Two Simple Composition Theorems with H-coefficients

Jacques Patarin

Two Simple Composition Theorems with H-coefficients

Jacques Patarin

Laboratoire de Mathématiques de Versailles, UVSQ,
CNRS, Université Paris-Saclay, 78035 Versailles, France
jpatarin@club-internet.fr

Abstract. We will present two new and simple theorems that show that when we compose permutation generators with independent keys, then the "quality" of CCA security increases. These theorems (Theorems 2 and 5 of this paper) are written in terms of H-coefficients (which are nothing else, up to some normalization factors, than transition probabilities). Then we will use these theorems on the classical analysis of Random Feistel Schemes (i.e. Luby-Rackoff constructions) and we will compare the results with the coupling technique. Finally, we will show an interesting difference between 5 and 6 Random Feistel Schemes. With 5 rounds on $2n$ bits $\rightarrow$ $2n$ bits, when the number of q queries satisfies $\sqrt{2^n} \ll q \ll 2^n$, we have some "holes" in the H-coefficient values, i.e. some H values are much smaller than the average value of H. This property for 5 rounds does not exist anymore on 6 rounds.

1 Introduction

Security amplification results for block ciphers typically state that cascading (i.e. composing with independent keys) two, or more, block ciphers gives a new block cipher that offers better security against some classes of adversaries. One of the most important composition results is the so-called "two weak make one strong" theorem. This theorem was first established up to logarithmic terms by Maurer and Pietrzak [11]. It was later tightened by Maurer, Pietrzak and Renner [12]. In 2010, Cogliati, Patarin and Seurin have obtained simpler proofs of this result by using the so-called "H-coefficient technique" (cf [2]). In this paper, we will prove two new, and relatively simple, composition theorems: Theorems 2 and 5 of this paper.
These theorems are written directly in term of "H-coefficients", i.e. in term of the number of generic keys that send some plaintexts on some ciphertexts. (This is the same, up to some normalization factors, than transition probabilities). We will then show how there theorems can be useful in term of classical cryptographic security (such as CCA: adaptive chosen plaintext and ciphertext attack). We work here in term of information theory for security, i.e. the adversary can ask only for a limited number q of queries, but the number of his (or her) computations is not limited. Interestingly, Stefano Tessaro has obtained [20] very

similar composition results in term of improved security. However, Stefano Tessaro works with complexity theory (instead of information theory), so the results and the proofs of [20] are in fact very different from the results and the proofs of this paper. Then we will apply our new theorems on random Feistel schemes, and show an interesting difference between 5 and 6 rounds.

2 A simple mathematical property

Theorem 1. *Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be real numbers and let α and β be real numbers, $\alpha \geq 0$, $\beta \geq 0$ such that:*
• $\sum_{i=0}^{n} x_i = 0$.
• $\sum_{i=0}^{n} y_i = 0$.
•$\forall i$, $1 \leq i \leq n$, $x_i \geq -\alpha$.
•$\forall i$, $1 \leq i \leq n$, $y_i \geq -\beta$.
Then: $\sum_{i=1}^{n} x_i y_i \geq -n\alpha\beta$.

Proof.

$$\sum_{i=1}^{n} (x_i + \alpha)(y_i + \beta) \geq 0$$

$$\sum_{i=1}^{n} x_i y_i + \beta \sum_{i=1}^{n} x_i + \alpha \sum_{i=1}^{n} y_i + n\alpha\beta \geq 0$$

Now since $\sum_{i=1}^{n} x_i = 0$ and $\sum_{i=1}^{n} y_i = 0$, we obtain $\sum_{i=1}^{n} x_i y_i \geq -n\alpha\beta$. □

3 A composition Theorem in CCA with H-coefficients

Definition 1. *Let G be a permutation generator that generates permutations from $\{0,1\}^N$ to $\{0,1\}^N$ from a large set of parameters K. The values of K will be called "keys", despite the fact that they are generally defined with much more bits than usual cryptographic keys, and therefore G is considered as a "generic generator". Let q be an integer (called the "number of queries"). Let $a = (a_i)$, $1 \leq i \leq q$, be q pairwise distinct elements of $\{0,1\}^N$, and similarly let $b = (b_i)$, $1 \leq i \leq q$, be q pairwise distinct elements of $\{0,1\}^N$. Then, by definition, $H(a,b)$ denotes the number of keys $k \in K$ such that: $\forall i$, $1 \leq i \leq q$, $G_k(a_i) = b_i$. $H(a,b)$ is simply denoted by H when there is no risk of confusion about the values of a and b, or when we want to speak of all these coefficients $H(a,b)$.*

Definition 2. *With the same notations as above, if there exist values (a_i) pairwise distinct, and values (b_i) pairwise distinct, $1 \leq i \leq q$, such that $H(a,b)$ (for these a and b) is much smaller than the average value of H, then we say that there is a "Hole" in the H-coefficient values with q queries.*

Theorem 2. *Let G_1 and G_2 be two permutation generators (with the same key space K) such that:*

(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: $H_1 \geq \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha_1)$ and similarly $H_2 \geq \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha_2)$ where H_1 denotes the H coefficient for G_1 and H_2 the H coefficient for G_2. Then:
(2) If we compose 2 such generators G_1 and G_2 with random independent keys, for the composition generator $G' = G_2 \circ G_1$, we have: for all sequences of pairwise distinct elements a_i, $1 \leq i \leq q$, and for all sequences of pairwise distinct elements b_i, $1 \leq i \leq q$, $H' \geq \frac{|K|^2}{2^N(2^N-1)\ldots(2^N-q+1)}(1-\alpha_1\alpha_2)$, where H' denotes the H coefficient for G'.

