The Security of Many-Round Luby-Rackoff Pseudo-Random Permutations

Ueli Maurer Krzysztof Pietrzak

Department of Computer Science
Swiss Federal Institute of Technology (ETH), Zurich
CH-8092 Zurich, Switzerland
{maurer,pietrzak}@inf.ethz.ch

Abstract. Luby and Rackoff showed how to construct a (super-)pseudo-random permutation $\{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$ from some number r of pseudo-random functions $\{0,1\}^n \rightarrow \{0,1\}^n$. Their construction, motivated by DES, consists of a cascade of r Feistel permutations. A Feistel permutation 1for a pseudo-random function f is defined as $(L, R) \rightarrow (R, L \oplus f(R))$, where L and R are the left and right part of the input and $\oplus$ denotes bitwise XOR or, in this paper, any other group operation on $\{0,1\}^n$. The only non-trivial step of the security proof consists of proving that the cascade of r Feistel permutations with independent uniform random functions $\{0,1\}^n \rightarrow \{0,1\}^n$, denoted Ψ_{2n}^r, is indistinguishable from a uniform random permutation $\{0,1\}^{2n} \rightarrow \{0,1\}^{2n}$ by any computationally unbounded adaptive distinguisher making at most $O(2^{cn})$ combined chosen plaintext/ciphertext queries for any $c < \alpha$, where α is a security parameter.

Luby and Rackoff proved $\alpha = 1/2$ for $r = 4$. A natural problem, proposed by Pieprzyk is to improve on α for larger r. The best known result, $\alpha = 3/4$ for $r = 6$, is due to Patarin. In this paper we prove $\alpha = 1 - O(1/r)$, i.e., the trivial upper bound $\alpha = 1$ can be approached. The proof uses some new techniques that can be of independent interest.

1 Introduction

The security of many cryptographic systems (e.g., block ciphers and message authentication codes) is based on the assumption that a certain component (e.g. DES or Rijndael) used in the construction is a pseudo-random function (PRF) [2]. Such systems are proven secure, relative to this assumption, by showing that any efficient algorithm for breaking the system can be transformed into an efficient distinguisher for the PRF from a uniform random function (URF).

1.1 Constructing Pseudorandom Permutations

There is a long line of research based on this paradigm, initiated in the seminal paper of Luby and Rackoff [5] who showed how to construct a pseudo-random permutation (PRP) from any PRF. That paper is not only of interest because it introduced the paradigm, but also because it proposed a very natural construction, motivated by DES, consisting of r Feistel permutations involving independent invocations of a PRF.

Usually, the only non-trivial step in a security proof of a construction based on PRF's is a purely probability-theoretic step, namely the analysis of the idealised construction when the PRF's are replaced by URF's, and proving that it is information-theoretically indistinguishable from, in our case, a uniform random permutation (URP), when the number of allowed queries is bounded (by a large, usually exponential bound).[1] The strength of the security proof depends on the size of this bound. Ideally, the number of allowed queries should be close to the trivial information-theoretic upper bound: If, for some sufficiently large number of queries, the expected entropy contained in the answers from the perfect system exceeds the entire internal randomness of the construction, then a (computationally unbounded) distinguisher trivially exists.

More concretely, the r-round Luby-Rackoff construction of a permutation $\{0,1\}^{2n} \to \{0,1\}^{2n}$ (hereafter denoted by Ψ_{2n}^r) consists of a cascade of r Feistel permutations involving independent URF's $\{0,1\}^n \to \{0,1\}^n$, where a Feistel permutation for a URF f is defined as $(L,R) \to (R, L \oplus f(R))$. Here L and R are the left and right part of the input and $\oplus$ denotes bitwise XOR or, in this paper, any other group operation on $\{0,1\}^n$.

Two versions of a PRP were considered in [5], namely when queries from only one side are allowed, and when queries from both sides are allowed. When the PRP is considered as a block cipher, these two variants correspond to chosen-plaintext and to combined chosen-plaintext/ciphertext attacks respectively . The latter variant was referred to in [5] as a super-PRP. In this paper we will only consider this stronger variant.

The problem we hence address is to prove that Ψ_{2n}^r is indistinguishable from a uniform random permutation $\{0,1\}^{2n} \to \{0,1\}^{2n}$ by any computationally unbounded adaptive distinguisher making a certain number of queries. To compare results, it makes sense to measure the number of queries on an logarithmic scale, i.e., by the number c when $O(2^{cn})$ queries are allowed. More precisely, one can state a result in terms of a constant α such that for all $c < \alpha$, indistinguishability holds for all sufficiently large n.

Luby and Rackoff proved that $\alpha = 1/2$ for $r = 4$. Since then several simplifications (e.g. [6],[8]) and generalisations of this result have appeared. Ramzan and Reyzin [13] proved that even if the adversary has black-box access to the middle two functions of Ψ_{2n}^4, the security is maintained. Naor and Reingold [8] showed that the security is maintained if one replaces the first and last round of Ψ_{2n}^4 with pairwise independent permutations, and even weaker constructions were proven secure in [11].

[1] Two systems **F** and **G** are indistinguishable by a distinguisher **D** making at most k queries if the following holds: The expected advantage of **D**, after making k queries to a black-box containing either **F** or **G** with equal probability, in guessing which system is in the black-box, is negligible.

A natural problem, proposed by Pieprzyk [12], is to improve on α for larger m. The best known result, $\alpha = 3/4$ for $r = 6$, is due to Patarin [10]. He also conjectured that better bounds hold. In this paper we address this problem and prove $\alpha = 1 - O(1/r)$, i.e., that the optimal upper bound[2] $\alpha = 1$ can be approached for increasing r.

The proof uses some new techniques that appear to be of independent interest. In many of the literature (e.g. [6],[7],[8]) on security proofs based on PRF's, one considers a (bad) event such that if the event does not occur, then the system behaves identically to the system it should be distinguished from, thus one can concentrate on the (simpler) problem of provoking the bad event. However, this approach has so far only been successful in analysing Ψ_{2n}^r for (small) constant r. Therefore, as a new technique, we extend the system Ψ_{2n}^r to a more sophisticated construction that offers new possibilities to define such events.

1.2 The Main Theorem

Our main theorem states that any computationally unlimited distinguisher, making at most k chosen plaintext/ciphertext queries has advantage at most $\frac{k^2}{2^{2n-2}} + \frac{k^{r+1}}{2^{r(n-3)-1}}$ in distinguishing Ψ_{2n}^{6r-1} from a uniform random permutation (URP).

As a corollary[3] we get that any distinguisher (as above), making at most $O(2^{cn})$ queries has exponentially small (in n) advantage in distinguishing Ψ_{2n}^r (here $r+1$ must be a multiple of 6) from a URP if $c < 1 - 6/(r+7)$.[4] This beats the best known bound $\alpha = 3/4$ for $r = 23$ where we get $\alpha = 4/5$.

1.3 Related Work

Different constructions of PRP's, so called unbalanced Feistel schemes, were investigated in ,[4],[8]. An unbalanced Feistel scheme over $\{0,1\}^n$, for $k \geq 2$ and $\ell = n/k$, is a generalisation of the original scheme, where the URF's in each round are unbalanced, i.e. $\{0,1\}^{(k-1)\ell} \rightarrow \{0,1\}^{\ell}$. For $k = 2$ one gets the original Feistel scheme. In [8] the security of those schemes (where the number of rounds is $k+2$) up to $2^{n(1-1/k)/2}$ queries is shown.[5] We approach the same bound (for

[2] This upper bound can be seen as follows: The internal randomness of Ψ_{2n}^r are the r function tables for the URF's, each containing $n2^n$ bits. The entropy of an output of a URP on $2n$ bits is $\log(2^{2n}) = 2n$ bits for the first, $\log(2^{2n} - 1)$ for the second output and so on, which is more that n for the first $2^{2n} - 2^n$ queries. So after $r2^n$ (this is in $O(2^n)$ for any fixed r) queries the entropy of the output of a URP is larger than the entropy in Ψ_{2n}^r, and thus a computationally unbounded distinguisher exists.

[3] We use that $\frac{k^2}{2^{2n-2}} + \frac{k^{r+1}}{2^{r(n-3)-1}} \in O\left(k^{r+1}/2^{rn}\right)$, and this is in $O(1)$ if k is in the order of $2^{n\left(1-\frac{1}{r+1}\right)}$.

[4] If 6 does not divide $r+1$, then this must be replaced by $c < 1 - 1/(\lfloor \frac{r+1}{6} \rfloor + 1)$.

[5] This bound is basically the birthday bound for the URF's used, and thus the square root of the information theoretic upper bound $O(2^{n(1-1/k)})$ for this construction.

a permutation over the same domain $\{0,1\}^n$) for the original Feistel scheme (i.e. fixed $k=2$) as the number of rounds is increased.

Another line of research concentrated on schemes where the number of different PRF's used in the construction is minimised (see [12],[9]). Knudsen [3] considered a different settings where the functions are not URF's but chosen from a family of size 2^k.

A generic way to strengthen the security of random permutations was given by Vaudenay [14] who showed that the advantage (for computationally unlimited distinguishers) in distinguishing a cascade of independent random permutations from a URP is basically only the product of the advantages of distinguishing each random permutation separately.

1.4 Organisation of the Paper

The proof of our main theorem is based on the framework of [7]; the required definitions and results from [7] are summarised in Section 2. In Section 3 we propose a new lemma which reduces the task of upper bounding the indistinguishability of a random permutation from a URP by adaptive combined chosen plaintext/ciphertext strategies, to the task of upper bounding the probability that a non-adaptive chosen plaintext only strategy succeeds in provoking some event. In Section 4 we first give a formal definition of the (many-round) Luby-Rackoff construction and then state the main theorem (Section 4.1). An outline of the proof is given in Section 4.2. The full proof is shown in Section 5. Section 6 summarises some conclusions.

2 Indistinguishability of Random Systems

This section summaries some definitions and results from [7], sometimes in a less general form. We also propose two simple new lemmas.

2.1 Notation

We denote sets by capital calligraphic letters (e.g. $\mathcal{X}$) and the corresponding capital letter X denotes a random variable taking values in $\mathcal{X}$. Concrete values for X are usually denoted by the corresponding small letter x. For a set $\mathcal{X}$ we denote by $\mathcal{X}^k$ the set of ordered k-tuples of elements from $\mathcal{X}$. $X^k=(X_1,X_2,\ldots,X_k)$ denotes a random variable taking values in $\mathcal{X}^k$ and a concrete value is usually denoted by $x^k=(x_1,x_2,\ldots,x_k)$.

Because we will consider different random experiments where the same random variables appear, we extend the standard notation for probabilities (e.g. $\mathsf{P}_V(v)$,($\mathsf{P}_{V|W}(v,w)$) by explicitly writing the random experiment $\mathcal{E}$ considered as a superscript, e.g. $\mathsf{P}_V^{\mathcal{E}}(v)$. Equivalence of distributions means equivalence on all inputs, i.e.

$$\mathsf{P}_V^{\mathcal{E}_1}=\mathsf{P}_V^{\mathcal{E}_2} \iff \forall v\in\mathcal{V}:\mathsf{P}_V^{\mathcal{E}_1}(v)=\mathsf{P}_V^{\mathcal{E}_2}(v)$$

If a denotes the event that a random variable A takes some specific value, say $a \iff A=1$, then we write $\mathsf{P}_a^{\mathcal{E}}$ to denote $\mathsf{P}_A^{\mathcal{E}}(1)$.

2.2 Random Automata and Random Systems

A central concept in the framework are systems, which take inputs (or queries) $X_1, X_2, \ldots \in \mathcal{X}$ and generate, for each new input X_i, an output $Y_i \in \mathcal{Y}$. Such a system can be deterministic or probabilistic, and it can be stateless or contain internal memory. A stateless deterministic system is simply a function $\mathcal{X} \to \mathcal{Y}$.

Definition 1 A *random function* $\mathcal{X} \to \mathcal{Y}$ (*random permutation* on $\mathcal{X}$) is a random variable which takes as values functions $\mathcal{X} \to \mathcal{Y}$ (permutations on $\mathcal{X}$). A deterministic system with state space Σ is called an $(\mathcal{X}, \mathcal{Y})$*-automaton* and is described by an infinite sequence $f_1, f_2, \ldots$ of functions, with $f_i : \mathcal{X} \times \Sigma \to \mathcal{Y} \times \Sigma$, where $(Y_i, S_i) = f_i(X_i, S_{i-1})$, S_i is the state at time i, and an initial state S_0 is fixed. An $(\mathcal{X}, \mathcal{Y})$*-random automaton* $\mathbf{F}$ is like an automaton but $f_i : \mathcal{X} \times \Sigma \times \mathcal{R} \to \mathcal{Y} \times \Sigma$ (where $\mathcal{R}$ is the space of the internal randomness), together with a probability distribution over $\mathcal{R} \times \Sigma$ specifying the internal randomness and the initial state.[6]

Definition 2 A *uniform random function (URF)* $\mathbf{R} : \mathcal{X} \to \mathcal{Y}$ (A *uniform random permutation (URP)* $\mathbf{P}$ on $\mathcal{X}$) is a random function with uniform distribution over all functions from $\mathcal{X}$ to $\mathcal{Y}$ (permutations on $\mathcal{X}$). Throughout, the symbols $\mathbf{P}$ and $\mathbf{R}$ are used for the systems defined above.

A large variety of constructions and definitions in the cryptographic literature can be interpreted as random functions, including pseudo-random functions. The more general concept of a (stateful) random system is considered because this is just as simple and because distinguishers can also be modelled as random systems.

The observable input-output behaviour of a random automaton $\mathbf{F}$ *is referred to as a random system.* In the following we use the terms random automaton and random system interchangeably when no confusion is possible.

Definition 3 An $(\mathcal{X}, \mathcal{Y})$*-random system* $\mathbf{F}$ is an infinite[7] sequence of conditional probability distributions $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}}$ for $i \geq 1$. Two random automata $\mathbf{F}$ and $\mathbf{G}$ are *equivalent*, denoted $\mathbf{F} \equiv \mathbf{G}$, if they correspond to the same random system, i.e., if $\mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}} = \mathsf{P}^{\mathbf{G}}_{Y_i|X^iY^{i-1}}$ for $i \geq 1$ or, equivalently, $\mathsf{P}^{\mathbf{F}}_{Y^i|X^i} = \mathsf{P}^{\mathbf{G}}_{Y^i|X^i}$ for $i \geq 1$.

2.3 Monotone Conditions for Random Systems

In the sequel it will be very useful to consider conditions defined for a random system $\mathbf{F}$. Loosely speaking, a random system $\mathbf{F}$ with a condition $\mathcal{A}$, denoted $\mathbf{F}^{\mathcal{A}}$, is the random system $\mathbf{F}$, but with an additional binary output $A_1, A_2, \ldots$, where $A_i = 1$ means that the condition holds after the ith query to $\mathbf{F}$. Throughout we will only consider *monotone* conditions, this is to say that if a condition fails to hold at some point, it never hold again.

[6] $\mathbf{F}$ can also be considered as a random variable taking on as values $(\mathcal{X}, \mathcal{Y})$-automata.
[7] Random systems with finite-length input sequences could also be defined.

Definition 4 Let $\mathbf{F}$ be any $(\mathcal{X},\mathcal{Y})$-random system. We use the notation $\mathbf{F}^{\mathcal{A}}$ to denote a $(\mathcal{X},\{0,1\}\times\mathcal{Y})$-random system ($\mathbf{F}^{\mathcal{A}}$ is defined the sequence of distributions $\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{A_iY_i|X^iY^{i-1}A^{i-1}}$ where $A_i \in \{0,1\}$) for which the following holds:

$$\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{Y_i|X^iY^{i-1}} = \mathsf{P}^{\mathbf{F}}_{Y_i|X^iY^{i-1}} \text{ for } i \geq 1$$

This simply says, that if we ignore the A_i's in $\mathbf{F}^{\mathcal{A}}$, we get a random system which is equivalent to $\mathbf{F}$. Moreover the A_i's are monotone, this is to say that $(A_i = 0) \Rightarrow (A_j = 0)$ for $j > i$ or equivalently $(A_i = 1) \Rightarrow (A_j = 1)$ for $j < i$. To save on notation we will denote an event $A_i = 1$ by a_i and $A_i = 0$ by $\overline{a}_i$. "A_i holds" means $A_i = 1$.

One way to define a random system $\mathbf{F}^{\mathcal{A}}$ is to explicitly give a distribution as in Definition 4. If $\mathbf{F}$ is given as a a description of a random automaton one can also define a "natural" condition directly on the random automaton.

As an example consider the random automaton Ψ^r_{2n} as described in the introduction. The evaluation of Ψ^r_{2n} requires the evaluation of the r internal URF's. Let ${}_jU_\ell$ denote the input to the URF in round j in the ℓth query (i.e. as Ψ^r_{2n} is queried with X_ℓ). Now we could for example define a condition $\mathcal{A}$ for Ψ^r_{2n} such that A_i holds if, for some fixed j, all the ${}_jU_\ell$ are distinct for $1 \leq \ell \leq i$. The proof of our main theorem will require a (more complicated) condition of this kind.

We will now define what we mean by equivalence of random systems with conditions.

Definition 5 $\mathbf{F}^{\mathcal{A}} \triangleq \mathbf{G}^{\mathcal{B}}$ means

$$\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{a_iY_i|X^ia_{i-1}Y^{i-1}} = \mathsf{P}^{\mathbf{G}^{\mathcal{B}}}_{b_iY_i|X^ib_{i-1}Y^{i-1}} \tag{1}$$

or, equivalently, $\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{a_iY^i|X^i} = \mathsf{P}^{\mathbf{G}^{\mathcal{B}}}_{b_iY^i|X^i}$ for all $i \geq 1$.

So $\mathbf{F}^{\mathcal{A}} \triangleq \mathbf{G}^{\mathcal{B}}$ if the systems are defined by the same distribution whenever the condition holds, i.e. for all x^i, y^i : $\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{A_iY^i|X^i}(1, y^i, x^i) = \mathsf{P}^{\mathbf{G}^{\mathcal{B}}}_{B_iY^i|X^i}(1, y^i, x^i)$. However, if the condition does not hold, we may have $\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{A_iY^i|X^i}(0, y^i, x^i) \neq \mathsf{P}^{\mathbf{G}^{\mathcal{B}}}_{B_iY^i|X^i}(0, y^i, x^i)$.

2.4 Distinguishers for Random Systems

We consider the problem of distinguishing two $(\mathcal{X},\mathcal{Y})$-random systems $\mathbf{F}$ and $\mathbf{G}$ by means of a computationally unbounded, possibly probabilistic adaptive distinguisher algorithm (or simply distinguisher) $\mathbf{D}$ asking at most k queries, for some k. The distinguisher generates X_1 as an input to $\mathbf{F}$ (or $\mathbf{G}$), receives the output Y_1, then generates X_2, receives Y_2, etc. Finally, after receiving Y_k, it outputs a binary decision bit. More formally:

Definition 6 A *distinguisher for $(\mathcal{X},\mathcal{Y})$-random systems* is a $(\mathcal{Y},\mathcal{X})$-random system $\mathbf{D}$ together with an initial value $X_1 \in \mathcal{X}$, which outputs a binary decision value E_k after some specified number k of queries to the system. By $\mathbf{D} \diamond \mathbf{F}$ we denote the random experiment where $\mathbf{D}$ is querying $\mathbf{F}$.

Throughout we will only consider distinguishers who never ask the same query twice. For bidirectional permutations (see Definition 11) we additionally require that distinguishers do not query with a Y_i (X_i) when they already received Y_i (X_i) on a query X_j (Y_j) for a $j < i$. Because we will only apply distinguishers to stateless systems (e.g. random permutations), this can be done without loss of generality. Such queries yield no information and thus there always is an an optimal distinguisher (see definition below) who never makes such queries.

Definition 7 The maximal advantage, of any adaptive distinguisher $\mathbf{D}$ issuing k queries, for distinguishing $\mathbf{F}$ and $\mathbf{G}$, is

$$\Delta_k(\mathbf{F},\mathbf{G}) \;:=\; \max_{\mathbf{D}} \left|\mathsf{P}^{\mathbf{D}\diamond\mathbf{F}}(E_k) - \mathsf{P}^{\mathbf{D}\diamond\mathbf{G}}(E_k)\right|.$$

A distinguisher that achieves the above maximum is called an *optimal distinguisher* for $\mathbf{F}$ and $\mathbf{G}$.

We now consider a distinguisher $\mathbf{D}$ (according to the view described above) which queries a system $\mathbf{F}^{\mathcal{A}}$ and whose aim it is to make $\mathcal{A}$ fail, making k queries to $\mathbf{F}^{\mathcal{A}}$ (i.e. to provoke the event $\overline{a}_k$).

Definition 8 For a random system $\mathbf{F}^{\mathcal{A}}$, let

$$\nu(\mathbf{F}^{\mathcal{A}},\overline{a}_k) \;:=\; \max_{\mathbf{D}} \mathsf{P}^{\mathbf{D}\diamond\mathbf{F}^{\mathcal{A}}}(\overline{a}_k)$$

be the maximal probability, for any adaptive distinguisher $\mathbf{D}$, of provoking $\overline{a}_k$ in $\mathbf{F}$. A distinguisher that achieves the above maximum is called an *optimal provoker* for $\overline{a}_k$ in $\mathbf{F}$. Moreover, let

$$\mu(\mathbf{F}^{\mathcal{A}},\overline{a}_k) \;:=\; \max_{x^k} \mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{\overline{a}_k|X^k}(x^k)$$

be the maximal probability of any non-adaptive distinguisher in provoking $\overline{a}_k$.

The following proposition (Theorem 1 from [7]) states that, for two random systems $\mathbf{F}$ and $\mathbf{G}$, we can upper bound $\Delta_k(\mathbf{F},\mathbf{G})$ by $\nu(\mathbf{F}^{\mathcal{A}},\overline{a}_k)$ or $\nu(\mathbf{G}^{\mathcal{B}},\overline{b}_k)$ if there is any $\mathbf{F}^{\mathcal{A}}$ and $\mathbf{G}^{\mathcal{B}}$ such that $\mathbf{F}^{\mathcal{A}} \mathrel{\hat{=}} \mathbf{G}^{\mathcal{B}}$.

Proposition 1 If $\mathbf{F}^{\mathcal{A}} \mathrel{\hat{=}} \mathbf{G}^{\mathcal{B}}$ then $\Delta_k(\mathbf{F},\mathbf{G}) \le \nu(\mathbf{F}^{\mathcal{A}},\overline{a}_k) = \nu(\mathbf{G}^{\mathcal{B}},\overline{b}_k)$. Moreover, $\mu(\mathbf{F}^{\mathcal{A}},\overline{a}_k) = \mu(\mathbf{G}^{\mathcal{B}},\overline{b}_k)$.

2.5 Some Useful Propositions and Definitions

Definition 9 The *cascade* of an $(\mathcal{X},\mathcal{Z})$-random permutation $\mathbf{E}$ and a $(\mathcal{Z},\mathcal{Y})$-random permutation $\mathbf{F}$, denoted $\mathbf{EF}$, is the $(\mathcal{X},\mathcal{Y})$-random permutation defined as applying $\mathbf{E}$ to the input sequence and $\mathbf{F}$ to the output of $\mathbf{E}$.

For $\mathbf{E}^{\mathcal{A}}$ and $\mathbf{F}^{\mathcal{B}}$ (where $\mathbf{E}$ and $\mathbf{F}$ are as above), the $(\mathcal{X},\mathcal{Y}\times\{0,1\}^2)$-random system $\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}}$ can be defined naturally.[8] $\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}} \mathrel{\hat{=}} \mathbf{G}^{\mathcal{C}}\mathbf{H}^{\mathcal{D}}$ means $\mathsf{P}^{\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}}}_{a_i b_i Y^i|X^i} = \mathsf{P}^{\mathbf{G}^{\mathcal{C}}\mathbf{H}^{\mathcal{D}}}_{c_i d_i Y^i|X^i}$, i.e., equivalence of the distributions if both conditions hold.

[8] Formally $\mathsf{P}^{\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}}}_{A_i B_i Y^i|X^i} = \sum_{\mathcal{Z}^i} \mathsf{P}^{\mathbf{E}^{\mathcal{A}}}_{A_i Z^i|X^i}\mathsf{P}^{\mathbf{F}^{\mathcal{B}}}_{B_i Y^i|Z^i}$.

Lemma 1 If $\mathbf{E}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{G}^{\mathcal{C}}$ and $\mathbf{F}^{\mathcal{B}} \stackrel{\circ}{=} \mathbf{H}^{\mathcal{D}}$ then $\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}} \stackrel{\circ}{=} \mathbf{G}^{\mathcal{C}}\mathbf{H}^{\mathcal{D}}$.

Proof. Let X and Z (Z and Y) denote the input and output of the first (second) permutation in the cascade. For any x^i and y^i we have

$$\mathsf{P}^{\mathbf{E}^{\mathcal{A}}\mathbf{F}^{\mathcal{B}}}_{a_i b_i Y^i|X^i}(y^i, x^i) = \sum_{z^i \in \mathcal{Z}^i} \mathsf{P}^{\mathbf{E}^{\mathcal{A}}}_{a_i Z^i|X^i}(z^i, x^i)\mathsf{P}^{\mathbf{F}^{\mathcal{B}}}_{b_i Y^i|Z^i}(y^i, z^i) =$$
$$\sum_{z^i \in \mathcal{Z}^i} \mathsf{P}^{\mathbf{G}^{\mathcal{C}}}_{c_i Z^i|X^i}(z^i, x^i)\mathsf{P}^{\mathbf{H}^{\mathcal{D}}}_{d_i Y^i|Z^i}(y^i, z^i) = \mathsf{P}^{\mathbf{G}^{\mathcal{C}}\mathbf{H}^{\mathcal{D}}}_{c_i d_i Y^i|X^i}(y^i, x^i).$$

For any $z^i \in \mathcal{Z}^i$, equality of the first (second) factor in the sums of the second and third term above holds because we have $\mathbf{E}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{G}^{\mathcal{C}}$ ($\mathbf{F}^{\mathcal{B}} \stackrel{\circ}{=} \mathbf{H}^{\mathcal{D}}$). □

The following proposition (Theorem 2 from [7]) states, that if the probability of a_i (i.e. $A_i = 1$), conditioned on a_{i-1}, X^i and Y^{i-1}, does not depend on Y^{i-1} (which is the output seen so far), then adaptive strategies are no better than non-adaptive ones in making $\mathcal{A}$ fail.

Proposition 2 If $\mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{a_i|X^i a_{i-1} Y^{i-1}} = \mathsf{P}^{\mathbf{F}^{\mathcal{A}}}_{a_i|X^i a_{i-1}}$, then $\nu(\mathbf{F}^{\mathcal{A}}, \overline{a}_k) = \mu(\mathbf{F}^{\mathcal{A}}, \overline{a}_k)$.

Definition 10 For a random permutation $\mathbf{Q}$, the inverse is also a random permutation and is denoted by $\mathbf{Q}^{-1}$. If the $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-random system $\mathbf{Q}^{\mathcal{A}}$ is defined, the $(\mathcal{Y}, \mathcal{X} \times \{0,1\})$-random system $\mathbf{Q}^{\mathcal{A}^{-1}}$ can be defined naturally: Let $\mathsf{P}^{\mathbf{Q}^{\mathcal{A}^{-1}}}_{A^i X^i|Y^i} = \mathsf{P}^{\mathbf{Q}^{-1}}_{X^i|Y^i}\mathsf{P}^{\mathbf{Q}^{\mathcal{A}}}_{A^i|X^i Y^i}$, i.e., we let $\mathsf{P}^{\mathbf{Q}^{\mathcal{A}^{-1}}}_{A^i|X^i Y^i} \stackrel{\text{def}}{=} \mathsf{P}^{\mathbf{Q}^{\mathcal{A}}}_{A^i|X^i Y^i}$.

Definition 11 For an $\mathcal{X}$-random permutation $\mathbf{Q}$, let $\langle\mathbf{Q}\rangle$ be the *bidirectional permutation*[9] $\mathbf{Q}$ with access from both sides (i.e., one can query both $\mathbf{Q}$ and $\mathbf{Q}^{-1}$). More precisely, $\langle\mathbf{Q}\rangle$ is the random function $\mathcal{X} \times \{0,1\} \to \mathcal{X}$ defined as follows:

$$\langle\mathbf{Q}\rangle(U_i, D_i) = \begin{cases} \mathbf{Q}(U_i) & \text{if } D_i = 0 \\ \mathbf{Q}^{-1}(U_i) & \text{if } D_i = 1 \;. \end{cases}$$

If $\mathbf{Q}^{\mathcal{A}}$ is defined, $\langle\mathbf{Q}^{\mathcal{A}}\rangle$ can also be defined naturally: Let $V_i := \langle\mathbf{Q}\rangle(U_i, D_i)$, and let X_i and Y_i be the i-th input and output of $\mathbf{Q}$ (i.e., if $D_i = 0$, then $X_i = U_i$ and $Y_i = V_i$, and if $D_i = 1$, then $Y_i = U_i$ and $X_i = V_i$). Now we let $\mathsf{P}^{\langle\mathbf{Q}^{\mathcal{A}}\rangle}_{A^i|X^i Y^i} \stackrel{\text{def}}{=} \mathsf{P}^{\mathbf{Q}^{\mathcal{A}}}_{A^i|X^i Y^i}$.

The proposition below is Lemma 10 (iii) from [7].

Proposition 3 $\mathbf{F}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{G}^{\mathcal{B}} \iff \langle\mathbf{F}^{\mathcal{A}}\rangle \stackrel{\circ}{=} \langle\mathbf{G}^{\mathcal{B}}\rangle$.

We will also need the following

Lemma 2 $\nu(\langle\mathbf{P}^{\mathcal{C}}\mathbf{P}^{\mathcal{D}^{-1}}\rangle, \overline{c}_k \vee \overline{d}_k) \leq \nu(\langle\mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}\rangle, \overline{c}_k) + \nu(\langle\mathbf{P}\mathbf{P}^{\mathcal{D}^{-1}}\rangle, \overline{d}_k)$.

Proof. Consider $\mathbf{D}$, the optimal provoker for $\overline{c}_k \vee \overline{d}_k$ in $\langle\mathbf{P}^{\mathcal{C}}\mathbf{P}^{\mathcal{D}^{-1}}\rangle$. $\mathbf{D}$ can be used to provoke $\overline{c}_k$ resp. $\overline{d}_k$ separately (though here it may not be optimal), the lemma now follows by application of the union bound and the observation that using optimal provokers for $\overline{c}_k$ resp. $\overline{d}_k$ can only increase the success probability. □

[9] This definition is motivated by considering a block cipher which, in a mixed chosen-plaintext and chosen-ciphertext attack, can be queried from both sides.

3 Cascades of Two Random Permutations

Consider a cascade of two independent URP's (recall by the symbol $\mathbf{P}$ we denote a URP) where a condition $\mathcal{C}$ is defined on the first URP. Eq. (2) from Lemma 3 states that making $\mathcal{C}$ fail is equally hard for *adaptive distinguishers* which may access the cascade from *both sides* as for an *non-adaptive distinguishers* who may access the URP (on which $\mathcal{C}$ is defined) only from *one side*.

Lemma 3

$$\nu(\langle \mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}\rangle, \overline{c}_k) = \nu(\mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}, \overline{c}_k) = \mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k) \tag{2}$$

$$\nu(\langle \mathbf{P}\mathbf{P}^{\mathcal{D}^{-1}}\rangle, \overline{d}_k) = \nu(\mathbf{P}^{\mathcal{D}}\mathbf{P}^{-1}, \overline{d}_k) = \mu(\mathbf{P}^{\mathcal{D}}, \overline{d}_k). \tag{3}$$

Proof. We first show that

$$\nu(\mathbf{P}^{\mathcal{C}}\mathbf{P}, \overline{c}_k) = \mu(\mathbf{P}^{\mathcal{C}}\mathbf{P}, \overline{c}_k) = \mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k). \tag{4}$$

Here $\mathsf{P}^{\mathbf{P}^{\mathcal{C}}\mathbf{P}}_{c_i|X^i c_{i-1} Y^{i-1}} = \mathsf{P}^{\mathbf{P}^{\mathcal{C}}\mathbf{P}}_{c_i|X^i c_{i-1}}$ holds because $\mathcal{C}$ is defined on the first $\mathbf{P}$ in the cascade, but Y^{i-1} gives no information (is independent) of the output of the first $\mathbf{P}$. The first step now follows from Proposition 2. The last step is trivial.

We now prove (2). Consider an optimal provoker $\mathbf{D}$ for $\overline{c}_k$ in $\langle \mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}\rangle$. Let Z_i denote the random variable denoting the value appearing between the two $\mathbf{P}$'s in the ith query. $\mathbf{D}$ may query either $\mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}$ or $\mathbf{P}\mathbf{P}^{\mathcal{C}^{-1}}$, a query to the latter results in a uniform random[10] Z_i and thus also a uniform random value at the input to $\mathbf{P}^{\mathcal{C}}$, we see that querying $\mathbf{P}\mathbf{P}^{\mathcal{C}^{-1}}$ can be no better to provoke $\overline{c}_k$ than querying $\mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}$ with a random value, thus there always is an optimal provoker for $\overline{c}_k$ which never chooses to query $\mathbf{P}\mathbf{P}^{\mathcal{C}^{-1}}$ and the first step of (2) follows. The second step follows from (4) using $\mathbf{P} \equiv \mathbf{P}^{-1}$. Eq.(3) follows by symmetry. □

Lemma 4 Let $\mathbf{F}$ and $\mathbf{G}$ be random permutations and let $\mathbf{P}$ be the uniform random permutation on the same domain as $\mathbf{F}$ and $\mathbf{G}$. If there are any $\mathbf{F}^{\mathcal{A}}, \mathbf{P}^{\mathcal{C}}, \mathbf{G}^{\mathcal{B}}$ and $\mathbf{P}^{\mathcal{D}}$ such that $\mathbf{F}^{\mathcal{A}} \doteq \mathbf{P}^{\mathcal{C}}$ and $\mathbf{G}^{\mathcal{B}} \doteq \mathbf{P}^{\mathcal{D}}$ holds, then[11]

$$\Delta_k(\langle \mathbf{F}\mathbf{G}^{-1}\rangle, \langle \mathbf{P}\rangle) \leq \mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k) + \mu(\mathbf{P}^{\mathcal{D}}, \overline{d}_k). \tag{5}$$

Proof. The first step below uses the simple fact that $\mathbf{P}\mathbf{P}^{-1} \equiv \mathbf{P}$, the second step below follows from Proposition 1 using $\mathbf{F}^{\mathcal{A}}\mathbf{G}^{-1^{\mathcal{B}}} \doteq \mathbf{P}^{\mathcal{C}}\mathbf{P}^{\mathcal{D}^{-1}}$, which follows from Lemma 1. The fourth step follows from Lemma 2.

$$\begin{aligned}
\Delta_k(\langle \mathbf{F}\mathbf{G}^{-1}\rangle, \langle \mathbf{P}\rangle) &= \Delta_k(\langle \mathbf{F}\mathbf{G}^{-1}\rangle, \langle \mathbf{P}\mathbf{P}^{-1}\rangle) \leq \\
\nu(\langle \mathbf{P}^{\mathcal{C}}\mathbf{P}^{\mathcal{D}^{-1}}\rangle, \overline{c_k \wedge d_k}) &= \nu(\langle \mathbf{P}^{\mathcal{C}}\mathbf{P}^{\mathcal{D}^{-1}}\rangle, \overline{c}_k \vee \overline{d}_k) \leq \\
\nu(\langle \mathbf{P}^{\mathcal{C}}\mathbf{P}^{-1}\rangle, \overline{c}_k) + \nu(\langle \mathbf{P}\mathbf{P}^{\mathcal{C}^{-1}}\rangle, \overline{d}_k) &= \mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k) + \mu(\mathbf{P}^{\mathcal{D}}, \overline{d}_k).
\end{aligned}$$

The last step follows from Lemma 3. □

[10] Meaning uniform random in $\mathcal{Z} \setminus \{Z_1, \ldots, Z_{i-1}\}$, see the comment at the end of Definition 6.

[11] Note that by Proposition 1 we have $\mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k) = \mu(\mathbf{F}^{\mathcal{A}}, \overline{a}_k)$ and $\mu(\mathbf{P}^{\mathcal{D}}, \overline{d}_k) + \mu(\mathbf{F}^{\mathcal{B}}, \overline{b}_k)$, so one could also write the second term of (5) as $\mu(\mathbf{F}^{\mathcal{A}}, \overline{a}_k) + \mu(\mathbf{F}^{\mathcal{B}}, \overline{b}_k)$.

Corollary 1 If $\mathbf{F}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{P}^{\mathcal{C}}$ then $\Delta_k(\langle \mathbf{F}\mathbf{F}^{-1} \rangle, \langle \mathbf{P} \rangle) \leq 2\mu(\mathbf{P}^{\mathcal{C}}, \overline{c}_k)$, where the two $\mathbf{F}$'s in the cascade $\mathbf{F}\mathbf{F}^{-1}$ must be independent.

Note that with this corollary we have reduced the problem of upper bounding the indistinguishability of a cascade $\mathbf{F}\mathbf{F}^{-1}$ of two random permutations from $\mathbf{P}$ by any *adaptive* strategy who may access the cascade from *both* sides to the task of finding a $\mathbf{F}^{\mathcal{A}}$ and a $\mathbf{P}^{\mathcal{C}}$ such that $\mathbf{F}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{P}^{\mathcal{C}}$ and upper bounding $\mu(\mathbf{P}, \overline{c}_k)$, i.e., the maximal probability of making $\mathcal{C}$ fail in $\mathbf{P}$ by any *non-adaptive* strategy who may access the permutation only from *one* side.

4 The Main Theorem

In Section 4.1 we give a formal definition of the (many-round) Luby-Rackoff construction and state the main theorem. In Section 4.2 an outline of the proof is given. The full proof is given in Section 5.

4.1 Statement of the Main Theorem

We denote by I_ℓ the set $\{0,1\}^\ell$ of bit-strings of length ℓ. For a random permutation (random function) $\mathbf{Q}_\ell$ the subscript denotes that it is a permutation on I_ℓ (a function $I_\ell \to I_\ell$).

Definition 12 For $n \in \mathbb{N}$, let[12] $\psi(\mathbf{R}_n)$ be the Feistel-permutation on I_{2n} defined by $\psi(\mathbf{R}_n)(L, R) \stackrel{\text{def}}{=} (R, L \oplus \mathbf{R}_n(R))$, where $L, R \in I_n$. By Ψ_{2n}^r we denote the permutation which is defined as a cascade of r permutations $\psi(\mathbf{R}_n)$, where the $\mathbf{R}_n$ are all independent.

Our main theorem states that

Theorem 1

$$\Delta_k\left(\langle \Psi_{2n}^{6r-1} \rangle, \langle \mathbf{P}_{2n} \rangle\right) \leq \frac{k^2}{2^{2n-2}} + \frac{k^{r+1}}{2^{r(n-3)-1}}.$$

4.2 Outline of the Proof

In this section we will see how the results from Sections 2 and 3 can be used to upper-bound the indistinguishability of any random permutation (and Ψ_{2n}^r in particular) from a URP. We stress here that all statements in those sections are about random systems, hence also about any particular realisation as random automata. Below we will also need the following simple

Lemma 5

$$\Delta_k\left(\langle \Psi_{2n}^r (\Psi_{2n}^r)^{-1} \rangle, \langle \mathbf{P}_{2n} \rangle\right) = \Delta_k\left(\langle \Psi_{2n}^{2r-1} \rangle, \langle \mathbf{P}_{2n} \rangle\right).$$

[12] Recall that the symbol $\mathbf{R}$ denotes a URF.

Proof. Let $\Pi(L,R) \stackrel{\text{def}}{=} (R,L)$, i.e. Π simply exchanges the right and left half of the input. We have $(\Psi_{2n}^r)^{-1} \equiv \Pi\Psi_{2n}^r\Pi$, with this and $\Psi_{2n}^1\Pi\Psi_{2n}^1 \equiv \Psi_{2n}^1$, which holds because the XOR of two independent URF's is again a URF, we get $\Psi_{2n}^r(\Psi_{2n}^r)^{-1} \equiv \Psi_{2n}^{r-1}\Psi_{2n}^1\Pi\Psi_{2n}^1\Psi_{2n}^{r-1}\Pi \equiv \Psi_{2n}^{2r-1}\Pi$. This and $\mathbf{P}_{2n} \equiv \mathbf{P}_{2n}\Pi$ proves the first step of $\Delta_k\left(\langle\Psi_{2n}^r(\Psi_{2n}^r)^{-1}\rangle, \langle\mathbf{P}_{2n}\rangle\right) = \Delta_k\left(\langle\Psi_{2n}^{2r-1}\Pi\rangle, \langle\mathbf{P}_{2n}\Pi\rangle\right) = \Delta_k\left(\langle\Psi_{2n}^{2r-1}\rangle, \langle\mathbf{P}_{2n}\rangle\right)$, the second step is trivial. □

Our aim is to upper-bound $\Delta_k(\langle\Psi_{2n}^r\rangle, \langle\mathbf{P}_{2n}\rangle)$. With Propositions 1 and 3 this can be done as follows: Define some random automata $\mathbf{F}$ and $\mathbf{G}$ such that $\mathbf{F} \equiv \Psi_{2n}^r$ and $\mathbf{G} \equiv \mathbf{P}_{2n}$. Then define conditions $\mathcal{A}$ and $\mathcal{B}$ on $\mathbf{F}$ and $\mathbf{G}$ respectively, such that $\mathbf{F}^{\mathcal{A}} \stackrel{\circ}{=} \mathbf{G}^{\mathcal{B}}$ and prove an upper bound $\varepsilon(k,r,n)$ for $\nu(\langle\mathbf{G}^{\mathcal{B}}\rangle, \overline{b}_k)$ (or $\nu(\langle\mathbf{F}^{\mathcal{A}}\rangle, \overline{a}_k)$, which is the same). It follows that $\Delta_k(\langle\Psi_{2n}^r\rangle, \langle\mathbf{P}_{2n}\rangle) \leq \varepsilon(k,r,n)$.

With Corollary 1 we can bypass the task of upper bounding $\nu(\langle\mathbf{G}^{\mathcal{B}}\rangle, \overline{b}_k)$ and give an upper bound $\varepsilon'(k,r,n)$ for $\mu(\mathbf{G}^{\mathcal{B}}, \overline{b}_k)$ instead.[13] Then, using Lemma 5 in the first and Corollary 1 in the second step, we get

$$\Delta_k(\langle\Psi_{2n}^{2r-1}\rangle, \langle\mathbf{P}_{2n}\rangle) = \Delta_k(\langle\Psi_{2n}^r(\Psi_{2n}^r)^{-1}\rangle, \langle\mathbf{P}_{2n}\rangle) \leq \varepsilon'(k,r,n).$$

Note that the price we payed for this simplification is that our bound now applies for Feistel-permutations with $2r-1$ instead of r rounds. This is basically the idea underlying the proof of the main theorem. Though in the proof the two random automata $\mathbf{F}$ and $\mathbf{G}$ (as discussed above) are not defined separately, we only give one random automaton $\mathbf{H}$ with two outputs, where the random system $\mathbf{H}$ is equivalent to[14] Ψ_{2n}^{3r} or $\mathbf{P}_{2n}$ if we ignore the right or left output of $\mathbf{H}$, respectively. For this $\mathbf{H}$, a condition $\mathcal{M}$ is given such that the two outputs of $\mathbf{H}^{\mathcal{M}}$ have the same distribution whenever the condition holds.

5 Proof of the Main Theorem

In this section we propose three lemmas used in the proof (as outlined in the previous section) of our main theorem. The proof itself is given at the end of this section.

For the sequel fix any $r \in \mathbb{N}$ and $n \in \mathbb{N}$, and let $\mathbf{H}$ denote the $(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$-random automaton (where $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = I_{2n} \stackrel{\text{def}}{=} \{0,1\}^{2n}$) as shown in Figure 1. Let $\mathbf{H}_L$ ($\mathbf{H}_R$) be equivalent to $\mathbf{H}$, but where the output is only the left (right) half of the output of $\mathbf{H}$.

$\mathbf{H}$ is constructed by combining a $\mathbf{P}_{2n}$ and a Ψ_{2n}^{3r} (the two Ψ_{2n}^{3r}'s drawn in the figure are one and the same). In $\mathbf{H}$, the output $Y_i \times Z_i$ on the ith query X_i is determined as follows: Ψ_{2n}^{3r} is queried with X_i, this gives Y_i. The X_i is also applied to $\mathbf{P}_{2n}$ to get a value $\tilde{X}_i$. Then Ψ_{2n}^{3r} is queried with $\tilde{X}_i$ which gives the Z_i.

[13] This is likely to be much easier because we need to consider only non-adaptive chosen-plaintext strategies instead of adaptive combined chosen plaintext and ciphertext strategies.

[14] for technical reasons, we will need the number of rounds to be a multiple of 3 here

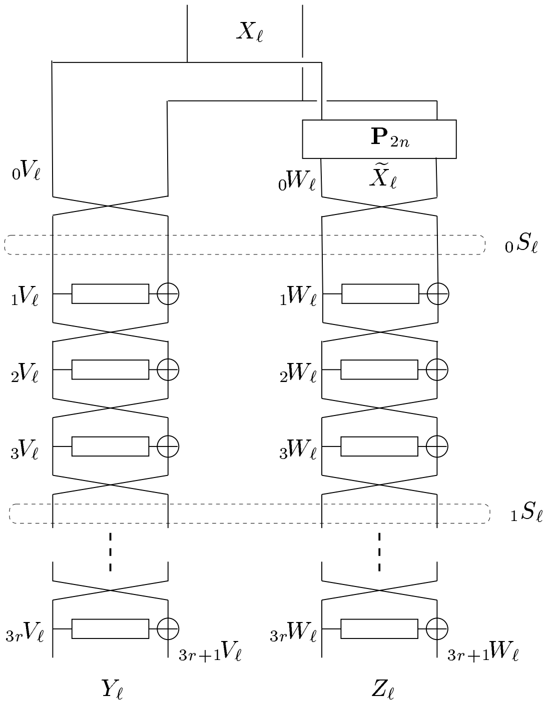

Fig. 1. The random automaton $\mathbf{H}$. The labelling refers to the random variables as used in Section 5. On input X_ℓ, Ψ_{2n}^{3r} is invoked twice, once with X_ℓ and once with $\widetilde{X}_\ell$. Ψ_{2n}^{3r} and $\mathbf{P}_{2n}$ are independent, and also all $3r$ URF's are independent. Note that in the figure the two Ψ_{2n}^{3r} permutations are one and the same, i.e., URF's at the same level are identical.

Lemma 6 $\mathbf{H}_L \equiv \Psi_{2n}^{3r}$ *and* $\mathbf{H}_R \equiv \mathbf{P}_{2n}$.

Proof. $\mathbf{H}_L \equiv \Psi_{2n}^{3r}$ can be seen directly from the definition of $\mathbf{H}_L$. We have $\mathbf{H}_R \equiv \mathbf{P}_{2n}\Psi_{2n}^{3r} \equiv \mathbf{P}_{2n}$ since the cascade of a URP with any other permutation (which is independent of the URP) is again a URP. □

Consider $\mathbf{H}$ being queried with some $x^k = \{x_1, \dots, x_k\} \in \mathcal{X}^k$. The ℓth query x_ℓ results in two invocations of Ψ_{2n}^{3r}, once with x_ℓ and once with $\widetilde{X}_\ell$, where $\widetilde{X}_\ell$ is the random variable denoting the output of $\mathbf{P}_{2n}$ on input x_ℓ (cf. