Security of Feistel Schemes with New and Various Tools

Rodolphe LAMPE and Jacques PATARIN

Abstract: We combine the H Coefficients technique and the Coupling technique to improve security bounds of balanced Feistel schemes. For q queries and round functions of $n-$bits to $n-$bits, we find that the CCA Security of $4+2r$ rounds Feistel schemes is upperbounded by $\frac{2q}{r+3}\left(\frac{4q}{2^n}\right)^{\frac{r+1}{2}} + \frac{q(q-1)}{2 \cdot 2^{2n}}$. This divides by roughly 1.5 the number of needed rounds for a given CCA Security, compared to the previous results of Hoang and Rogaway [?] who found an advantage of $\frac{2q}{r+1}\left(\frac{4q}{2^n}\right)^r$ for $6r-1$ rounds Feistel schemes . Independently of this result, using a new theorem on H Coefficients, we compose 6 rounds Feistel schemes to upperbound the CCA security of $6r$ rounds Feistel schemes: $\left(\frac{8q}{2^n}\right)^r + \frac{q(q-1)}{2 \cdot 2^{2n}}$ when $q \leq \frac{2^n}{67n}$.

Keywords: Feistel schemes, Coupling, H coefficients, Security proof, Luby-Rackoff construction.

1 Introduction and Previous Results

1.1 Introduction

Since the seminal article of Luby and Rackoff [?] in 1989, security proofs of Feistel schemes have been extensively studied ([?], [?], [?], [?], [?], [?], [?]). 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. 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 [?], Hoang/Rogaway [?]). The second one, the H Coefficients technique, gives better results when the number of rounds is relatively small (Patarin [?] [?] [?]). Nevertheless, this technique leads sometimes to complex computations. In this article, we use the Coupling ideas to improve the H Coefficients technique in two different ways. First, we use intertwined conditions inspired by the Coupling technique to count H Coefficients and we find that the CCA Security of $4+2r$ rounds Feistel schemes verifies:

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

Then, we introduce a new theorem: the "H Coefficients Composition Theorem". We use this theorem to study the security of $6r$ rounds Feistel schemes using previous results of Patarin [?] for 6 rounds Feistel schemes. For $q \leq \frac{2^n}{67n}$, the CCA Security of $6r$ rounds Feistel schemes verifies

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

These methods can also be applied to many other schemes such as unbalanced Feistel schemes ([?]), alternating Feistel schemes ([?], [?]), $type-1, type-2$ and $type-3$ Feistel schemes ([?]), Benes schemes ([?]), MISTY's schemes ([?]), Feistel's with bijective round functions or format-preserving encryption ([?],[?]) which are beyond the scope of this work.

In a first section, we will present the H Coefficients technique (including proofs). We will use this technique in Section 2 and 3. In Section 2, we introduce our new technique, using intertwined conditions. This technique is inspired by the Coupling technique and the work of Hoang-Rogaway [?]. This way, we prove that we get the same CCA security than the previous best known bound of [?] using roughly 1.5 times fewer rounds. In Section 3, we introduce a new theorem on H Coefficients and apply it to prove CCA security of $6r$ rounds Feistel schemes when the number of queries is not too big.

1.2 Notations

Let n be an integer, F_n be the set of all functions from $\{0;1\}^n$ to $\{0;1\}^n$ and B_n the set of all permutations from $\{0;1\}^n$. Let f_1 be a function of F_n. Let L, R be two $n-$bit strings in $\{0;1\}^n$. Let $\Psi(f_1)$ 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, let $\Psi^r(f_1, ..., f_r)$ 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, in short, Ψ^r. When the functions $f_1, ..., f_r$ are randomly chosen in F_n^r, Ψ^r is called a "generic" Feistel scheme with r rounds, or a Luby-Rackoff construction.

