conditions: $R_{i_\lambda} = R_{i_{\lambda+1}}$ or $(Y_{i_\lambda} = Y_{i_{\lambda+1}}$ is one of the equalities of $\mathcal{F})$.
Example. If $R_1 = R_2$, $R_2 = R_3$, $(Y_3 = Y_4$ is in $\mathcal{F})$, and $R_4 = R_5$, then we have a line of length 4 in $R, Y_\mathcal{F}$.

Lemma 3 $\forall\theta \in \mathbb{N}$, *when the values* $[L_i, R_i]$ *are randomly chosen, pairwise distinct, and when the framework* $\mathcal{F}$ *is randomly chosen (i.e. with a distribution proportional to* $Weight(\mathcal{F})$*), then the probability to have a line of length* $\geq \theta$ *in* $R, Y_\mathcal{F}$ *is* $\leq \frac{q^{\theta+1}\cdot 2^\theta}{2^{n\theta}} + \frac{q^2}{2\cdot 2^{2n}}$.

Proof of Lemma 3 The proof is easy, we proceed as we did for the circles. The term $q^{\theta+1}$ comes from the possibilities for the $\theta+1$ indices $i_1, i_2, \ldots, i_{\theta+1}$, the term 2^θ comes from the fact that from i_λ to $i_{\lambda+1}$ we can have 2 possibilities: one equation in R, or one equation in Y, $1 \leq \lambda \leq \theta$, and the term $\frac{q^2}{2\cdot 2^{2n}}$ comes from the property seen in the proof of Lemma 2.

Lemma 4 *When the values* $[L_i, R_i]$ *are randomly chosen, pairwise distinct, the values* $[S_i, T_i]$ *are randomly chosen, and* $\mathcal{F}$ *is randomly chosen (i.e. with a distribution of probability in* $Weight(\mathcal{F})$*), then the probability* p *to have a circle in* $R, Y_\mathcal{F}$*, or to have a circle in* $S, X_\mathcal{F}$*, or to have a line in* $R, Y_\mathcal{F}$ *of length* $\geq \theta$*, or to have a line in* $S, X_\mathcal{F}$ *f length* $\geq \theta$ *satisfies:*

$$p \leq \frac{3q^2}{2^{2n}(1-\frac{q^2}{2\cdot 2^{2n}})} + \frac{2q^{\theta+1}\cdot 2^\theta}{2^{n\theta}}$$

Moreover, if $\theta \geq n$ *and* $8 \leq q \leq \frac{2^n}{8}$*, we have* $p \leq \frac{4q^2}{2^{2n}}$ *and* $\frac{2q^{\theta+1}\cdot 2^\theta}{2^{n\theta}} \leq \frac{2q}{2^{2n}}$*. (Therefore we can assume that we have no line of length* $\geq n$ *if* $q \ll 2^n$*).*

Proof of Lemma 4 The probability to have a circle in $R, Y_\mathcal{F}$ is $\leq \frac{q^2}{2^{2n}(1-\frac{q^2}{2\cdot 2^{2n}})}$. (cf Lemma 2). Similarly, the probability to have a circle in $S, X_\mathcal{F}$is $\leq \frac{q^2}{2^{2n}(1-\frac{q^2}{2\cdot 2^{2n}})}$. (symmetry of the hypothesis). The probability to have a line in $R, Y_\mathcal{F}$ of length

$\geq \theta$ is $\leq \frac{q^{\theta+1} \cdot 2^\theta}{2^{n\theta}} + \frac{q^2}{2 \cdot 2^{2n}}$ (cf Lemma 3). Similarly for a line in $S, X_\mathcal{F}$ of length $\geq \theta$. Therefore $p \leq \frac{3q^2}{2^{2n}(1-\frac{q^2}{2\cdot 2^{2n}})} + \frac{2q^{\theta+1}\cdot 2^\theta}{2^{n\theta}}$ as claimed. Moreover, if $\theta \geq n$ and $q \leq \frac{2^n}{8}$, we have $\frac{2q^{\theta+1}\cdot 2^\theta}{2^{n\theta}} \leq 2q \cdot (\frac{2q}{2^n})^\theta \leq 2q(\frac{1}{4})^\theta \leq \frac{2q}{2^{2n}}$. Then $p \leq \frac{3q^2}{2^{2n}(1-\frac{1}{128})} + \frac{2q}{2^{2n}}$ and this is $\leq \frac{4q^2}{2^{2n}}$ if $q \geq 8$ as claimed.
From now on, we assume $8 \leq q \leq \frac{2^n}{8}$. We take $\theta = n$, and we will say that a framework $\mathcal{F}$ is in "Case 5" if $\mathcal{F}$ has no circle in $R, Y_\mathcal{F}$, no circle in $S, X_\mathcal{F}$, no line of length $\geq n$ in $R, Y_\mathcal{F}$ and no line of length $\geq n$ in $S, X_\mathcal{F}$. Then,

