Composition Theorems for CCA Cryptographic Security

No Institute Given

Abstract. We present two new theorems to analyze the indistinguishability of the composition of cryptographic permutations and the indistinguishability of the XOR of cryptographic functions. Using the H Coefficients technique of [Pat01], for any two families of permutations F and G with CCA distinghuishability advantage $\leq \alpha_F$ and $\leq \alpha_G$, we prove that the set of permutations $f \circ g, f \in F, g \in G$ has CCA distinguishability advantage $\leq \alpha_F \times \alpha_G$. This simple composition result gives a CCA indistinguishability geometric gain when composing blockciphers (unlike previously known clasical composition theorems). As an example, we apply this new theorem to analyze $4r$ and $6r$ rounds Feistel schemes with $r \geq 1$ and we improve previous best known bounds for a certain range of queries. Similarly, for any two families of functions F and G with distinghuishability advantage $\leq \alpha_F$ and $\leq \alpha_G$, we prove that the set of functions $f \oplus g, f \in F, g \in G$ has distinguishability advantage $\leq \alpha_F \times \alpha_G$. As an example, we apply this new theorem to analyze the XOR of $2r$ permutations and we improve the previous best known bounds for certain range of queries.

Keywords: H coefficients, Security proof, Composition, XOR of permutations, Feistel Schemes, Luby-Rackoff construction.

1 Introduction

The indistinguishability of cryptographic permutations is a well-known problem. Since the seminal article of Luby and Rackoff [LR85] in 1989, security proofs of Feistel schemes have been extensively studied ([HR10], [MP03], [NR99], [Mau92], [Vau98], [DP10]). It is particularly interesting and difficult to obtain such proofs beyond the birthday bound and ideally to the bound of "the information theory". After this bound, we can no longer hope to prove security against an attacker with an unbounded power of computation. Essentially two techniques have been developed to give beyond the birthday bound proofs. The first one, the Coupling technique, gives very good results when we study schemes with many rounds (Maurer [MP03], Hoang/Rogaway [HR10]). The second one, the H Coefficients technique, gives better results when the number of rounds is relatively small (Patarin [Pat04] [Pat08] [Pat10]) but, until recently with the indistinguishability proof of the key-alternating Cipher [CS13], this technique was not applied to many rounds schemes. We introduce a new theorem: the "H Coefficients Composition Theorem". This theorem says that one has a geometric gain of the distinguishability advantage when composing two families of permutations: for any two families of permutations F and G with distinghuishability advantage α_F and α_G, we prove that the set of permutations $f \circ g, f \in F, g \in G$ has distinguishability advantage $\alpha_F \times \alpha_G$. So far, the theory of Random Systems of Maurer [Mau02] was the only one to have composition theorems that yield a geometric gain ([GM09],[MP04],[MPR07]). Unfortunately, none of these composition theorems yield a geometric gain for CCA Indistinguishability (the queries can be direct or inverse). Our composition theorem complete this gap. As an example, we apply this new theorem to analyze $4r$ and $6r$ rounds Feistel schemes with $r \geq 1$. We obtain the same kind of geometric gain when xoring two families of functions: for any two families of

functions F and G with distinghuishability advantage $\leq \alpha_F$ and $\leq \alpha_G$, we prove that the set of functions $f \oplus g, f \in F, g \in G$ has distinguishability advantage $\leq \alpha_F \times \alpha_G$. As an example, we apply this new theorem to analyze the XOR of $2r$ permutations.

Notations. We denote F_n the set of all functions from $\{0;1\}^n$ to $\{0;1\}^n$ and B_n the set of all permutations of $\{0;1\}^n$. For any $f \in F_n$ and $L, R \in \{0;1\}^n$, $\Psi(f)$ denotes the permutation of B_{2n} defined by: $\Psi_f([L,R]) = [R, L \oplus f(R)]$. More generally, if $f_1, ..., f_r$ are r functions of F_n, $\Psi^r(f_1, ..., f_r)$ denotes the permutation of B_{2n} defined by: $\Psi^r(f_1, ..., f_r) = \Psi(f_r) \circ \cdots \circ \Psi(f_1)$. This permutation is called a balanced Feistel scheme with r rounds or, shortly, Ψ^r. When the functions $f_1, ..., f_r$ are uniformly random, Ψ^r is called a "generic" Feistel scheme with r rounds, or a Luby-Rackoff construction. For any integer q, we denote J_q the set of $(X_1, \ldots, X_q) \in (\{0;1\}^{2n})^q$ such that $X_i \neq X_j$ for every $i \neq j$ and we define $j_q = |J_q| = 2^{2n} \times (2^{2n} - 1) \times \cdots \times (2^{2n} - q + 1)$.

2 The H coefficients technique

Indistinguishability Let F be a set of permuations, O an oracle which acts either like a random permutation or like a random permutation in F. Let A be an attacker which can send direct and inverse queries to O. After sending q queries (receiving each time an answer from the oracle), the attacker outputs a bit b. The advantage of A to distinguish F from a random permutations is defined as

$$\left| \Pr[b = 1 | O \text{ acts like a random } f \in F] - \Pr[b = 1 | O \text{ acts like a random permutation}] \right| .$$

We denote $\mathbf{Adv}_F^{\mathrm{PRP}}(q)$ the maximum advantage over any attacker to distinguish F from random permutations and $\mathbf{Adv}_F^{\mathrm{PRF}}(q)$ the maximum advantage over any attacker (only making direct queries) to distinguish F from random functions. In this article, we will prove security bounds using the general framework given by the "H Coefficients technique" of Patarin [Pat08].

