

# A Proof of security in $O(2^n)$ for the Benes Scheme

Jacques Patarin

Université de Versailles  
45 avenue des Etats-Unis, 78035 Versailles Cedex, France  
jacques.patarin@prism.uvsq.fr

**Abstract.** In [1], W. Aiello and R. Venkatesan have shown how to construct pseudorandom functions of  $2n$  bits  $\rightarrow 2n$  bits from pseudorandom functions of  $n$  bits  $\rightarrow n$  bits. They claimed that their construction, called “Benes” reaches the optimal bound ( $m \ll 2^n$ ) of security against adversaries with unlimited computing power but limited by  $m$  queries in an Adaptive Chosen Plaintext Attack (CPA-2). This result may have many applications in Cryptography (cf [1, 19, 18] for example). However, as pointed out in [18] a complete proof of this result is not given in [1] since one of the assertions in [1] is wrong. It is not easy to fix the proof and in [18], only a weaker result was proved, i.e. that in the Benes Schemes we have security when  $m \ll f(\epsilon) \cdot 2^{n-\epsilon}$ , where  $f$  is a function such that  $\lim_{\epsilon \rightarrow 0} f(\epsilon) = +\infty$  ( $f$  depends only of  $\epsilon$ , not of  $n$ ). Nevertheless, no attack better than in  $O(2^n)$  was found. In this paper we will in fact present a complete proof of security when  $m \ll O(2^n)$  for the Benes Scheme, with an explicit  $O$  function. Therefore it is possible to improve all the security bounds on the cryptographic constructions based on Benes (such as in [19]) by using our  $O(2^n)$  instead of  $f(\epsilon) \cdot 2^{n-\epsilon}$  of [18].

*Key words:* Pseudorandom function, unconditional security, information-theoretic primitive, design of keyed hash functions, security above the birthday bound.

## 1 Introduction

In this paper we will study again the “Benes” Schemes of [1] and [18]. (The definition of the “Benes” Schemes will be given in Section 2). More precisely, the aim of this paper is to present a complete proof of security for the Benes schemes when  $m \ll O(2^n)$  where  $m$  denotes the number of queries in an Adaptive Chosen Plaintext Attack (CPA-2) with an explicit  $O$  function. With this security result we will obtain a proof for the result claimed in [1] and this will also solve an open problem of [18], since in [18] only a weaker result was proved (security when  $m \ll f(\epsilon) \cdot 2^{n-\epsilon}$  where  $f$  is a function such that  $\lim_{\epsilon \rightarrow 0} f(\epsilon) = +\infty$ ). It is important to get precise security results for these schemes, since they may have many applications in Cryptography, for example in order to design keyed hash functions (cf [1]) or in order to design Information-theoretic schemes (cf [18]).

Here we will prove security “above the birthday bound”, i.e. here we will prove security when  $m \ll 2^n$  instead of the “birthday bound”  $m \ll \sqrt{2^n}$  where  $m$  denotes the number of queries in an Adaptive Chosen Plaintext Attack (CPA-2).  $\sqrt{2^n}$  is called the ‘birthday bound’ since when  $m \ll \sqrt{2^n}$ , if we have  $m$  random strings of  $n$  bits, the probability that two strings are equal is negligible.  $2^n$  is sometimes called the ‘Information bound’ since security when  $m \ll 2^n$  is the best possible security against an adversary that can have access to infinite computing power. In fact, in [18], it is shown that Benes schemes can be broken with  $m = O(2^n)$  and with  $O(2^n)$  computations. Therefore security when  $m \ll O(2^n)$  is really the best security result that we can have with Benes schemes.

In [2], Bellare, Goldreich and Krawczyk present a similar construction that provides length-doubling for the input. However their construction is secure only against random queries and not against adaptively chosen queries. Benes schemes, in contrast, produce pseudorandom functions secure against adaptively chosen queries.

It is interesting to notice that there are many similarities between this problem and the security of Feistel schemes built with random round functions (also called Luby-Rackoff constructions), or the security of the Xor of two random permutations (in order to build a pseudorandom function from two pseudorandom permutations). The security of random Feistel schemes above the birthday bound has been studied for example in [13], [15], [17], and the security of the Xor of two random permutations above the birthday bound has been studied for example in [3], [8]. However the analysis of the security of the Benes schemes requires a specific analysis and the proof strategy used for Benes schemes is significantly different than for Feistel or the Xor of random permutations. In fact, our proof of security for Benes schemes in  $m \ll O(2^n)$  is more simple than the proofs of security in  $m \ll O(2^n)$  for Feistel schemes or the Xor of random permutations, since we will be able, as we will see, to use a special property of Benes schemes.