Lemma 5 *When the values $[L_i, R_i]$ are randomly chosen, pairwise distinct, and the values $[S_i, T_i]$ are randomly chosen, then the probability p that*

$$[\sum_{\textit{all frameworks } \mathcal{F} \textit{ of Case } 5} Weight(\mathcal{F}) \leq (2^{2nq})(1 - \frac{2q}{2^n})]$$

satisfy $p \leq \frac{2q}{2^n}$.

Proof of Lemma 5. This comes immediately from Lemma 4 and

$$\sum_{\text{all frameworks } \mathcal{F}} Weight(\mathcal{F}) = \text{ Number of all sequences } (X_i, Y_i),\ 1 \leq i \leq q$$

$$= 2^{2nq}$$

We are now (at last) ready to prove Theorem 3, with explicit values for α et β.
Proof of Theorem 3. Let $\mathcal{F}$ be a framework of Case 5.We will use this Theorem of [2].

Theorem 16 *("Theorem $P_i \oplus P_j$" for any ξ_{max})*
let (A) be a set of a equation $P_i \oplus P_j = \lambda_k$ with α variables such that:
1. We have no circle in P in the equations (A).
2. We have no more than ξ_{max} indices in the same block.
3. By linearity from (A) we cannot generate an equation $P_i = P_j$ with $i \neq j$. (This means that if i and j are in the same block, then the expression in $\lambda_1, \lambda_2, \ldots, \lambda_a$ for $P_i \oplus P_j$ is $\neq 0$.
Then: if $\xi_{max}\alpha \ll 2^n$, we have $H_\alpha \geq J_\alpha$. More precisely the fuzzy condtion $\xi_{max}\alpha \ll 2^n$ can be written with the explicit bound: $(\xi_{max} - 1)\alpha \leq \frac{2^n}{67}$.

Then, from this Theorem 16 of [2], if $(67q)n \leq 2^n$, since $\xi_X \leq n$, we have at least $\frac{J_{q-x}}{2^{n(r+y)}}$ solutions $(X_1, \ldots, X_q)$ that satisfy $(C1)$, and at least $\frac{J_{q-y}}{2^{n(s+x)}}$ solutions $(Y_1, \ldots, Y_q)$ that satisfy $(C2)$. Since $Weight(\mathcal{F}) = J_{q-x} \cdot J_{q-y}$ we see that the number of $(X_1, \ldots, X_q), (Y_1, \ldots, Y_q)$ solutions of $(C1)$ and $(C2)$ is $\geq \frac{Weight(\mathcal{F})}{2^{n(r+s)} \cdot 2^{n(x+y)}}$. Therefore, from Lemma 4 and Theorem 2 we obtain: if $8 \leq q \leq \frac{2^n}{128n}$, then when the $[L_i, R_i]$ are randomly chosen pairwise distinct, and when the $[S_i, T_i]$ are randomly chosen, we have with a probability $\geq 1 - \frac{2q}{2^n}$ that

$$H \geq \frac{|F_n|^4}{2^{4nq}} \sum_{\text{all frameworks } \mathcal{F} \text{ of Case } 5} Weight(\mathcal{F})$$

$$H \geq \frac{|F_n|^4}{2^{2nq}}(1 - \frac{2q}{2^n})$$

Therefore we have proved Theorem 3 with $\alpha = \frac{2q}{2^n}$, $\beta = \frac{2q}{2^n}$ and $8 \leq q \leq \frac{2^n}{67n}$.
Example. Let $\mathcal{F}$ be this framework of equalities: $(Y_2 = Y_3)$ and $(X_2 = X_4 = X_5)$. Let assume $R_1 = R_2$. Then for $(C1)$ we have (cf Theorem 2):

$$\begin{cases} X_1 \oplus X_2 = L_1 \oplus L_2 \\ X_2 \oplus X_3 = S_2 \oplus S_3 \end{cases}$$

And for $(C2)$ we have

$$\begin{cases} Y_2 \oplus Y_4 = R_2 \oplus R_4 \\ Y_2 \oplus Y_5 = R_2 \oplus R_5 \end{cases}$$

We have here $\geq \frac{J_{q-2}}{2^{2n}}$ solutions for $X_1, \ldots, X_q$ and $\geq \frac{J_{q-1}}{2^{2n}}$ solutions for $Y_1, \ldots, Y_q$, and here $r = 1$, $s = 0$, $x = 2$, $y = 1$.