Theorem 1 (H Coefficients Theorem, 1991). *Let F be a subset of B_{2n} indexed by a set of keys K: $F = \{f_k, k \in K\}$. If there exists a real number $\alpha > 0$ such that, for any $(Y_1, ..., Y_q) \in J_q$ and any $(X_1, ..., X_q) \in J_q$, the number $H(X, Y)$ of keys k, such that, for all i, f_k sends X_i to Y_i, verifies:*

$$H(X, Y) \geq (1 - \alpha) \frac{|K|}{2^{2nq}},$$

Then the advantage to distinguish between permutations f_k of F, with $k \in_R K$, and random permutations verifies:

$$\mathbf{Adv}_F^{\mathrm{PRP}}(q) \leq \alpha + \frac{q(q-1)}{2 \cdot 2^{2n}} .$$

Proof. The proof is in [Pat08] and in Appendix A. □

There are many variants of this H Coefficients Theorem (cf [Pat08]). For example:

Theorem 2 (1991). *Let F be a subset of B_{2n} indexed by a set of keys K: $F = \{f_k, k \in K\}$. If there exists a real number $\alpha > 0$ such that, for any $(Y_1, ..., Y_q) \in J_q$ and any $(X_1, ..., X_q) \in J_q$, the number $H(X, Y)$ of keys k, such that, for all i, f_k sends X_i to Y_i, verifies:*

$$H(X, Y) \geq (1 - \alpha)\frac{|K|}{j_q},$$

Then the advantage of any CCA attacker to distinguish between permutations f_k of F, with $k \in_R K$, and random permutations verifies:

$$\mathbf{Adv}_F^{\mathrm{PRP}}(q) \leq \alpha \ .$$

Proof. See Appendix B □

Theorem 3 (1991). *Let F be a subset of F_{2n} indexed by a set of keys K: $F = \{f_k, k \in K\}$. If there exists a real number $\alpha > 0$ such that, for any $(Y_1, ..., Y_q) \in (\{0,1\}^n)^q$ and any $(X_1, ..., X_q) \in J_q$, the number $H(X, Y)$ of keys k, such that, for all i, f_k sends X_i to Y_i, verifies:*

$$H(X, Y) \geq (1 - \alpha)\frac{|K|}{2^{2nq}},$$

Then the advantage of any CCA attacker to distinguish between functions f_k of F, with $k \in_R K$, and random functions verifies:

$$\mathbf{Adv}_F^{\mathrm{PRF}}(q) \leq \alpha \ .$$

Proof. See [Pat08]. □

3 The composition theorem (composition with ∘)

Let F and G be two subsets of B_{2n} and we denote $F \circ G$ the set $\{f \circ g, f \in F, g \in G\}$. For any subset F of B_{2n}, we denote $H^F(X, Y)$ the number of functions in F sending X_i to Y_i for all $i \leq q$.

Theorem 4. *If it exists α_F and α_G in $[0; 1]$ such that, for any $(X_1, ..., X_q) \in J_q$ and any $(Y_1, ..., Y_q) \in J_q$:*

$$H^F(X, Y) \geq (1 - \alpha_F) \times \frac{|F|}{2^{2nq}} \text{ and } H^G(X, Y) \geq (1 - \alpha_G) \times \frac{|G|}{2^{2nq}},$$

then, for any $(X_1, ..., X_q) \in J_q$ and any $(Y_1, ..., Y_q) \in J_q$:

$$H^{F \circ G}(X, Y) \geq (1 - \alpha_F \alpha_G) \times \frac{|F| \times |G|}{2^{2nq}} \ .$$

Proof. For any pairwise distinct $X_1, ..., X_q$ and pairwise distinct $Y_1, ..., Y_q$. One has

$$H^{F \circ G}(X, Y) = \sum_T H^G(X, T) \times H^F(T, Y), \tag{1}$$

the sum being taken over the pairwise distinct $T_1, ..., T_q$. We notice that one has:

$$\sum_T 1 = j_q \leq 2^{2nq} \tag{2}$$

$$\sum_T H^F(T,Y) = |F| \tag{3}$$

$$\sum_T H^G(X,T) = |G| \tag{4}$$

We compute the right part of equality (1) by introducing the values $m_F = (1-\alpha_F) \times \frac{|F|}{2^{2nq}}$ and $m_G = (1-\alpha_G) \times \frac{|G|}{2^{2nq}}$:

$$\begin{aligned}
&\sum_T H^G(X,T) \times H^F(T,Y) \\
= &\sum_T ((H^G(X,T) - m_G) + m_G) \times ((H^F(T,Y) - m_F) + m_F) \\
= &\sum_T (H^G(X,T) - m_G) \times (H^F(T,Y) - m_F) \\
&+ \sum_T m_G \times (H^F(T,Y) - m_F) + \sum_T (H^G(X,T) - m_G) \times m_F + \sum_T m_G m_F.
\end{aligned}$$

By hypothesis, the first term is positive. We now compute the second term:

$$\begin{aligned}
&\sum_T m_G \times (H^F(T,Y) - m_F) \\
= &\ (1-\alpha_G) \times \frac{|G|}{2^{2nq}} \left(\sum_T H^F(T,Y) - \sum_T (1-\alpha_F) \times \frac{|F|}{2^{2nq}} \right) \\
\geq &\ (1-\alpha_G) \times \frac{|G|}{2^{2nq}} \left(|F| - (1-\alpha_F)|F| \Big(\sum_T \frac{1}{2^{2nq}} \Big) \right) \text{ from (3)} \\
\geq &\ \frac{|G| \times |F|}{2^{2nq}} \left((1-\alpha_G) - (1-\alpha_G)(1-\alpha_F) \Big(\sum_T \frac{1}{2^{2nq}} \Big) \right)
\end{aligned}$$

The third term is computed in the same way. The fourth term gives:

$$\sum_T m_G m_F = \frac{|G||F|}{2^{2nq}} \left((1-\alpha_G)(1-\alpha_F) \Big(\sum_T \frac{1}{2^{2nq}} \Big) \right) .$$

Summing the four terms, one has:

$$\sum_T H^G(X,T) \times H^F(T,Y) \geq \frac{|G||F|}{2^{2nq}} \left((1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) \Big(\sum_T \frac{1}{2^{2nq}} \Big) \right) .$$

There is at most 2^{2nq} choices for T (cf (2)) so

$$\begin{aligned}
(1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) \left(\sum_T \frac{1}{2^{2nq}} \right) &\geq (1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) \\
&= 1 - \alpha_G \alpha_F.
\end{aligned}$$

□

The next theorem is a variant of the previous theorem. This variant is interesting to understand the geometric gain we obtain by composing Feistel schemes. One remind that $j_q = 2^{2n} \times (2^{2n} - 1) \times \cdots \times (2^{2n} - q + 1)$.