Proof. Let $\tilde{H}_1$ (respectively $\tilde{H}_2$) denote the mean value of H_1 (respectively H_2). We have:

$$\tilde{H}_1 = \tilde{H}_2 = \frac{|K|}{2^N(2^N-1)\ldots(2^N-q+1)}$$

Let denote by $\tilde{H}'$ the mean value of H for $G' = G_2 \circ G_1$. We have

$$\tilde{H}' = \frac{|K|^2}{2^N(2^N-1)\ldots(2^N-q+1)}$$

Let $a = (a_1, \ldots, a_q)$ be q pairwise distinct plaintexts, and $b = (b_1, \ldots, b_q)$ be q ciphertexts of G'. Let J be the set of all $(t_1, \ldots, t_q)$ pairwise distinct values of $\{0,1\}^N$. We have $|J| = 2^N(2^N-1)\ldots(2^N-q+1)$. For $G' = G_2 \circ G_1$, we have:

$$H(a,b) = \sum_{t \in J} H_1(a,t)H_2(t,b)$$

We also have $\sum_{t\in J} H_1(a,t) = |K|$ and $\sum_{t\in J} H_2(t,b) = |K|$ since each key sends a value a to a specific value t. We also have $|K| = \tilde{H}_1 \cdot |J| = \tilde{H}_2 \cdot |J|$. By hypothesis, we also have:

$$\forall t \in J,\ H_1(a,t) \geq \tilde{H}_1(1-\alpha_1) \ \text{ and } H_2(a,t) \geq \tilde{H}_2(1-\alpha_2)$$

$\forall t \in J$, let $x_t = \frac{H_1(a,t)}{\tilde{H}_1} - 1$ and $y_t = \frac{H_2(t,b)}{\tilde{H}_2} - 1$. $\forall t \in J$, we have $x_t \geq -\alpha_1$, and $y_t \geq -\alpha_2$, $\sum_{t\in J} x_t = 0$ and $\sum_{t\in J} y_t = 0$. Therefore, from Theorem 1, we have $\sum_{t\in J} x_t y_t \geq -|J|\alpha_1\alpha_2$. For $G' = G_2 \circ G_1$, we have:

$$\begin{aligned} H(a,b) &= \textstyle\sum_{t\in J} H_1(a,t) \cdot H_2(t,b) \\ &= \textstyle\sum_{t\in J} \left(\tilde{H}_1 x_t + \tilde{H}_1\right)\left(\tilde{H}_2 y_t + \tilde{H}_2\right) \\ &= \textstyle\sum_{t\in J} \tilde{H}_1\tilde{H}_2 x_t y_t + \tilde{H}_1\tilde{H}_2 y_t + \tilde{H}_1\tilde{H}_2 x_t + \tilde{H}_1\tilde{H}_2 \\ &\geq -\tilde{H}_1\tilde{H}_2|J|\alpha_1\alpha_2 + |J|\tilde{H}_1\tilde{H}_2 \end{aligned}$$

Moreover $\tilde{H}' = \frac{|K|^2}{|J|} = |J|\tilde{H}_1\tilde{H}_2$. We have proved: $H(a,b) \geq \tilde{H}'(1-\alpha_1\alpha_2)$ as claimed. □

Theorem 3. *(H-coefficient technique, sufficient condition for security against CCA)*

Let α and β be real numbers, $\alpha > 0$ and $\beta > 0$
If: There exists a subset E of $(\{0,1\}^{qN})^2$ such that
(1a) For all $(a,b) \in E$, we have:

$$H(a,b) \geq \frac{|K|}{2^{Nq}}(1-\alpha)\overset{\circ}{1}$$

with

$$\overset{\circ}{1} \overset{def}{=} \frac{1}{(1-\frac{1}{2^N})(1-\frac{2}{2^N})\dots(1-\frac{q-1}{2^N})}$$

(1b) For all CCA acting on a random permutation f of $\mathcal{P}_N$, 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 CCA with q queries (i.e. q chosen plaintexts or ciphertexts) we have: $\mathbf{Adv}^{PRP} \leq \alpha + \beta$ where $\mathbf{Adv}^{PRP}$ denotes the probability to distinguish $G(f_1,\dots,f_r)$ when $(f_1,\dots,f_r) \in_R K$ from a permutation $f \in_R \mathcal{P}_N$.

Proof. This theorem is proved in [16, 17]. □

Corollary 1. *From theorem 3 (H-coefficients in CCA) with $\beta = 0$, we see that we have: $\mathbf{Adv}^{PRP} \leq \alpha_1\alpha_2$ where $\mathbf{Adv}^{PRP}$ denotes the advantage in CCA to distinguish $G_2 \circ G_1$ (when the keys are independently and randomly chosen) from a permutation $f \in_R \mathcal{P}_n$.*

By induction, we see:

Theorem 4. *Let q and k be two integers. Let $\alpha_1,\dots,\alpha_k$ be k real values. Let $G_1,\dots,G_k$ be k permutation generators such that: for all sequences of pairwise distinct elements a_i, and for all sequences of pairwise distinct elements b_i, $1 \leq i \leq q$, we have:*

$$H \geq \frac{|K|}{2^N(2^N-1)\dots(2^N-q+1)}(1-\alpha_j)$$

If we compose k such generators $G_1,\dots,G_k$ with random and independent keys, for the composition generator $G' = G_k \circ \dots \circ G_1$, we have: for all sequences of pairwise distinct elements a_i, $1 \leq i \leq q$ and for all sequences of pairwise distinct elements b_i, $1 \leq i \leq q$, $H \geq \frac{|K|}{2^N(2^N-1)\dots(2^N-q+1)}(1-\alpha_1\dots\alpha_k)$. Therefore, from theorem 3 with $\beta = 0$, we see that we have: $\mathbf{Adv}^{PRP} \leq \alpha_1\dots\alpha_k$

4 A composition theorem to eliminate a "hole"

J denotes, as above, the set of all q pairwise distinct values of $\{0,1\}^N$.