C Proof of CPA-2 security for Ψ^5

Our results on Ψ^5 for proving CPA-2 security will be based on this Theorem:

Theorem 17 *There are some values $\alpha > 0$ and $\beta > 0$ and there is a subset $E \subset I_{2n}^q$ such that:*
1. $|E| \geq (1 - \beta)2^{2nq}$
2. For all sequences $[L_i, R_i]$, $1 \leq i \leq q$ of pairwise distinct element of I_{2n} and for all sequences $[S_i, T_i]$, $1 \leq i \leq q$, of E we have: $H \geq \frac{|F_n|^5}{2^{2nq}}(1 - \alpha)$ where H denotes the number of $(f_1, f_2, f_3, f_4, f_5) \in F_n^5$ such that $\forall i$, $1 \leq i \leq q$, $\Psi^5(f_1, f_2, f_3, f_4, f_5)[L_i, R_i] = [S_i, T_i]$.
3. α and β can be chosen $\ll 1$ when $q \ll \frac{2^n}{n}$. (Moreover we will obtain explicit values for α and β: $\alpha = \frac{4q}{2^n}$, $\beta = \frac{q}{2^n}$ when $q \leq \frac{2^n}{67n}$).

Remark. It is equivalent to speak of such a subset E, or to say that when $[S_i, T_i]$, $1 \leq i \leq q$ are randomly chosen, the probability to have 2. is $\geq 1 - \beta$.
To prove Theorem 17, we will still use the formula given for H_4 (i.e. for Ψ^4) by Theorem 2, but we will perform one more round at the beginning. More precisely, we write $\Psi^5 = \Psi^4 \circ \Psi^1$.
The q chosen cleartexts are $[L_i, R_i]$, $1 \leq i \leq q$.
After one round they become $[R_i, X'_i]$, $1 \leq i \leq q$, with $X'_i = L_i \oplus f_1(R_i)$.
After two rounds they become $[X'_i, X_i]$, $1 \leq i \leq q$.
After three rounds they become $[X_i, Y_i]$, $1 \leq i \leq q$.
After four rounds they become $[Y_i, S_i]$, $1 \leq i \leq q$.
Finally after 5 rounds they become $[S_i, T_i]$, $1 \leq i \leq q$.
We have: $H_5([L_i, R_i, S_i, T_i], 1 \leq i \leq q) = \sum_{f_1 \in F_n}(H_4([R_i, X'_i, S_i, T_i], 1 \leq i \leq q))$ with $\forall i$, $1 \leq i \leq q$, $X'_i = L_i \oplus f_1(R_i)$. When a framework $\mathcal{F}$ for Ψ^4 is fixed (in $\Psi^5 = \Psi^4 \circ \Psi^1$), 5 cases occur:
Case 1. A contradiction appears by linearity in the equations generated by $(C1)$. In this case we have a “circle” of equalities in $X', Y_{\mathcal{F}}$ that generates a “circle”

in X by $(C1)$.
Case 2. A contradiction appears by linearity in the equations generated by $(C2)$. In this case we have a "circle" of equalities in $S, X_{\mathcal{F}}$ that generates a "circle" in Y by $(C2)$.
Case 3. No contradiction appears by linearity but $\xi_X > n$. In this case we have a line of equalities in $X', Y_{\mathcal{F}}$ of length $> n$ that generates a line of equalities in X by $(C1)$ of length $> n$.
Case 4. No contradiction appears by linearity but $\xi_Y > n$. In this case we have a line of equalities in $S, X_{\mathcal{F}}$ of length $> n$ that generates a line of equalities in Y by $(C2)$ of length $> n$.
Case 5. No contradiction occurs by linearity and $\xi_X \leq n$ and $\xi_Y \leq n$
To prove Theorem 17 we will first prove that Cases 1, 2, 3, 4 appear with a negligible probability when the $[S_i, T_i]$ variables are randomly chosen, when f_1 is randomly chosen in F_n, and when $\mathcal{F}$ is randomly chosen (this means as usual with a distribution of probability proportional to $Weight(\mathcal{F})$). This is what we will do now.
Circles in $X', Y_{\mathcal{F}}$