Theorem 5. *If it exists α_F and α_G in $[0;1]$ such that, for all pairwise distinct $X_1, ..., X_q$ and for all pairwise distinct $Y_1, ..., Y_q$:*

$$H^F(X,Y) \geq (1-\alpha_F) \times \frac{|F|}{j_q} \text{ and } H^G(X,Y) \geq (1-\alpha_G) \times \frac{|G|}{j_q},$$

then, for all pairwise distinct $X_1, ..., X_q$ and for all pairwise distinct $Y_1, ..., Y_q$:

$$H^{F \circ G}(X,Y) \geq (1-\alpha_F\alpha_G) \times \frac{|F| \times |G|}{j_q} \ .$$

Proof. See Appendix C □

The Theorem 5 and Theorem 2 imply that the advantage to distinguish functions $f \circ g$ is less or equal to $\alpha_F\alpha_G$ when the advantage to distinguish functions f is α_F and the advantage to distinguish functions g is α_G. We obtain a geometric gain when we compose functions.

Comparison with Maurer's composition theorems In [MPR07], Maurer Renner and Pietrzak proved that composing two family of permutations NCPA-secure (indistinguishable against non-adaptive chosen plaintext attack) yield a family of permutation CCA-secure (indistinguishable against adaptative chosen plaintext and ciphertext attack). While this theorem is very useful to obtain CCA-proofs from NCPA-proofs of security ([HR10], [LPS12]), the gain is not geometric. The CCA advantage of the composed permutations is upperbounded by the sum of the NCPA advantage of the two family of permutations. In this paper, we upperbound the composed permutations by the product of the CCA advantage of the two family of permutations.

4 The XOR Theorem (composition with $\oplus$)

Let F and G be two subsets of F_n and we denote $F \oplus G$ the set $\{f \oplus g, f \in F, g \in G\}$. For any subset F of F_n, we denote $H^F(X,Y)$ the number of functions in F sending X_i to Y_i for all $i \leq q$.

Theorem 6. *If it exists α_F and α_G in $[0;1]$ such that, for all pairwise distinct $X_1, ..., X_q$ and for all $Y_1, ..., Y_q$:*

$$H^F(X,Y) \geq (1-\alpha_F) \times \frac{|F|}{2^{nq}} \text{ and } H^G(X,Y) \geq (1-\alpha_G) \times \frac{|G|}{2^{nq}},$$

then, for all pairwise distinct $X_1, ..., X_q$ and for all $Y_1, ..., Y_q$:

$$H^{F \oplus G}(X,Y) \geq (1-\alpha_F\alpha_G) \times \frac{|F| \times |G|}{2^{nq}} \ .$$

Proof. For any pairwise distinct $X_1, ..., X_q$ and any $Y_1, ..., Y_q$. One has

$$H^{F\oplus G}(X,Y) = \sum_T H^G(X,T) \times H^F(X, Y \oplus T), \tag{5}$$

the sum being taken over all $T_1, ..., T_q$ and $Y \oplus T$ denoting the vector $(Y_i \oplus T_i)_i$. We notice that one has:

$$\sum_T 1 = 2^{nq} \tag{6}$$

$$\sum_T H^F(T,Y) = |F| \tag{7}$$

$$\sum_T H^G(X,T) = |G| \tag{8}$$

We compute the right part of equality (5) by introducing the values $m_F = (1-\alpha_F) \times \frac{|F|}{2^{nq}}$ and $m_G = (1-\alpha_G) \times \frac{|G|}{2^{nq}}$:

$$\begin{aligned}
&\sum_T H^G(X,T) \times H^F(X, Y \oplus T) \\
= &\sum_T ((H^G(X,T) - m_G) + m_G) \times ((H^F(X, Y \oplus T) - m_F) + m_F) \\
= &\sum_T (H^G(X,T) - m_G) \times (H^F(X, Y \oplus T) - m_F) \\
&+ \sum_T m_G \times (H^F(X, Y \oplus T) - m_F) + \sum_T (H^G(X,T) - m_G) \times m_F + \sum_T m_G m_F.
\end{aligned}$$

By hypothesis, the first term is positive. We now compute the second term:

$$\begin{aligned}
&\sum_T m_G \times (H^F(X, Y \oplus T) - m_F) \\
&= (1-\alpha_G) \times \frac{|G|}{2^{nq}} \left(\sum_T H^F(X, Y \oplus T) - \sum_T (1-\alpha_F) \times \frac{|F|}{2^{nq}} \right) \\
&\geq \ (1-\alpha_G) \times \frac{|G|}{2^{nq}} \left(|F| - (1-\alpha_F)|F| \Big(\sum_T \frac{1}{2^{nq}} \Big) \right) \text{ from (7)} \\
&\geq \ \frac{|G| \times |F|}{2^{nq}} \left((1-\alpha_G) - (1-\alpha_G)(1-\alpha_F) \right) \text{ from (6)}
\end{aligned}$$

The third term is computed in the same way. The fourth term gives:

$$\sum_T m_G m_F = \frac{|G||F|}{2^{nq}} \left((1-\alpha_G)(1-\alpha_F) \right) \ .$$

Summing the four terms, one has:

$$\begin{aligned}
\sum_T H^G(X,T) \times H^F(X, Y \oplus T) &\geq \frac{|G||F|}{2^{nq}} \left((1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) \right) \\
&\geq \frac{|G||F|}{2^{nq}} \Big(1 - \alpha_F \alpha_G \Big) \ .
\end{aligned}$$

□

Remark: This theorem is generalizable to any group law $\star$.

