Design of near - optimal pseudorandom functions and pseudorandom permutations in the information - theoretic model

Jacques Patarin, Paul Camion
Université de Versailles
45 avenue des Etats-Unis
78035 Versailles Cedex - France

Abstract

In this paper we will extend the Benes and Luby-Rackoff constructions to design various pseudo-random functions and pseudo-random permutations with near optimal information-theoretic properties. An example of application is when Alice wants to transmit to Bob some messages against Charlie, an adversary with unlimited computing power, when Charlie can receive only a percentage τ of the transmitted bits. By using Benes, Luby-Rackoff iterations, concatenations and fixing at 0 some values, we will show in this paper how to design near optimal pseudo-random functions for all values of τ. Moreover we will show how to design near optimal pseudo-random permutations when τ can have any value such that the number of bits obtained by Charlie is smaller than the square root of all the transmitted bits.

1 Introduction

In their famous paper [6], M. Luby and C. Rackoff have shown that in adaptive plaintext attack (CPA-2) with m queries, the probability p to distinguish a 3-round random Feistel scheme (i.e. a Feistel scheme of $2n$ bits $\rightarrow 2n$ bits made by using 3 random round functions of n bits $\rightarrow n$ bits) from a truly random permutation of $2n$ bits $\rightarrow 2n$ bits, is always $p \leq \frac{m^2}{2^n}$, i.e. we have CPA-2 security when $m \ll \sqrt{2^n}$.
Similarly, the probability p to distinguish a 4-round random Feistel scheme from a truly random permutation of $2n$ bits $\rightarrow 2n$ bits in an adaptive chosen plaintext and chosen ciphertext attack (CPCA-2) with m queries, is always $p \leq \frac{m^2}{2^n}$, i.e. we have CPCA-2 security when $m \ll \sqrt{2^n}$.
These results are valid if the adversary has unbounded computing power as long as he does only m queries. The bound $m \ll \sqrt{2^n}$ is called the "birthday

bound".
These results of Luby and Rackoff have inspired a considerable amount of research. In [11] a summary of existing works on this topic is given.
One direction of research, that we will also followed in the present paper, is to improve the birthday bound, i.e. the security bound $m \ll \sqrt{2^n}$ for various permutations or functions generators.
In [13], and independently in [1], it is shown that for the Luby-Rackoff theorems for 3 or 4 rounds, the bound $m \ll \sqrt{2^n}$ is optimal.
In [1], W. Aiello and R. Venkatesan have found a construction of locally random functions, called "Benes", where the optimal bound ($m \ll 2^n$, or $m \ll O\left(2^{n(1-\varepsilon)}\right)$ with a different constant in the O for all $\varepsilon > 0$ see [16]) can be obtained instead of the birthday bound. This bound $m \ll 2^n$ is called "optimal" or "optimal in information-theoretic cryptography", since it is the best possible bound against an adversary with unlimited computing power. However here the functions are not permutations.
Similarly , U. Maurer in [7] and J. Patarin in [14] have found some other construction of locally random functions (not permutations) where they can get as close as wanted to the optimal bound: $m \ll 2^{n(1-\varepsilon)}$ and for all $\varepsilon > 0$ they have a construction. However, unlike Benes and unlike the schemes that we will study in this paper, we have here a different construction for each $\varepsilon > 0$. Moreover, with Benes or with the function Ω_p of [14], the bound of the number of bits of security (i.e. the maximum number of bits of queries) is always about $\leq$ the square root of the total number of possible inputs (the functions are from $2n$ bits $\rightarrow 2n$ bits, so the total number of inputs is $\leq 2^{2n}$). With one-time-pad the number of bits of security is equal with the total number of bits sent.
In this paper, we will look for constructions of pseudorandom functions or pseudorandom permutations, where we will still be very near the optimal information-theoretic cryptography bound, but where the number of bits of security may have different values than the square root of the total number of possible inputs (as with Benes) or the number of total bits sent (as with one-time-pad). It can be noticed that our proof will be simple, since we will use the (relatively difficult) results on Benes and Feistel schemes to design new schemes by using only simple operations: concatenations, composition of functions, and fixing at 0 some values.

Remark In [9], U. Maurer and J. Massey have also design some schemes with, as the schemes of this paper, near-perfect local randomness. However the designs of the schemes of [9] are completely different from the schemes of this paper, since they use error correcting code theory, instead of variations of Feistel and Benes constructions. The properties of the schemes of [9] are also very different from the schemes of this paper. The schemes of [9] have a uniform property of perfectly local randomness unlike the schemes of this paper where the schemes have a probabilistic security: the probability to distinguish them from perfectly random schemes is negligible. The schemes of [9] are linear, while the schemes of this paper are not linear. The complexity of computation of the schemes of [9] is at least quadratic in the number of bits sent while the complexity of the schemes of this paper will be linear in the number of bits sent (when the keys are generated and stored). So the schemes of [9] and the schemes of this paper are completely different, and we believe that both are interesting for cryptography.

2 Notations and first examples

• CPA-2 means "adaptive chosen plaintext attack".
• CPCA-2 means "adaptive chosen plaintext and chosen ciphertext attack".
• $I_n = \{0, 1\}^n$ is the set of the 2^n binary strings of length n.
• F_n is the set of all functions $f : I_n \to I_n$. Thus $|F_n| = 2^{n \cdot 2^n}$.
• For a, $b \in I_n$, $a \oplus b$ stands for bit by bit exclusive or of a and b.
• For a, $b \in I_n$, $a||b$ stands for the concatenation of a and b.
• For $a, b \in I_n$, we also denote by $[a, b]$ the concatenation $a||b$ of a and b.
• For any $f, g \in F_n$, $f \circ g$ denotes the usual composition of functions.
• Let f_1 be a function of F_n. Let L_i, R_i, S_i and T_i be four n-bit strings in I_n. Then by definition

$$\Psi(f_1)[L_i, R_i] = [S_i, T_i] \overset{\text{def}}{\Leftrightarrow} \begin{cases} S_i = R_i \\ T_i = L_i \oplus f_1(R_i) \end{cases}$$

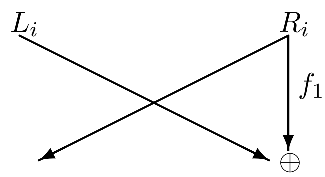

Figure 1: One round of Feistel Transformation ψ.