Lemma 6 *$\forall\lambda > 0$, for all pairwise distinct $[L_i, R_i]$, $1 \leq i \leq q$, when f_1 is randomly chosen in F_n we have with probability $\geq 1 - \frac{1}{\lambda}$ that the number N of (i, j), $i \neq j$ such that $X'_i = X'_j$ satisfies: $N \leq \frac{\lambda q(q-1)}{2^n}$.*

Proof. This comes immediately from this lemma:

Lemma 7 *For all pairwise distinct $[L_i, R_i]$, $1 \leq i \leq q$ (of course pairwise distinct means $i \neq i \Rightarrow L_i \neq L_j$* or *$R_i \neq R_j$) the number of (f_1, i, j) such that $X'_i = X'_j$, $i \neq j$ is $\leq |F_n|\frac{q(q-1)}{2^n}$.*

Proof of Lemma 7. $X'_i = X'_j$ means $L_i \oplus f_1(R_i) = L_j \oplus f_1(R_j)$. This implies $R_i \neq R_j$ (because $L_i = L_j$ and $R_i = R_j$ implies $i = j$). Thus, when (i, j) is fixed, the number of f_1 such that $X'_i = X'_j$ is exactly $\frac{|F_n|}{2^n}$ if $R_i \neq R_j$, and is exactly 0 if $R_i = R_j$. Therefore, since we have at most $q(q-1)$ values (i, j), $i \neq j$, $R_i \neq R_j$, the total number of (f_1, i, j) such that $X'_i = X'_j$ is $\leq \frac{|F_n|q(q-1)}{2^n}$ as claimed.

Lemma 8 *For all $\lambda > 0$, for all values $[L_i, R_i]$, $1 \leq i \leq q$ (no matter how cleverly chosen they are), when $f_1 \in_R F_n$, and $\mathcal{F}$ is randomly chosen, the probability p to have a circle in $X', Y_{\mathcal{F}}$ satisfies:*

$$p \leq \frac{1}{\lambda} + \frac{q^2}{2 \cdot 2^{2n}} + \frac{\lambda^2 q^4}{2^{4n}(1 - \frac{\lambda q^2}{2^{2n}})}$$

For $\lambda = \frac{2^n}{q}$ this gives:

$$p \leq \frac{q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}} + \frac{q^2}{2^{2n}(1 - \frac{q}{2^n})}$$

Therefore, if $q \leq \frac{2^n}{4}$, we have: $p \leq \frac{2q}{2^n}$.

Proof.

Circles of length 2.

To have a circle of length 2, we must find two indices i and j, $i < j$, such that: $(X'_i = X'_j)$ and ($Y_i = Y_j$ is in $\mathcal{F}$). For (i, j), $i < j$, we have $\frac{q(q-1)}{2}$ possibilities. Now when i and j are fixed, the probability to have $X'_i = X'_j$, i.e. to have $L_i \oplus f_1(R_i) = L_j \oplus f_1(R_j)$, is $\frac{1}{2^n}$ when $f_1 \in_R F_n$, since here we must have $R_i \neq R_j$ (because $R_i = R_j$ and $L_i = L_j$ imply $i = j$). Similarly, the probability to have ($Y_i = Y_j$ is in $\mathcal{F}$) when $\mathcal{F}$ is randomly chosen (i.e. with a distribution in $Weight(\mathcal{F})$) is $\frac{1}{2^n}$. Therefore the probability to have $(X'_i = X'_j)$ and ($Y_i = Y_j$ is in $\mathcal{F}$) is $\leq \frac{q(q-1)}{2 \cdot 2^{2n}}$.