Theorem 5. *Let G_1 and G_2 be two permutation generators with the same key space K. Let H_1 (respectively H_2) denotes the H-coefficients for G_1 (respectively G_2).*

If:

(1) For all sequences of pairwise distinct elements a_i, $1 \leq i \leq q$, and for all sequences of pairwise distinct $b_i \in E_1$, $1 \leq i \leq q$, we have

$$H_1 \geq \frac{|K|}{2^N(2^N-1)\dots(2^N-q+1)}(1-\alpha_1)$$

with $|E_1| \geq |J|(1-\epsilon_1)$.

(2) Similarly, for all sequences of pairwise distinct elements a_i, $1 \leq i \leq q$, and for all sequences of pairwise distinct $b_i \in E_2$, $1 \leq i \leq q$, we have

$$H_2 \geq \frac{|K|}{2^N(2^N-1)\dots(2^N-q+1)}(1-\alpha_2)$$

with $|E_2| \geq |J|(1-\epsilon_2)$.

Then: for the composition generator $G_2^{-1} \circ G_1$, for all sequences of pairwise distinct elements a_i, and for all sequences of pairwise distinct b_i, we have

$$H' \geq \frac{|K|^2}{2^N(2^N-1)\dots(2^N-q+1)}(1-\epsilon_1-\epsilon_2)(1-\alpha_1)(1-\alpha_2)$$

where H' denotes the H-coefficients for $G_2^{-1} \circ G_1$ (wa have no hole). Moreover, if $E_1 = E_2$, then

$$H' \geq \frac{|K|^2}{2^N(2^N-1)\dots(2^N-q+1)}(1-\epsilon_1)(1-\alpha_1)(1-\alpha_2)$$

Proof. For $G' = G_2^{-1} \circ G_1$, we have: $H'(a,b) = \sum_{t \in J} H_1(a,t)H_2(t,b)$, with $\sum_{t \in J} H_1(a,t) = |K|$ and $\sum_{t \in J} H_2(t,b) = |K|$. Let $\tilde{H_1} = \frac{|K|}{|J|}$, $\tilde{H_2} = \frac{|K|}{|J|}$, and $\tilde{H'} = \frac{|K|^2}{|J|} = \tilde{H_1}\tilde{H_2}|J|$. We have: $|J| = 2^N(2^N-1)\dots(2^N-q+1)$. Let $P_1 = J \backslash E_1$ and $P_2 = J \setminus E_2$. Then

$$\begin{aligned}
H'(a,b) &\geq \sum_{t \in J \backslash P_1 \backslash P_2} H_1(a,t)H_2(t,b) \\
&\geq \sum_{t \in J \backslash P_1 \backslash P_2} \tilde{H_1}(1-\alpha_1)\tilde{H_2}(1-\alpha_2) \\
&\geq |J \setminus P_1 \setminus P_2|\tilde{H_1}(1-\alpha_1)\tilde{H_2}(1-\alpha_2) \\
&\geq |J|(1-\epsilon_1-\epsilon_2)\tilde{H_1}(1-\alpha_1)\tilde{H_2}(1-\alpha_2) \\
&\geq \frac{|K|^2}{|J|}(1-\epsilon_1-\epsilon_2)(1-\alpha_1)(1-\alpha_2)
\end{aligned}$$

as claimed. □

5 Comments about the composition theorems

These very simple theorems of composition (Theorem 2 and Theorem 5) are not very well known because the classical theorems of composition (with more difficult proofs) usually do not consider hypothesis in term of the values on the H coefficients. (Sometimes, as in [2], H-coefficients are used for the proofs of the Theorems, but not in the terms of the Theorems). For example, the famous "two weak make one strong" theorem of Maurer and Pietrzak [9, 12] says that if F and G are NCPA secure, then the composition $G^{-1} \circ F$ is CCA secure. This result only holds in the information-theoretic setting, not in the computational setting (cf [15, 19]). Another example is this theorem of [2]:

Theorem 6. *(i.e. [2] Theorem 5 p.17)*
Let E, F and G be 3 block ciphers with the same message space M. Denote $\epsilon_E = \mathbf{Adv}_E^{NCPA}(q)$, $\epsilon_F = \mathbf{Adv}_F^{NCPA}(q)$, $\epsilon_{F^{-1}} = \mathbf{Adv}_{F^{-1}}^{NCPA}(q)$ and $\epsilon_{G^{-1}} = \mathbf{Adv}_{G^{-1}}^{NCPA}(q)$, where q is the number of queries. We have:

$$\mathbf{Adv}_{G\circ F\circ E}^{CCA}(q) \leq \epsilon_E\epsilon_F + \epsilon_E\epsilon_{G^{-1}} + \epsilon_{F^{-1}}\epsilon_{G^{-1}} + \min\left\{\epsilon_E\epsilon_F, \epsilon_E\epsilon_{G^{-1}}, \epsilon_{F^{-1}}\epsilon_{G^{-1}}\right\}$$