## 2 Notation

We will use the same notation as in [18].

- $I_n = \{0, 1\}^n$  is the set of the  $2^n$  binary strings of length  $n$ .
- $F_n$  is the set of all functions  $f : I_n \rightarrow I_n$ . 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  $b$ .
- Given four functions from  $n$  bits to  $n$  bits,  $f_1, \dots, f_4$ , 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:

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

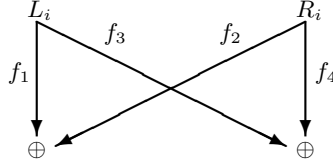


Fig. 1. Butterfly transformation

- Given eight functions from  $n$  bits to  $n$  bits,  $f_1, \dots, f_8$ , we use them to define the **Benes transformation** (see [1]) (back-to-back Butterfly) as the composition of two Butterfly transformations. On input  $[L_i, R_i]$ , the output is given by  $[S_i, T_i]$ , with:

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

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

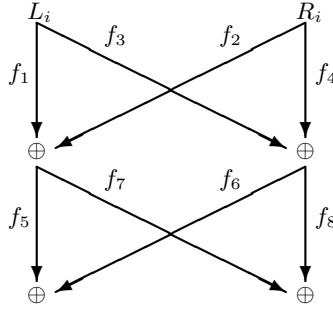


Fig. 2. Benes transformation (back-to-back Butterfly)

### 3 A problem in the Proof of [1]

As showed in [18], there is a problem in the security proof of [1]. Let us recall what the problem is.

**Definition 1** We will say that we have “a circle in  $X, Y$  of length  $k$ ” if we have  $k$  pairwise distinct indices such that  $X_{i_1} = X_{i_2}, Y_{i_2} = Y_{i_3}, X_{i_3} = X_{i_4}, \dots, X_{i_{k-1}} = X_{i_k}, Y_{i_k} = Y_{i_1}$ . 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  $k$ .

Let  $[L_1, R_1], [L_2, R_2], [L_3, R_3]$  and  $[L_4, R_4]$  be four chosen inputs such that  $L_1 = L_2, R_2 = R_3, L_3 = L_4$  and  $R_4 = R_1$  (and  $R_1 \neq R_2$  and  $L_1 \neq L_3$ ). (Here we will say that we have “a circle in  $L, R$ ” of length 4). Let  $p$  be the probability for these inputs to produce “a circle in  $X, Y$ ” (or, in the language of [1], an “alternating cycle”) after a Butterfly. In [1], page 318, it is claimed that “the probability that the top Butterfly produces an alternating cycle of length  $2j$  is  $\leq 2^{-2jn}$ ”. So here this means  $p \leq \frac{1}{2^{4n}}$ . However we will see that  $p \geq \frac{1}{2^{2n}}$ . We have:

$$\begin{aligned}
X_1 &= f_1(L_1) \oplus f_2(R_1) & Y_1 &= f_3(L_1) \oplus f_4(R_1) \\
X_2 &= f_1(L_2) \oplus f_2(R_2) = f_1(L_1) \oplus f_2(R_2) & Y_2 &= f_3(L_2) \oplus f_4(R_2) = f_3(L_1) \oplus f_4(R_2) \\
X_3 &= f_1(L_3) \oplus f_2(R_3) = f_1(L_3) \oplus f_2(R_2) & Y_3 &= f_3(L_3) \oplus f_4(R_3) = f_3(L_3) \oplus f_4(R_2) \\
X_4 &= f_1(L_4) \oplus f_2(R_4) = f_1(L_3) \oplus f_2(R_1) & Y_4 &= f_3(L_4) \oplus f_4(R_4) = f_3(L_3) \oplus f_4(R_1)
\end{aligned}$$

**First possible circle in  $X, Y$**  We will get the circle  $X_1 = X_2, Y_2 = Y_3, X_3 = X_4$  and  $Y_4 = Y_1$  if and only if  $f_2(R_1) = f_2(R_2)$  and  $f_3(L_1) = f_3(L_3)$  and the probability for this is exactly  $\frac{1}{2^{2n}}$  (since  $R_1 \neq R_2$  and  $L_1 \neq L_3$ ).

**Conclusion** The probability  $p$  to have a circle in  $X, Y$  of length 4 (i.e. the probability that the top Butterfly produces an alternating cycle of length 4 in the language of [1]) is  $\geq \frac{1}{2^{2n}}$ , so it is not  $\leq \frac{1}{2^{4n}}$  as claimed in [1].