Circles of length μ, $\mu \geq 4$, μ even

To have a circle of lengthμ, we must find μ pairwise distinct indices $i_1, i_2, \ldots, i_\mu$ such that $(X'_{i_1} = X'_{i_2})$, ($Y_{i_2} = Y_{i_3}$ is in $\mathcal{F}$), $(X'_{i_3} = X'_{i_4})$, …, ($Y_{i_\mu} = Y_{i_1}$ is in $\mathcal{F}$). From Lemma 6 we know that $\forall \lambda > 0$ we have a probability $\geq 1 - \frac{1}{\lambda}$ that for $(i_1, i_2, \ldots, i_\mu)$ we have at most $\left(\frac{\lambda q^2}{2^n}\right)^{\frac{\mu}{2}}$ possibilities, when f_1 is randomly chosen in F_n. Now when $(i_1, i_2, \ldots, i_\mu)$ are fixed, when $\mathcal{F}$ is randomly chosen, we have a probability $\frac{1}{2^{\frac{\mu}{2}n}}$ to have ($Y_{i_2} = Y_{i_3}$ is in $\mathcal{F}$), …, ($Y_{i_\mu} = Y_{i_1}$ is in $\mathcal{F}$). Therefore, $\forall \lambda > 0$ the probability to have a circle of length ≥ 4 is $\leq \frac{1}{\lambda} + \sum_{\mu=4,\, \mu \text{ even}}^{=\infty} \frac{\lambda^{\frac{\mu}{2}} q^\mu}{2^{n\mu}} \leq \frac{1}{\lambda} + \frac{\lambda^2 q^4}{2^{4n}(1 - \frac{\lambda q^2}{2^{2n}})}$. Since we have seen that the probability to have a circle of length 2 is $\leq \frac{q^2}{2 \cdot 2^{2n}}$ we obtain Lemma 8.

Lines in $X', Y_{\mathcal{F}}$

Lemma 9 $\forall \lambda > 0$, $\forall k \in \mathbb{N}^*$, *for all pairwise distinct* $[L_i, R_i]$, $1 \leq i \leq q$, *when* f_1 *is randomly chosen in* F_n *we have a probability* $\geq 1 - \frac{1}{\lambda}$ *that the number* N *of* $(i_1, i_2, \ldots, i_k)$ *such that* $(i_1, i_2, \ldots, i_k)$ *are pairwise distinct and* $X'_{i_1} = X'_{i_2} = \ldots = X'_{i_k}$ *satisfies:* $N \leq \frac{\lambda q^k}{2^{(k-1)n}}$.

Remark. Lemma 6 was a special case of Lemma 9 with $k = 2$.

Proof. This result comes essentially from this lemma:

Lemma 10 *For all pairwise distinct* $[L_i, R_i]$, $1 \leq i \leq q$, *the number of* $(f_1, i_1, i_2, \ldots, i_k)$ *such that* $i_1, i_2, \ldots, i_k$ *are pairwise distinct and* $X'_{i_1} = X'_{i_2} = \ldots = X'_{i_k}$ *is* $\leq |F_n| \frac{q^k}{2^{(k-1)n}}$.

Proof of Lemma 10 $X'_{i_1} = X'_{i_2} = \ldots = X'_{i_k}$ means $L_{i_1} \oplus f_1(R_{i_1}) = L_{i_2} \oplus f_1(R_{i_2}) = \ldots = L_{i_k} \oplus f_1(R_{i_k})$ (1). Since $i_1, i_2, \ldots, i_k$ are pairwise distinct, this implies here $R_{i_1}, R_{i_2}, \ldots R_{i_k}$ are pairwise distinct (because $L_i = L_j$ and $R_i = R_j \Rightarrow i = j$). Thus, when $i_1, i_2, \ldots, i_k$ are fixed, the number of f_1 that satisfy (1) is exactly $\frac{|F_n|}{2^{(k-1)n}}$ if $R_{i_1}, R_{i_2}, \ldots R_{i_k}$ are pairwise distinct and exactly 0 if $R_{i_1}, R_{i_2}, \ldots R_{i_k}$ are not pairwise distinct. Since we have at most q^k values $(i_1, i_2, \ldots, i_k)$, we obtain Lemma 10.

Lemma 11 $\forall \lambda > 0$, $\forall \theta \in \mathbb{N}^*$, *for all values* $[L_i, R_i]$, $1 \leq i \leq q$ *(no matter how cleverly chosen they are), when* $f_1 \in_R F_n$, *and* $\mathcal{F}$ *is randomly chosen, the* *probability p to have a line of length $\geq \theta$ in $X', Y_{\mathcal{F}}$ satisfies:*

$$p \leq \frac{1}{\lambda} + 2^{\theta} \frac{\lambda^{\frac{\theta+1}{2}} q^{\theta+1}}{2^{n\theta}}$$

For $\lambda = \frac{2^n}{q}$, we see that if $q \leq \frac{2^n}{64}$, and $\theta \geq n$, we have: $p \leq \frac{2q}{2^n}$.