Why do we have 3 rounds in this theorem and only 2 rounds in theorem 2 for the product of the advantages? (Moreover theorem 6 was also proved by using the H-coefficient technique [2]). This is because in theorem 2, we used the additional property that there are no "holes" in the hypothesis that H is greater than or equal to the mean value $H(1-\epsilon)$, i.e. that this property was true for any q pairwise distinct inputs and q pairwise distinct outputs.
It is also interesting to compare our new Theorem 4 ($\mathbf{Adv}^{PRP} \leq \alpha_1 \dots \alpha_k$) with these theorems of [2]:

Theorem 7. *(i.e. [2] Theorem 2 p.10)*
Let $E_1, \dots, E_n$ be n block ciphers with the same message space $\mathcal{M}$. For any integer q, one has

$$\mathbf{Adv}_{E_n\circ\cdots\circ E_1}^{cca}(q) \leq 2^{n-1} \max_{1\leq i\leq n} \left(\prod_{1\leq j\leq i-1} \mathbf{Adv}_{E_j}^{ncpa}(q) \times \prod_{i+1\leq j\leq n} \mathbf{Adv}_{E_j^{-1}}^{ncpa}(q) \right).$$

Corollary 2. *(i.e. [2] Corollary 1 p.11)*
Let $E_1, \dots, E_n$ be n block ciphers with the same message space $\mathcal{M}$. Fix $q \geq 1$. For $i = 1, \dots, n$, let $\epsilon_i = \max\{\mathbf{Adv}_{E_i}^{ncpa}(q), \mathbf{Adv}_{E_i^{-1}}^{ncpa}(q)\}$. Then one has

$$\mathbf{Adv}_{E_n\circ\cdots\circ E_1}^{cca}(q) \leq 2^{n-1} \max_{1\leq i\leq n} \prod_{\substack{1\leq j\leq n \\ j\neq i}} \epsilon_j.$$

We see that with our new Theorem 4, we do not have the coefficient 2^{n-1}, and also we do not lost one of the n products. Therefore, if all the $\epsilon_i = \epsilon$ for example, we will get $\mathbf{Adv}^{CCA} \leq \epsilon^n$ instead of $\mathbf{Adv}^{CCA} \leq 2^{n-1}\epsilon^{n-1}$. However, in order to use our new Theorem 4, we need two conditions that were not in Theorem 7: the fact that we have "no hole" and an expression of ϵ directly in terms of the H-coefficients instead of $\mathbf{Adv}^{CCA}$. Therefore our Theorems and the theorems of [2] are both useful.

6 Application to Feistel Ciphers

We denote by Ψ^k a generic balanced Feistel Cipher with k rounds, i.e. a balanced Feistel cipher from $\{0,1\}^{2n}$ to $\{0,1\}^{2n}$ with k rounds, where the round functions are k random functions from $\{0,1\}^n$ to $\{0,1\}^n$. Ψ^k is also called a Luby-Rackoff's construction. We will show here how our new theorems can be useful for cryptographic security results on Ψ^k. (However, our new theorems are also interesting independently of these problems). The generic security problem has been intensively studied by many authors (for example [3, 10, 18]) since Luby and Rackoff major paper [8]. In [10], it was proved that when $k \to +\infty$, we have CCA security on Ψ^k when the number of queries q satisfies $q \ll 2^n$, and some explicit bounds for the Advantage in CCA are given. These bounds were later improved and at present, the best security bounds are obtained via the "H-coefficient technique", or via the coupling technique. These two techniques are very different and, interestingly, they give slightly different results.

Results with the H-coefficient technique
A general view of the H-coefficient technique is given in [16, 17] with the connections between these H-coefficients and various cryptographic securities (KPA, CPA, CCA,...). In 2016, in [4], another general H-coefficient theorem for CCA was proved. Essentially, the idea (of the results of [4]) is that, instead of introducing some sets E with good or bad properties (as in [17]), a computation of the mean value (computed with the probability on random permutations) is introduced. This is called the "Expectation Method" in [4].
In [18], the H-coefficient technique was used to study the security of Ψ^k. The main result was that we have CCA security for $q \ll 2^n$ not only when $k \to \infty$, but already after a finite number of rounds. More precisely, this property occurs for Ψ^k when $k \geq 5$ and an explicit bound for the Advantage in CCA is given in [18] for Ψ^6. In [1], the H-coefficient technique was used to obtain tight security bounds on Even-Mansour Cipers. From [18], we have (cf theorem 6 p.8):

Theorem 8. *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* $\alpha = \frac{8q}{2^n}$ *if* $q \leq \frac{2^n}{67n}$.

From this and Theorem 3, we obtain:

Theorem 9. *When* $q \leq \frac{2^n}{67n}$,

$$\mathbf{Adv}^{CCA}(\Psi^6) \leq \frac{8q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}}$$

Proof.

$$2^N(2^N-1)\dots(2^N-q+1) \geq 2^{qN}\left(1 - \frac{1+2+\dots+(q-1)}{2^N}\right) \geq 2^{qN}\left(1 - \frac{q(q-1)}{2 \cdot 2^N}\right)$$

Therefore, for Ψ^6, when $q \leq \frac{2^n}{67n}$

$$H \geq \frac{|\mathcal{F}_n|^6}{2^{2n}(2^{2n}-1)\dots(2^{2n}-q+1)} \left(1 - \frac{q^2}{2 \cdot 2^{2n}}\right) \left(1 - \frac{8q}{2^n}\right)$$

where $\mathcal{F}_n$ denotes the set of all functions from $\{0,1\}^n$ to $\{0,1\}^n$.

$$H \geq \frac{|\mathcal{F}_n|^6}{2^{2n}(2^{2n}-1)\dots(2^{2n}-q+1)} \left(1 - \frac{q^2}{2 \cdot 2^{2n}} - \frac{8q}{2^n}\right)$$

Now from this and Theorem 3 (with $\beta = 0$), we obtain:

$$\mathbf{Adv}^{CCA}(\Psi^6) \leq \frac{8q}{2^n} + \frac{q^2}{2 \cdot 2^{2n}}$$

as claimed.