As we will see in this paper, this problem is not easily solved: a precise analysis will be needed in order to prove the security result  $m \ll 2^n$ .

## 4 “Lines” and “Circles” in $X, Y$

“Circles” in  $X, Y$  have been defined in Section 3. Similarly, (as in [18] p.104) we can define “Lines” in  $X, Y$  like this:

**Definition 2** *If  $k$  is odd, we will say that we have “a line in  $X, Y$  of length  $k$ ” if we have  $k + 1$  pairwise distinct indices such that  $X_{i_1} = X_{i_2}, Y_{i_2} = Y_{i_3}, X_{i_3} = X_{i_4}, \dots, Y_{i_{k-1}} = Y_{i_k}, X_{i_k} = X_{i_{k+1}}$ . Similarly, if  $k$  is even, we will say that we have “a line in  $X, Y$  of length  $k$ ” if we have  $k + 1$  pairwise distinct indices such that  $X_{i_1} = X_{i_2}, Y_{i_2} = Y_{i_3}, X_{i_3} = X_{i_4}, \dots, X_{i_{k-1}} = X_{i_k}, Y_{i_k} = Y_{i_{k+1}}$ . So in a line in  $X, Y$  we have  $k + 1$  indices, and  $k$  equations, in  $X$  or in  $Y$ , and these equations can be written “in a line” from the indices.*

**Remark:** with this definition, a “line in  $X, Y$ ” always starts with a first equation in  $X$ . This will not be a limitation in our proofs. Of course we could also have defined lines in  $X, Y$  by accepting the first equation to be in  $X$  or in  $Y$  and then to alternate  $X$  and  $Y$  equations.

To get our security results, as for [1] and [18], we will start from this theorem:

**Theorem 1** *The probability to distinguish Benes schemes, when  $f_1, \dots, f_8$  are randomly chosen in  $F_n$ , from random functions of  $2n$  bits  $\rightarrow 2n$  bits in CPA-2 is always less than or equal to  $p$ , where  $p$  is the probability to have a circle in  $X, Y$ .*

*Proof of theorem 1*

A proof of Theorem 1 can be found in [1] written in the language of “alternating cycles”, or in [18] p.97, written with exactly these notations of “circles”. In fact, this result can easily be proved like this:

With Benes, we have:

$$\begin{aligned}
\forall i, 1 \leq i \leq m, \text{Benes}(f_1, \dots, f_8)[L_i, R_i] = [S_i, T_i] &\Leftrightarrow \\
\begin{cases} S_i = f_5(X_i) \oplus f_6(Y_i) \\ T_i = f_7(X_i) \oplus f_8(Y_i) \end{cases} & \quad (1)
\end{aligned}$$

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

When there are no circles in  $X, Y$  in each equation (1), we have a new variable  $f_5(X_i)$  or  $f_6(Y_i)$ , and a new variable  $f_7(X_i)$  or  $f_8(Y_i)$ , so if  $f_5, f_6, f_7, f_8$  are random functions, the outputs  $S_i$  and  $T_i$  are perfectly random and independent from the previous  $S_j, T_j, i < j$ .

In this paper we will now evaluate  $p$  in a new way, in order to get stronger security result. For this we will introduce and study the properties of “first dependency lines”.

## 5 First dependencies

**Definition 3** A line in  $X, Y$  of length  $k$  will be called a “first dependency” line when all the equations in  $X, Y$  except the last one are independent and when the last one (i.e. the equation number  $k$ ) is a consequence of the previous equations in  $X, Y$ .

**Example:** If  $L_1 = L_3, L_2 = L_4, R_1 = R_2, R_3 = R_4$ , then  $(X_1 = X_2), (Y_2 = Y_3), (X_3 = X_4)$  is a “first dependency line”, since  $(X_1 = X_2)$  and  $(Y_2 = Y_3)$  are independent, but  $(X_3 = X_4)$  is a consequence of  $(X_1 = X_2)$ .

**Definition 4** A circle in  $X, Y$  will be called a “circle with one dependency” when all the equations in the circle, except one are independent from the others, and when exactly one is a consequence of the others equations in  $X, Y$ .

The key argument in our proof will be this (new) Theorem:

**Theorem 2** When  $f_1, f_2, f_3, f_4$  are randomly chosen in  $F_n$ , the probability  $q_k$  to have a “first dependency line” in  $X, Y$  of length  $k$  satisfies  $q_k \leq k^5 \frac{m^{k-1}}{2^{(k-1)n}}$

**Remark.** Some possible improvements of this Theorem 2 (with a better coefficient than  $k^5$ ) will be given in Section 7. However this version with a coefficient  $k^5$  will be enough for us, in order to get a security for Benes in  $O(2^n)$  as we will see in Section 6.