Let q be the number of queries. For a given Ψ^r, let $X_1, ..., X_q$ denote the q inputs and $Y_1, ..., Y_q$ the q outputs. For all $i \in [1, q]$ and $k \in [0; r]$, let X_i^k denote the first n bits of the outputs of X_i after k rounds and X_i^{r+1} the last n bits of Y_i. This means, for example, that $X_i = [X_i^0, X_i^1]$ and $Y_i = [X_i^r, X_i^{r+1}]$.

To simplify computations, we will note $J_q = 2^{2n} \times (2^{2n}-1) \times \cdots \times (2^{2n}-q+1)$.

1.3 The coefficients H technique

In this article, we will prove security bounds using the general framework given by the "H Coefficients technique" of Patarin [?][?].

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 all $Y_1, ..., Y_q \in \{0;1\}^{2n}$ pairwise distinct and for all $X_1, ..., X_q \in \{0;1\}^{2n}$ pairwise distinct, 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 permutations f_k of F, with $k \in_R K$, and random permutations verifies:

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

Proof: See Appendix **??**

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

Theorem 2 (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 all $Y_1, ..., Y_q \in \{0;1\}^{2n}$ pairwise distinct and for all $X_1, ..., X_q \in \{0;1\}^{2n}$ pairwise distinct, 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:

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

Proof: See Appendix **??**

In Section **??**, we will prove a new theorem on these H coefficients.

2 Proving Security of $4+2r$ rounds Feistel Schemes with intertwine conditions

2.1 The CCA Security of Ψ^{4+2r}

We will use the theorem of Patarin we just introduced in the previous Section **??** to find the CCA security of Ψ^{4+2r} for any positive integer r. We fix q inputs X_ℓ and q outputs Y_ℓ and our goal is to count the number of $(f_1, f_2, ..., f_{4+2r}) \in F_n^{4+2r}$ such that $\Psi(f_1, f_2, ..., f_{4+2r})(X_\ell) = Y_\ell$ for all $\ell \leq q$.

It means that we have to find the $(f_1, ..., f_{4+2r})$ such that, for all $\ell \leq q$, it exists $X_\ell^2, ..., X_\ell^{2+2r}$ verifying the $4+2r$ equations:

$$\begin{cases} X_\ell^2 = X_\ell^0 \oplus f_1(X_\ell^1) \\ X_\ell^3 = X_\ell^1 \oplus f_2(X_\ell^2) \\ \quad\vdots \\ X_\ell^{i+1} = X_\ell^{i-1} \oplus f_i(X_\ell^i) \\ \quad\vdots \\ X_\ell^{5+2r} = X_\ell^{3+2r} \oplus f_{4+2r}(X_\ell^{4+2r}) \end{cases}$$

For every $\ell \leq q$ and any $f_1, ..., f_{4+2r}$, we define $\overrightarrow{X}_\ell^0, ..., \overrightarrow{X}_\ell^{5+2r}$ by induction : $\overrightarrow{X}_\ell^0 = X_\ell^0, \overrightarrow{X}_\ell^1 = X_\ell^1$ and $\forall i \in \{2, ..., 5+2r\}$:

$$\overrightarrow{X}_\ell^i = \overrightarrow{X}_\ell^{i-2} \oplus f_{i-1}(\overrightarrow{X}_\ell^{i-1}).$$

More intuitively, we compute $\overrightarrow{X}_\ell^i$ "from the top": $\overrightarrow{X}_\ell^0$ and $\overrightarrow{X}_\ell^1$ are already defined, then f_1 will define $\overrightarrow{X}_\ell^2$, f_2 will define $\overrightarrow{X}_\ell^3$ and so on to $\overrightarrow{X}_\ell^i$. We defined $\overrightarrow{X}_\ell^i$ such that the first $i-1$ equations are trivially verified.

In a symmetric way, we define $\overleftarrow{X}_\ell^{5+2r}, ..., \overleftarrow{X}_\ell^0$ by induction : $\overleftarrow{X}_\ell^{5+2r} = X_\ell^{5+2r}, \overleftarrow{X}_\ell^{4+2r} = X_\ell^{4+2r}$ and $\forall i \in \{0, ..., 3+2r\}$:

$$\overleftarrow{X}_\ell^i = \overleftarrow{X}_\ell^{i+2} \oplus f_{i+1}(\overleftarrow{X}_\ell^{i+1}).$$

For any $k \in \{1, ..., r+1\}$, considering internal variables $\overrightarrow{X}_\ell^0, ..., \overrightarrow{X}_\ell^{2k}, \overleftarrow{X}_\ell^{2k+1}, ..., \overleftarrow{X}_\ell^{5+2r}$, we see that, by construction, the first $2k-1$ equations are verified and the last $(4+2r)-(2k+1)$ equations are verified.

We have:

$$\overrightarrow{X}_\ell^1 \xrightarrow[f_1]{} \cdots \xrightarrow[f_{2k-1}]{} \overrightarrow{X}_\ell^{2k} \overset{?}{\longleftrightarrow} \overleftarrow{X}_\ell^{2k+1} \xleftarrow[f_{2k+2}]{} \cdots \xleftarrow[f_{4+2r}]{} \overleftarrow{X}_\ell^{4+2r}$$

For any $f_1, ..., f_{4+2r}$, any ℓ and any $k \in \{1, ..., r+1\}$, we defined internal variables such that the first $2k-1$ equations are verified and the last $(4+2r)-(2k+1)$ equations are verified. We only need to verify two more equations:

$$\begin{cases} f_{2k}(\overrightarrow{X}_\ell^{2k}) = \overrightarrow{X}_\ell^{2k-1} \oplus \overleftarrow{X}_\ell^{2k+1} \\ f_{2k+1}(\overleftarrow{X}_\ell^{2k+1}) = \overrightarrow{X}_\ell^{2k} \oplus \overleftarrow{X}_\ell^{2k+2} \end{cases}$$

So far, we made no restrictions on the round functions f_i. Now, we need f_{2k} and f_{2k+1} to verify this two equations. We could just take the $(f_1, ..., f_{4+2r})$ such that the two equations are verified for the ℓ-th query. The problem is that, if $\overrightarrow{X}_\ell^{2k}$ collides with a previous query (ie $\overrightarrow{X}_\ell^{2k} = \overrightarrow{X}_i^{2k}$ for some $i < \ell$), the functions f_{2k} would already be defined in $\overrightarrow{X}_i^{2k}$ and we are not sure we can always choose functions f_{2k} verifying the first equation. We have the same problem for f_{2k+1}.

This is why we need to choose the functions $f_1, ..., f_{2k-1}, f_{2k+2}, ..., f_{4+2r}$ such that $\overrightarrow{X}_\ell^{2k}$ and $\overleftarrow{X}_\ell^{2k+1}$ don't collide. We will make this selection for every query. After that, we will connect $\overrightarrow{X}_\ell^{2k}$ to $\overleftarrow{X}_\ell^{2k+1}$ for each query. Note that k depends of ℓ, we will not always connect in the same place.

We explained our strategy, we now turn to details.

For any ℓ and any $k \in \{1, ..., r+1\}$, we note :
- Col_ℓ^{2k} the event "$\overrightarrow{X}_\ell^{2k} = \overrightarrow{X}_i^{2k}$ for some $i < \ell$".
- Col_ℓ^{2k+1} the event "$\overleftarrow{X}_\ell^{2k+1} = \overleftarrow{X}_i^{2k+1}$ for some $i < \ell$".

We want to select $f_1, ..., f_{4+2r}$ such that, for each query, it is possible to find k such that Col_ℓ^{2k} and Col_ℓ^{2k+1} are both wrong. We note:

$$C_\ell = \bigcup_{k=1}^{r+1} \left(\neg Col_\ell^{2k} \cap \neg Col_\ell^{2k+1} \right),$$

this is the event that the query ℓ is "connectable" : we can find k such that $\overrightarrow{X}_\ell^{2k}$ and $\overleftarrow{X}_\ell^{2k+1}$ don't collide so we can select f_{2k} and f_{2k+1} to verify the two "connection" equations.