Results with the Coupling Technique.
The coupling technique is a major tool from the theory of Markov chains that allows to conveniently upper bound the so-called mixing time of a chain, i.e. the number of steps it takes for the chain, starting from any distribution, to be at statistical distance at most ϵ from its stationary distribution. The first use of coupling in cryptography is due to Mironov [13], who used it to analyze the RC4 stream cipher. It was first applied to (maximally unbalanced) Feistel ciphers by Morris, Rogaway and Stegers [14]. This was generalized to other types of Feistel ciphers (including the balanced Feistel Ψ^k) by Hoang and Rogaway [3]. Subsequently, the coupling technique was used to analyze the iterated Even-Mansour Cipher [5], tweakable block ciphers constructions [6] and Feistel schemes where the round functions are of the form: $x \to F(x \oplus k)$ where F is a random oracle and k the secret key [7].
From [3], we have

Theorem 10. *With $k' = \lfloor \frac{(k-1)}{2} \rfloor$, we have:*

$$\mathbf{Adv}^{NCPA}(\Psi^k) \leq \frac{2^{k'}}{k'+1} \cdot \frac{q^{k'+1}}{2^{k'n}}$$

and

$$\mathbf{Adv}^{CCA}(\Psi^{2k-1}) \leq \frac{2^{k'}}{k'+1} \cdot \frac{q^{k'+1}}{2^{k'n}}$$

From Theorem 10, we see that with the coupling technique, we obtain:

NCPA: Ψ^3 has security when $q^{n/2}$
Ψ^5 has security when $q^{2n/3}$
Ψ^7 has security when $q^{3n/4}$
etc.

CCA: Ψ^5 has security when $q^{n/2}$
Ψ^7 has security when $q^{2n/3}$
Ψ^9 has security when $q^{3n/4}$
etc.

Therefore, in terms of queries, Theorem 2 (from H-coefficient technique) gives a better bound than Theorem 3 (from the coupling technique), since it gives CCA security for Ψ^6 when $q \ll 2^n$ (and therefore for Ψ^k, for all $k \geq 6$). However:

1. The proofs of Theorem 8 and Theorem 9 are much more complex than the proof of Theorem 10.
2. For a fixed value q, the **Adv** given in Theorem 10 is bounded by term that can be as small as wanted when k increases, unlike Theorem 2 where **Adv** is fixed when q and n are fixed.

Results with our new Theorems
In a way from our new Theorem 4, we can get "the best of the two worlds", since from it and Theorem 8, we obtain:

Theorem 11. *For all integer $k \geq 1$, when $q \leq \frac{2^n}{67n}$, we have:*

$$\mathbf{Adv}^{CCA}(\Psi^{6k}) \leq \left(\frac{8q}{2^n} + \frac{q^2}{2.2^{2n}}\right)^k$$

Proof. In the proof of Theorem 9, we have seen that for Ψ^6, we have, when $q \leq \frac{2^n}{67n}$,

$$H \geq \frac{|\mathcal{F}_n|^6}{2^{2n}(2^{2n}-1)\dots(2^{2n}-q+1)}\left(1 - \frac{8q}{2^n} - \frac{q^2}{2.2^{2n}}\right)$$

Therefore, from our new composition Theorem 4, we obtain that for Ψ^{6k}, when $q \leq \frac{2^n}{67n}$,

$$H \geq \frac{|\mathcal{F}_n|^{6k}}{2^{2n}(2^{2n}-1)\dots(2^{2n}-q+1)}\left(1 - \left(\frac{8q}{2^n} + \frac{q^2}{2.2^{2n}}\right)^k\right)$$

Theorem 11 is now obtained from this and Theorem 3 with $\beta = 0$.

This is the best bound known at present on Ψ^k: when $q \ll 2^n$, it gives CCA security, and when q and n are fixed such that $\frac{8q}{2^n} + \frac{q^2}{2.2^{2n}} < 1$, the bound can be as small as wanted by increasing k.

7 Other CCA bounds on Ψ^k

Worse bounds, but simpler proofs
When we look at the (difficult) proof of Theorem 8 on Ψ^6, we see that security when $q \ll 2^{3n/2}$ can easily be done. The security when $q \ll 2^{3n/4}$ is also relatively easy, and $q \ll 2^{4n/5}$ is a bit more complex.
Therefore, is is possible to stop the proof at, say, $q \ll 2^{4n/5}$ and then to use the coupling technique from Ψ^6 (instead of Ψ^3) or to use our new Theorem 4 in order to obtain a security bound. This bound will not be as good as the bound of Theorem 11, but the proof will be much simpler: we see that we have many possible tradeoffs between the quality of the bounds and the simplicity of the proofs.