• Let $f_1, f_2, \ldots, f_k$ be k functions of F_n. Then by definition:

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

The permutation $\Psi^k(f_1, \ldots, f_k)$ is called a "Feistel scheme with k rounds" or shortly Ψ^k. When $f_1, \ldots, f_k$ are randomly and independently chosen in F_n, then $\Psi^k(f_1, \ldots, f_k)$ is called a "random Feistel scheme with k rounds" or a "Luby-Rackoff construction with k rounds".
• Given four functions from n bits to n bits, $f_1, \ldots, f_4$, we use them to define the **Butterfly transformation** (see [1]) from $2n$ bits to $2n$ bits. On input $[L_i, R_i]$, the output is given by $[X_i, Y_i]$, with:

$$X_i = f_1(L_i) \oplus f_2(R_i) \text{ and } Y_i = f_3(L_i) \oplus f_4(R_i).$$

• Given eight functions from n bits to n bits, $f_1, \ldots, f_8$, we use them to define the **Benes transformation** (see [1]) (back-to-back Butterfly) as the

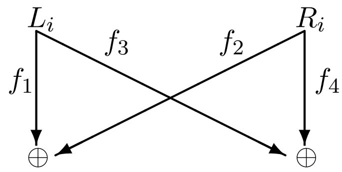

Figure 2: Butterfly transformation

composition of two Butterfly transformations. On input $[L_i, R_i]$, the output is given by $[S_i, T_i]$, with: $Benes(f_1, \cdots, f_8)[L_i, R_i] = [S_i, T_i]$ if and only if:

$$S_i = f_5(f_1(L_i) \oplus f_2(R_i)) \oplus f_6(f_3(L_i) \oplus f_4(R_i)) = f_5(X_i) \oplus f_6(Y_i)$$

$$T_i = f_7(f_1(L_i) \oplus f_2(R_i)) \oplus f_8(f_3(L_i) \oplus f_4(R_i)) = f_7(X_i) \oplus f_8(Y_i).$$

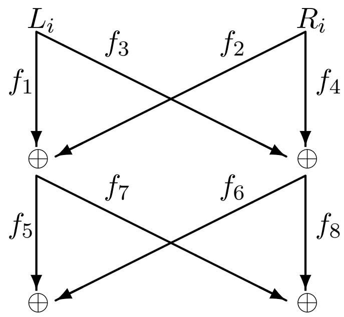

Figure 3: Benes transformation (back-to-back Butterfly)

When we will study a scheme we will denote by:
• K the number of bits of the keys.
• N the total number of bits of all the messages that we can send.
• τ the proportion of the bits sent that the enemy can obtain in an adaptive attack (all the other bits sent are not received by the enemy).
• m the number of messages that the enemy can obtain in an adaptive attack.
• n' the number of bits of each message.
• e the number of bits that the enemy can obtain in an adaptive attack. So we have $m \leq e \leq mn'$, and $\tau = \frac{e}{N}$.
• When we have security when $e \ll K^{1/2}$ we say that we have the "birthday bound". When we have security when $e \ll K^{2/3}$ we say that we have the "3-collision bound". When we have security when $e \ll K^{\theta-1/\theta}$ we say that we have the "θ-collision bound". If a scheme is such that we have for all integer $\theta \geq 1$ security with the "θ-collision bound", we say that the scheme is "near-optimal". Notice that the scheme does not change with θ: it is the same scheme with the "θ-collision bound" for all integer $\theta \geq 1$.
• More precisely we will say that a scheme is "near-optimal" or "near-optimal in the information-theoretic model" if the scheme is secure against cryptographic attacks when $\forall \varepsilon > 0$, $e \ll f(\varepsilon)K^{1-\varepsilon}$ where $f(\varepsilon)$ is a function of

ε only (not of K). By "cryptographic attacks" we mean here CPA-2 for pseudorandom functions, and CPCA-2 for pseudorandom permutations.
• As always in cryptography, we will assume that the enemy knows everything about the schemes, except the values of the secret keys (Kerckhoff's principle). However here K is as large as N.

Example 1 With one-time-pad, we have $e = K = N$, $\tau = 1$, so the scheme is "near-optimal" (with the definition above).

Example 2 If we send 2^n messages with the algorithm $S_i = X_i \oplus k$, where $1 \le i \le 2^n$, $X_i, S_i, k \in I_n$, k is the secret key (of n bits), we have $N = 2^n \cdot n$ (all the possible messages are all the elements of I_n, and we have here 2^n elements of n bits) and $K = n$. However, here there is a very simple known plaintext attack: let X_1 and X_2 be some known values, $X_1 \neq X_2$, and test if $S_1 \oplus S_2 = X_1 \oplus X_2$. For a random function, the probability to have $S_1 \oplus S_2 = X_1 \oplus X_2$ is $\frac{1}{2^n}$, and for this scheme the probability is 1. So we can distinguish this scheme from a random function with only $m = 2$ messages in a known plaintext attack. So here for security with this scheme we can send only one message. Moreover, even if we know only one bit (for example bit number b) of S_1, S_2, X_1, X_2, then the probability to have $(S_1 \oplus S_2)_b = (X_1 \oplus X_2)_b$ is $\frac{1}{2}$ for random functions and 1 for this scheme. So an enemy will be able to distinguish this scheme from a random scheme with a non negligible probability ($\frac{1}{2}$ here) with only two chosen bits (one bit of X_1 and one bit chosen at the same position b of X_2), and here $K = n$. Therefore this scheme is not "near-optimal": we do not even have the birthday bound here since 2 is smaller than $\sqrt{n}$.

3 The theorems that we will use

Theorem 3.1 (Luby and Rackoff) *The probability p to distinguish ψ^4 (i.e. a 4-round Feistel scheme with 4 random functions f_1, f_2, f_3, f_4 of $I_n \to I_n$ as round functions) from a truly random permutation of $I_{2n} \to I_{2n}$ in an adaptive chosen plaintext/chosen ciphertext attack (CPCA-2) always satisfies: $p \le \frac{m(m-1)}{2^n}$.*

Proof The theorem was originally given in [6]. Some simplified proof are given for example in [7] for non-adaptive attacks, and in [12] for adaptive attacks.

Information-theoretic properties With our notations of section 2, we have here $n' = 2n$, $N = 2^{2n} \cdot 2n$, $K = 4n \cdot 2^n$ (since f_1, f_2, f_3, f_4 are the secret key here), and we have security when $m \ll \sqrt{2^n}$. Here $e \le mn'$, so

$e \leq \sqrt{2^n} \cdot 2n$, and $K = 4n \cdot 2^n$. So this scheme ψ^4 is not "near-optimal" if the key is f_1, f_2, f_3, f_4 (we just have the "birthday bound" here).

Remark However, from ψ^4 we will design "near-optimal" schemes in this paper, but in these schemes f_1, f_2, f_3, f_4 will not be truly random functions of F_n.

Theorem 3.2 (Patarin) *The probability p to distinguish ψ^6 (i.e. a 6-round Feistel scheme with 6 random functions $f_1, f_2, f_3, f_4, f_5, f_6$ of $I_n \to I_n$ as round functions) from a truly random permutation of $I_{2n} \to I_{2n}$ in an adaptive chosen plaintext/chosen ciphertext attack (CPCA-2) always satisfies: $p \ll 1$ when $m \ll 2^n$. We will note: $p \leq$ Feistel 6 security $(m, 2^n)$, with Feistel 6 security $(x, y) \to 0$ when $x \ll y$.*

Proof This theorem is given in [15] p.110. Notice that we have here security with 6 rounds when $m \ll 2^n$ instead of $m \ll \sqrt{2^n}$ with 4 rounds.

Information-theoretic properties With our notations of section 2, we have here $n' = 2n$, $N = 2^{2n} \cdot 2n$, $K = 6n \cdot 2^n$, and we have security when $m \ll 2^n$. So we have security when $e \ll 2^n$, and $K = 6n \cdot 2^n$. Since $2^n \geq (6n \cdot 2^n)^(1 - \varepsilon)$ for all $\varepsilon > 0$ and sufficiently large n this scheme is "near-optimal" with our definition of section 2. However here, when τ is fixed, N is fixed, or when K is fixed, N is fixed ($N \simeq K^2$). In this paper we will design some solutions for various independent values of N and K.