Proof. The proof is easy from Lemma 9 since from Lemma 9, with $\lambda \geq 1$, we see that when θ is fixed the larger value for p is obtained when equalities in X' and equalities in $Y_{\mathcal{F}}$ alternate: $X'_{i_1} = X'_{i_2}$, ($Y_{i_2} = Y_{i_3}$ is in $\mathcal{F}$) etc. The coefficient 2^{θ} comes from the fact that for each indice i_{λ}, between i_{λ} and $i_{\lambda+1}$ we can have either an equality inX', or an equality in $Y_{\mathcal{F}}$. Now for $\lambda = \frac{2^n}{q}$, $\frac{1}{\lambda} + 2^{\theta} \frac{\lambda^{\frac{\theta+1}{2}} q^{\theta+1}}{2^{n\theta}} = \frac{q}{2^n} + \sqrt{2^n}\sqrt{q}\big(\frac{2\sqrt{q}}{\sqrt{2^n}}\big)^{\theta}$. If $q \leq \frac{2^n}{64}$ and $\theta \geq n$, this gives $p \leq \frac{q}{2^n} + \frac{\sqrt{2^n}\sqrt{q}}{2^{2n}}$ therefore we can write with simply $p \leq \frac{2q}{2^n}$ as claimed.
From now on, we assume $q \leq \frac{2^n}{64}$. We take $\theta = n$ and we will say that a framework $\mathcal{F}$ is in "Case 5" if $\mathcal{F}$ has no circle in $X', Y_{\mathcal{F}}$, no circle in $S, X_{\mathcal{F}}$, no line of length $\geq n$ in $X', Y_{\mathcal{F}}$ and no line of length $\geq n$ in $S, X_{\mathcal{F}}$.

Lemma 12 *If $q \leq \frac{2^n}{64}$, for all pairwise distinct values $[L_i, R_i]$, $1 \leq i \leq q$ when the values $[S_i, T_i]$ are randomly chosen, $1 \leq i \leq q$, the probability p that*

$$[\sum_{f_1 \in F_n} \sum_{\textit{all frameworks } \mathcal{F} \textit{ of Case 5}} Weight(\mathcal{F}) \leq |F_n| \cdot 2^{nq}(1 - \frac{4q}{2^n})]$$

satisfies: $p \leq \frac{q}{2^n}$.

Proof. This comes from Lemma 9, Lemma 8 and Lemma 11. The term in $p \leq \frac{q}{2^n}$ comes from Lemma 5, i.e. the circles and the lines in $S, X_{\mathcal{F}}$. These circles and lines in $S, X_{\mathcal{F}}$ are analyzed for Ψ^5 exactly as we did for Ψ^4 (only the first round has changed, not the last round). The term $|F_n| \cdot 2^{nq}(1 - \frac{4q}{2^n})$ comes from Lemma 8 (with $\frac{-2q}{2^n}$) and Lemma 11 (another $\frac{-2q}{2^n}$), and the fact that $\sum$ all frameworks $_{\mathcal{F}}\, Weight(\mathcal{F}) = 2^{2nq}$.
We are now ready to prove Theorem 17, with explicit values for α and β.
Proof of Theorem 17 From Lemma 12 and the theorem "$P_i \oplus P_j$", i.e. Theorem 6 of [2], we see that if $q \leq \frac{2^n}{67n}$, then for all pairwise distinct $[L_i, R_i]$, $1 \leq i \leq q$, we have a probability $\geq 1 - \frac{q}{2^n}$ that $H \geq \frac{|F_n|^5}{2^{2nq}}(1 - \frac{4q}{2^n})$ when the $[S_i, T_i]$ values, $1 \leq i \leq q$, are randomly chosen. Therefore, we have proved Theorem 17 with $\alpha = \frac{4q}{2^n}$ and $\beta = \frac{q}{2^n}$ when $q \leq \frac{2^n}{67n}$.

Theorem 18 *If $q \leq \frac{2^n}{67n}$, then for every CPA-2 with q adaptive chosen plaintexts, we have: $Adv^{PRF} \leq \frac{5q}{2^n}$ and $Adv^{PRR} \leq \frac{5q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}}$. Where Avd^{PRF} denotes the advantage to distinguish Ψ^5 from $f \in_R F_{2n}$, and Adv^{PRP} denotes the advantage to distinguish Ψ^5 from $f \in_R B_{2n}$.*

Proof. This comes immediately from Theorem 12 of Appendix A, Theorem 17, and the classical pseudorandom functions / pseudorandom permutations switching lemma.