We want all the queries connectable so we want to compute the probability of

$$\bigcap_{\ell=1}^{q} C_\ell.$$

We have:

$$P\left(\bigcap_{\ell=1}^{q} C_\ell\right) = 1 - P\left(\bigcup_{\ell=1}^{q} \neg C_\ell\right) \geq 1 - \sum_{\ell=1}^{q} P\left(\neg C_\ell\right) \tag{1}$$

Lemma 1. *For every $\ell \in \{1, ..., q\}$, we have:*

$$P\left(\neg C_\ell\right) \leq \left(\frac{4(\ell-1)}{2^n}\right)^{\frac{r+1}{2}}.$$

Proof: Let $\ell \in \{1, ..., q\}$, we have:

$$\neg C_\ell = \bigcap_{k=1}^{r+1} \left(Col_\ell^{2k} \cup Col_\ell^{2k+1}\right) = \bigcup_{\alpha \in \{0;1\}^{r+1}} \bigcap_{k=1}^{r+1} Col_\ell^{2k+\alpha(k)} \tag{2}$$

We noted $\alpha(k)$ the $k-$th bit of α. This equality implies that:

$$P(\neg C_\ell) \leq \sum_{\alpha \in \{0;1\}^{r+1}} P\left(\bigcap_{k=1}^{r+1} Col_\ell^{2k+\alpha(k)}\right) \leq \sum_{\alpha \in \{0;1\}^{r+1}} \min\left(P\left(\bigcap_{\alpha(k)=0} Col_\ell^{2k}\right), P\left(\bigcap_{\alpha(k)=1} Col_\ell^{2k+1}\right)\right) \tag{3}$$

It is easy to see that, for any k, the event Col_ℓ^{2k} depends of the functions $f_1, ..., f_{2k-1}$. For any fixed $f_1, ..., f_{2k-2}$, the values of X_ℓ^{2k-2} and X_ℓ^{2k-1} are fixed so $P(Col_\ell^{2k})$ is the probability that it exists $i < \ell$ such that f_{2k-1} verifies :

$$f_{2k-1}(X_i^{2k-1}) \oplus f_{2k-1}(X_\ell^{2k-1}) = X_i^{2k-2} \oplus X_\ell^{2k-2}.$$

We now use the reasoning of lemma 3 of [**?**]. We see that if $X_i^{2k-1} = X_\ell^{2k-1}$ then it is impossible to have $f_{2k-1}(X_i^{2k-1}) \oplus f_{2k-1}(X_\ell^{2k-1}) = X_i^{2k-2} \oplus X_\ell^{2k-2}$ because it would imply that $X_i^{2k-2} = X_\ell^{2k-2}$ which is impossible since the queries have been chosen pairwise distinct. If $X_i^{2k-1} \neq X_\ell^{2k-1}$ then the equation $f_{2k-1}(X_i^{2k-1}) \oplus f_{2k-1}(X_\ell^{2k-1}) = X_i^{2k-2} \oplus X_\ell^{2k-2}$ occurs with probability 2^{-n} because f_{2k-1} is uniformly random. We see that, for fixed $f_1, ..., f_{2k-2}$, the event Col_ℓ^{2k} occurs with probability less or equal to $\frac{\ell-1}{2^n}$. This implies that

$$P\left(\bigcap_{\alpha(k)=0} Col_\ell^{2k}\right) \leq \left(\frac{\ell-1}{2^n}\right)^{|\{k, \alpha(k)=0\}|}.$$

We use the same reasoning for Col_ℓ^{2k+1} and, using **??**, we have :

$$
\begin{aligned}
P(\neg C_\ell) &\leq \sum_{\alpha\in\{0;1\}^{r+1}} \min\left(\left(\frac{\ell-1}{2^n}\right)^{|\{k,\alpha(k)=0\}|}, \left(\frac{\ell-1}{2^n}\right)^{|\{k,\alpha(k)=1\}|}\right) \\
&\leq \sum_{\alpha\in\{0;1\}^{r+1}} \left(\frac{\ell-1}{2^n}\right)^{\frac{r+1}{2}} \\
&\leq \left(\frac{4(\ell-1)}{2^n}\right)^{\frac{r+1}{2}}
\end{aligned} \tag{4}
$$