Theorem 3.3 (Aiello and Venkatesen) *The probability p to distinguish Benes$(f_1, \cdots, f_8)$ with 8 random functions $(f_1, \cdots, f_8)$ of $I_n \to I_n$ from a truly random function of $I_{2n} \to I_{2n}$ in an adaptive chosen plaintext attack (CPA-2) always satisfies: $p \ll 1$ when $m \ll 2^n$. We will note: $p \leq$ Benes security $(m, 2^n)$, with Benes 8 security $(x, y) \to 0$ when $x \ll y$.*

Proof This theorem is given in [1]. However the proof given in [1] is valid for most attacks, but not for all CPA-2 attacks (see [16]). Nevertheless in [16] a complete proof is given, and it is shown that $p \leq \frac{m^2}{2 \cdot 2^n}$, $p \leq \frac{m^3}{2^{2n}}$, $p \leq \frac{m^4}{2^{3n}} + \frac{6m^2}{2^{2n}}$, and more generally, it is shown that for all integer $k \geq 1$, $p \leq \frac{k \cdot k^{2k} m^2}{2 \cdot 2^{2n}} + \frac{m^{k+1}}{2^{nk}}$. So for any $\varepsilon > 0$, for sufficiently large n, $m \ll 2^{n(1-\varepsilon)}$ gives CPA-2 security for Benes.

Information-theoretic properties With our notations of section 2, we have here $n' = 2n$, $N = 2^{2n} \cdot 2n$, $K = 8n \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so when $\tau \ll \frac{2^n \cdot 2n}{2^{2n} \cdot 2n} = \frac{1}{2^n}$. Here we have security when e
$ll2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 8n \cdot 2^n$. Since for all $\varepsilon' > 0$ we can find an

$\varepsilon > 0$ such that for sufficiently large n, $2^{n(1-\varepsilon)} \geq (8n \cdot 2^n)^{(1-\varepsilon)}$ this scheme is "near-optimal" with our definition of section 2. However here (as with ψ^6), when τ is fixed, N is fixed, or when K is fixed, N is fixed ($N \simeq K^2$). In this paper we will design some solutions for various independent values of N and K.

4 A stream cipher from a pseudorandom function

From a pseudorandom permutation G of α bits $\rightarrow$ α bits, we have immediately a block encryption scheme: each cleartext M_i will have α bits, $1 \leq i \leq m$, and the ciphertext C_i will be: $C_i = G(M_i)$. If the probability to distinguish G from a truly random permutation of α bits in CPCA-2 is negligible when the number m of messages is $\leq m_0$, then this scheme is secure against CPCA-2 when $m \leq m_0$.
From a pseudorandom function G of α bits $\rightarrow$ β bits, we can also easilly design a scheme to encrypt messages. We can proceed like this: each cleartext M_i will have β bits, $1 \leq i \leq m$, and the ciphertext C_i of the cleartext number i will be: $C_i = G(i) \oplus M_i$ (this is a stream cipher). If the probability to distinguish G from a truly random permutation of α bits $\rightarrow \beta$ bits in CPA-2 is negligible when the number m of messages is $\leq m_0$, then this scheme is secure against CPCA-2 when $m \leq m_0$. This comes from the fact that in CPCA-2, when M_i is chosen and C_i is given then this is equivalent to choose i and get $G(i)$, and when C_i is chosen and M_i is given, then this is again equivalent to choose i and get $G(i)$.

5 First variants of Benes, concatenations

5.1 Concatenation of two Benes: Benes from $2n$ bits $\rightarrow 4n$ bits

Let $[L_i, R_i]$, $1 \leq i \leq m$, be the inputs, $[L_i, R_i] \in I_{2n}$, let $f_1, \cdots, f_{16}$ be 16 functions of F_n, and let $G(f_1, \cdots, f_{16})[L_i, R_i] = [S_i, T_i, S'_i, T'_i]$ if and only if:

$$\begin{cases} Benes(f_1, \cdots, f_8)[L_i, R_i] = [S_i, T_i] \\ Benes(f_9, \cdots, f_{16})[L_i, R_i] = [S'_i, T'_i] \end{cases}$$

Theorem 5.1 *The probability p to distinguish $G(f_1, \cdots, f_{16})$ with 16 random functions $f_1, \cdots, f_{16}$ of $I_n \rightarrow I_n$ from a truly random function of $I_{2n} \rightarrow I_{2n}$ in CPA-2 always satisfies: $p \ll 1$ when $m \ll 2^n$. More precisely: $p \leq 2$ Benes security $(m, 2^n)$, where Benes security $(m, 2^n)$ represents the security bound of the original Benes function (as seen in section 3).*

Proof $G(f_1, \cdots, f_{16})$ is the concatenation of two Benes functions with independent keys. The probability p to distinguish the $[S_i, T_i, S'_i, T'_i]$ values from random values $[A_i, B_i, C_i, D_i]$ in CPA-2 is $p \leq p_1 + p_2$, where p_1 is the probability to distinguish the $[S_i, T_i, S'_i, T'_i]$ from $[S_i, T_i, C_i, D_i]$ in CPA-2 (where C_i, D_i are random values), and p_1 is the probability to distinguish the $[S_i, T_i, C_i, D_i]$ from $[A_i, B_i, C_i, D_i]$, $1 \leq i \leq m$, in CPA-2 (where A_i, B_i, C_i, D_i, $1 \leq i \leq m$, are random values). Now $p_1 \leq$ Benes security$(m, 2^n)$ (because if we want to distinguish $[S_i, T_i, S'_i, T'_i]$ from $[S_i, T_i, C_i, D_i]$ with probability p_1, we can distinguish $[S'_i, T'_i]$ from $[C_i, D_i]$, $1 \leq i \leq m$, with probability p_1) and similarly $p_2 \leq$ Benes security$(m, 2^n)$. So $p \leq 2$ Benes security$(m, 2^n)$ as claimed.

Information-theoretic properties of this scheme With our notations of section 2, we have here $n' = 4n$ (from a function generator of α bits $\rightarrow \beta$ bits we can build a stream cipher with messages of β bits as explained in section 4, and here $\beta = 4n$), $N = 2^{2n} \cdot 4n$, $K = 16n \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$. Here we have security when $e \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 16n \cdot 2^n$, so the scheme is "near-optimal" with our definition of section 2.

5.2 Other variants of Benes from $2n$ bits $\rightarrow 4n$ bits

Instead of using 16 functions of F_n as the key, we can use only 12 functions of F_n (or even less), as we will explain now. However this change (12 instead of 16) is not a very important change for us, since we want to obtain "near-optimal" schemes and with our definition of "near-optimal", a change by factor $\frac{12}{16}$, or 2, or any small constant, does not change the property of "near-optimal" (unlike the change by factor 2^n or $\sqrt{2^n}$), i.e. we concentrate the analysis on the dominant terms of the values K, N, τ. However in practical applications, to divide, the length of the key by a factor $\frac{12}{16}$ or by a factor 2 might be interesting.