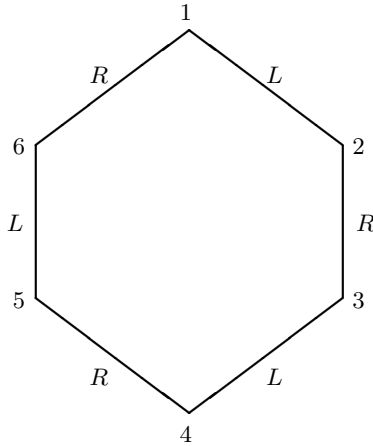
*Proof of theorem 2*

### a) Rough Evaluation

Since we have  $(k - 1)$  independent equations in  $X$  or  $Y$ , when all the indices are fixed the probability to have all these equations is  $\frac{1}{2^{(k-1)n}}$ . Now, in order to choose the  $k + 1$  indices of the messages, we have less than  $m^{k+1}$  possibilities. Therefore,  $q_k \leq \frac{m^{k+1}}{2^{(k-1)n}}$ . Moreover, the last equation (in  $X$  or  $Y$ ) is a consequence of the previous equations in  $X, Y$ . However, a dependency in these equations implies the existence of a circle in  $L, R$  on a subset of the indices involved in the

dependency. [The proof is exactly the same as for Theorem 1 except that here we use  $L, R$  instead of  $X, Y$  and  $X, Y$  instead of  $S, T$ ].

Now if we have a circle in  $L, R$  of length  $\alpha$ ,  $\alpha$  even, we know that  $\frac{\alpha}{2}$  of the messages in the circle come from the other  $\frac{\alpha}{2}$  messages.

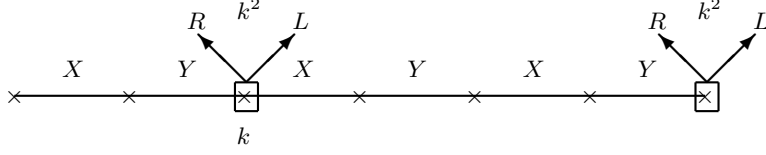


**Fig. 3.** Example of circle in  $L, R$

For example, if  $L_1 = L_2, R_2 = R_3, L_3 = L_4, R_4 = R_5, L_5 = L_6, R_6 = R_1$ , we have a circle in  $L, R$  of length 6, and if we know the messages 1, 3, 5, then we know  $(L_1, R_1), (L_3, R_3), (L_5, R_5)$ , and we can deduce  $(L_2, R_2), (L_4, R_4)$  and  $(L_6, R_6)$ , since  $(L_2, R_2) = (L_1, R_3), (L_4, R_4) = (L_3, R_5)$  and  $(L_6, R_6) = (L_5, R_1)$ . In a circle in  $L, R$  of length  $\alpha$ , we must have  $\alpha \geq 4$ , since  $\alpha = 2$  gives  $L_i = L_j$  and  $R_i = R_j$ , and therefore  $i = j$ . Therefore, if there is a circle in  $L, R$  we will be able to find  $\frac{\alpha}{2}$  messages,  $\frac{\alpha}{2} \geq 2$ , from the other messages of the circle. So, in order to choose  $k + 1$  indices of the messages in a first dependency line, we will have  $O(m^{k-1})$  possibilities (instead of  $m^{k+1}$  possibilities since at least 2 messages will be fixed from the others), and therefore  $q^k \leq \frac{O(m^{(k-1)})}{2^{(k-1)n}}$ . We will now evaluate the term  $O(m^{(k-1)})$  more precisely.

**b) More precise evaluation**

From a first dependency line in  $X, Y$  we have just seen that at least two messages of the line, let say messages  $[L_a, R_a]$  and  $[L_b, R_b]$  are such that  $L_a = L_i, R_a = R_j, L_b = R_k, R_b = R_l$  with  $i, j, k, l \notin \{a, b\}$ . Moreover, we can choose  $b$  to be the last message of the line (since between the two last messages we have a dependency in  $X$  or in  $Y$  from the other equations in  $X$  and  $Y$ ). Now for  $a$  we have less than  $k$  possibilities, and for  $i, j, k, l$  we have less than  $(k - 1)^4$  possibilities. Therefore, for the choice of the  $k + 1$  messages of the line we have



**Fig. 4.** An example of line in  $X, Y$

less than  $k(k-1)^4 m^{k-1}$  possibilities, which is less than  $k^5 m^{k-1}$ . Therefore,  $q_k \leq k^5 \frac{m^{k-1}}{2^{(k-1)n}}$  as claimed.

**Remark.** We can not always choose  $a$  and  $b$  to be the last two messages, because it is possible that we have an equality in  $L$ , or in  $R$ , between these two last messages. However, we can always choose  $b$  to be the last message, as we did here.

**Theorem 3** When  $f_1, f_2, f_3, f_4$  are randomly chosen in  $F_n$ , the probability  $q_k$  to have a “first dependency line” in  $X, Y$  of length  $k$ , or a “circle with one dependency” of length  $k-1$  ( $k$  odd) satisfies:  $q_k \leq k^5 \frac{m^{k-1}}{2^{(k-1)n}}$ .

*Proof of theorem 3*

This is just a simple extension of Theorem 2. A circle of length  $k-1$  with one dependency can be seen as a special line of length  $k$  with the first index equal to the index number  $k$ , and the proof given for Theorem 2 extended to the classical lines in  $X, Y$  and to these special lines gives immediately Theorem 3.

## 6 Security of the Benes schemes

**Theorem 4** When  $f_1, f_2, f_3, f_4$  are randomly chosen in  $F_n$ , the probability  $p$  to have a circle in  $X, Y$  satisfies, if  $m \leq \frac{2^n}{2}$

$$p \leq \frac{m^2}{2^{2n}} \left( \frac{1}{1 - \frac{m^2}{2^{2n}}} \right) + \frac{m^2}{2^{2n}} \left( \sum_{k=3}^{+\infty} \frac{k^5}{2^{(k-3)n}} \right)$$

$$\text{and } \sum_{k=3}^{+\infty} \frac{k^5}{2^{(k-3)n}} = 3^5 + \frac{4^5}{2} + \frac{5^5}{2^2} + \frac{6^5}{2^3} + \dots \text{ converges to a finite value.}$$

Therefore, when  $m \ll 2^n$ ,  $p \simeq 0$ , as wanted.

*Proof of theorem 4*

For each circle in  $X, Y$  of length  $k$ ,  $k$  even, we have three possibilities:

a) Either all the  $k$  equations in  $X, Y$  are independent. Then the probability to have a circle is less than or equal to  $\frac{m^k}{2^{kn}}$ .

b) Or there exists a first dependency line of length strictly less than  $k$  in the equations in  $X, Y$  of the circle.

c) Or the circle is a circle with exactly one dependency.

Now from Theorems 2 and 3, we get immediately:

$$p \leq \left( \frac{m^2}{2^{2n}} + \frac{m^4}{2^{4n}} + \frac{m^6}{2^{6n}} + \frac{m^8}{2^{8n}} + \dots \right) + \sum_{k=3}^{+\infty} \frac{k^5 m^{k-1}}{2^{(k-1)n}}$$

Therefore, if  $m \leq \frac{2^n}{2}$ ,

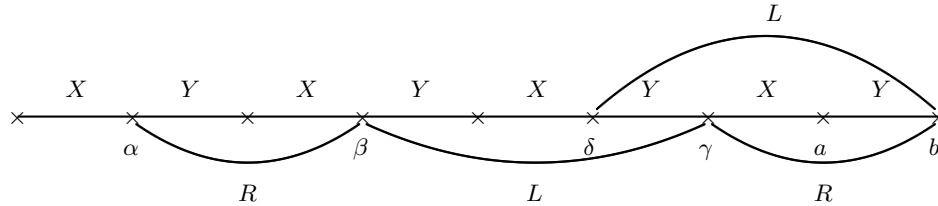
$$p \leq \frac{m^2}{2^{2n}} \left( \frac{1}{1 - \frac{m^2}{2^{2n}}} \right) + \frac{m^2}{2^{2n}} \left( \sum_{k=3}^{+\infty} \frac{k^5}{2^{(k-3)n}} \right)$$

as claimed (since  $\frac{m^{k-3}}{2^{(k-3)n}} \leq \frac{1}{2^{(k-3)n}}$ ). Therefore, from Theorem 1, we see that we have proved the security of Benes when  $m \ll O(2^n)$  against all CPA-2, with an explicit  $O$  function, as wanted.

## 7 Improving the $k^5$ coefficient

By working a little more it is possible, as we will see now, to improve the  $k^5$  coefficient in Theorem 2. First, we will see that it is possible to choose  $k^4$  instead of  $k^5$ .

**Theorem 5** *When  $f_1, f_2, f_3, f_4$  are randomly chosen in  $F_n$ , the probability  $q_k$  to have a “first dependency line” in  $X, Y$  of length  $k$  satisfies  $q_k \leq k^4 \frac{m^{k-1}}{2^{(k-1)n}}$ .*



**Fig. 5.** Illustration of the proof in  $k^4$  instead of  $k^5$ .

*Proof of theorem 5*

We will still denote by  $a$  and  $b$  the indices of the last equations (in  $X$  or in  $Y$  and dependent from the other equations). We can proceed like this:

a) We choose 4 indices  $\alpha, \beta, \gamma, \delta \notin \{b\}$  in the line  $X, Y$ . We have here less than  $k^4$  possibilities to choose  $\alpha, \beta, \gamma, \delta$ .



b) We choose all the  $k - 1$  messages of indices  $\notin \{a, b\}$  in the line of length  $k$ . We have here less than  $m^{k-1}$  possibilities.

c) The messages of indices  $\beta$  and  $b$  will be fixed from the previous values from these equations:  $R_\beta = R_\alpha, L_\beta = L_\gamma, R_b = R_\gamma, L_b = L_\delta$ .

Therefore we have less than  $k^4 m^{k-1}$  possibilities for the choice of the  $k + 1$  messages in the first dependency line, so  $q_k \leq k^4 \frac{m^{k-1}}{2^{(k-1)n}}$  as claimed.

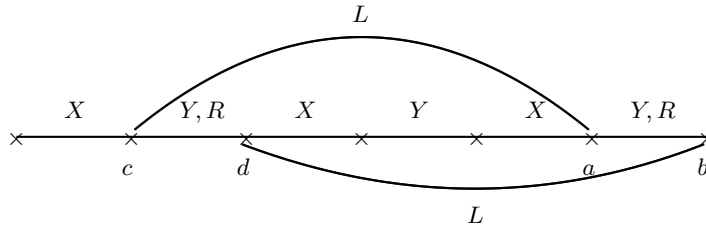
As we will see now, we can get further improvements on the coefficient  $k^4$  by looking at the type of circle in  $L, R$  that contains  $a$  and  $b$ .

**Theorem 6** *With the same notation as in Theorem 5, we have:  $q_k \leq \frac{1}{2^{(k-1)n}} (3km^{k-1} + k^6 m^{k-2})$ .*

*Proof of theorem 6*

We know that the last equation of the line ( $X_a = X_b$  or  $Y_a = Y_b$ ) is a consequence of the previous equations in  $X$  or  $Y$ . We also know that such a consequence is only possible if there is a circle in  $L, R$  that includes the two last points  $a$  and  $b$ . In a circle in  $L, R$  of length  $\alpha$ ,  $\alpha$  even, we have seen that  $\alpha \geq 4$  and that  $\frac{\alpha}{2}$  points can be fixed from the others. We will consider two cases:  $\alpha = 4$  and  $\alpha \geq 6$ .

**Case 1:**  $\alpha = 4$ . In this case, the circle in  $L, R$  is between  $a, b$  and two other points  $c, d$  such that the equation (in  $X$  or  $Y$ ) in  $a, b$  is a consequence of the equation in  $c, d$  (in  $X$  or  $Y$ ). Therefore, for  $\{c, d\}$  we have at most  $\frac{k}{2}$  possibilities (cf figure 6). Now when  $\{a, b, c, d\}$  are fixed, for the circle in  $L, R$  we have at most  $3 \times 2$  possibilities ( $R_a = R_b, R_c$  or  $R_d$  and when this equation in  $R$  is fixed, we have two possibilities for the equation in  $L$ ). Therefore, we have at most  $\frac{k}{2} \times 3 \times 2 \times m^{k-1}$  possibilities for a first dependency line in  $X, Y$  in this case 1.



**Fig. 6.** Example of dependency generated by a circle of length 4 in  $L, R$

**Case 2:**  $\alpha \geq 6$ . In this case, at least 2 indices can be fixed from the others, and by using exactly the same arguments as in the proof of Theorem 5 above, with two more points, we see immediately that we have at most  $k^6 \cdot m^{k-2}$  possibilities for a first dependency line in  $X, Y$  in this case 2. By combining case 1 and case 2, we get immediately Theorem 6.

**Theorem 7** *With the same notation as in Theorem 5, we have:*

$$q_k \leq \frac{1}{2^{(k-1)n}} (3km^{k-1} + 30k^2m^{k-2} + k^8m^{k-3})$$

*Proof of theorem 7*

The proof is exactly the same as above: the term in  $3km^{k-1}$  comes from circles in  $L, R$  of length 4, the term in  $30k^2m^{k-2}$  (i.e.  $5! \cdot \frac{k}{2} \cdot \frac{k}{2} \cdot m^{k-2}$ ) comes from circles in  $L, R$  of length 6, and the term in  $k^8m^{k-3}$  from circles in  $L, R$  of length greater than or equal to 8.