Better bounds
Our Theorem 11 is the best explicit bound known at present on Ψ^k. However, it is expected that this bound can still be improved (not in term of queries: the bound $q \ll 2^n$ already obtained on Ψ^k is optimal in information complexity, but this bound can be improved in term of smaller value for $\mathbf{Adv}^{CCA}$). One way to obtain better bounds would be to analyze Ψ^{5k} instead of Ψ^{6k}. Ψ^5 is CCA secure when $q \ll 2^n$ (cf [18]), but in Ψ^5 (unlike Ψ^6), we have "holes" when $\sqrt{2^n} \ll q \ll 2^n$ (cf Appendix B of this paper). Therefore, we cannot use our new composition Theorem 4 on Ψ^{5k} (unlike what we did on Ψ^{6k}). However, Theorem 7 and Corollary 2 of [2] can be used on Ψ^{5k}. Due to the coefficient 2^{n-1} and to the fact that we loose one term ϵ_i of the product in Theorem 7 and Corollary 2 (see section 5) we expect our results on Ψ^{6k} to be better than the results on $\Psi^{5k'}$ (obtained from Theorem 7) for small values of k and k'. However, for large values of k and k', the results on $\Psi^{5k'}$ should be better. We will not do it in this paper more precisely since we do not have an explicit bound for CCA security on Ψ^5 (but just the fuzzy bound $q \ll 2^n$). Moreover, in this paper, we study CCA security of Ψ^6 mainly to illustrate our new composition results.

References

1. Chen, S., Steinberger, J.P.: Tight Security Bounds for Key-Alternating Ciphers. Advances in Cryptology – EUROCRYPT '14, P.Q. Nguyen, E. Oswald (eds), 237-350, Springer-Verlag, Lecture Notes in Computer Science, 8441, (2014)
2. Cogliati, B., Patarin, J., Seurin, Y.: Security Amplification for the Composition of Block Cipher: Simpler Proofs and New Results. Selected Areas in Cryptography– SAC '14, A. Joux, A. Yousssef (eds), 129-146, Springer-Verlag, Lecture Notes in Computer Science, 8781, (2014)
3. Hoang, V.T., Rogaway, P.: On Generalized Feistel Networks. Advances in Cryptology – CRYPTO 2010, Tal Rabin (ed), 613-630, Springer-Verlag, Lecture Notes in Computer Science, 6223, (2010)
4. Hoang, V.T., Tessaro, S.: Key-Alternating Ciphers and Key-Length extension: Exact bounds and Multi-User Security. Advances in Cryptology – CRYPTO 2016, M. Robshaw, J. Katz (eds), 3-32, Springer-Verlag, Lecture Notes in Computer Science, 9814, (2016)
5. Lampe, R., Patarin, J., Seurin, Y.: An Asymptotically Tight Security Analysis of the Iterated Advances in Cryptology – ASIACRYPT 2012, X. Wang, K. Sako (eds), 278-295 Springer-Verlag, Lecture Notes in Computer Science, 7658, (2012)
6. Lampe, R., Seurin, Y.: Tweakable Blockciphers with Asymptotically Optimal Security. Fast Software Encryption - FSE 2013, S. Moriai, (ed), 133-146, Springer-Verlag, Lecture Notes in Computer Science, 8424, (2013)
7. Lampe, R., Seurin, Y.: Security Analysis of Key-Alternating Feistel Ciphers. Fast Software Encryption - FSE 2013, C. Cid, C. Rechberger (eds), 243-264, Springer-Verlag, Lecture Notes in Computer Science, 840, (2014)
8. Luby, M., Rackoff, C.: How to construct pseudo-random permutations from pseudo-random functions. SIAM Journal on Computing **17**, 373–386 (1988)
9. Maurer, U.: Indistinguishability of Random Systemes. Advances in Cryptology – EUROCRYPT '02, L.R. Knudsen (ed), 110-132, Springer-Verlag, Lecture Notes in Computer Science, 2332, (2002)

10. Maurer, U., Pietrzak, K.: The Security of Many-Round Luby-Rackoff Pseudo-Random Permutations. Advances in Cryptology – EUROCRYPT '03, E. Biham (ed), 544-561, Springer-Verlag, Lecture Notes in Computer Science, 2656, (2003)
11. Maurer, U., Pietrzak, K.: Composition of Random Systems: When Two Weak Make One Strong. Theory of Cryptography – TCC '04, Naor, M. (ed), 410-427, Springer-Verlag, Lecture Notes in Computer Science, 2951, (2004)
12. Maurer, U., Pietrzak, K., Renner, R.: Indistinguishability Amplification. Advances in Cryptology – CRYPTO '07, A. Menezes (ed), 130-149, Springer-Verlag, Lecture Notes in Computer Science, 4622, (2007)
13. Mironov, I.:. (Not So) Random Shuffles of RC4. Advances in Cryptology – CRYPTO '02, M. Yung (ed), 304-319, Springer-Verlag, Lecture Notes in Computer Science, 2442, (2002)
14. Morris, B., Rogaway, P., Stegers,T.: How to Encipher Messages on a Small Domain. Advances in Cryptology – CRYPTO '09, S. Halevi (ed), 286-302, Springer-Verlag, Lecture Notes in Computer Science, 5677, (2009)
15. Myers, S.: Black-Box Composition Does Not Imply Adaptive Security. Advances in Cryptology – EUROCRYPT '04, C. Cachin, J.L. Camenisch (eds), 189-206, Springer-Verlag, Lecture Notes in Computer Science, 3027, (2004)
16. J. Patarin, *Étude des Générateurs de Permutations Pseudo-aléatoires basés sur le schéma du D.E.S.*, PhD, November 1991.
17. Patarin, J.: The "coefficient H" technique. Selected Areas in Cryptography – SAC '08, R. Avanzi, L. Keliher, F. Sica (eds), 328-345, Springer-Verlag, Lecture Notes in Computer Science, 5381, (2009)
18. Patarin, J.: Security of Balanced and Unbalanced Feistel Schemes with Linear Non Equalities , in *Cryptology ePrint Archive: Report 2010/293.*
19. Pietrzak, K.: Composition Does Not Imply Adaptive Security. Advances in Cryptology – CRYPTO '05, V. Shoup (ed), 55-65, Springer-Verlag, Lecture Notes in Computer Science, 3621, (2005)
20. Tessaro, S.: Security Amplification for the Cascade of Arbitrarily Weak PRPs: Tight Bounds via the Interactive Hardcore Lemma. Theory of Cryptography – TCC '11, Ishai, Y. (eds), 37-54, Springer-Verlag, Lecture Notes in Computer Science, 6597, (2011)

A An exact formula for the H-coefficient for $\Psi^k, 1 \leq k \leq 5$

The aim of this Appendix A is to prove Theorem 16, i.e. to obtain an exact formula H for Ψ^5. (A similar formula was already mentioned in [18]). We will need this Theorem 16 in Appendix B.

Definition of Ψ^k
We recall the definition of the balanced Feistel Schemes, i.e. the classical Feistel schemes. Let $\mathcal{P}_{2n}$ be the set of all permutations from $\{0,1\}^{2n}$ to $\{0,1\}^{2n}$. Let $\mathcal{F}_n$ be the set of all functions from $\{0,1\}^n$ to $\{0,1\}^n$. Let L, R, S and T be four n-bit strings in $\{0,1\}^n$. Let $\Psi(f_1)$ denotes the permutation of $\mathcal{P}_{2n}$ 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 functions of $\mathcal{F}_n$, let $\Psi^k(f_1, \ldots, f_k)$ denotes the permutation of $\mathcal{P}_{2n}$ 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 $\mathcal{F}_n$, 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'.

Definition 3. 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 $\{0,1\}^{2n}$, 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:

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

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

Let $[L_i, R_i], [S_i, T_i], 1 \leq i \leq q$, be a given sequence of $2q$ values of $\{0,1\}^{2n}$. Let r be the number of independent equalities $R_i = R_j$, $i \neq j$, and let s be the number of independent equalities $S_i = S_j$, $i \neq j$.

Theorem 12. *The exact formula for H_1 (i.e. for Ψ^1) is:*

$$H_1 = 0 \text{ if } (C) \text{ is not satisfied}$$
$$H_1 = \frac{|\mathcal{F}_n|}{2^{nq}} \cdot 2^{nr} \text{ if } (C) \text{ is satisfied}$$

where (C) is this set of conditions:

1. $\forall i,\ 1 \leq i \leq q,\ R_i = S_i$
2. $\forall i, j\ 1 \leq i \leq q,\ 1 \leq j \leq q,\ R_i = R_j \Rightarrow T_i \oplus L_i = T_j \oplus L_j$

Proof. For one round, we have $\Psi^1([L_i, R_i]) = [S_i, Y_i] \Leftrightarrow S_i = R_i$ and $T_i = L_i \oplus f_1(R_i)$. Therefore, if (C) is not satisfied, $H_1 = 0$. Now if (C) is satisfied, then f_1 is fixed on exactly $q - r$ points by $f_1(R_i) = T_i \oplus L_i$, and we obtain theorem 12 as claimed. □

Theorem 13. *The exact formula for H_2 (i.e. for Ψ^2) is:*

$$H_2 = 0 \text{ if } (C) \text{ is not satisfied}$$
$$H_2 = \frac{|\mathcal{F}_n|^2}{2^{2nq}} \cdot 2^{n(r+s)} \text{ if } (C) \text{ is satisfied}$$

where (C) is this set of conditions:

1. $\forall i,j\, 1 \leq i \leq q,\, 1 \leq j \leq q,\, R_i = R_j \Rightarrow L_i \oplus L_j = S_i \oplus S_j$
2. $\forall i,j\, 1 \leq i \leq q,\, 1 \leq j \leq q,\, S_i = S_j \Rightarrow R_i \oplus R_j = T_i \oplus T_j$

Proof. For two rounds we have $\psi^2([L_i, R_i]) = [S_i, T_i] \Leftrightarrow S_i = L_i \oplus f_1(R_i)$ and $T_i = R_i \oplus f_2(S_i)$. Therefore if (C) is not satisfied, $H_2 = 0$. Now if (C) is satisfied then (f_1, f_2) is fixed on exactly $2q - r - s$ points, and we obtain theorem 13 as claimed. □

Definition 4. *(Framework for Ψ^3)*
For 3 rounds, Ψ^3, we define a "framework" as a set of equations $X_i = X_j$. We will say that two frameworks are equal if they imply exactly the same set of equations in X.

Theorem 14. *The exact formula for H_3 (i.e. for Ψ^3) is:*

$$H_3 = \frac{|\mathcal{F}_n|^3 \cdot 2^{n(r+s)}}{2^{3nq}} \sum_{\substack{\text{all frameworks } \mathcal{F} \\ \text{that satisfy } (F1)}} 2^{nx}[\text{ Number of } X_i \text{ satisfying } (C1)]$$

where:

- *x is the number of independent equalities $X_i = X_j$ for a framework $\mathcal{F}$.*
- *$(F1) : X_i = X_j$ is in $\mathcal{F} \Rightarrow S_i \oplus S_j = R_i \oplus R_j$*

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

Proof. We write $\Psi^3 = \Psi \circ \Psi^2$ with $\Psi^2([L_i, R_i]) = [X_i, S_i]$ and $\Psi([X_i, S_i]) = [S_i, T_i]$. For Ψ^2, we obtain from theorem 13, $2^{n(r+x)} \frac{|\mathcal{F}_n|^2}{2^{2nq}}$ solutions when $(C1)$ is satisfied. For Ψ, we obtain from theorem 12, $2^{ns} \frac{|\mathcal{F}_n|}{2^{nq}}$ solutions when $(C1)$ is satisfied. Thus, we obtain theorem 14 as claimed. □

Definition 5. *(Framework for Ψ^4)*
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. For a framework $\mathcal{F}$, we denote by x the number of independent equalitites $X_i = X_j$, and by y the number of independent equalitites $Y_i = Y_j$.