Let $[L_i, R_i]$, $1 \leq i \leq m$, be the inputs, $[L_i, R_i] \in I_{2n}$, let $f_1, \cdots, f_{12}$ be 12 functions of F_n, and let $G(f_1, \cdots, f_{12})[L_i, R_i] = [S_i, T_i, S'_i, T'_i]$ if and only if:

$$\begin{cases} S_i = f_5(X_i) \oplus f_6(Y_i) \\ T_i = f_7(X_i) \oplus f_8(Y_i) \\ S'_i = f_9(X_i) \oplus f_{10}(Y_i) \\ T'_i = f_{11}(X_i) \oplus f_{12}(Y_i) \end{cases}$$

with

$$\begin{cases} X_i = f_1(L_i) \oplus f_2(R_i) \\ Y_i = f_3(L_i) \oplus f_4(R_i) \end{cases}$$

Theorem 5.2 *The probability p to distinguish $G(f_1, \cdots, f_{12})$ with 12 random functions $f_1, \cdots, f_{12}$ of $I_n \rightarrow I_n$ from a truly random function of*

$I_{2n} \to I_{2n}$ in CPA-2 always satisfies: $p \ll 1$ when $m \ll 2^n$. More precisely: $p \leq$ the probability q to have a "circle in X, Y" when f_1, f_2, f_3, f_4 are randomly chosen, and this probability q is $\ll 1$ when $m \ll 2^n$ (or $m \ll 2^{n(1-\varepsilon)}$ for any fixed $\varepsilon > 0$

Definition 5.1 *We will say that we have "a circle in X, Y of length k" if we have k pairwise distinct indices such that $X_{i_1} = X_{i_2}$, $Y_{i_2} = Y_{i_3}$, $X_{i_3} = X_{i_4}$,..., $X_{i_{k-1}} = X_{i_k}$, $Y_{i_k} = Y_{i_1}$. We will say that we have "a circle in X, Y" if there is an even integer k, $k \geq 2$, such that we have a circle in X, Y of length k.*

Proof of theorem 5.2 (We give here only the main idea since for our purpose theorem 5.1 is sufficient as explained above). In [1] and in [16] it is explained that if f_1, f_2, f_3, f_4 are such that we have no circle in X, Y, then f_5, f_6, f_7, f_8 will make S_i, T_i perfectly random, because in each new equation we have f_5 or f_6 or f_7 or f_8 on a new variable. With exactly the same argument, we see that if f_1, f_2, f_3, f_4 are such that we have no circle in X, Y, $f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{12}$ will make S_i, T_i, S'_i, T'_i perfectly random. Moreover in [16] it is proved that the probability q to have a circle in X, Y when f_1, f_2, f_3, f_4 are randomly chosen in F_n is $\ll 1$ when $m \ll 2^{n(1-\varepsilon)}$ (for any fixed $\varepsilon > 0$).

5.3 Concatenation of λ Benes: Benes from $2n$ bits $\to \lambda(2n)$ bits

What we have done in sections 5.1 and 5.2 for two Benes, we can do it for any number λ of Benes. We obtain like this from 8λ (or $4\lambda + 4$) functions f_i of n bits $\to n$ bits, a function G of $2n$ bits $\to \lambda(2n)$ bits. For each fixed value of λ, G is with $K = 8\lambda \cdot n \cdot 2^n$ (or $K = (4\lambda + 4)n2^n$), $n' = 2\lambda n$, $N = 2^{2n} \cdot 2\lambda n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$. Here we have security when $e \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 8\lambda n \cdot 2^n$, so for each fixed value of λ, the scheme is "near-optimal" with our definition of section 2.

6 First variants of Benes, fixing some output bits

6.1 Benes from $2n$ bits $\to n$ bits

Here we can decide to use a Benes scheme such that the output is only S_i (we do not need T_i anymore). We obtain like this a pseudorandom function G of $2n$ bits $\to n$ bits, the secret key is made of 6 random functions f_i (we do not need f_7 and f_8 anymore). From theorem 3.3 we have:

Theorem 6.1 *The probability p to distinguish this scheme G with 6 random functions f_i in CPA-2 always satisfies: $p \ll 1$ when $m \ll 2^n$. More* *precisely:* $p \leq$ *Benes security* $(m, 2^n)$*, with the same notation for Benes security* $(m, 2^n)$ *as above.*

Proof Each CPA-2 on G with probability p gives a CPA-2 on the original Benes with probability p. So theorem 6.1 is immediately implied by theorem 3.3.

Information-theoretic properties of this scheme Here $n' = n$, $N = 2^{2n} \cdot n$, $K = 6n \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$.Here we have security when $e \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 6n\lambda n \cdot 2^n$, so the scheme is "near-optimal" with our definition of section 2.

6.2 Benes from $2n$ bits $\to \beta$ bits, $1 \leq \beta \leq n$

Let $[L_i, R_i]$, $1 \leq i \leq m$, be the inputs, $[L_i, R_i] \in I_{2n}$, let $f_1, \cdots, f_4$ be 4 random functions of $I_n \to I_n$, let $f_5, \cdots, f_8$ be 4 random functions of $I_n \to I_\beta$, and let $G(f_1, \cdots, f_8)[L_i, R_i] = [S_i]$ if and only if: $S_i = f_5(f_1(L_i) \oplus f_2(R_i)) \oplus f_6(f_3(L_i) \oplus f_4(R_i))$. Here G is the restriction of Benes to the first β bits.

Theorem 6.2 *The probability* p *to distinguish this scheme* G *with 4 random functions of* $I_n \to I_n$*, and 4 random functions of* $I_n \to I_\beta$ *in CPA-2 always satisfies:* $p \ll 1$ *when* $m \ll 2^n$*. More precisely:* $p \leq$ *Benes security* $(m, 2^n)$*, with the same notation for Benes security* $(m, 2^n)$ *as above.*

Proof Each CPA-2 on G with probability p gives a CPA-2 on the original Benes with probability p. So theorem 6.2 is immediately implied by theorem 3.3.

Information-theoretic properties of this scheme Here $n' = \beta$, $N = 2^{2n} \cdot \beta$, $K = 4n \cdot 2^n + 4\beta \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$. Here we have security when $e \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 4n \cdot 2^n + 4\beta \cdot 2^n \leq 8n \cdot 2^n$, so the scheme is "near-optimal" with our definition of section 2.

Conclusion of sections 5 and 6 By combining the constructions of sections 5 and 6, we can design "near-optimal" pseudorandom functions generator G of $2n$ bits $\to \beta$ bits, for all integer β. So it is relatively easy to modify the length of the outputs of G. Let $F_{\alpha,\beta}$ be the set of all functions if $I_\alpha \to I_\beta$. $|F_{\alpha,\beta}| = (2^\beta)^{2^\alpha} = 2^{\beta \cdot 2^\alpha}$. So a random element of $F_{\alpha,\beta}$ is given by $\beta \cdot 2^\alpha$ bits. This value increases only linearly in β, but exponentially in α. This explains, in a way why, when we design pseudorandom functions generator G of α bits $\to \beta$ bits, to modify β is relatively easier than to modify α. In the next sections, we will modify α.

7 "Concentration" of the key for pseudorandom functions

7.1 Benes from α bits $\rightarrow 2n$ bits, $n \leq \alpha \leq 2n$

Let $[L_i, r_i]$, $1 \leq i \leq m$, be the inputs, $L_i \in I_n$, $r_i \in I_{\alpha-n}$. Let $R_i = r_i || 0_{2n-\alpha}$ where $0_{2n-\alpha}$ is $2n - \alpha$ bits at 0. So $R_i \in I_n$ (since r_i has $\alpha - n$ bits and $0_{2n-\alpha}$ has $2n - \alpha$ bits). Let $f_1, \cdots, f_8$ be 8 random functions of F_n, and let $G(f_1, \cdots, f_8)[L_i, r_i] = Benes(f_1, \cdots, f_8)[L_i, R_i]$.

Theorem 7.1 *The probability p to distinguish this scheme $G(f_1, \cdots, f_8)$ with 8 random functions $f_1, \cdots, f_8$ of $I_n \rightarrow I_n$ from truly random functions of $I_\alpha \rightarrow I_{2n}$ in CPA-2 always satisfies: $p \ll 1$ when $m \ll 2^n$. More precisely: $p \leq$ Benes security $(m, 2^n)$, where Benes security $(m, 2^n)$ represents the security bound of the original Benes function.*

Proof Each CPA-2 on G with probability p gives immediately a CPA-2 on the original Benes with probability p, since in a CPA-2 on the original Benes we can always decide to choose R_i values with the $2n - \alpha$ last bits at 0. So theorem 7.1 is immediately implied by theorem 3.3

Remark This construction of "near-optimal" pseudorandom function of α bits $\rightarrow 2n$ bits from a "near-optimal" pseudorandom function of $2n$ bits $\rightarrow 2n$ bits is very easy (even obvious), but with pseudorandom permutation we will not be able to get such a simple construction (because if we fix at 0 some bits of the input we do not obtain a permutation anymore). The "concentration" of the key for pseudorandom permutations is a much more difficult problem than with pseudorandom functions.

Information-theoretic properties of this scheme Here $n' = 2n$, $N = 2^\alpha \cdot 2n$, $K = 8n \cdot 2^n$ (or $K = 7n \cdot 2^n + n \cdot 2^{\alpha-n}$ since we can define f_2 from $I_{\alpha-n} \rightarrow I_n$), and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so we have security when $e \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, and $K = 8n \cdot 2^n$. So this scheme is "near-optimal" with our definition of section 2.

7.2 Pseudorandom functions of α bits $\rightarrow 2n$ bits, $1 \leq \alpha \leq n$

Let l_i, $1 \leq i \leq m$, be the inputs, $l_i \in I_\alpha$. Let f_1, f_2 be 2 random functions of $I_\alpha \rightarrow I_n$, and let G be the function of $I_\alpha \rightarrow I_{2n}$ defined by: $G(l_i) = f_1(l_i) || f_2(l_i)$.

Theorem 7.2 *If f_1, f_2 are 2 perfectly random functions of $I_\alpha \rightarrow I_n$, independently chosen, then G is a perfectly random function of $I_\alpha \rightarrow I_{2n}$.*

Proof The proof is obvious: for each new value l_i, the value $f_1(l_i)$ and $f_2(l_i)$ are perfectly random variables of I_n, independently chosen, so $f_1(l_i)||f_2(l_i)$ is a perfectly random variable of I_{2n}.

Remark The only interest of this theorem 7.2 is to illustrate the fact that in theorem 7.1 the condition $n \leq \alpha$ is not a real restriction in the design of our pseudorandom functions. Or, equivalently, that when the length K of the key becomes larger than the number of bits to define G, we can have a perfectly random function G.

7.3 Pseudorandom functions of α bits $\to \beta$ bits, $n \leq \alpha \leq 2n$

By combining the construction of section 7.1 and of section 5.3 we get immediately a "near-optimal" function from α bits $\to \beta$ bits, $n \leq \alpha \leq 2n$, for all value α, with security against CPA-2 when $m \ll 2^n$.

8 "Dilution" of the key for pseudorandom functions

8.1 Pseudorandom functions of $4n$ bits $\to 4n$ bits with security when $m \ll 2^n$

Definition of G Let $f_1^{(1)}, \cdots, f_8^{(1)},\ f_1^{(2)}, \cdots f_8^{(2)},\ \cdots,\ f_1^{(8)}, \cdots f_8^{(8)}$ be 64 random functions of F_n, independently chosen. Let $F_1, \cdots, F_8$ be 8 random functions of $I_{2n} \to I_{2n}$ such that:
$F_1 = Benes(f_1^{(1)}, \cdots, f_8^{(1)})$
$F_2 = Benes(f_1^{(2)}, \cdots, f_8^{(2)})$
$\vdots$
$F_8 = Benes(f_1^{(8)}, \cdots, f_8^{(8)})$.
Let $G(f_1^{(1)}, \cdots, f_8^{(8)})$ be the function of $I_{4n} \to I_{4n}$ defined by

$$G = Benes(F_1, \cdots, F_8).$$

Expression of G Let $[L_i, R_i]$, $1 \leq i \leq m$, be the inputs of G, $L_i, R_i \in I_{2n}$, with $L_i = [l_i, r_i]$, $R_i = [l'_i, r'_i]$, $l_i, r_i, l'_i, r'_i \in I_n$.

Then $\forall i$, $1 \leq i \leq m$, $\forall [S_i, T_i] \in I_{4n}$, $G[L_i, R_i] = [S_i, T_i]$ if and only if:

$$\begin{cases} S_i = F_5(F_1(L_i) \oplus F_2(R_i)) \oplus F_6(F_3(L_i) \oplus F_4(R_i)) \\ T_i = F_7(F_1(L_i) \oplus F_2(R_i)) \oplus F_8(F_3(L_i) \oplus F_4(R_i)) \end{cases}$$

with, for example,

$$F_1(L_i) = F_1[l_i, r_i] = [f_5^{(1)}(f_1^{(1)}(l_i) \oplus f_2^{(1)}(r_i)) \oplus f_6^{(1)}(f_3^{(1)}(l_i) \oplus f_4^{(1)}(r_i)), \\ f_7^{(1)}(f_1^{(1)}(l_i) \oplus f_2^{(1)}(r_i)) \oplus f_8^{(1)}(f_3^{(1)}(l_i) \oplus f_4^{(1)}(r_i))]$$

Theorem 8.1 *The probability p to distinguish $G(f_1^{(1)}, \cdots, f_8^{(8)})$ with 64 random functions $f_1^{(1)}, \cdots, f_8^{(8)}$ randomly and independently chosen in F_n from truly random functions of $I_{4n} \to I_{4n}$ in CPA-2 always satisfies: $p \ll 1$ when $m \ll 2^n$. More precisely: $p \leq$ Benes security $(m, 2^{2n})$+8 Benes security $(m, 2^n)$, where Benes security $(m, 2^n)$ represents the security bound of the original Benes function.*

Proof $p \leq q + r_1 + r_2 + \cdots + r_8$

- where q iq the probability to distinguish $Benes(F_1, \cdots, F_8)$ from a random function of F_{4n} when $F_1, \cdots, F_8$ are 8 random functions independently chosen in F_{2n}.
- where r_i, $1 \leq i \leq 8$, is the probability to distinguish $F_i = Benes(f_1^{(i)}, \cdots, f_8^{(i)})$ from a random function of F_{2n} when $f_1^{(i)}, \cdots, f_8^{(i)}$ are 8 random functions independently chosen in F_{2n}.

So $q =$ Benes security $(m, 2^{2n})$, and $\forall i$, $1 \leq i \leq 8$, $r_i =$ Benes security $(m, 2^n)$, so we have $p \leq$ Benes security $(m, 2^{2n}) + 8$ Benes security $(m, 2^n)$, as claimed.

Remark Here Benes security $(m, 2^{2n})$ is much smaller than Benes security $(m, 2^n)$ but we do not need this property. More precisely, in a construction $G = H(F_1, \cdots, F_8)$ with $F_1, \cdots, F_8$ build with Benes as above, the birthday bound for H is sufficient to prove that G will be near-optimal (this comes from the fact that here the F_i are of $2n$ bits $\to 2n$ bits and G of $4n$ bits $\to 4n$ bits with security when $m \ll 2^n$). We will use this property in section 9.1 where H will be, in this example, a Feistel scheme with 4 rounds.

Information-theoretic properties of this scheme Here $n' = 4n$, $N = 2^{4n} \cdot 4n$, $K = 64n \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so when $e \ll 2^{n(1-\varepsilon)}$, with $K = 64n \cdot 2^n$. So this scheme is "near-optimal" with our definition of section 2.

8.2 Pseudorandom functions of $2^\alpha n$ bits $\to 2^\alpha n$ bits with security when $m \ll 2^n$

We can build a "near-optimal" function G of $8n$ bits $\to 8n$ bits with $G = Benes(F_1, \cdots, F_8)$ where $F_1, \cdots, F_8$ are 8 functions of $4n$ bits $\to 4n$ bits generated as in section 8.1. So here the secret key is made by $8 \times 64 = 512$ random functions f_i of n bits $\to n$ bits, and the proof that G is "near-optimal" is obtained as we did for the proof of theorem 8.1. This construction can be generalized immediately for each fixed value α, to get a "near-optimal" function of $2^\alpha n$ bits $\to 2^\alpha n$ bits with security when $m \ll 2^n$ (however α must be small if we want a key not too large).

8.3 Pseudorandom functions of α bits $\to \beta$ bits

By combining the constructions of sections 5, 6, 7 and 8, we can build a "near-optimal" function of α bits $\to \beta$ bits, for all values α and β, with security against CPA-2 when $m \ll 2^n$ (here unlike with the original Benes or Feistel schemes, α and β are not fixed when n is fixed).
For example, to get a function of $3n$ bits $\to n$ bits, with security when m 2^n we will start from the construction of section 8.2 of $4n$ bits $\to 4n$ bits and then by fixing some input/output values as we did in sections 5, 6, we get a quasi-optimal function of $3n$ bits $\to n$ bits (here instead of 64 functions we need only 60 functions since we do not need $f_2^{(2)}, f_4^{(2)}, f_2^{(4)}$ and $f_4^{(4)}$ since we start from $3n$ bits instead of $4n$ bits).

8.4 A natural scheme that is not near-optimal here: the Benes/Damgård scheme

To obtain a pseudorandom function of $3n$ bits $\to n$ bits with security when $m \ll 2^n$, we have seen a solution in section 8.3 above. This solution uses 64-4=60 random functions of n bits $\to n$ bits. We might think of a simpler construction, for example the scheme H below. Let $[l_i, r_i, l'_i]$, $1 \leq i \leq m$, be the inputs, $l_i, r_i, l'_i \in I_n$. Benes is a function of $2n$ bits $\to 2n$ bits. Let Benes* denotes the function of $2n$ bits $\to n$ bits that is made of the first n bits of Benes.
Let $f_1, \cdots, f_6, f'_1, \cdots, f'_6$ be 12 random functions of F_n. Let H be the function of $3n$ bits $\to n$ bits such that: $H[l_i, r_i, l'_i] = t_i$ if and only if: $t_i = Benes^*(f'_1, \cdots, f'_6)(Benes^*(f_1, \cdots, f_6)[l_i, r_i], l'_i)$. So we have: $t_i = f'_5(f'_1(s_i) \oplus f'_2(l'_i)) \oplus f'_6(f'_3(s_i) \oplus f'_4(l'_i))$ with $s_i = f_5(f_1(l_i) \oplus f_2(r_i)) \oplus f_6(f_3(l_i) \oplus f_4(r_i))$.
We call this scheme H a "Benes/Damgård" scheme, since the construction is very similar to the Damgård construction of [3] when we apply this construction to build a Hash function of $3n$ bits $\to n$ bits from a Hash function of $2n$ bits $\to n$ bits. the construction of [3] is proved secure in the following sense: from a function of $2n$ bits $\to n$ bits resistant to collisions, then the construction of $3n$ bits $\to n$ bits will also be resistant to collisions. However the property to be resistant to CPA-2 when $m \ll 2^n$ is not equivalent with the property to be resistant to collision: as we will see below, this property is true for Benes* of $2n$ bits $\to n$ bits, and will be wrong for H of $3n$ bits $\to n$ bits.

A (non-adaptive) chosen plaintext attack on H when $m \simeq \sqrt{2^n}$
Let us choose l'_i = constant, and let N be the number of (i, j), $1 \leq i < j \leq m$, such that $t_i = t_j$.

- For random values t_i, we will have $N \simeq \frac{m(m-1)}{2 \cdot 2^n}$ (1) since $t_i \in I_n$.

- For $t_i = H[l_i, r_i, l'_i]$, since $l'_i = l'_j$, we can have $t_i = t_j$ if $s_i = s_j$, or if $s_i \neq s_j$ and $f'_5(f'_1(s_i) \oplus f'_2(l'_i)) \oplus f'_6(f'_3(s_i) \oplus f'_4(l'_i)) = f'_5(f'_1(s_j) \oplus f'_2(l'_j)) \oplus f'_6(f'_3(s_j) \oplus f'_4(l'_j))$. So we will have $N \simeq 2\frac{m(m-1)}{2 \cdot 2^n}$ (2). When $m \simeq \sqrt{2^n}$, we will be able to distinguish if we are in case (1) or (2), so when $m \simeq \sqrt{2^n}$we can distinguish H from a truly random function of $3n$ bits $\rightarrow n$ bits.

So H is not near-optimal (we just have the birthday bound for H). This example shows that not all the simple constructions from Benes are near-optimal, and that our results of sections 8.1, 8.2 and 8.3 are not obvious.

9 "Dilution" of the key for pseudorandom permutations

9.1 Pseudorandom permutation of $4n$ bits $\rightarrow$ $4n$ bits with security when $m \ll 2^n$

Definition of G Let $f_1^{(1)}, \cdots, f_8^{(1)}$, $f_1^{(2)}, \cdots f_8^{(2)}$, $\cdots$, $f_1^{(4)}, \cdots f_8^{(4)}$ be 32 random functions of F_n, independently chosen. Let F_1, F_2, F_3, F_4 be 4 random functions of $I_{2n} \rightarrow I_{2n}$ such that:
$F_1 = Benes(f_1^{(1)}, \cdots, f_8^{(1)})$
$F_2 = Benes(f_1^{(2)}, \cdots, f_8^{(2)})$
$F_3 = Benes(f_1^{(3)}, \cdots, f_8^{(3)})$
$F_4 = Benes(f_1^{(4)}, \cdots, f_8^{(4)})$.
Let $G(f_1^{(1)}, \cdots, f_8^{(4)})$ be the function of $I_{4n} \rightarrow I_{4n}$ defined by

$$G = \psi^4(F_1, F_2, F_3, F_4).$$

Expression of G Let $[L_i, R_i]$, $1 \leq i \leq m$, be the inputs of G, $L_i, R_i \in I_{2n}$, with $L_i = [l_i, r_i]$, $R_i = [l'_i, r'_i]$, $l_i, r_i, l'_i, r'_i \in I_n$.