**Theorem 8** *With the same notation as in Theorem 5, we have: for all integer  $\mu$ :*

$$q_k \leq \frac{1}{2^{(k-1)n}} \left( 3km^{k-1} + 5! \left(\frac{k}{2}\right)^2 m^{k-2} + 7! \left(\frac{k}{2}\right)^3 m^{k-3} + 9! \left(\frac{k}{2}\right)^4 m^{k-4} + \dots \right. \\ \left. + (2\mu + 1)! \left(\frac{k}{2}\right)^\mu m^{k-\mu} + k^{2\mu+4} m^{k-\mu-1} \right)$$

*Alternatively, we also have:*

$$q_k \leq \frac{1}{2^{(k-1)n}} \left( \sum_{\mu=1}^{+\infty} (2\mu + 1)! \left(\frac{k}{2}\right)^\mu m^{k-\mu} \right)$$

*Proof of theorem 8*

The proof is exactly the same as above. The term  $(2\mu + 1)! \left(\frac{k}{2}\right)^\mu m^{k-\mu}$  comes from the circles in  $L, R$  of length  $2\mu + 2$ , and the term  $k^{2\mu+4} m^{k-\mu-1}$  from the circles in  $L, R$  of length greater than or equal to  $2\mu + 4$ , such that these circles in  $L, R$  generate the dependency  $X_a = X_b$  (or  $Y_a = Y_b$ ) from the previous equations in  $X, Y$ .

#### **Application to the Benes schemes**

We can immediately apply these results to the Benes schemes, by using these improved results instead of Theorem 2. For example, from Theorem 6 and Theorem 1 we get:

**Theorem 9** *The probability  $p$  to distinguish Benes schemes from truly random functions of  $F_{2n}$  satisfies:*

$$p \leq \frac{m^2}{2^{2n}} \left( \frac{1}{1 - \frac{m^2}{2^{2n}}} \right) + \sum_{k=3}^{+\infty} \frac{3km^{k-1}}{2^{(k-1)n}} + \sum_{k=5}^{+\infty} \frac{k^6 m^{k-2}}{2^{(k-1)n}}$$

and therefore if  $m \leq \frac{2^n}{2}$  we get:

$$p \leq \frac{m^2}{2^{2n}} \left( \frac{1}{1 - \frac{m^2}{2^{2n}}} \right) + \frac{m^2}{2^{2n}} \left( \sum_{k=3}^{+\infty} \frac{3k}{2^{(k-3)}} \right) + \frac{m^3}{2^{4n}} \left( \sum_{k=5}^{+\infty} \frac{k^6}{2^{(k-5)}} \right) \quad (2)$$

In (2), we have again obtained a proof of security for the Benes schemes against all CPA-2 when  $m \ll O(2^n)$ . Moreover the  $O$  function obtained here is slightly better compared with the  $O$  function obtained with Theorem 4.

## 8 Modified Benes, i.e. Benes with $f_2 = f_3 = \text{Id}$

If we take  $f_2 = f_3 = \text{Id}$  in the Benes schemes, we obtain a scheme called ‘‘Modified Benes’’ (see [1, ?]). Then we have:  $X_i = f_1(L_i) \oplus R_i$ ,  $Y_i = L_i \oplus f_4(R_i)$  and the output  $[S_i, T_i]$  is such that  $S_i = f_5(X_i) \oplus f_6(Y_i)$  and  $T_i = f_7(X_i) \oplus f_8(Y_i)$ . It is conjectured that the security for Modified Benes is also in  $O(2^n)$  but so far we just have a proof of security in  $O(2^{n-\epsilon})$  for all  $\epsilon > 0$  (see [18]). It is interesting to notice that the proof technique used in this paper for the regular Benes cannot be used for the Modified Benes, since, as we will see in the example below, for Modified Benes, unlike for regular Benes, the first ‘dependent’ equation can fix only one index instead of two. Example: If we have  $L_1 = L_3$ ,  $L_2 = L_4$ ,  $R_1 \oplus R_2 \oplus R_3 \oplus R_4 = 0$ , then we will get the ‘line’,  $X_1 = X_2$ ,  $Y_3 = Y_4$ ,  $X_3 = X_4$  from only two independent equations in  $f$ , ( $X_1 = X_2$  and  $Y_3 = Y_4$ ), and the first ‘dependent’ equation, here  $X_3 = X_4$ , fixes only the index 4 from the previous indices (since  $L_4 = L_2$  and  $R_4 = R_1 \oplus R_2 \oplus R_3$ ). Therefore, a proof of security in  $O(2^n)$  for the Modified Benes will be different, and probably more complex than our proof of security on  $O(2^n)$  for the regular Benes.

## 9 Conclusion

W. Aiello and R. Venkatesan did a wonderful work by pointing out the great potentialities of the Benes schemes and by giving some very important parts of a possible proof. Unfortunately the complete proof of security when  $m \ll 2^n$  for CPA-2 is more complex than what they published in [1] due to some possible attacks in L,R. However, in this paper we have been able to solve this open problem by improving the analysis and the results of [18]. The key point in our improved proof was to analyse more precisely what happens just after the first ‘dependent’ equations in X,Y (with the notation of Section 3), and to use the fact that in this case two ‘indices’ are fixed from the others. Therefore we have obtained the optimal security bound (in  $O(2^n)$ ) with an explicit  $O$  function. This automatically improves the proved security of many schemes based on Benes, for example the schemes of [19].

## References

1. W. Aiello and R. Venkatesan, *Foiling Birthday Attacks in Length-Doubling Transformations - Benes: a non-reversible alternative to Feistel*, Eurocrypt '96, LNCS 1070, pp. 307–320, Springer.
2. M. Bellare, O. Goldreich and H. Krawczyk *Stateless evaluation of pseudorandom functions: Security beyond the birthday barrier*, Crypto '99, LNCS 1666, Springer-Verlag.
3. M. Bellare and R. Impagliazzio *A Tool for Obtaining Tighter Security Analysis of Pseudorandom Based Constructions, with Applications to PRP to PRF Conversion*, Cryptology ePrint archive: 19995/024: Listing for 1999.
4. I. Damgård, *Design Principles of Hash Functions*, Crypto '89, Springer-Verlag.
5. O. Goldreich, S. Goldwasser and S. Micali, *How to Construct Random Functions*, *JACM*, 33, pp 792–807, 1986.
6. M. Luby. *Pseudorandomness and Its Cryptographic Applications*, *Princeton Computer Science Notes*, Princeton University Press.
7. M. Luby and C. Rackoff. *How to construct pseudorandom permutations from pseudorandom functions*, *SIAM Journal on Computing*, vol. 17, nb 2, pp. 373–386, April 1988.
8. S. Lucks. *The Sum of PRP Is a Secure PRF*, *Eurocrypt '00*, Lecture Notes in Computer Science 1807, pp. 470–487, Springer-Verlag.
9. U. Maurer. *A simplified and Generalized Treatment of Luby-Rackoff Pseudorandom Permutation Generators*, *Eurocrypt '92*, Lecture Notes in Computer Science 658, pp. 239–255, Springer-Verlag.
10. U. Maurer. *Information-Theoretic Cryptography*, *Crypto '99*, Lecture Notes in Computer Science 1666, pp. 47–64, Springer.
11. U. Maurer. *Indistinguishability of Random Systems*, *Eurocrypt '02*, Lecture Notes in Computer Science 2332, pp. 110–132, Springer.
12. U. Maurer and K. Pietrzak. *The security of Many-Round Luby-Rackoff Pseudo-Random Permutations*, *Eurocrypt '03*, pp. 544–561, Springer.
13. 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 ACM Symp. on Theory of Computing*, 1997, pp. 189–199.
14. J. Patarin, *New results on pseudo-random permutation generators based on the DES scheme*, *Crypto '91*, Lecture Notes in Computer Science 576, pp. 301–312, Springer-Verlag.
15. 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.
16. J. Patarin, *Luby-Rackoff: 7 rounds are Enough for  $2^{n(1-\epsilon)}$  Security*, *Crypto '03*, Lecture Notes in Computer Science 2729, pp. 513–529, Springer.
17. J. Patarin, *Security of Random Feistel Schemes with 5 or more rounds*, *Crypto '04*, Lecture Notes in Computer Science 3152, pp. 106–122, Springer.
18. J. Patarin and A. Montreuil, *Benes and Butterfly Schemes Revisited*, *ICISC '05*, Lecture Notes in Computer Science 3935, pp. 92–116, Springer-Verlag.
19. J. Patarin and P. Camion, *Design of near-optimal pseudorandom permutations in the information-theoretic model*, Cryptology ePrint archive: 2005/153: Listing for 2005.