New Results on the Pseudorandomness of Some Blockcipher Constructions

Henri Gilbert and Marine Minier

France Télécom R&D
38-40, rue du Général Leclerc
92794 Issy les Moulineaux Cedex 9
France

Abstract. In this paper, we describe new results on the security, in the Luby-Rackoff paradigm, of two modified Feistel constructions, namely the L-scheme, a construction used at various levels of the MISTY blockcipher which allows to derive a $2n$-bit permutation from several n-bit permutations, and a slightly different construction named the R-scheme. We obtain pseudorandomness and super-pseudorandomness proofs for L-schemes and R-schemes with a sufficient number of rounds, which extend the pseudorandomness and non superpseudorandomness results on the 4-round L-scheme previously established by Sugita [Su96] and Sakurai et al. [Sa97]. In particular, we show that unlike the 3-round L-scheme, the 3-round R-scheme is pseudorandom, and that both the 5-round L scheme and the 5-round R scheme are super pseudorandom (whereas the 4 round versions of both schemes are not super pseudorandom). The security bounds obtained here are close to those established by Luby and Rackoff for the three round version of the original Feistel scheme.

1 Introduction

A key dependent cryptographic function such as a blockcipher can be viewed as a random function associated with a randomly selected key value. It is generally defined using a recursive construction process. Each step of the recursion consists of deriving a random function (or permutation) f from r previously defined random functions (or permutations) $f_1, .., f_r$, and can be represented by a relation of the form $f = \Phi(f_1, .., f_r)$. The most studied example so far is the $f = \Psi(f_1, .., f_r)$ r-round Feistel construction, which allows to derive a $2n$-bit to $2n$-bit random permutation from r n-bit to n-bit functions. But there exist other well known constructions such as for instance Massey and Lai's alternative to the Feistel scheme used in IDEA [La90] and the constructions allowing to deduce a $2n$-bit permutation from several n-bit permutations used in Matsui's MISTY blockcipher [Ma93].

The strongest security requirement one can put on a f random function or permutation representing a key dependent cryptographic function is (informally speaking) that f be undistinguishable with a non negligible success probability from a perfect random function f^* or permutation c^*, even if a probabilistic

M. Matsui (Ed.): FSE 2001, LNCS 2355, pp. 248–266, 2002.

testing algorithm A of unlimited power is used for that purpose and if the q number of adaptively chosen queries of A to the random instance of f or f^* to be tested is large.

It is generally not possible to prove undistiguishability properties for "real life" cryptologic random function f and large q numbers of queries, because this would require a far too long key length. However, it is often possible to prove or disprove that if a random function f encountered at a given level of a cryptologic function construction is related to random functions encountered at the lower recursion level by a relation of the form $f = \Phi(f_1, .., f_r)$, then if we replace the actual f_1 to f_r random functions of the cipher by independent perfect random functions or permutations f_1^* to f_r^* (or, in a more sophisticated version of the same approach, by f_1' to f_r' functions which are sufficiently undistinguishable from f_1^* to f_r^*), the resulting modified f random function is undistinguishable from a random function (or permutation). This provides a useful method for assessing the soundness of blockcipher constructions. For instance, in the case of a three-round Feistel construction, a well known theorem first proved by Luby and Rackoff [Lu88] provides upper bounds on the $|p - p^*|$ advantage of any A testing algorithm in distinguishing the $f = \Psi(f_1^*, f_2^*, f_3^*)$ $2n$-bit random permutation deduced from three independent ideal random functions f_1^*, f_2^* and f_3^* from a $2n$-bit perfect random permutation c^* with q adaptively chosen queries to the tested instance of f or f^*. This advantage is bounded over by $\frac{q^2}{2^n}$.

The research on pseudorandomness properties of cryptographic constructions initiated Luby and Rackoff's seminal paper [Lu88] has represented a very active research study for the last decade. Just to mention a few examples, Zheng, Matsumoto and Imai and later on Sugita and Sakurai et al. investigated generalised Feistel constructions [Zh89],[Su96],[Su97], Patarin explicited the link between the best advantage of a q-queries distinguisher and the q-ary transition probabilities associated with f and proved undistinguihability bounds for numerous r-round Feistel constructions [Pa91], Maurer showed how to generalise undistinguishability results related to perfect random functions to undistinguishability results related to nearly perfect random functions (e.g. locally random functions)[Ma92], Bellare, Kilian, Rogaway et al. [Be94] investigated the application of similar techniques to modes of operation such as CBC MACs, Aiello and al. proved undistiguihability results on some parallelizable alternatives to the Feistel construction [Ai96] , Vaudenay embedded techniques for deriving undistinguishability bounds into a broader framework he named the decorrelation theory, and applied bounds provided by decorrelation techniques to proving the resistance of actual ciphers, e.g. DFC, against differential and linear cryptanalysis.

In this paper, we describe new results on the security of some blockcipher constructions in the above described paradigm, i.e. we investigate some $f = \Phi(f_1, .., f_k)$ constructions and upper bound the probability of distinguishing f from a perfect random function when Φ is applied to perfect random functions f_i^* or to perfect random permutations c_i^*. We consider alternatives to the Feistel construction allowing to derive a $2n$-bit permutation from several n-bit permutations, namely the so-called L-scheme and R-scheme constructions. The L-type

construction is used for instance at various levels of the construction of Matsui and al. Misty blockcipher [Ma93], as well as in the Kasumi variant of Misty recently adopted as the standard blockcipher for encryption and integrity protection in third generation mobile systems [Ka]. We obtain pseudorandomness and superpseudorandomness proofs for L-scheme and R-scheme constructions with a sufficient number of rounds, which extend the results on the pseudo randomness of the 4-round L-scheme previously established by Sugita [Su96] and Sakurai et al. [Sa97]. In particular, we show that unlike the 3-round L scheme, the 3-round R scheme is pseudorandom, and that both the 5-round L scheme and the 5-round R scheme are super pseudorandom (whereas the 4 round versions of both schemes are not super pseudorandom).

This paper organised as follows: Section 2 introduces basic definitions and useful general results on random functions and techniques for proving that two random functions are undistiguishable. Section 3 describes the R and L schemes. Sections 4 and 5 present our results on the pseudo-randomness and the superpseudorandomness of the L-scheme and the R-scheme respectively, for various numbers of rounds, and Section 6 concludes the paper.

2 Preliminaries

2.1 Notation

Through this paper we are using the following notation: I_n denotes the $\{0,1\}^n$ set. $F_{n,m}$ denotes the ${I_n}^{I_m}$ set of functions from I_n into I_m: thus $|F_{n,m}| = 2^{m.2^n}$. F_n denotes the $F_{n,n}$ set: thus $|F_n| = 2^{n.2^n}$. P_n denotes the set of permutations on I_n: thus $|P_n| = 2^n!$.