Comparison with Maurer's composition theorems In [MPR07], Maurer Renner and Pietrzak proved that xoring two family of functions CPA-secure (indistinguishable against adaptive chosen plaintext attack) yield a family of function CPA-secure with a geometric gain. In this paper, we prove the very same thing but in the framework of the H Coefficients Technique. We notice that our two composition theorems are extremely similar.
In Appendix D, we prove an other composition theorem. This theorem follows the strategy of the coupling technique in which we make security proofs analyzing queries after queries in a local way.

5 Applications

5.1 Security of Ψ^{4r}

In [Pat98], J. Patarin proved the following:

Theorem 7 (1991). *For Ψ^4, for all pairwise distinct $X_1, ..., X_q \in \{0;1\}^{2n}$ and for all pairwise distinct $Y_1, ..., Y_q \in \{0;1\}^{2n}$:*

$$H(X,Y) \geq \left(1 - \frac{q^2}{2^n}\right) \times \frac{|F_n|^4}{2^{2nq}} \ .$$

Combining this theorem and using our new composition theorem of Section 3 yield:

Theorem 8. *For any integer $r \geq 1$:*

$$\mathbf{Adv}^{\mathrm{PRP}}_{\Psi^{4r}}(q) \leq \left(\frac{q^2}{2^n}\right)^r + \frac{q(q-1)}{2 \cdot 2^{2n}} \ .$$

5.2 Security of Ψ^{5r}

In [Pat98], J. Patarin proved the following:

Theorem 9 (1991). *For Ψ^5, for all pairwise distinct $X_1, ..., X_q \in \{0;1\}^{2n}$ and for all pairwise distinct $Y_1, ..., Y_q \in \{0;1\}^{2n}$:*

$$H(X,Y) \geq \left(1 - \frac{q^3}{2^{2n}}\right) \times \frac{|F_n|^5}{2^{2nq}} \ .$$

Combining this theorem and using our new composition theorem of Section 3 yield:

Theorem 10. *For any integer $r \geq 1$:*

$$\mathbf{Adv}^{\mathrm{PRP}}_{\Psi^{5r}}(q) \leq \left(\frac{q^3}{2^{2n}}\right)^r + \frac{q(q-1)}{2 \cdot 2^{2n}} \ .$$

5.3 Security of Ψ^{6r}

We can apply the same reasoning on Ψ^{6r} using the result of [Pat98]:

Theorem 11 (1998). *For Ψ^6, for all $q \leq 2^n$, for all $X_1, ..., X_q \in J_q$ and for all $Y_1, ..., Y_q \in J_q$:*

$$H(X,Y) \geq \left(1 - \frac{47q^4}{2^{3n}} - \frac{16q^2}{2^{2n}}\right) \times \frac{|F_n|^6}{2^{2nq}} \ .$$

Then, using the composition theorem, this imply:

Theorem 12. *For any integer $r \geq 1$ and any $q \leq 2^n$:*

$$\mathbf{Adv}^{\mathrm{PRP}}_{\Psi^{6r}}(q) \leq \left(\frac{47q^4}{2^{3n}} + \frac{16q^2}{2^{2n}}\right)^r + \frac{q(q-1)}{2 \cdot 2^{2n}} \ .$$

In [Pat10] p.8, J. Patarin proved this theorem:

Theorem 13 (2010). *For Ψ^6 and $q \leq \frac{2^n}{67n}$, for all $X_1, ..., X_q \in J_q$ and for all $Y_1, ..., Y_q \in J_q$:*

$$H(X,Y) \geq \left(1 - \frac{8q}{2^n}\right) \times \frac{|F_n|^6}{2^{2nq}} \ .$$

Again, combining this theorem and using our new composition theorem of Section 3 yield:

Theorem 14. *For any integer $r \geq 1$ and any $q \leq \frac{2^n}{67n}$:*

$$\mathbf{Adv}^{\mathrm{PRP}}_{\Psi^{6r}}(q) \leq \left(\frac{8q}{2^n}\right)^r + \frac{q(q-1)}{2 \cdot 2^{2n}} \ .$$

This Theorem 14 is, so far, the best security bound known for Feistel schemes when $q \leq \frac{2^n}{67n}$.

Results This new bound is the best one when we can use it (we need $q \leq \frac{2^n}{67n}$). In the following table, we give the advantage for $q = \frac{2^n}{67n}$ with our bound and with the previous best known bound of [HR10].

Table 1. The advantage for $q = \frac{2^n}{67n}$ with our bound and with the previous best known bound of [HR10].

n	rounds	$\log_2(q)$	Our advantage	HR's advantage
32	18	20.9	$2^{-24.2}$	$2^{-8.8}$
64	18	51.9	$2^{-27.2}$	1
128	18	114.9	$2^{-30.2}$	1
32	48	20.9	$2^{-64.5}$	$2^{-55.3}$
64	48	51.9	$2^{-72.5}$	$2^{-32.5}$
128	48	114.9	$2^{-80.5}$	1

For the values of this array, we notice that our results are much better than previous best known bounds. Especially for cases 2, 3 and 6, we prove that the advantage is small.

5.4 Security of the XOR of 2r permutations

In this section, we consider the indistinguishability of the XOR of $2r$ random permutations from a random permutation. We use the following result for the XOR of 2 random permutations.