Then $\forall i$, $1 \leq i \leq m$, $\forall [S_i, T_i] \in I_{4n}$, $G[L_i, R_i] = [S_i, T_i]$ if and only if:

$$(1) \begin{cases} S_i = L_i \oplus F_1(R_i) \oplus F_3(R_i \oplus F_2(L_i \oplus F_1(R_i))) \\ T_i = R_i \oplus F_2(L_i \oplus F_1(R_i)) \oplus F_4(S_i) \end{cases}$$

We also have this expression of the $[L_i, R_i]$ from the $[S_i, T_i]$:

$$(2) \begin{cases} R_i = T_i \oplus F_4(S_i) \oplus F_2(S_i \oplus F_3(T_i \oplus F_4(S_i))) \\ L_i = S_i \oplus F_3(T_i \oplus F_4(S_i)) \oplus F_1(R_i) \end{cases}$$

Theorem 9.1 *The probability p to distinguish $G(f_1^{(1)}, \cdots, f_8^{(4)})$ with 32 random functions $f_1^{(1)}, \cdots, f_8^{(4)}$ randomly and independently chosen in F_n from truly random permutations of $I_{4n} \rightarrow I_{4n}$ in CPCA-2 always satisfies: $p \ll 1$*

when $m \ll 2^n$. More precisely: $p \leq \frac{m(m-1)}{2^{2n}} + 4$ Benes security $(m, 2^{2n})$, where Benes security $(m, 2^n)$ represents the security bound of the original Benes function.

Remark As already noticed in [14], here we have CPCA-2 security for G, from CPCA-2 security for ψ^4 and only CPA-2 security (not CPCA-2) for the F_i functions.

Proof We give here the main ideas since a similar proof was done in [14] p.147. In CPCA-2 we can have two types of queries: direct or inverse.

First case: direct query If $[L_i, R_i]$ is the input of a direct query, then we will get $[S_i, T_i]$ with the expression (1). So $[S_i, T_i]$ can be computed from the values $F_1(R_i)$, $F_2(L_i \oplus F_1(R_i))$, $F_3(R_i \oplus F_2(L_i \oplus F_1(R_i)))$ and $F_4(L_i \oplus F_1(R_i) \oplus F_3(R_i \oplus F_2(L_i \oplus F_1(R_i))))$.

Second case: inverse query If $[S_i, T_i]$ is the input of an inverse query, then we will get $[L_i, R_i]$ with the expression (2). So $[L_i, R_i]$ can be computed from the values $F_4(S_i)$, $F_3(T_i \oplus F_4(S_i))$, $F_2(S_i \oplus F_3(T_i \oplus F_4(S_i)))$ and $F_1(T_i \oplus F_4(S_i) \oplus F_2(S_i \oplus F_3(T_i \oplus F_4(S_i))))$.

Let us assume that we have a CPCA-2 to distinguish G from a truly random permutation of $4n$ bits $\to$ $4n$ bits with probability p. From theorem 3.1 (Luby and Rackoff) and the analysis above of direct and inverse queries, we know that $p \leq \frac{m(m-1)}{2^{2n}} + q$, where q is the probability to distinguish F_1, F_2, F_3, F_4 from 4 truly random functions of $2n$ bits $\to$ $2n$ bits in CPA-2. Now from theorem 3.3 (Security of Benes) we know that $q \ll 1$ when $m \ll 2^n$, and more precisely that $q \leq 4$ Benes security $(m, 2^n)$. So we have $p \leq \frac{m(m-1)}{2^{2n}} + 4$ Benes security $(m, 2^n)$, as claimed.

Information-theoretic properties of this scheme Here $n' = 4n$, $N = 2^{4n} \cdot 4n$, $K = 32n \cdot 2^n$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so when $e \ll 2^{n(1-\varepsilon)}$, with $K = 32n \cdot 2^n$. So this scheme is "near-optimal" with our definition of section 2.

Remark Here as noticed in section 8.1, when $G = H(F_1, \cdots, F_4)$, the birthday bound for H is sufficient to prove that G will be near-optimal, because G is a function of $4n$ bits $\to$ $4n$ bits, the F_i are from $2n$ bits $\to$ $2n$ bits, and the key is made from functions $f_i^{(j)}$ of n bits $\to$ n bits.

9.2 Pseudorandom permutations of 2α bits $\to 2\alpha$ bits with security when $m \ll 2^n$

Let F_1, F_2, F_3, F_4 be 4 pseudorandom functions of α bits $\to \alpha$ bits, $\alpha \geq 2n$, build with independent keys, such that the probability q to distinguish these functions from truly random functions of $I_\alpha \to I_\alpha$ satisfies $q \ll 1$ when $m \ll 2^n$, and such that these functions are near-optimal. We know from section 8 how to build such functions. Let G be the pseudorandom permutation of 2α bits $\to 2\alpha$ bits such that: $G = \psi^4(F_1, F_2, F_3, F_4)$.

Theorem 9.2 *The probability p to distinguish G from a truly random permutation of $I_{2\alpha} \to I_{2\alpha}$ in CPCA-2 satisfies: $p \ll 1$ when $m \ll 2^n$.*

Proof The proof is the same as the proof of theorem 9.1: the probability p to distinguish G from a truly random permutation of $I_{2\alpha}$ in CPCA-2 satisfies $p \leq q + r$, where q is the probability to distinguish F_1, F_2, F_3, F_4 from truly random functions of $I_\alpha \to I_\alpha$ in CPA-2, and where r is the probability to distinguish $\psi^4(F_1, F_2, F_3, F_4)$ from a truly random permutation of $I_{2\alpha}$ when F_1, F_2, F_3, F_4 are randomly and independently chosen in the set of all functions of $I_\alpha \to I_\alpha$. From Luby-Rackoff theorem we have $r \leq \frac{m(m-1)}{2^\alpha}$. Since here by hypothesis $\alpha \geq 2n$, q and r are negligible if $m \ll 2^n$, so p is negligible if $m \ll 2^n$.

Information-theoretic properties of this scheme Here $n' = 2\alpha$, $N = 2^{2\alpha} \cdot 2\alpha$, $K = O(n \cdot 2^n)$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so when $e \ll 2^{n(1-\varepsilon)}$. So this scheme is "near-optimal" with our definition of section 2.

9.3 Pseudorandom permutation of 2α bits $\to 2\alpha$ bits, $2\alpha \geq 2n$ with security when $m \ll 2^n$

Let $F_1, F_2, F_3, F_4, F_5, F_6$ be 6 pseudorandom functions of α bits $\to \alpha$ bits, $\alpha \geq n$, build with independent keys, such that the probability q to distinguish these functions from truly random functions of F_α satisfies $q \ll 1$ when $m \ll 2^n$, and such that these functions are near-optimal. We know from sections 7 and 8 how to build such functions. Let G be the pseudorandom permutation of 2α bits $\to 2\alpha$ bits such that: $G = \psi^6(F_1, F_2, F_3, F_4, F_5, F_6)$.

Theorem 9.3 *The probability p to distinguish G from a truly random permutation of $I_{2\alpha} \to I_{2\alpha}$ in CPCA-2 satisfies: $p \ll 1$ when $m \ll 2^n$.*

Proof The proof is the same as the proof of theorem 9.2, except that instead of using theorem 3.1 for ψ^4, we use now theorem 3.2 for ψ^6. The probability p to distinguish G from a truly random permutation of $I_{2\alpha}$

in CPCA-2 satisfies $p \leq q + r$, where q is the probability to distinguish $F_1, F_2, F_3, F_4, F_5, F_6$ from truly random functions of $I_\alpha \to I_\alpha$ in CPA-2, and where r is the probability to distinguish $\psi^6(F_1, F_2, F_3, F_4, F_5, F_6)$ from a truly random permutation of $I_{2\alpha}$ when $F_1, F_2, F_3, F_4, F_5, F_6$ are randomly and independently chosen in the set of all functions of $I_\alpha \to I_\alpha$. From theorem 3.2 we have $r \ll 1$ when $m \ll 2^\alpha$ (instead of $m \ll 2^{\alpha/2}$ for ψ^4). Since here by hypothesis $\alpha \geq n$, q and r are negligible if $m \ll 2^n$, so p is negligible if $m \ll 2^n$.

Information-theoretic properties of this scheme Here $n' = 2\alpha$, $N = 2^{2\alpha} \cdot 2\alpha$, $K = O(n \cdot 2^n)$, and we have security when $m \ll 2^{n(1-\varepsilon)}$ for any $\varepsilon > 0$, so when $e \ll 2^{n(1-\varepsilon)}$. So this scheme is "near-optimal" with our definition of section 2.

10 "Concentration" of the key for pseudorandom permutations

To build a pseudorandom permutation of 2α bits $\to$ 2α bits, $n \leq \alpha < 2n$, with security when $m \ll 2^n$ with near-optimal security is still an open problem. We suggest that the analysis of unbalanced Feistel schemes of $2n$ bits $\to$ $2n$ bits built from random round functions of α bits $\to$ α bits, $n < \alpha < 2n$ might be useful, but such an analysis has only been done so far up to birthday bound (cf. [11]).

11 Examples of applications

- As mentioned in [1], the Benes scheme can be useful to design keyed hash functions. The variants given in this paper can also be used to design keyed hash functions, with a compression that can be chosen independently from the length of the key.
- As mentioned in [9], schemes with local randomness properties can be excellent building blocks within practical ciphers for spreading local randomness when used together with compressing transformations that guarantee confusion.
- As also mentioned in [9], they are very useful wherever a secret key must be expanded, for example in key scheduling within block ciphers.
- Finally, they can be used to send message with unconditional security when the number of bits obtained by the enemy is smaller compared with the number of bits of the secret key, as explained in this paper.

12 Conclusion

In this paper, we have seen how to design pseudorandom functions and pseudorandom permutations with a security bound near the optimal information-theoretic security bound, and with various density of the secret keys.
For pseudorandom functions, we have shown some solutions for all the possible densities of the secret keys. For pseudorandom permutations, we have also shown some solutions for all possible densities of the secret keys when the number of bits of the keys is about $\leq$ the square root of the total number of possible inputs. If the densities of secret keys is such that the number of bits of the keys is larger than the square root of the total number of possible inputs (i.e. "densification of the keys" for pseudorandom permutations) then, to obtain similar result, we suggest to study unbalanced Feistel schemes with rounds of functions of α bits $\rightarrow \beta$ bits, with $\alpha > \beta$ (the analysis of such schemes beyong the birthday bound is still an open problem. The security up to the birthday bound was proved in [11]).
We can notice that all our constructions use only the original Benes and Feistel constructions with very simple changes: concatenations, composition of functions and fixing at 0 some values. Our schemes are very fast to compute (the complexity is linear in the number of bits of the messages to be sent) when the keys have been generated and stored. The schemes are also very flexible, since Alice can send as many messages as wanted one day, and these messages will be decrypted by Bob with the keys, and then Alice can send some other messages the other days (with the same keys), with still the same global security property (i.e. we have security against Charlie with unbounded computing power if the number of bits of information obtained by the enemy in an adaptive attack is very small compared with the number of bits of the key). These schemes can also be seen as generalizations of the one-time-pad but here instead of a number of bits of key equal to the number of bits of the message sent, we need a number of bits of key about equal with the number of bits of information obtained by the enemy in an adaptive attack.

References

[1] W. Aiello and R. Venkatesan, *Foiling Birthday Attacks in Lenght-Doubling Transformations - Benes: a non-reversible alternative to Feistel*, LNCS n1070, Eurocrypt '96, pp. 307–320, Springer.

[2] M. Blum and S. Micali, *How to Generate Cryptographically Strong Sequences of Pseudorandom Bits*, *SIAM Journal on Computing*, 13, pp 850–864, 1994.

[3] I. Damgård, *Design Principles of Hash Functions. Crypto '89*, LNCS n435, pp. 416–427, Springer-Verlag.

[4] O. Goldreich, S. Goldwasser and S. Micali, *How to Construct Random Functions*, *JACM*, 33, pp 792–807, 1986.

[5] M. Luby. *Pseudorandomness and Its Cryptographic Applications, Princeton Computer Science Notes, Princeton University Press.*

[6] M. Luby and C. Rackoff. *How to construct pseudorandom permutations from pseudorandom functions. SIAM Journal on Computing*, vol. 17, n2, pp. 373–386, April 1988.

[7] U. Maurer. *A simplified and Generalized Treatment of Luby-Rackoff Pseudorandom Permutation Generators. Eurocrypt '92* (Lecture Notes in Computer Science n658), pp. 239–255, Springer-Verlag.

[8] U. Maurer. *Indistinguishability of Random Systems. Eurocrypt '02* (Lecture Notes in Computer Science n2332), pp. 110–132, Springer.

[9] U. Maurer and J. Massey, *Perfect local randomness in pseudo-random sequences. Crypto '89*, LNCS n435, pp. 100–112, Springer-Verlag.

[10] U. Maurer and K. Pietrzak. *The security of Many-Round Luby-Rackoff Pseudo-Random Permutations. Eurocrypt '03*, LNCS n2656, pp. 544–561, Springer.

[11] M. Naor and O. Reingold, *On the construction of pseudo-random permutations: Luby-Rackoff revisited*, Journal of Cryptology, vol. 12, 1999, pp. 29–66. Extended abstract was published in Proc. 29th Ann. ACM Symp. on Theory of Computing, 1997, pp. 189–199.

[12] J. Patarin, *Pseudorandom permutations based on the DES scheme. Eurocode '90* (Lecture Notes in Computer Science 514), pp. 193–204, Springer.

[13] J. Patarin, *New Results on Pseudorandom Permutation generators based on the DES Scheme,Crypto '91*, LNCS n576, pp 301–312, Springer-Verlag.

[14] J. Patarin, *Improved security bounds for pseudorandom permutations*, 4th ACM Conference on Computer and Communications Security, April 1-4, 1997, Zurich, ACM Press, pp. 142–150.

[15] J. Patarin, *Security of Random Feistel Scemes with 5 or more rounds. Crypto '04*(Lecture Notes in Computer Science 3152), pp. 106–122, Springer.

[16] J. Patarin and A. Montreuil, *Benes and Butterfly schemes revisited*, paper avaiable on ePrint 2005/004.