Theorem 15. *The exact formula for H_4 (i.e. for Ψ^4) is:*

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

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

where

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

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

Proof. We write $\psi^4 = \Psi \circ \Psi^3$ with $\Psi^3([L_i, R_i]) = [Y_i, S_i]$ and $\Psi([Y_i, S_i]) = [S_i, T_i]$, and we sum over all possible Y. Then from theorems 12 and 14, we obtain theorem 15. □

Definition 6. *(Framework for Ψ^5)*
For 5 rounds, Ψ^5, a "framework" is a set of equations $X_i = X_j$ or $Y_i = Y_j$, or $Z_i = Z_j$. We will say that two frameworks are equal if they imply exactly the same set of equalities in X, Y and Z. For a framework $\mathcal{F}$, we denote by x the number of independent equalitites $X_i = X_j$, by y the number of independent equalitites $Y_i = Y_j$, and by z the number of independent equalitites $Z_i = Z_j$.

Theorem 16. *The exact formula for H_5 (i.e. for Ψ^5) is:*

$$H_5 = \frac{|\mathcal{F}_n|^5 \cdot 2^{n(r+s)}}{2^{5nq}} \sum_{\text{all frameworks } \mathcal{F}} 2^{n(x+y+z)} [\text{ Number of } X_i, Z_i \text{ satisfying } (C1)]$$

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

where

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

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

Proof. We write $\Psi^5 = \Psi \circ \Psi^4$ with $\Psi^4([L_i, R_i]) = [Z_i, S_i]$ and $\Psi([Z_i, S_i]) = [S_i, T_i]$, and we sum over all possible Z. Then from theorems 12 and 15, we obtain theorem 16. □

B "Holes" on Ψ^5 when $\sqrt{2^n} \ll q \ll 2^n$

We will present here a "structural" difference between Ψ^5 and Ψ^6: in Ψ^5, we have "holes" when $\sqrt{2^n} \ll q \ll 2^n$ (but not in Ψ^6: cf Theorem 8).

5 rounds.
For Ψ^5, with $q \simeq \sqrt{2^n}$, we can choose all the R_i with the same value, all the S_i with the same value and the property: $\forall i, j,\ 1 \leq i \leq q,\ 1 \leq j \leq q,\ T_i \oplus T_j \neq L_i \oplus L_j$. For example, the first $\frac{n}{2}$ bits of the L_i values are always 0 and the last $\frac{n}{2}$ bits of the T_i values are always 0. Since all the R_i values are equal, then all the L_i values are pairwise distinct (because we want pairwise distinct $[L_i, R_i]$) and all the X_i values are pairwise distinct (because $R_i = R_j \Rightarrow X_i \oplus X_j = L_i \oplus L_j$. Similarly, since all the S_i values are equal, then all the T_i values are distinct (because we want pairwise distinct $[S_i, T_i]$) and all the Z_i values are pairwise distinct (because $S_i = S_j \Rightarrow Z_i \oplus Z_j = T_i \oplus T_j$). Moreover all the Y_i values are also pairwise distinct, because $Y_i = Y_j \Rightarrow X_i \oplus X_j = Z_i \oplus Z_j \Rightarrow L_i \oplus L_j = T_i \oplus T_j$, but we always have: $L_i \oplus L_j \neq T_i \oplus T_j$.
We know (cf Appendix A, Theorem 16) that the exact formula for H is:

$$H_5 = \frac{|\mathcal{F}_n|^5 \cdot 2^{n(r+s)}}{2^{5nq}} \sum_{\text{all frameworks } \mathcal{F}} 2^{n(x+y+z)} [\text{ Number of } X_i, Z_i \text{ satisfying } (C1)]$$

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

Here we have only one framework (all the X_i are pairwise distinct, Y_i pairwise distinct, Z_i pairwise distinct) with $r = q-1, s = q-1, x = y = z = 0$, [Number of X_i satisfying $(C1)$] $= 2^n$, [Number of Z_i satisfying $(C1)$] $= 2^n$, and [Number of Y_i satisfying $(C2)$] $= 2^n(2^n-1)\ldots(2^n-q+1)$. we obtain:

$$H_5 = \frac{|\mathcal{F}_n|^5}{2^{2nq}} \cdot \left(1 - \frac{1}{2^n}\right)\left(1 - \frac{2}{2^n}\right)\ldots\left(1 - \frac{q-1}{2^n}\right) \ll \frac{|F_n|^5}{2^{2nq}}$$

when $q \ll \sqrt{2^n}$. However $\tilde{H}_5 = \frac{|\mathcal{F}_n|^5}{(2^n)(2^n-1)\ldots(2^n-q+1)} \simeq \frac{|\mathcal{F}_n|^5}{2^{2nq}}$. Therefore here we have $H_5 \ll \tilde{H}_5$, i.e. a "hole" of length $\sqrt{2^n}$.

This result is not in contradiction with the act that Ψ^5 is CCA secure when $q \ll 2^n$ because it is not possible in a CCA attack with q queries to obtain $R_1 = R_2 = \ldots = R_m$ and $S_1 = S_2 = \ldots = S_m$ with $m \simeq \sqrt{2^n}$.