Theorem 15. *For all pairwise distinct $X_1, \ldots, X_q \in \{0,1\}^n$ and all $Y_1, \ldots, Y_q \in \{0,1\}^n$, the number H of $(f,g) \in (B_n)^2$ such that $f(X_i) \oplus g(X_i) = Y_i$ satisfy:*

$$H(X,Y) \geq \left(1 - \frac{q^3}{2^{2n} - 2q \cdot 2^n + q^2}\right) \frac{|B_n|^2}{2^{nq}} \quad .$$

Proof. Let $X_1, ..., X_q \in \{0,1\}^n$ be pairwise distinct and $Y_1, ..., Y_q \in \{0,1\}^n$. For the first equation $f(X_1) \oplus g(X_1) = Y_1$, there is 2^n choices for $f(X_1)$ and one for $g(X_1)$. For the second equation, $f(X_2)$ can take $2^n - 1$ values (since f is a permutation) and it has to be such that $g(X_2) = f(X_2) \oplus Y_2$ is different of $g(X_1)$ so it can take at least $2^n - 2$ values and $g(X_2)$ takes one value. For the third equation, applying the same reasoning, we see that $f(X_3)$ can take at least $2^n - 4$ values. Applying the same reasoning for all equations, one has:

$$H(X,Y) \geq \prod_{i=0}^{q-1} (2^n - 2i) \quad .$$

The proof ends after some computations:

$$\begin{aligned}
\prod_{i=0}^{q-1} (2^n - 2i) \times \frac{2^{nq}}{|B_n|^2} &= \prod_{i=0}^{q-1} \frac{(2^n - 2i)2^n}{(2^n - i)^2} \\
&= \prod_{i=0}^{q-1} \left(1 - \frac{i^2}{2^{2n} - 2i \cdot 2^n + i^2}\right) \\
&\geq \prod_{i=0}^{q-1} \left(1 - \frac{q^2}{2^{2n} - 2q \cdot 2^n + q^2}\right) \\
&\geq 1 - \frac{q^3}{2^{2n} - 2q \cdot 2^n + q^2}
\end{aligned}$$

□

Using theorem 6 and 3 yield:

Theorem 16. *For any integer $r \geq 1$, the advantage to distinguish between the XOR of $2r$ random permutations and a random function satisfy:*

$$\mathbf{Adv}^{\mathrm{PRF}}_{f_1 \oplus \cdots \oplus f_{2r}}(q) \leq \left(\frac{q^3}{2^{2n} - 2q \cdot 2^n + q^2}\right)^r \quad .$$

If one improve Theorem 15, this induce an improvement for Theorem 16. This is what we do, using [NP13], which yield a better bound but the proof is a lot more complicated.

Theorem 17. *For all pairwise distinct $X_1, \ldots, X_q \in \{0,1\}^n$ and all $Y_1, \ldots, Y_q \in \{0,1\}^n$, the number H of $(f,g) \in (B_n)^2$ such that $f(X_i) \oplus g(X_i) = Y_i$ satisfy:*

$$H(X,Y) \geq \left(1 - \frac{q^2}{(2^n-q)^2} - \frac{4q^4}{2^n(2^n-q)^2}\right) \frac{|B_n|^2}{2^{nq}} \quad .$$

This theorem, as always, yield an upperbound on the advantage to distinguish the XOR of $2r$ random functions:

Theorem 18. *For any integer $r \geq 1$, the advantage to distinguish between the XOR of $2r$ random permutations and a random function satisfy:*

$$\mathbf{Adv}^{\mathrm{PRF}}_{f_1 \oplus \cdots \oplus f_{2r}}(q) \leq \left(\frac{q^2}{(2^n-q)^2} + \frac{4q^4}{2^n(2^n-q)^2}\right)^r \quad .$$

So far, the best known bound is proven by Lucks [Luc00]:

Theorem 19. *For any integer $r \geq 1$, the advantage to distinguish between the XOR of $2r$ random permutations and a random function satisfy:*

$$\mathbf{Adv}^{\mathrm{PRF}}_{f_1 \oplus \cdots \oplus f_{2r}}(q) \leq \frac{q^{2r+1}}{2^{2rn}} \quad .$$

Comparison between the best known bound of Lucks and our bound We see that the bound of Lucks is small when $q << 2^{\frac{2r}{2r+1}n}$ while our bound is small when $q << 2^{\frac{3}{4}n}$. This means that the bound of Lucks is better when the number of queries is not negligible from $2^{\frac{3}{4}n}$. However, our bound is better when $q << 2^{\frac{1}{2}n}$. Indeed, in that case, the term $\frac{4q^4}{2^n(2^n-q)^2}$ is negligible compared to $\frac{q^2}{(2^n-q)^2}$ which imply that our bound is close to $\left(\frac{q^2}{(2^n-q)^2}\right)^r$ which is about q times smaller than Lucks's bound.

References

[CS13] Shan Chen and John P. Steinberger, *Tight security bounds for key-alternating ciphers*, IACR Cryptology ePrint Archive **2013** (2013), 222.

[DP10] Yevgeniy Dodis and Krzysztof Pietrzak, *Leakage-resilient pseudorandom functions and side-channel attacks on Feistel networks*, CRYPTO, 2010, pp. 21–40.

[GM09] Peter Gazi and Ueli M. Maurer, *Cascade encryption revisited*, ASIACRYPT, 2009, pp. 37–51.

[HR10] Viet Tung Hoang and Phillip Rogaway, *On generalized Feistel networks*, CRYPTO, 2010, pp. 613–630.

[LPS12] Rodolphe Lampe, Jacques Patarin, and Yannick Seurin, *An asymptotically tight security analysis of the iterated even-mansour cipher*, ASIACRYPT, 2012, pp. 278–295.

[LR85] Michael Luby and Charles Rackoff, *How to construct pseudo-random permutations from pseudo-random functions (abstract)*, CRYPTO, 1985, p. 447.

[Luc00] Stefan Lucks, *The sum of prps is a secure prf*, EUROCRYPT, 2000, pp. 470–484.

[Mau92] Ueli M. Maurer, *A simplified and generalized treatment of Luby-Rackoff pseudorandom permutation generator*, EUROCRYPT, 1992, pp. 239–255.

[Mau02] ———, *Indistinguishability of random systems*, EUROCRYPT, 2002, pp. 110–132.

[MP03] Ueli M. Maurer and Krzysztof Pietrzak, *The security of many-round Luby-Rackoff pseudo-random permutations*, EUROCRYPT, 2003, pp. 544–561.

[MP04] ———, *Composition of random systems: When two weak make one strong*, TCC, 2004, pp. 410–427.

[MPR07] Ueli M. Maurer, Krzysztof Pietrzak, and Renato Renner, *Indistinguishability amplification*, CRYPTO, 2007, pp. 130–149.

[NP13] Valerie Nachef and Jacques Patarin, *A proof of security in o(2n) for the xor of two random permutations*, IACR Cryptology ePrint Archive **2013** (2013).

[NR99] Moni Naor and Omer Reingold, *On the construction of pseudorandom permutations: Luby-Rackoff revisited*, J. Cryptology **12** (1999), no. 1, 29–66.

[Pat98] Jacques Patarin, *About Feistel schemes with six (or more) rounds*, FSE, 1998, pp. 103–121.

[Pat01] ———, *Generic attacks on Feistel schemes*, ASIACRYPT, 2001, pp. 222–238.

[Pat04] ———, *Security of random Feistel schemes with 5 or more rounds*, CRYPTO, 2004, pp. 106–122.

[Pat08] ———, *The "Coefficients H" Technique*, Selected Areas in Cryptography, 2008, pp. 328–345.

[Pat10] ———, *Security of balanced and unbalanced Feistel schemes with linear non equalities*, IACR Cryptology ePrint Archive **2010** (2010), 293.

[Vau98] Serge Vaudenay, *Provable security for block ciphers by decorrelation*, STACS, 1998, pp. 249–275.

Appendices

A Proof of Theorem 1

Consider an attacker A who can query q times an oracle O. The oracle O acts all the time like a random $f \in F$ or like a random permutation. The attacker can make direct queries or inverse queries. After q queries, the attacker outputs 1 or 0. We denote P_1 the probability that A outputs 1 if O is $f \in F$ and P_1^* the probability that A outputs 1 if O is a random permutation. Our goal is to upperbound $|P_1 - P_1^*|$. We denote $\gamma_1, ..., \gamma_q$ the q queries and $\delta_1, ..., \delta_q$ the q answers. If the ith query is direct then $\delta_i = O(\gamma_i)$ and if the ith query is inverse then $\delta_i = O^{-1}(\gamma_i)$. If you know $\delta_1, ..., \delta_q$, you have uniquely defined the q inputs $X_1, ..., X_q$ and the q outputs $Y_1, ..., Y_q$. For all $\delta = (\delta_1, ..., \delta_q)$, we denote $X(\delta) = (X_1, ..., X_q)$ and $Y(\delta) = (Y_1, ..., Y_q)$. We denote $\Sigma = \{\delta$ such that A outputs $1\}$. For a fixed δ, a random permutation send γ on δ with probability

$$\frac{1}{2^{2nq}(1 - \frac{q(q-1)}{2 \cdot 2^{2n}})} \quad .$$

Indeed, there is q outputs of $2n$ bits so there is 2^{2nq} different outputs. We need them to be pairwise distinct and the probability that it exists $i < j$ such that $Y_i = Y_j$ is $\frac{q(q-1)}{2 \cdot 2^{2n}}$ because there is $\frac{q(q-1)}{2}$ possibilities for the choices of i and j and a probability $\frac{1}{2^{2n}}$ that these two queries are equal. Then:

$$P_1^* = \frac{|\Sigma|}{2^{2nq}(1 - \frac{q(q-1)}{2 \cdot 2^{2n}})} \quad .$$

We define C the set of $f \in F$ such that A outsputs 1 if $O = f$. Then $P_1 = \frac{|C|}{|F|}$. For all $\delta \in \Sigma$, we denote C_δ the set of functions f such that f sends $X(\delta)$ on $Y(\delta)$. Then

$$P_1 = \sum_{\delta \in \Sigma} \frac{|C_\delta|}{|F|} = \sum_{\delta \in \Sigma} \frac{H(X(\delta), Y(\delta))}{|F|} \quad .$$

Now, remember the hypothesis : $\frac{H(X(\delta), Y(\delta))}{|F|} \geq (1 - \alpha) \times \frac{1}{2^{2nq}}$. So $P_1 \geq \frac{|\Sigma|(1-\alpha)}{2^{2nq}}$ and $P_1 \geq P_1^*(1 - \alpha)(1 - \frac{q(q-1)}{2 \cdot 2^{2n}})$ which imply

$$P_1 - P_1^* \geq -\alpha - \frac{q(q-1)}{2 \cdot 2^{2n}} \quad .$$

Doing all the same reasoning for the 0 output:

$$(1 - P_1) - (1 - P_1^*) \geq -\alpha - \frac{q(q-1)}{2 \cdot 2^{2n}}$$

which is equivalent to

$$P_1 - P_1^* \leq \alpha + \frac{q(q-1)}{2 \cdot 2^{2n}} \quad .$$

This proves

$$|P_1 - P_1^*| \leq \alpha + \frac{q(q-1)}{2 \cdot 2^{2n}} \quad .$$

B Proof of Theorem 2

Consider an attacker A who can query q times an oracle O. The oracle O acts all the time like a random $f \in F$ or like a random permutation. The attacker can make direct queries or inverse queries. After q queries, the attacker outputs 1 or 0.
we denote

$$P_1 = \text{Probability that A outputs 1 if } O \text{ is } f \in F$$

and

$$P_1^* = \text{Probability that A outputs 1 if } O \text{ is a random permutation } .$$

Our goal is to upperbound $|P_1 - P_1^*|$.
we denote $\gamma_1, ..., \gamma_q$ the q queries and $\delta_1, ..., \delta_q$ the q answers. If the ith query is direct, one has $\delta_i = O(\gamma_i)$, if the ith query is inverse, one has $\delta_i = O^{-1}(\gamma_i)$.
If you know $\delta_1, ..., \delta_q$, you have uniquely defined the q inputs $X_1, ..., X_q$ and the q outputs $Y_1, ..., Y_q$.
For all $\delta = (\delta_1, ..., \delta_q)$, we denote $X(\delta) = (X_1, ..., X_q)$ and
$Y(\delta) = (Y_1, ..., Y_q)$.
we denote $\Sigma = \{\delta \text{ such that } A \text{ outputs } 1\}$. For a fixed δ, a random permutation send γ on δ with probability

$$\frac{1}{2^{2n}(2^{2n}-1) \times \cdots \times (2^{2n}-q+1)}$$

so

$$P_1^* = \frac{|\Sigma|}{2^{2n}(2^{2n}-1) \times \cdots \times (2^{2n}-q+1)} \ .$$

we denote $C = \{f \in F \text{ such that } A \text{ outsputs } 1 \text{ if } O = f\}$. One has

$$P_1 = \frac{|C|}{|F|} \ .$$

If we denote, for all $\delta \in \Sigma$, C_δ the set of functions f such that f sends $X(\delta)$ on $Y(\delta)$ then

$$P_1 = \sum_{\delta \in \Sigma} \frac{|C_\delta|}{|F|} = \sum_{\delta \in \Sigma} \frac{H(X(\delta), Y(\delta))}{|F|} \ .$$

Now, remember the hypothesis :

$$\frac{H(X(\delta), Y(\delta))}{|F|} \geq (1-\alpha) \times \frac{1}{2^{2n}(2^{2n}-1) \times \cdots \times (2^{2n}-q+1)} \ .$$

So

$$P_1 \geq \frac{|\Sigma|(1-\alpha)}{2^{2n}(2^{2n}-1) \times \cdots \times (2^{2n}-q+1)} \ .$$

So

$$P_1 \geq P_1^*(1-\alpha)$$
$$\Rightarrow P_1 - P_1^* \geq -\alpha \ .$$

Doing all the same reasoning for the 0 output, one has

$$(1-P_1) - (1-P_1^*) \geq -\alpha$$

which is equivalent to

$$P_1 - P_1^* \leq \alpha \ .$$

This proves

$$|P_1 - P_1^*| \leq \alpha \ .$$

C Proof of Theorem 5

We fix $X_1, ..., X_q$ pairwise distinct and $Y_1, ..., Y_q$ pairwise distinct. One has

$$H^{F \circ G}(X, Y) = \sum_T H^G(X, T) \times H^F(T, Y), \tag{9}$$

the sum being taken over the $T_1, ..., T_q$ pairwise distinct. We notice that one has:

$$\sum_T 1 = j_q \tag{10}$$

$$\sum_T H^F(T, Y) = |F| \tag{11}$$

$$\sum_T H^G(X, T) = |G| \tag{12}$$

We compute the right part of equality (9) by introducing the values $m_F = (1 - \alpha_F) \times \frac{|F|}{j_q}$ and $m_G = (1 - \alpha_G) \times \frac{|G|}{j_q}$:

$$\begin{aligned}
& \sum_T H^G(X, T) \times H^F(T, Y) \\
= & \sum_T ((H^G(X, T) - m_G) + m_G) \times ((H^F(T, Y) - m_F) + m_F) \\
= & \sum_T (H^G(X, T) - m_G) \times (H^F(T, Y) - m_F) \\
& + \sum_T m_G \times (H^F(T, Y) - m_F) + \sum_T (H^G(X, T) - m_G) \times m_F + \sum_T m_G m_F.
\end{aligned}$$

By hypothesis, the first term is positive. We now compute the second term:

$$\begin{aligned}
& \sum_T m_G \times (H^F(T, Y) - m_F) \\
= & \ (1 - \alpha_G) \times \frac{|G|}{j_q} \left(\sum_T H^F(T, Y) - \sum_T (1 - \alpha_F) \times \frac{|F|}{j_q} \right) \\
\geq & (1 - \alpha_G) \times \frac{|G|}{j_q} \left(|F| - (1 - \alpha_F)|F| \Big(\sum_T \frac{1}{j_q} \Big) \right) \text{ from (11)} \\
\geq & \frac{|G| \times |F|}{j_q} \times (1 - \alpha_G)\alpha_F \text{ from (10)}
\end{aligned}$$

The third term is computed the same way. The fourth term gives:

$$\sum_T m_G m_F = \frac{|G||F|}{j_q} \times (1 - \alpha_G)(1 - \alpha_F) \ .$$

Summing the four terms, one has:

$$\sum_T H^G(X,T) \times H^F(T,Y) \geq \frac{|G||F|}{j_q} \times (1 - \alpha_F \alpha_G) \ .$$

D Proof of the H Coefficients Local Composition Theorem

For any family of permutations F of B_n, any $\ell \leq q \leq 2^n$, any pairwise distinct $X_1, \ldots, X_q \in \{0,1\}^n$ and any pairwise distinct $Y_1, \ldots, Y_q \in \{0,1\}^n$, we denote H_ℓ^F the number of permutations $f \in F$ sending X_i to Y_i for all i from 1 to ℓ.

Theorem 20. *If it exists α_F and α_G in $[0;1]$ such that, for all $X_1, ..., X_{\ell+1}$ pairwise distinct and for all $Y_1, ..., Y_{\ell+1}$ pairwise distinct:*

$$H_{\ell+1}^F(X,Y) \geq (1-\alpha_F) \times \frac{H_\ell^F(X,Y)}{2^n - \ell} \text{ and } H_{\ell+1}^G(X,Y) \geq (1-\alpha_G) \times \frac{H_\ell^G(X,Y)}{2^n - \ell},$$

then, for all $X_1, ..., X_{\ell+1}$ pairwise distinct and for all $Y_1, ..., Y_{\ell+1}$ pairwise distinct:

$$H_{\ell+1}^{F \circ G}(X,Y) \geq (1 - \alpha_F \alpha_G) \times \frac{H_\ell^{F \circ G}(X,Y)}{2^n - \ell} \ .$$

Proof. We fix $X_1, ..., X_{\ell+1}$ pairwise distinct and $Y_1, ..., Y_{\ell+1}$ pairwise distinct. One has

$$H_{\ell+1}^{F \circ G}(X,Y) = \sum_{T_1, \ldots, T_l, T_{\ell+1}} H_{\ell+1}^G(X,T) \times H_{\ell+1}^F(T,Y)$$

and

$$H_l^{F \circ G}(X,Y) = \sum_{T_1, \ldots, T_l} H_l^G(X,T) \times H_l^F(T,Y)$$

with $T_1, ..., T_{\ell+1}$ pairwise distinct. so we prove the theorem if we prove that, for every $T_1, ..., T_l$ pairwise distinct, one has

$$\sum_{T_{\ell+1}} H_{\ell+1}^G(X,T) \times H_{\ell+1}^F(T,Y) \geq (1 - \alpha_F \alpha_G) \frac{H_\ell^G(X,T) \times H_\ell^F(T,Y)}{2^n - \ell}$$

with the sum taken over the choices of $T_{\ell+1}$ such that $T_1, ..., T_{\ell+1}$ are pairwise distinct.
We compute the left part of this inequality by introducing the values $m_F = (1-\alpha_F) \times \frac{H_\ell^F(T,Y)}{2^n - \ell}$ and $m_G = (1 - \alpha_G) \times \frac{H_\ell^G(X,T)}{2^n - \ell}$:

$$\begin{aligned}
& \sum_{T_{\ell+1}} H_{\ell+1}^G(X,T) \times H_{\ell+1}^F(T,Y) \\
= & \sum_{T_{\ell+1}} ((H_{\ell+1}^G(X,T) - m_G) + m_G) \times ((H_{\ell+1}^F(T,Y) - m_F) + m_F) \\
= & \sum_{T_{\ell+1}} (H_{\ell+1}^G(X,T) - m_G) \times (H_{\ell+1}^F(T,Y) - m_F) \\
& + \sum_{T_{\ell+1}} m_G \times (H_{\ell+1}^F(T,Y) - m_F) + \sum_{T_{\ell+1}} (H_{\ell+1}^G(X,T) - m_G) \times m_F + \sum_{T_{\ell+1}} m_G m_F.
\end{aligned}$$

By hypothesis, the first term is positive. We now compute the second term:

$$\begin{aligned}
&\sum_{T_{\ell+1}} m_G \times (H^F_{\ell+1}(T,Y) - m_F) \\
&= (1-\alpha_G) \times \frac{H^G_\ell(X,T)}{2^n-\ell}\left(\sum_{T_{\ell+1}} H^F_{\ell+1}(T,Y) - \sum_{T_{\ell+1}} (1-\alpha_F) \times \frac{H^F_\ell(X,T)}{2^n-\ell}\right) \\
&\geq (1-\alpha_G) \times \frac{H^G_\ell(X,T)}{2^n-\ell}\left(H^F_\ell(T,Y) - (1-\alpha_F)H^F_\ell(T,Y)\Big(\sum_{T_{\ell+1}} \frac{1}{2^n-\ell}\Big)\right) \\
&\geq \quad \frac{H^G_\ell(X,T)H^F_\ell(T,Y)}{2^n-\ell}\left((1-\alpha_G) - (1-\alpha_G)(1-\alpha_F)\Big(\sum_{T_{\ell+1}} \frac{1}{2^n-\ell}\Big)\right)
\end{aligned}$$

The third term is computed the same way. The fourth term gives:

$$\sum_{T_{\ell+1}} m_G m_F = \frac{H^G_\ell(X,T)H^F_\ell(T,Y)}{2^n-\ell}\left((1-\alpha_G)(1-\alpha_F)\Big(\sum_{T_{\ell+1}} \frac{1}{2^n-\ell}\Big)\right) \ .$$

Summing the four terms, one has:

$$\sum_{T_{\ell+1}} H^G_{\ell+1}(X,T) \times H^F_{\ell+1}(T,Y) \geq \frac{H^G_\ell(X,T)H^F_\ell(T,Y)}{2^n-\ell}\left((1-\alpha_G)+(1-\alpha_F)-(1-\alpha_G)(1-\alpha_F)\Big(\sum_{T_{\ell+1}} \frac{1}{2^n-\ell}\Big)\right) \ .$$

There is at most $2^n - \ell$ choices for $T_{\ell+1}$ because $T_1, ..., T_{\ell+1}$ are pairwise distinct. So

$$(1-\alpha_G)+(1-\alpha_F)-(1-\alpha_G)(1-\alpha_F)\Big(\sum_{T_{\ell+1}} \frac{1}{2^n-\ell}\Big) \geq (1-\alpha_G)+(1-\alpha_F)-(1-\alpha_G)(1-\alpha_F) = 1-\alpha_G\alpha_F,$$

which ends the proof. □

Our two main theorems (Theorems 4 and 6) are called global because it compares the number H of permutations (or functions) sending all inputs to all outputs to a constant number. In theorem 20 however, we call it local because we compare the number H for the first $\ell + 1$ queries to number H for the first ℓ queries. Intuitively, the idea is to analyze the probability to connect the $\ell + 1$ query when you have already connected the first ℓ queries. This is the strategy used in many security proofs using the Coupling technique (for example [HR10] and [LPS12]).