Using the inequation **??** and the previous Lemma, we have:

$$
\begin{aligned}
P\left(\bigcap_{\ell=1}^{q} C_\ell\right) &\geq 1 - \sum_{\ell=1}^{q}\left(\frac{4(\ell-1)}{2^n}\right)^{\frac{r+1}{2}} \\
&\geq 1 - \left(\frac{4}{2^n}\right)^{\frac{r+1}{2}} \times \sum_{\ell=0}^{q-1} \ell^{\frac{r+1}{2}} \\
&\geq 1 - \left(\frac{4}{2^n}\right)^{\frac{r+1}{2}} \times \int_0^q \ell^{\frac{r+1}{2}} d\ell \\
&\geq 1 - \left(\frac{4}{2^n}\right)^{\frac{r+1}{2}} \times \frac{q^{\frac{r+1}{2}+1}}{\frac{r+1}{2}+1} \\
&\geq 1 - \frac{2q}{r+3}\left(\frac{4q}{2^n}\right)^{\frac{r+1}{2}}.
\end{aligned} \tag{5}
$$

Now that we compute the probability of making all queries connectable, we need to actually connect them. For every ℓ, let note D_ℓ the event "$\Psi(f_1, ..., f_{4+2r})$ sends X_ℓ to Y_ℓ". We have:

$$
\begin{aligned}
P\left(\bigcap_{\ell=1}^{q} D_\ell \bigcap \bigcap_{\ell=1}^{q} C_\ell\right) &= P\left(\bigcap_{\ell=1}^{q} C_\ell\right) \times P\left(\bigcap_{\ell=1}^{q} D_\ell | \bigcap_{\ell=1}^{q} C_\ell\right) \\
&= P\left(\bigcap_{\ell=1}^{q} C_\ell\right) \times \prod_{\ell=1}^{q} P\left(D_\ell | \bigcap_{\ell=1}^{q} C_\ell \bigcap \bigcap_{i<\ell} D_i\right)
\end{aligned} \tag{6}
$$

Lemma 2. *For every $\ell \in \{1, ..., q\}$, we have:*

$$
P\left(D_\ell | \bigcap_{\ell=1}^{q} C_\ell \bigcap \bigcap_{i<\ell} D_i\right) \geq \frac{1}{2^{2n}}.
$$

Proof: If C_ℓ is true, it exists $k \in \{1, ..., r+1\}$ such that $\overrightarrow{X}_\ell^{2k}$ and $\overleftarrow{X}_\ell^{2k+1}$ don't collide. Since the round functions are independent and uniformly random, the following equations happen with probability $\frac{1}{2^{2n}}$:

$$\begin{cases} f_{2k}(\overrightarrow{X}_\ell^{2k}) = \overrightarrow{X}_\ell^{2k-1} \oplus \overleftarrow{X}_\ell^{2k+1} \\ f_{2k+1}(\overleftarrow{X}_\ell^{2k+1}) = \overrightarrow{X}_\ell^{2k} \oplus \overleftarrow{X}_\ell^{2k+2} \end{cases}$$

With probability $\frac{1}{2^{2n}}$, we have connected X_ℓ to Y_ℓ. Indeed, we have choosen the first $2k-1$ internal variables to trivially verify the first $2k-1$ equations. We then choose the last $(4+2r)-(2k+1)$ internal variables to trivially verify the last $(4+2r)-(2k+1)$ equations. As we just proved, the last two equations are true with probability $\frac{1}{2^{2n}}$ so all needed equations are verified and D_ℓ is true. $\square$

From the previous lemmma, the equation **??** and the inequation **??**, we have:

$$P\left(\bigcap_{\ell=1}^{q} D_\ell \bigcap \bigcap_{\ell=1}^{q} C_\ell\right) \geq \left(1 - \frac{2q}{r+3}\left(\frac{4q}{2^n}\right)^{\frac{r+1}{2}}\right) \times \frac{1}{2^{2nq}}. \tag{7}$$

Remember that we defined H the number of $(f_1, ..., f_{4+2r})$ such that $\Psi(f_1, ..., f_{4+2r})(X_\ell) = Y_\ell$ for every ℓ. It implies that

$$P\left(\bigcap_{\ell=1}^{q} D_\ell \bigcap \bigcap_{\ell=1}^{q} C_\ell\right) \leq \frac{H}{|F_n|^{4+2r}}.$$

This two inequations imply that:

$$H \geq \left(1 - \frac{2q}{r+3}\left(\frac{4q}{2^n}\right)^{\frac{r+1}{2}}\right) \times \frac{|F_n|^{4+2r}}{2^{2nq}}. \tag{8}$$

Using the Theorem **??** of Patarin, we have:

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

2.2 Results

We recall the bound of Hoang-Rogaway and our bound :

Hoang-Rogaway: $\mathbf{Adv}^{\mathrm{cca}}_{\Psi^{6r-1}}(q) \leq \frac{2q}{r+1}\left(\frac{4q}{2^n}\right)^{r}.$

Our bound: $\mathbf{Adv}^{\mathrm{cca}}_{\Psi^{4+2r}}(q) \leq \frac{2q}{r+3}\left(\frac{4q}{2^n}\right)^{\frac{r+1}{2}} + \frac{q(q-1)}{2 \cdot 2^{2n}}.$

In the following table, we give, for various values of n and number of rounds, the $\log_2$ of the number of queries giving an advantage of $\frac{1}{2}$ using our bound and the previous best known bound from [?]. In the last column, we give the ratio of the number of queries giving an advantage of $\frac{1}{2}$ with our bound divided by the number of queries giving the same advantage with the previous bound.

Table 1. $\log_2$ of the number of queries to get an advantage of 1/2

n	rounds	Our bound	HR's bound	Ratio
32	18	24.66	23.35	2.5
64	18	50.26	47.67	6.0
128	18	101.46	96.31	35.5
32	48	28.92	28.11	1.8
64	48	58.36	56.62	3.4
128	48	117.24	113.63	12.2

3 Proving Security of Feistel Schemes with a new Theorem on H Coefficients

3.1 The H Coefficients Composition Theorem

We fix q inputs $X_1, ..., X_q$ and q outputs $Y_1, ..., Y_q$. Let F and G be two subsets of B_{2n} and we note $F \circ G$ the set $\{f \circ g, f \in F, g \in G\}$. For any subset F of B_{2n}, we note $H^F(X, Y)$ the number of functions in F sending X_i to Y_i for all $i \leq q$.

Theorem 3 (2012). *If it exists α_F and α_G in $[0; 1]$ such that, for all $X_1, ..., X_q$ pairwise distinct and for all $Y_1, ..., Y_q$ pairwise distinct:*

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

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

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

For any $X_1, ..., X_q$ pairwise distinct and $Y_1, ..., Y_q$ pairwise distinct. We have

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

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

$$\sum_T 1 = J_q \leq 2^{2nq} \tag{10}$$

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

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

We compute the right part of equality (**??**) 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:

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

The third term is computed 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, we have:

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

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

$$(1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) \Big(\sum_T \frac{1}{2^{2nq}} \Big) \geq (1-\alpha_G) + (1-\alpha_F) - (1-\alpha_G)(1-\alpha_F) = 1 - \alpha_G \alpha_F,$$

which ends the proof. $\square$

We will use this theorem in the next section to study $6r$ rounds Feistel scheme.

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.

We recall that $J_q = 2^{2n} \times (2^{2n} - 1) \times \cdots \times (2^{2n} - q + 1)$.

Theorem 4 (2012). *If it exists α_F and α_G in $[0;1]$ such that, for all $X_1, ..., X_q$ pairwise distinct and for all $Y_1, ..., Y_q$ pairwise distinct:*

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

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

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

Proof: See Appendix **??**

From this Theorem **??** and Theorem **??**, we see 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.

3.2 Proving Security with mirror theory and the Global H theorem

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

Theorem 5 (2010). *For Ψ^6, for all $X_1, ..., X_q \in \{0;1\}^{2n}$ pairwise distinct and for all $Y_1, ..., Y_q \in \{0;1\}^{2n}$ pairwise distinct:*

$$H(X,Y) \geq \left(1 - \frac{8q}{2^n}\right) \times \frac{|F_n|^6}{2^{2nq}} \text{ if } q \leq \frac{2^n}{67n}.$$

Therefore, combining this theorem and using our new "Global H Theorem" of Section **??**, we obtain:

Theorem 6 (2012). *For any $r \geq 1$ and any $q \leq \frac{2^n}{67n}$:*

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

This Theorem **??** is, so far, the best security bound known for Feistel schemes when $q \leq \frac{2^n}{67n}$. However, Patarin's proof of Theorem **??** is difficult (cf [?], [?]). Nevertheless, some variants of Theorem **??** are much easier to prove: for example, instead of $\alpha = \frac{8q}{2^n}$ we can use $\alpha = \frac{q^3}{2^{2n}}, \alpha = \frac{q^4}{2^{3n}}$ or $\alpha = \frac{q^5}{2^{4n}}$ (see [?], [?] for more details). Then, from each of these variants, our Global H Theorem will immediately give a geometrical improvement on the advantage when we multiply the number of rounds.

4 Results

This new bound is very good when we can use it (we need $q \leq \frac{2^n}{67n}$). In the following table, we give the $\log_2$ of the advantage for $q = \frac{2^n}{67n}$ with our bound and with the previous best known bound of [?].

Table 2. $\log_2$ of the advantage for $q = \frac{2^n}{67n}$ with our bound and with the previous best known bound of [?].

n	rounds	$\log_2(q)$	Our advantage	HR's advantage
32	18	20.9	-24.2	-8.8
64	18	51.9	-27.2	0
128	18	114.9	-30.2	0
32	48	20.9	-64.5	-55.3
64	48	51.9	-72.5	-32.5
128	48	114.9	-80.5	0

5 Conclusion

In this paper, we combine ideas from two different proof techniques: the Coupling technique and the H Coefficients technique. We introduce a new Theorem: the "H Coefficients Composition Theorem". From this new theorem, we are able to obtain security proofs that combine the efficiency of the H Coefficients for small rounds Feistel schemes and the geometric gain of the Coupling technique. We apply these results only on the classical balanced generic Feistel schemes but the technique can also be applied to many different schemes like unbalanced Feistel schemes or MISTY schemes for example. Independently of that, we have also combined ideas of the Coupling technique and the H Coefficients technique to study intertwined conditions. This new approach lead to significant improvements, we divide by roughly 1.5 the number of needed rounds to obtain a given CCA security.

Appendices

A Proof of the Coefficients H Theorem first variant ([?], 1991, p.38)

Consider an attacker A who can query q times an oracle O. The oracle O acts all the time like a Feistel scheme Ψ^r or like a random permutation. The attacker can make direct queries or inverse queries. After q queries, the attacker outputs 1 or 0.

We note

$$P_1 = \text{Probability that A outputs 1 if } O \text{ is } \Psi^r$$

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 note $\gamma_1, ..., \gamma_q$ the q queries and $\delta_1, ..., \delta_q$ the q answers. If the ith query is direct, we have $\delta_i = O(\gamma_i)$, if the ith query is inverse, we have $\delta_i = O^{-1}(\gamma_i)$.

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 note $X(\delta) = (X_1, ..., X_q)$ and
$Y(\delta) = (Y_1, ..., Y_q)$.

We note $\Sigma = \{\delta \text{ such that } A \text{ outputs } 1\}$. For a fixed δ, a random permutation send γ on δ with probability

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

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.

so

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

We note $C = \{(f_1, ..., f_r)$ such that A outsputs 1 if $O = \Psi(f_1, ..., f_r)\}$. We have

$$P_1 = \frac{|C|}{|F_n|^r}.$$

If we note, for all $\delta \in \Sigma$, C_δ the set of round functions $f_1, ..., f_r$ such that $\Psi(f_1, ..., f_r)$ sends $X(\delta)$ on $Y(\delta)$ then

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

Now, remember the hypothesis :

$$\frac{H(X(\delta), Y(\delta))}{|F_n|^r} \geq (1 - \alpha) \times \frac{1}{2^{2nq}}.$$

So

$$P_1 \geq \frac{|\Sigma|(1 - \alpha)}{2^{2nq}}.$$

So

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

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

Doing all the same reasoning for the 0 output, we have

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

which is equivalent to

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

This proves

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

□

B Proof of the Coefficients H Theorem second variant ([?], 1991, p.38)

Consider an attacker A who can query q times an oracle O. The oracle O acts all the time like a Feistel scheme Ψ^r or like a random permutation. The attacker can make direct queries or inverse queries. After q queries, the attacker outputs 1 or 0.

We note

$$P_1 = \text{Probability that A outputs 1 if } O \text{ is } \Psi^r$$

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 note $\gamma_1, ..., \gamma_q$ the q queries and $\delta_1, ..., \delta_q$ the q answers. If the ith query is direct, we have $\delta_i = O(\gamma_i)$, if the ith query is inverse, we have $\delta_i = O^{-1}(\gamma_i)$.

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 note $X(\delta) = (X_1, ..., X_q)$ and
$Y(\delta) = (Y_1, ..., Y_q)$.

We note $\Sigma = \{\delta \text{ such that } A \text{ outputs } 1\}$. 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 note $C = \{(f_1, ..., f_r) \text{ such that } A \text{ outsputs 1 if } O = \Psi(f_1, ..., f_r)\}$. We have

$$P_1 = \frac{|C|}{|F_n|^r}.$$

If we note, for all $\delta \in \Sigma$, C_δ the set of round functions $f_1, ..., f_r$ such that $\Psi(f_1, ..., f_r)$ sends $X(\delta)$ on $Y(\delta)$ then

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

Now, remember the hypothesis :

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

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, we have

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

which is equivalent to

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

This proves

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

□

C Proof of the variant of the H Coefficients Composition Theorem (2012)

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

$$H^{F \circ G}(X, Y) = \sum_T H^G(X, T) \times H^F(T, Y), \tag{13}$$

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

$$\sum_T 1 = J_q \tag{14}$$

$$\sum_T H^F(T, Y) = |F| \tag{15}$$

$$\sum_T H^G(X, T) = |G| \tag{16}$$

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

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

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

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

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, we have:

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

□

Theorem:
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:

$$H^F_{\ell+1}(X,Y) \geq (1-\alpha_F) \times \frac{H^F_\ell(X,Y)}{2^{2n}-\ell} \text{ and } H^G_{\ell+1}(X,Y) \geq (1-\alpha_G) \times \frac{H^G_\ell(X,Y)}{2^{2n}-\ell},$$

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

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

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

$$H^{F\circ G}_{\ell+1}(X,Y) = \sum_{T_1,...,T_l,T_{\ell+1}} H^G_{\ell+1}(X,T) \times H^F_{\ell+1}(T,Y)$$

and

$$H^{F\circ G}_{l}(X,Y) = \sum_{T_1,...,T_l} H^G_{l}(X,T) \times H^F_{l}(T,Y)$$

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, we have

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

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 $m_F = (1-\alpha_F) \times \frac{H_\ell^F(T,Y)}{2^{2n}-\ell}$ and $m_G = (1-\alpha_G) \times \frac{H_\ell^G(X,T)}{2^{2n}-\ell}$:

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

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

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

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

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

Summing the four terms, we have:

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

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

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

which ends the proof. $\square$