2.2 Random Functions

A random function of $F_{n,m}$ is defined as a random variable f of $F_{n,m}$, and can be viewed as a probability distribution $(Pr[f = \varphi])_{\varphi \in F_{n,m}}$ over $F_{n,m}$, or equivalently as a $(f_\omega)_{\omega \in \Omega}$ family of $F_{n,m}$ elements. In particular:

- A n-bit to m-bit key dependent cryptographic function is determined by a randomly selected key value $K \in \mathcal{K}$, and can thus be represented by the random function $f = (f_K)_{K \in \mathcal{K}}$ of $F_{n,m}$.
- A cryptographic construction of the form $f = \Phi(f_1, f_2, .., f_r)$ can be viewed as a random function of $F_{n,m}$ determined by r random functions $f_i \in F_{n_i,m_i}, i = 1..r$

Definition 1. *We define a perfect random function f^* of $F_{n,m}$ as a uniformly drawn element of $F_{n,m}$. In other words, f^* is associated with the uniform probability distribution over $F_{n,m}$. We define a c^* perfect random permutation on I_n as a uniformly drawn element of P_n. In other words, c^* is associated with the uniform probability distribution over P_n.*

Definition 2. *(t-ary transition probabilities associated with f). Given a random function f of $F_{n,m}$, we define the $Pr[x \stackrel{f}{\mapsto} y]$ transition probability associated with a x t-uple of I_n inputs and a y t-uple of I_m outputs as*

$Pr[x \stackrel{f}{\mapsto} y] = Pr[f(x_1) = y_1 \wedge f(x_2) = y_2 \wedge ... \wedge f(x_t) = y_t]$
$= Pr_{\omega \in \Omega}[f_\omega(x_1) = y_1 \wedge f_\omega(x_2) = y_2 \wedge ... \wedge f_\omega(x_t) = y_t]$

In the sequel we will use the following simple properties:

Property 1 *If f^* is a perfect random function $F_{n,m}$ and if $x = (x_1, ..., x_t)$ is a t-uple of pairwise distinct I_n values and if y is any t-uple of I_m values, then $Pr[x \stackrel{f^*}{\mapsto} y] = \frac{1}{|I_m|^t} = 2^{-m.t}$*

Property 2 *Let c^* be a perfect random permutation on I_n. If $x = (x_1, ..., x_t)$ is a t-uple of pairwise distinct I_n values $y = (y_1, ..., y_t)$ is a t-uple of pairwise distinct I_n values then $Pr[x \stackrel{c^*}{\mapsto} y] = (I_n - t)!/|In|! = \frac{(2^n - t)!}{(2^n)!}$*

Property 3 *Let c^* be a perfect random permutation on I_n. If x and x' are two distinct elements of I_n and δ is a given value of I_n then $Pr[c^*(x) \oplus c^*(x') = \delta] \leq \frac{2}{2^n}$.*

Proof: $\Pr[c^*(x) \oplus c^*(x') = 0] = 0$ since $x \neq x'$. If $\delta \neq 0$, $\Pr[c^*(x) \oplus c^*(x') = \delta] = \frac{2^n \cdot 2^{n-2} \cdots 1}{2^n!} = \frac{1}{2^n - 1} \leq \frac{2}{2^n}$. So, $\Pr[c^*(x) \oplus c^*(x') = \delta] \leq \frac{2}{2^n}$.

2.3 Distinguishing Two Random Functions

In proofs of security such as the one presented here, we want to upper bound the probability of any algorithm to distinguish whether a given fixed function φ is an instance of a random function $f = \Phi(f_1^*, f_2^*, .., f_r^*)$ of $F_{n,m}$ or an instance of the perfect random function f^*, using less than q queries to φ.

Let A be any distinguishing algorithm of unlimited power that, when input with a function φ of $F_{n,m}$ (which can be modeled as an " oracle tape " in the probabilistic Turing Machine associated with A) selects a fixed number q of distinct chosen or adaptively chosen input values X_i (the queries), obtains the q corresponding output values $Y_i = f(X_i)$, and based on these results outputs 0 or 1. Denote by p (resp by p^*) the probability for A to answer 1 when fed with a random instance of f (resp of f^*). We want to find upper bounds on the $Adv_A(f, f^*) = |p - p^*|$ advantage of A in distinguishing f from f^* in q queries.

As first noticed by Patarin [Pa91], the best $Adv_A(f, f^*)$ advantage of any A distinguishing algorithm for distinguishing f from f^* is entirely determined by the $Pr[x \stackrel{f}{\mapsto} y]$ q-ary transition probabilities associated with each $X = (X_1, \cdots, X_q)$ q-uple of pairwise distinct I_n values and each $Y = (Y_1, \cdots, Y_q)$ q-uple of I_m values. The following Theorem, which was first proved in [Pa91], and equivalent versions of which can be found in [Va99], is a very useful tool for deriving establishing upper bounds on the $Adv_A(f, f^*)$ based on properties of the $Pr[x \stackrel{f}{\mapsto} y]$ q-ary transition probabilities.

Theorem 1 *Let f be a random function of $F_{n,m}$ and f^* a perfect random function representing a uniformly drawn random element of $F_{n,m}$. Let q be an integer. Denote by $\mathcal{X}$ the ${I_n}^q$ set of all $X = (X_1, \cdots, X_q)$ q-tuples of pairwise distinct elements. If there exists a $\mathcal{Y}$ subset of ${I_m}^q$ and two positive real numbers ϵ_1 and ϵ_2 such that*

1) $|\mathcal{Y}| > (1 - \epsilon_1) \cdot |I_m|^q$ *(i)*

2) $\forall X \in \mathcal{X} \ \forall Y \in \mathcal{Y} Pr[X \stackrel{f}{\mapsto} Y] \geq (1 - \epsilon_2) \cdot \frac{1}{|I_m|^q}$ *(ii)*

then for any A distinguishing using q queries

$$Adv_A(f, f^*) \leq \epsilon_1 + \epsilon_2$$

In order to improve the selfreadability of this paper, a short proof of Theorem 1 is provided in appendix at the end of this paper.

3 Description of the L- and R-Schemes

We now describe two simple variants of the Feistel scheme, that we propose to name L-scheme and R-scheme, following the terminology proposed by Kaneko and al. in their paper on the provable security against differential and linear cryptanalysis of generalised Feistel ciphers [Ka97].

The L-scheme and R-scheme both allow to derive a 2n-bit to 2n-bit permutation from several n-bit to n bit permutations (not only n-bit to n-bit functions as in the Feistel scheme), using only one n-bit to n bit permutation per round.

The 1-round L-scheme is depicted in Figure 1. It transforms a c_1 permutation of I_n into the $\psi_L(c_1)$ permutation of I_{2n} defined by

$$\psi_L(c_1)(x^1, x^0) = (x^0, c_1(x^1) \oplus x^0)$$

The extension to r rounds is straightforward: the r-round L-scheme transforms r I_n permutations c_1 to c_r into the I_{2n} permutation defined by

$$\psi_L(c_1, c_2, ..., c_r) = \psi_L(c_r) \circ \cdots \circ \psi_L(c_1)$$

The L-scheme is used at several levels of the construction of the MISTY and KASUMI ciphers, namely the derivation of the so-called FI and FO functions, and also the upper level of the construction in the case of MISTY2. One remarkable feature of the r-round L-scheme is that two c_i permutations can be processed in paralell.

The 1-round R-scheme is depicted in Figure 1 too. It transforms a $c1$ permutation of I_n into the $\psi_R(c_1)$ permutation of I_{2n} defined by

$$\psi_R(c_1)(x^1, x^0) = (c_1(x^1) \oplus x^0, c_1(x^1))$$

The r-round R-scheme transforms r I_n permutations c_1 to c_r into the I_{2n} permutation defined by $\psi_R(c_1, c_2, ..., c_r) = \psi_R(c_r) \circ \cdots \circ \psi_R(c_1)$.

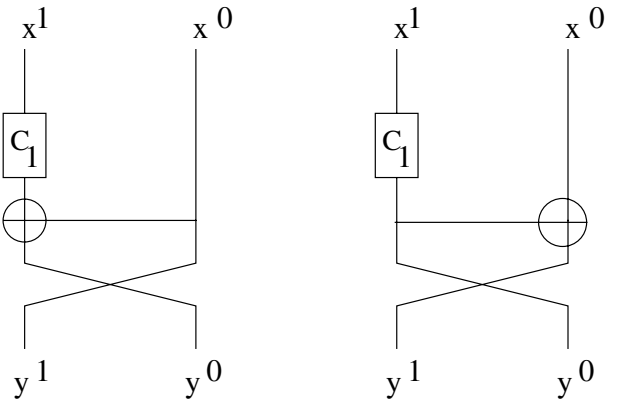

Fig. 1. L-scheme one round at left and R-scheme one round at right

In the sequel, we will several times consider the slightly simplified versions $\psi'_L(c_1, c_2, ..., c_r)$ and $\psi'_R(c_1, c_2, ..., c_r)$ of $\psi_L(c_1, c_2, ..., c_r)$ and $\psi_R(c_1, c_2, ..., c_r)$ obtained by omitting the XOR operation and the exchange of the left and right halves in the final round. We will sometimes analyse such simplified variants, whose pseudorandomness properties are obviously the same as those of the full r-round L or R scheme from which they are derived, instead of the full r-round L or R scheme, in order to simplify some discussions. We can notice that the $\psi'_R(c_1, c_2, ..., c_r)$ and $\psi'_L(c_r^{-1}, c_{r-1}^{-1}, ..., c_1^{-1})$ permutations are inverse of each other. This remark will be useful when it comes to analysing the super pseudorandomness properties of the L and R schemes.

Through the two next Sections, especially in proofs, we are using the following additional notation:

- I is an abreviation for the $(I_n)^q$ set.
- $I^{\neq} = ((I_n)^q))^{\neq}$ denotes the subset of $(I_n)^q$ consisting of all the q-tuples of pairwise distinct I_n values and $I^{=} = (I_n)^q \backslash I^{\neq}$.
- $\mathcal{X}$ denotes the subset of $(I_{2n})^q$ consisting of all $(x_1, \cdots, x_q)$ q-tuples of pairwise distinct I_{2n} values.
- $\mathcal{Y}$ will denote a subset of $(I_{2n})^q$ consisting of $(y_1, \cdots, y_q)$ q-tuples of I_{2n} values. The exact definition of $\mathcal{Y}$ will vary. This $\mathcal{Y}$ will be redefined in each Section where this notation is needed.

4 Analysis of the L-Scheme

In this Section, we compare, for various values of the r number of rounds of an L-scheme, the $f = \psi_L(c_1^*, c_2^*, ..., c_r^*)$ $2n$-bit random permutation deduced from r independent perfect random n-bit permutations $c_1^*, c_2^*, ..., c_r^*$ with a perfect $2n$-bit function f^* or a $2n$-bit perfect permutation c^*.

4.1 Three-Round L-Scheme: $\psi_L(c_1^*, c_2^*, c_3^*)$ Is Not a Pseudo-Random Function

As already noticed by several authors [Zh89], the function $f = \psi_L(c_1^*, c_2^*, c_3^*)$ associated with the three-round L-scheme is not pseudo-random.

Since the omission of the final XOR and the final exchange of the left and right output halves does not affect the pseudorandomness properties of f, we can consider the function $f = \psi'_L(c_1^*, c_2^*, c_3^*)$, instead of $\psi_L(c_1^*, c_2^*, c_3^*)$.

Let us show that 4 chosen input queries suffice to distinguish f from a the perfect random function f^* with a very large probability. Let us consider the encryption, under f, of two distinct $2n$-bit (x^1, x^0) plaintext blocks (a, b) and (a', b) which right halves are equal, and denote by (c, d) and (c', d') the two corresponding (y^1, y^0) ciphertext blocks. We can notice that $d \oplus d'$ is equal to $c_1^*(a) \oplus c_1^*(a')$, and thus independent of b. Therefore if we replace b by any other value b' and do the same computation as before, the new obtained value of $d \oplus d'$ will be left unchanged. This property allows to distinguish f from a perfect random function of I_{2n} with an advantage close to 1.

4.2 Four-Round L-Scheme: $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*)$ Is a Pseudo-Random Function

As already established by Sakurai et al. [Sa97], the four-round version of the L-scheme is indistinguishable from a perfect pseudo-random function. In order for this paper to provide a self contained summary of the properties of the L and R schemes inside the security framework introduced in Section 2, we restate this result as follows.

Theorem 2 *Let n be an integer, c_1^*, c_2^*, c_3^*, c_4^* be four independent random function from I_n to I_n and f^* be the perfect random function on the I_{2n} set. Let $f = \psi_L(f_1^*, f_2^*, f_3^*, f_4^*)$ denote the random permutation associated with the four rounds of L-scheme. For any adaptative distinguisher $\mathcal{A}$ with q queries we have:*

$$Adv_{\mathcal{A}}^{q}(f, f^*) \leq \frac{7}{2} q^2 2^{-n}$$

A short proof for Theorem 2 is provided in appendix at the end of this paper. Since the proof technique is rather similar to the one used in the more detailed proof of Theorem 5 on the pseudorandomness of the 3-round R-scheme, we omitted some details in the proof of Theorem 2.

4.3 Four-Round L-Scheme: $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*)$ Is Not a Super Pseudo-Random Permutation

As already established by Sakurai and al. [Sa97], the 4-round L-scheme does not provide a super pseudo random permutation, i.e. it is possible with a small number of encryption and decryption queries to distinguish $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*)$ from a perfect random permutation.

Instead of providing here a direct proof of this property, let us show that this is a straightforward consequence of the fact (established in the next Section) that the 4-round R-scheme does not provide a super pseudo random function. As a matter of fact, $\psi'_R(c_1, c_2, c_3, c_4)$ and $\psi'_L(c_4^{-1}, c_3^{-1}, c_2^{-1}, c_1^{-1})$ are inverse of each other, as stated in Section 2 and therefore, the distinguisher for $\psi_R(c_1^*, c_2^*, c_3^*, c_4^*)$ can be converted in a distinguisher for $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*)$ with the same number of queries and the same advantage.

4.4 Five-Round L-Scheme: $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ Is a Super Pseudo-Random Permutation

Recall that a super pseudo random distinguisher is an adaptative distinguisher which can call at one and the same time the cipher c and the cipher c^{-1}. The following Theorem shows that the five-round version of the L-scheme provides a super pseudorandom permutation.

Theorem 3 *Let n be an integer, c_1^*, c_2^*, c_3^*, c_4^*, c_5^* be five independent random functions from I_n to I_n and c^* be the perfect random permutation on the I_{2n} set. Let $c = \psi_L(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ denote the random permutation associated with the five round L scheme. For any adaptative super pseudorandom permutation distinguisher $\mathcal{A}$ with q queries, we have:*

$$Adv_{\mathcal{A}}^{q}(c, c^*) \leq \frac{9}{2} \cdot \frac{q^2}{2^n}$$

To prove this theorem, we need to use a variant of Theorem 1 due to Patarin in [Pa91] concerning permutations (that we provide here without proof):

Theorem 4 *Let m be an integer, ϵ be a positive real number, c be a random permutation on the $\{0,1\}^m$ set, c^* be the perfect random permutation on the same set. We denote by $\mathcal{X}$ the subset of $(X_1, \cdots, X_q)$ q-tuples that are pairwise distinct. Let $\mathcal{A}$ be any super pseudo random distinguisher with q queries. If $Pr[X \overset{c}{\mapsto} Y] \geq (1-\epsilon) \cdot \frac{1}{|I_m|^q}$ for all X and Y q-tuples in $\mathcal{X}$ then $Adv_{\mathcal{A}}^{q}(c, c^*) \leq \epsilon + \frac{q(q-1)}{2 \cdot 2^m}$*

Proof of Theorem 3: We will compare the $c = \psi'_L(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ permutation generator of Figure 2 (wich superpseudorandomness properties are exactly the same as for $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$) with the perfect random permutation c^* of I_{2n}. For that purpose, let us consider any $X = (x_i^1, x_i^0) \in \mathcal{X}$ q-tuple of pairwise distinct values of I_{2n} and any $Y = (y_i^1, y_i^0)$ q-tuple of pairwise distinct values of I_{2n}. We want to establish lower bound on $\Pr[X \overset{c}{\mapsto} Y]$ and then apply Theorem 4 above. We are using the notation $x^2 = (x_i^2)_{i=1..q}$, $x^3 = (x_i^3)_{i=1..q}$, $x^4 = (x_i^4)_{i=1..q}$ to refer to the q-tuples of I_n intermediate words induced by the q considered f computations, at the locations marked in Figure 2.

$$\Pr[X \overset{c}{\mapsto} Y] = \sum_{x^2, x^3, x^4} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_2^*(x^0) \oplus x^2 = x^3) \wedge (c_3^*(x^2) \oplus x^3 = x^4)$$

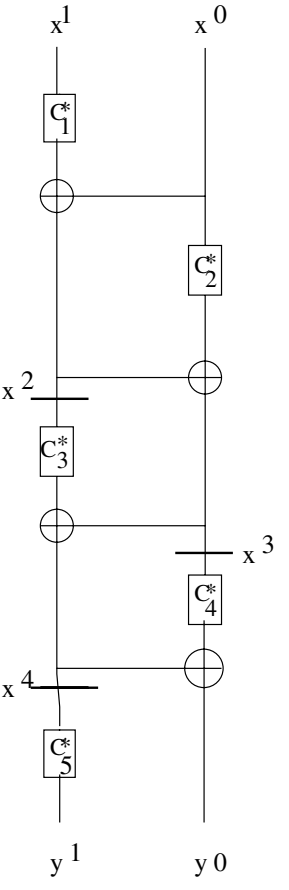

Fig. 2. L-scheme five rounds

$$\begin{aligned} & \wedge(c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)] \\ \geq \sum_{x^2,x^3 \in I^{\neq}} & \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_2^*(x^0) \oplus x^2 = x^3)] \\ \cdot \sum_{x^4} & \Pr\left[\begin{array}{c}(c_3^*(x^2) \oplus x^3 = x^4) \wedge \\ (c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)\end{array}\right] \end{aligned} \quad (1)$$

Let us consider any fixed x^2, x^3 q-tuples of $I^{\neq}$. In order to establish a lower bound on the $\sum_{x^4} \Pr[(c_3^*(x^2) \oplus x^3 = x^4) \wedge (c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)]$ factor in (2), we define the following set Z of x^4 q-tuples:

$$Z = \{x^4 | x^4 \sim y^1 \wedge x^4 \oplus y^0 \in I^{\neq} \wedge x^4 \oplus x^3 \in I^{\neq}\}$$

where $x^4 \sim y^1$ means that $\forall i, j \; x_i^4 = x_j^4$ if and only if $y_i^1 = y_j^1$. Let us denote by $q_1 \leq q$, the number of distinct y_i^1 values. there exist $i_1, \cdots, i_{q_1}$ indexes such that $y_{i_1}^1, \cdots, y_{i_{q_1}}^1$ are pairwise distinct. Each $i_k \in \{i_1, \cdots, i_{q_1}\}$ index determine a class such that for all elements i of this class, $y_i^1 = y_{i_k}^1$. So, $\forall i \in [1, .., q], \exists! i_k \in \{i_1, \cdots, i_{q_1}\} / y_i^1 = y_{i_k}^1, Cl(i) =_{def} i_k$.

There exist $\alpha = \frac{2^n!}{(2^n-q_1)!}$ x^4 values such that $x^4 \sim y^1$ (as a matter of fact such an x^4 is entirely determined by q_1 distinct values). Now:

$$\begin{aligned}|Z| &\geq |\{x^4|x^4 \sim y^1\}| - |\{x^4|x^4 \sim y^1 \wedge x^4 \oplus y^0 \notin I^{\neq}\}| \\ &\quad -|\{x^4|x^4 \sim y^1 \wedge x^4 \oplus x^3 \notin I^{\neq}\}| \\ &\geq |\{x^4|x^4 \sim y^1\}| - \sum_{i \neq j} |\{x^4|x^4 \sim y^1 \wedge x_i^4 \oplus x_j^4 = y_i^0 \oplus y_j^0\}| \\ &\quad - \sum_{i \neq j} |\{x^4|x^4 \sim y^1 \wedge x_i^4 \oplus x_j^4 = x_i^3 \oplus x_j^3\}|\end{aligned}$$

Given $i \neq j$, we can upper bound the size of $S_{ij} = \{x^4|x^4 \sim y^1 \wedge x_i^4 \oplus x_j^4 = y_i^0 \oplus y_j^0\}$ by $\frac{2\alpha}{2^n}$.
As a matter of fact:

- if $y_i^1 = y_j^1$, then $y_i^0 \oplus y_j^0 \neq 0$ (because otherwise the (y_i^1, y_i^0) word would be equal to (y_j^1, y_j^0)), but $x^4 \sim y^1$ implies $x_i^4 = x_j^4$ and thus $x_i^4 \oplus x_j^4$ cannot be equal to $y_i^0 \oplus y_j^0$. So, $|S_{ij}| = 0$
- $y_i^1 \neq y_j^1$, then if $x^4 \in S_{ij}$, $x^4_{Cl(j)}$ is entirely determined by $x^4_{Cl(i)}$ since $x^4_{Cl(j)} = x^4_{Cl(i)} \oplus y_i^0 \oplus y_j^0$. Thus $|S_{ij}|$ contains at most $2^n(2^n-2)\cdots(2^n-q_1) = \frac{\alpha}{2^n-1} \leq \frac{2\alpha}{2^n}$ elements.

Similary, using the fact that $x_i^3 \neq x_j^3$, we can upper bound the size of $\{x^4|x^4 \sim y^1 \wedge x_i^4 \oplus x_j^4 = x_i^3 \oplus x_j^3\}|$ by $\frac{2\alpha}{2^n}$. So, we have:

$$|Z| \geq \alpha \left[1 - \frac{q(q-1)}{2}\frac{2}{2^n} - \frac{q(q-1)}{2}\frac{2}{2^n}\right] \text{ i.e. } |Z| \geq \frac{2^n!}{(2^n-q_1)!}[1 - \frac{2q^2}{2^n}]$$

Now, $\sum_{x^4} \Pr[(c_3^*(x^2) \oplus x^3 = x^4) \wedge (c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)] \geq \sum_{x^4 \in Z} \Pr[(c_3^*(x^2) \oplus x^3 = x^4) \wedge (c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)]$. But, for any $x^4 \in Z$, we have $\Pr[c_5^*(x^4) = y^1] = \frac{(2^n-q_1)!}{2^n!} = \frac{1}{\alpha}$ and $\Pr[(c_3^*(x^2) = x^4 \oplus x^3)] = \frac{(2^n-q)!}{2^n!}$ due to Property 2 and the fact that the x_i^2 and the $x_i^4 \oplus x_i^3$ are pairwise distinct. We also have $\Pr[(c_3^*(x^2) = x^4 \oplus x^3)] = \frac{(2^n-q)!}{2^n!}$ for the same reasons, so that:
$\sum_{x^4} \Pr[(c_3^*(x^2) \oplus x^3 = x^4) \wedge (c_4^*(x^3) \oplus x^4 = y^0) \wedge (c_5^*(x^4) = y^1)]$

$$\begin{aligned}&\geq |Z| \cdot \left(\frac{(2^n-q)!}{2^n!}\right)^2 \cdot \frac{1}{\alpha} \\ &\geq \alpha \cdot \left(1 - \frac{2q^2}{2^n}\right)\left(\frac{(2^n-q)!}{2^n!}\right)^2 \cdot \frac{1}{\alpha} \\ &\geq \left(1 - \frac{2q^2}{2^n}\right)\left(\frac{(2^n-q)!}{2^n!}\right)^2\end{aligned}$$

If we now come back to inequality (2), we thus have:

$$\Pr[X \overset{c}{\mapsto} Y] \geq \sum_{x^2,x^3 \in I^{\neq}} \Pr\begin{bmatrix}(c_1^*(x^1) \oplus x^0 = x^2)\\ \wedge \\ (c_2^*(x^0) \oplus x^2 = x^3)\end{bmatrix} \cdot \left(1 - \frac{2q^2}{2^n}\right)\left(\frac{(2^n-q)!}{2^n!}\right)^2 \quad \text{(i)}$$

Let us now establish a lower bound on

$$\begin{aligned} B &= \sum_{x^2,x^3 \in I^{\neq}} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_2^*(x^0) \oplus x^2 = x^3)] \\ &= \Pr[x^2 \in I^{\neq} \wedge x^3 \in I^{\neq}] \\ &\quad \cdot \Pr[(c_1^*(x^1) \oplus x^0) \in I^{\neq} \wedge (c_2^*(x^0) \oplus x^2) \in I^{\neq} | x^2 \in I^{\neq}] \end{aligned}$$

But $\Pr[(c_1^*(x^1) \oplus x^0) \in I^{\neq}] \geq 1 - \sum_{i \neq j} \Pr[c_1^*(x_i^1) \oplus c_1^*(x_j^1) = x_i^0 \oplus x_j^0]$ and it is easy to establish (using the fact that $(x_i^1, x_i^0) \neq (x_j^1, x_j^0)$ and Property 3), that for any two distinct indexes i and j, $\Pr[c_1^*(x_i^1) \oplus c_1^*(x_j^1) = x_i^0 \oplus x_j^0] \leq \frac{1}{2^n - 1} \leq \frac{2}{2^n}$ Thus $\Pr[(c_1^*(x^1) \oplus x^0) \in I^{\neq}] \geq 1 - \frac{q(q-1)}{2} \cdot \frac{2}{2^n} \geq 1 - \frac{q^2}{2^n}$. for similar reasons, $\Pr[(c_2^*(x^0) \oplus x^2) \in I^{\neq} | x^2 \in I^{\neq}] \geq 1 - \frac{q^2}{2^n}$. Thus $B \geq \left(1 - \frac{q^2}{2^n}\right)^2 \geq 1 - \frac{2q^2}{2^n}$ (ii) . Now, by combinig (i) and (ii), we obtain:

$$\Pr[X \overset{c}{\mapsto} Y] \geq \left(1 - \frac{2q^2}{2^n}\right)^2 \left(\frac{(2^n-q)!}{2^n!}\right)^2$$

Now, $\left(\frac{(2^n-q)!}{2^n!}\right)^2 \geq \frac{1}{2^{2nq}}$ and $\left(1 - \frac{2q^2}{2^n}\right)^2 \geq 1 - \frac{4q^2}{2^n}$, so that

$$\Pr[X \overset{c}{\mapsto} Y] \geq \left(1 - \frac{4q^2}{2^n}\right) \cdot \frac{1}{2^{2nq}}$$

We can now apply Theorem 4 with $\epsilon = \frac{4q^2}{2^n}$ and we obtain:

$$\begin{aligned} \mathrm{Adv}_{\mathcal{A}}^q(c, c^*) &\leq \tfrac{4q^2}{2^n} + \tfrac{q(q-1)}{2_2 2^{2n}} \\ \mathrm{Adv}_{\mathcal{A}}^q(c, c^*) &\leq \tfrac{9}{2} \cdot \tfrac{q^2}{2^n} \end{aligned}$$

□

5 Analysis of the R-Scheme

In this Section, we compare, for various values of the r number of rounds of an R-scheme, the $f = \psi_R(c_1^*, c_2^*, ..., c_r^*)$ $2n$-bit random permutation deduced from r independent perfect random n-bit permutations $c_1^*, c_2^*, ..., c_r^*$ with a perfect $2n$-bit function f^*.

5.1 Three-Round R-Scheme

We first establish the following theorem for a 3-round version of the R-scheme.

Theorem 5 *Let n be an integer, c_1^*, c_2^*, c_3^* be three independent perfect random permutation from I_n to I_n and f^* be the perfect random function on the I_{2n} set. Let $f = \psi_R(c_1^*, c_2^*, c_3^*)$ denote the random permutation associated with the 3-rounds R-scheme. For any adaptive distinguisher $\mathcal{A}$ with q queries, we have:*

$$Adv_{\mathcal{A}}^{q}(f, f^*) \leq 3q^2 2^{-n}$$

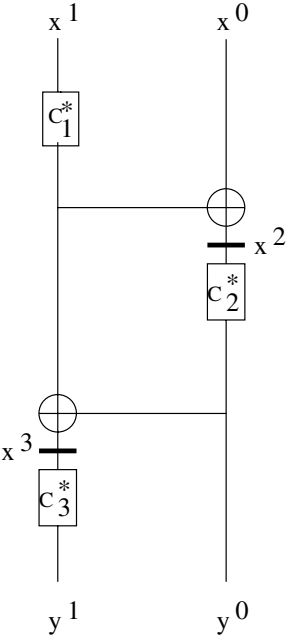

Fig. 3. R-scheme three rounds

Proof: We will compare the $f = \psi'_R(c_1^*, c_2^*, c_3^*)$ permutation generator of Figure 3 (which pseudorandomness properties are exactly the same as for $\psi_R(c_1^*, c_2^*, c_3^*)$) with the perfect random function f^*. Let us first introduce some notation. We consider a $X = (X_i)_{i\in[1..q]} = (x_i^1, x_i^0)$ q-tuple of $2n$-bit f input words. We denote the corresponding q-tuple of f output words by $Y = (Y_i)_{i\in[1..q]} = (y_i^1, y_i^0)_{i\in[1..q]}$. For each $(y_i^1, y_i^0) = \psi'_R(c_1^*, c_2^*, c_3^*)(x_i^1, x_i^0)$ computation, we denote by x_i^2 and x_i^3 the intermediate values which locations are marked in Figure 3. More explicitly:

$$x_i^2 = c_1^*(x_i^1) \oplus x_i^0$$
$$x_i^3 = c_1^*(x_i^1) \oplus c_2^*(x_i^2) = y_i^0 \oplus c_1^*(x_i^1)$$

Finally, we denote the $(x_i^0)_{i\in[1..q]}$, $(x_i^1)_{i\in[1..q]}$, $(x_i^2)_{i\in[1..q]}$, $(x_i^3)_{i\in[1..q]}$, $(y_i^0)_{i\in[1..q]}$ and $(y_i^1)_{i\in[1..q]}$ q-tuples of n-bit words by x^0, x^1, x^2, x^3, y^0, y^1 respectively.

We now define $\mathcal{X}$ as the set of X q-tuples of pairwise distinct I_{2n} words (i.e. such that for any distinct i,j numbers in $[1..q]$, $x_i^1 \neq x_j^1$ or $x_i^0 \neq x_j^0$), and define $\mathcal{Y}$ as the set of those Y q-tuples of I_{2n} words such that the corresponding y^1 and y^0 q-tuples both consist of pairwise distinct I_n words: $\mathcal{Y} = \{(Y_1, \cdots, Y_q) \in (I_{2n})^q \ / \ y^1 \in I^{\neq},\ y^0 \in I^{\neq}\}$.

We want to establish a lower bound on the size of $\mathcal{Y}$ and the $\Pr[X \overset{f}{\mapsto} Y]$ transition probability associated with any X q-tuple in $\mathcal{X}$ and any Y q-tuple in $\mathcal{Y}$ and show that there exists ϵ_1 and ϵ_2 real numbers satisfying conditions of Theorem 1.

Let us first establish a lower bound on $|\mathcal{Y}|$. We have:

$$\begin{aligned}
|\mathcal{Y}| &= |I^{2n}|^q \cdot (1 - \Pr[y^1 \notin I^{\neq} \vee y^0 \notin I^{\neq}]) \\
&\geq |I^{2n}|^q (1 - \sum_{i,j\in[1..q], i\neq j} \Pr[y_i^1 = y_j^1] - \sum_{i,j\in[1..q], i\neq j} \Pr[y_i^0 = y_j^0]) \\
&\geq |I^{2n}|^q (1 - \frac{q(q-1)}{2} \cdot 2^{-n} - \frac{q(q-1)}{2} \cdot 2^{-n}) \\
&\geq |I^{2n}|^q (1 - \frac{q(q-1)}{2^n})
\end{aligned}$$

So, we can take $\epsilon_1 = \frac{q(q-1)}{2^n}$.

Now, given any X q-tuple of $\mathcal{X}$ and any Y q-tuple of $\mathcal{Y}$ let us establish a lower bound on $\Pr[X \overset{f}{\mapsto} Y]$.

$$\begin{aligned}
\Pr[X \overset{f}{\mapsto} Y] &= \textstyle\sum_{x^2,x^3\in I} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_1^*(x^1) \oplus y^0 = x^3) \\
&\qquad \wedge (c_2^*(x^2) = y^0) \wedge (c_3^*(x^3) = y^1)] \\
&\geq \textstyle\sum_{x^2,x^3\in I^{\neq}} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_1^*(x^1) \oplus y^0 = x^3)] \\
&\qquad \cdot \Pr[(c_2^*(x^2) = y^0) \wedge (c_3^*(x^3) = y^1)] \qquad (2)
\end{aligned}$$

First, for any x^2 q-tuple of $I^{\neq}$ and any x^3 q-tuple of $I^{\neq}$, let us compute $\Pr[(c_2^*(x^2) = y^0) \wedge (c_3^*(x^3) = y^1)] = \Pr[(c_2^*(x^2) = y^0)] \cdot \Pr[(c_3^*(x^3) = y^1)]$. Since x^2 and y^0 both belong to $I^{\neq}$, we can can apply Property 2 of Section 2 concerning random permutations, so that $\Pr[(c_2^*(x^2) = y^0)] = \frac{(2^n-q)!}{2^n!}$. For the same reason, we also have $\Pr[(c_3^*(x^3) = y^1)] = \frac{(2^n-q)!}{2^n!}$.

So, we have $\Pr[(c_2^*(x^2) = y^0) \wedge (c_3^*(x^3) = y^1)] = \left(\frac{(2^n-q)!}{2^n!}\right)^2$.

Now, $\left(\frac{(2^n-q)!}{2^n!}\right)^2 \geq \frac{1}{2^{2nq}}$.

Therefore, inequality (2) implies:

$$\Pr[X \overset{f}{\mapsto} Y] \geq \sum_{x^2,x^3\in I^{\neq}} \frac{1}{2^{2nq}} \cdot \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_1^*(x^1) \oplus y^0 = x^3)] \quad (i)$$

Let us now estimate $B = \sum_{x^2,x^3 \in I^{\neq}} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_1^*(x^1) \oplus y^0 = x^3)]$

We have:

$$\begin{aligned} B &= \Pr[(c_1^*(x^1) \oplus x^0 \in I^{\neq} \wedge (c_1^*(x^1) \oplus y^0 \in I^{\neq}] \\ &= 1 - \Pr[(c_1^*(x^1) \oplus x^0 \notin I^{\neq} \vee (c_1^*(x^1) \oplus y^0 \notin I^{\neq}] \\ &\geq 1 - \sum_{i,j,i\neq j} Pr[c_1^*(x_i^1) \oplus x_i^0 = c_1^*(x_j^1) \oplus x_j^0] \\ &\quad - \sum_{i,j,i\neq j} Pr[c_1^*(x_i^1) \oplus y_i^0 = c_1^*(x_j^1) \oplus y_j^0] \end{aligned}$$

Let us evaluate $\Pr[c_1^*(x_i^1) \oplus x_i^0 = c_1^*(x_j^1) \oplus x_j^0]$ and $\Pr[c_1^*(x_i^1) \oplus c_1^*(x_j^1) = x_i^0 \oplus x_j^0]$. Due to Property 3 of Section 2, if $x_i^1 \neq x_j^1$ given any fixed difference δ (here equal to $x_i^0 \oplus x_j^0$), $\Pr[c_1^*(x_i^1) \oplus c_1^*(x_j^1) = \delta] \leq \frac{2}{2^n}$. On the other hand, if $x_i^1 = x_j^1$, then $x_i^0 \neq x_j^0$, so that $\Pr[c_1^*(x_i^1) \oplus x_i^0 = c_1^*(x_j^1) \oplus x_j^0]$. In all cases, $\Pr[c_1^*(x_i^1) \oplus x_i^0 = c_1^*(x_j^1) \oplus x_j^0] \leq \frac{2}{2^n}$. Applying this property to the $\frac{q(q-1)}{2}$ $(i,j), i \neq j$ pairs of $[1..q]$ indexes, we obtain $\sum_{i,j} Pr[c_1^*(x_i^1) \oplus x_i^0 = c_1^*(x_j^1) \oplus x_j^0] \leq \frac{q(q-1)}{2^n}$. For the same reasons, $\sum_{i,j,i\neq j} Pr[c_1^*(x_i^1) \oplus y_i^0 = c_1^*(x_j^1) \oplus y_j^0] \leq \frac{q(q-1)}{2^n}$. Thus:

$$B \geq 1 - \frac{2q(q-1)}{2^n} \quad (ii)$$

By using inequalities (i) and (ii), we obtain:

$$\Pr[X \overset{f}{\mapsto} Y] \geq \left(1 - \frac{2q(q-1)}{2^n}\right) \cdot \frac{1}{2^{2nq}}$$

We can notice that $Pr[X \overset{f^*}{\mapsto} Y] = \frac{1}{2^{2nq}}$. So we can apply Theorem 1 with $\epsilon_1 = \frac{q(q-1)}{|I^n|}$ and $\epsilon_2 = \frac{2q(q-1)}{|I^n|}$. We obtain:

$$\begin{gathered} \mathrm{Adv}_{\mathcal{A}}^q(f, f^*) \leq \frac{3q(q-1)}{2^n} \\ \mathrm{Adv}_{\mathcal{A}}^q(f, f^*) \leq \frac{3q^2}{2^n} \end{gathered}$$

□

5.2 Four-Round R-Scheme: $\psi_R(c_1^*, c_2^*, c_3^*, c_4^*)$ Is Not a Super Pseudo-Random Permutation

We consider the four-round permutation generator f deduced from $\psi_R(c_1^*, c_2^*, c_3^*, c_4^*)$ by omitting the final XOR (this does obviously not matter for the super pseudo randomness issue considered here). The random function f can be represented by extending the 3-round function of Figure 3 above by one round.

Let us show that 2 chosen encryption and two chosen decryption queries suffice to distinguish f from a the perfect random permutation $c*$ with a very large probability. Let us consider the encryption, under the function f, of two

distinct $2n$-bit (x^1, x^0) plaintext blocks (a, b) and (a, b') which left halves are equal, and denote by (c, d) and (c', d') the two obtained (y^1, y^0) ciphertext blocks. It is easy to check that if we swap the left halves of the two obtained ciphertexts, thus obtaining two modified ciphertexts (c', d) and (c, d') and if we decrypt (c', d) and (c, d') under f^{-1}, we obtain two plantext values α, β and α', β' which left halves are equal: $\alpha = \alpha'$. This would be extremely unlikely to happen if f was replaced by a perfect random permutation $c*$.

The above test allows to distinguish f from a perfect random permutation of I_{2n} with a probability close to 1.

5.3 Five-Round R-Scheme: $\psi_R(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ Is a Super Pseudo-Random Permutation

The following theorem is a direct consequence of Theorem 3 due to the fact that $\psi'_R(c_1, c_2, c_3, c_4, c_5)$ and $\psi'_L(c_1^{-1}, c_2^{-1}, c_3^{-1}, c_4^{-1}, c_5^{-1})$ are inverse of each other, every distinguisher for $\psi_R(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ can be converted into a distinguisher for $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ with the same number of encryption and decryption queries. Therefore, Theorem 3 results in the following analogue theorem for the 5-round R-scheme.

Theorem 6 *Let n be an integer, c_1^*, c_2^*, c_3^*, c_4^*, c_5^* be five independent random functions from I_n to I_n and c^* be the perfect random permutation on the I_{2n} set. Let $c = \psi_R(c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ denote the random permutation associated with the five round R-scheme. For any adaptive super pseudorandom permutation distinguisher $\mathcal{A}$ with q queries, we have:*

$$Adv_{\mathcal{A}}^{q}(c, c^*) \leq \frac{9}{2} \cdot \frac{q^2}{2^n}$$

6 Conclusion

As a consequence of previous results, the security properties of the L-scheme and the R-scheme are distinct when it comes to chosen plaintext attacks, but equivalent when it comes to chosen plaintext or ciphertext attacks. As a matter of fact, the minimal number of rounds required in order of the R-scheme to be undistinguishable from a pseudorandom function with adaptively chosen encryption queries is less than for the L-scheme (3 rounds instead of 4), whereas the minimal numbers of rounds required by the R-scheme and the L-scheme in order to be undistinguishable from a pseudorandom permutation with adaptively chosen encryption or decryption queries are equal (5 rounds for both schemes).

A Appendix

A.1 A Short Proof of Theorem 1

Let us restrict ourselves to the case of any fixed deterministic algorithm A which uses q adaptively chosen queries (the generalisation to the case of a probabilistic algorithm is easy).

A has the property that if the q-uple of outputs encountered during an A computation is $Y = (Y_1, \cdots, Y_q)$, the value of the $X = (X_1, \cdots, X_q)$ q-uple n is of query inputs encountered during this computation is entirely determined (this is easy to prove by induction: the initial query input X_1 is fixed ; if for a given A computation the first query output is Y_1, then X_2 is determined, etc.). We denote by $X(Y)$ the single q-uple of query inputs corresponding to any possible Y q-uple of query outputs, and we denote by S_A the subset of those $Y \in {I_m}^q$ values such that if the $X(Y)$ and Y q-uples query inputs and outputs are encountered in a A computation, then A outputs a 1 answer.

The p and p^* probabilities can be expressed using S_A as

$$p = \sum_{Y \in S_A} Pr[X(Y) \stackrel{f}{\mapsto} Y] \text{ and } p^* = \sum_{Y \in S_A} Pr[X(Y) \stackrel{f^*}{\mapsto} Y]$$

We can now lower bound p using the following inequalities:

$p \geq \sum_{Y \in S_A \cap \mathcal{Y}} (1 - \epsilon_2) \cdot Pr[X(Y) \stackrel{f^*}{\mapsto} Y]$ (due to inequality (ii))

$\geq \sum_{Y \in S_A} (1 - \epsilon_2) Pr[X(Y) \stackrel{f^*}{\mapsto} Y] - \sum_{Y \in {I_m}^{q-} mathcal Y} (1 - \epsilon_2) \cdot Pr[X(Y) \stackrel{f^*}{\mapsto} Y]$

But

$$\sum_{Y \in S_A} (1 - \epsilon_2) \cdot Pr[X(Y) \stackrel{f^*}{\mapsto} Y] = (1 - \epsilon_2) p^*$$

and

$$\sum_{Y \in {I_m}^q - \mathcal{Y}} (1 - \epsilon_2) \cdot Pr[X(Y) \stackrel{f^*}{\mapsto} Y] = (1 - \epsilon_2) \cdot \frac{|I_m|^q - |\mathcal{Y}|}{|I_m|} \leq (1 - \epsilon_2) \cdot \epsilon_1 \text{ (due to inequality (i)).}$$

Therefore $p \geq (1 - \epsilon_2)(p^* - \epsilon_1) = p^* - \epsilon_1 - \epsilon_2 \cdot p^* + \epsilon_1 \cdot \epsilon_2$

thus finally (using $p^* \leq 1$ and $\epsilon_1 \cdot \epsilon_2 \geq 0$)

$p \geq p* - \epsilon_1 - \epsilon_2$ (a)

If we now consider the A' distinguisher which outputs are the inverse of those of A (i.e. A' answers 0 iff A answers 1), we obtain an inequality involving this time $1 - p$ and $1 - p*$:

$(1 - p) \geq (1 - p*) - \epsilon_1 - \epsilon_2$ (b)

Combining inequalities (a) and (b), we obtain $|p - p*| \leq \epsilon_1 + \epsilon_2$ QED.

A.2 A Proof Sketch for Theorem 2

We will compare the $f = \psi'_L(c_1^*, c_2^*, c_3^*, c_4^*)$ permutation generator of Figure 4 (which pseudorandomness properties are exactly the same as for $\psi_L(c_1^*, c_2^*, c_3^*, c_4^*)$) with the perfect random function f^*. This proof is near to the proof of Section 5.1. That's why we do not detail some computations that are the same that in Section 5.1.

Let us first introduce some notation. We consider a $X = (X_i)_{i \in [1..q]} = (x_i^1, x_i^0)$ q-tuple of 2n-bit f input words. We denote the corresponding q-tuple of f output words by $Y = (Y_i)_{i \in [1..q]} = (y_i^1, y_i^0)_{i \in [1..q]}$. For each $(y_i^1, y_i^0) = \psi'_L(c_1^*, c_2^*, c_3^*, c_4^*)(x_i^1, x_i^0)$ computation, we denote by x_i^2 and x_i^3 the intermediate values which locations are marked in Figure 4. More explicitly:

$$x_i^2 = c_1^*(x_i^1) \oplus x_i^0$$
$$x_i^3 = c_2^*(x_i^0) \oplus x_i^2$$

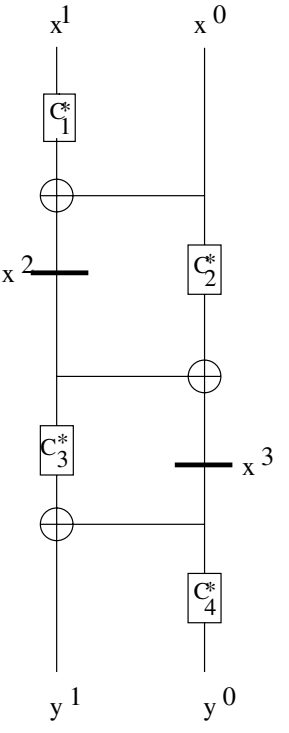

Fig. 4. L-scheme four rounds

Finally, we denote the $(x_i^0)_{i\in[1..q]}$, $(x_i^1)_{i\in[1..q]}$, $(x_i^2)_{i\in[1..q]}$, $(x_i^3)_{i\in[1..q]}$, $(y_i^0)_{i\in[1..q]}$ and $(y_i^1)_{i\in[1..q]}$ q-tuples of n-bit words by x^0, x^1, x^2, x^3, y^0, y^1 respectively.

We now define $\mathcal{X}$ as the set of X q-tuples of pairwise distinct I_{2n} words (i.e. such that for any distinct i,j numbers in $[1..q]$, $x_i^1 \neq x_j^1$ or $x_i^0 \neq x_j^0$), and define $\mathcal{Y}$ as the set of those Y q-tuples of I_{2n} words such that the corresponding y^1 q-tuples consists of pairwise distinct I_n words: $\mathcal{Y} = \{(Y_1, \cdots, Y_q) \in (I^{2n})^q \ / \ y^1 \in I^{\neq},\ y^0 \in I^{\neq}\}$.

We want to lower bound the size of $\mathcal{Y}$ and the $\Pr[X \overset{f}{\mapsto} Y]$ transition probability associated with any X q-tuple in $\mathcal{X}$ and any Y q-tuple in $\mathcal{Y}$ and show that there exists ϵ_1 and ϵ_2 real numbers satisfying conditions of Theorem 1.

We have (for more details, see section 5.1):

$$\begin{aligned} |\mathcal{Y}| &= |I^{2n}|^q \cdot (1 - \Pr[y^1 \notin I^{\neq}]) \\ &\geq |I^{2n}|^q (1 - \frac{q(q-1)}{2.2^n}) \end{aligned}$$

So, we can take $\epsilon_1 = \frac{q(q-1)}{2.2^n}$.

Now, given any X q-tuple of $\mathcal{X}$ and any Y q-tuple of $\mathcal{Y}$ let us establish a lower bound on $\Pr[X \overset{f}{\mapsto} Y]$ (for more details, see section 5.1).

$$\Pr[X \overset{f}{\mapsto} Y] \geq \sum_{x^2,x^3,x^3\oplus y^1 \in I^{\neq}} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (x^3 = c_2^*(x^0))] \oplus x^2) \cdot \Pr[(c_3^*(x^2) \oplus x^3 = y^0) \wedge (c_4^*(x^3) = y^1)] \quad (3)$$

First, for any x^2, x^3, $x^3 \oplus y^1$ q-tuple of $I^{\neq}$, we have $\Pr[(c_3^*(x^2) \oplus x^3 = y^0) \wedge (c_3^*(x^3) = y^1)] = \left(\frac{(2^n-q)!}{2^n!}\right)^2 \geq \frac{1}{2^{2nq}}$ (for more details, see section 5.1). Therefore, inequality (1) implies:

$$\Pr[X \overset{f}{\mapsto} Y] \geq \sum_{x^2,x^3,x^3\oplus y^1 \in I^{\neq}} \frac{1}{2^{2nq}} \cdot \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_1^*(x^1) \oplus y^0 = x^3)] \quad (i)$$

Let us now estimate:

$$B = \sum_{x^2,x^3,x^3\oplus y^1 \in I^{\neq}} \Pr[(c_1^*(x^1) \oplus x^0 = x^2) \wedge (c_2^*(x^0) \oplus x^2 = x^3)]$$

We have:

$$\begin{aligned} B &= \Pr[(c_1^*(x^1) \oplus x^0) \in I^{\neq} \wedge (c_2^*(x^0) \oplus c_1^*(x^1) \oplus x^0) \in I^{\neq} \wedge (x^3 \oplus y^1) \in I^{\neq}] \\ &= 1 - \Pr[(c_1^*(x^1) \oplus x^0 \in I^{=}) \vee (c_2^*(x^0) \oplus c_1^*(x^1) \oplus x^0) \in I^{=} \vee (x^3 \oplus y^1) \in I^{=}] \\ &\geq 1 - 3.\frac{q(q-1)}{2}.\frac{2}{2^n} \end{aligned}$$

By using inequalities (i) and (ii), we obtain:

$$\Pr[X \overset{f}{\mapsto} Y] \geq \left(1 - \frac{3q(q-1)}{2^n}\right) \cdot \frac{1}{2^{2nq}}$$

We can notice that $Pr[X \overset{f^*}{\mapsto} Y] = \frac{1}{2^{2nq}}$. So we can apply Theorem 1 with $\epsilon_1 = \frac{q(q-1)}{2|I^n|}$ and $\epsilon_2 = \frac{3q(q-1)}{|I^n|}$. We obtain:

$$\begin{aligned} \mathrm{Adv}_{\mathcal{A}}^q(f, f^*) &\leq \tfrac{7q(q-1)}{2.2^n} \\ \mathrm{Adv}_{\mathcal{A}}^q(f, f^*) &\leq \tfrac{7q^2}{2.2^n} \end{aligned}$$

□

References

[Ai96] W. Aiello, R. Venkatesan, "Foiling Birthday Attacks in Length-Doubling Transformations". In *Advances in Cryptology - Eurocrypt'96*, LNCS 1070, p. 307, Springer Verlag, Saragossa, Spain, May 1996.

[Be94] M. Bellare, J. Kilian, P. Rogaway, "The Security of Cipher Block Chaining". In *Advances in Cryptology - CRYPTO'94*,LNCS 839, p. 341, Springer-Verlag, Santa Barbara, U.S.A., 1994.

[Ka] Specification of the 3GPP confidentiality and Integrity algorithm KASUMI. Documentation available on http://www.etsi.org/

[Ka97] Y. Kaneko, F. Sano, K. Sakurai, "On Provable Security against Differential and Linear Cryptanalysis in Generalized Feistel Ciphers with Multiple Random Functions". In *Selected Areas in Cryptography - SAC'97*, Ottawa, Canada, August 1997.

[La90] X. Lai, J.L. Massey, "A Proposal for a New Block Encryption Standard". In *Advances in Cryptology - Eurocrypt'90*, LNCS 473 , p. 389, Springer Verlag, Aarhus, Denemark, 1991.

[Lu88] M. Luby, C. Rackoff, "How to Construct Pseudorandom Permutations from Pseudorandom Function". In *Siam Journal on Computing* , vol. 17, p. 373, 1988.

[Ma92] U. Maurer, "A Simplified and generalised treatment of Luby-Rackoff Pseudorandom Permutation Generators", In *Advances in Cryptology - Eurocrypt'92*, LNCS 658 , p. 239, Springer Verlag, New York, USA, 1992.

[Ma93] M. Matsui, "New Block Encryption Algorithm MISTY", In *Fast Software Encryption - FSE'97*, LNCS 1267, p. 54, Springer Verlag, Haifa, Israel, 1997.

[Pa91] J. Patarin, "Etude de Générateurs de Permutation Basés sur le Schéma du D.E.S. ", Phd. Thesis, University of Paris VI, 1991.

[Sa97] K. Sakurai, Y. Zheng, "On Non-Pseudorandomness from Block Ciphers with Provable Immunity Against Linear Cryptanalysis, In *IEICE Trans. Fundamentals*, vol. E80-A, n. 1, January 1997.

[Su96] M. Sugita, "Pseudorandomness of a Block Cipher MISTY", Technical Reporrt of IEICE, ISEC96-9.

[Su97] M. Sugita, "Pseudorandomness of a Block Cipher with Recursive Structures", Technical Report of IEICE, ISEC97-9.

[Va99] S. Vaudenay, "On Provable Security for Conventional Cryptography", In *ICISC'99*, invited lecture.

[Zh89] Y. Zheng, T. Matsumoto, H. Imai, "On the Construction of Block Ciphers Provably Secure and Not Relying on Any Unproved Hypotheses". In *Advances in Cryptology - CRYPTO'89*, LNCS 435, p. 461, Springer-Verlag, Santa Barbara, U.S.A., 1990.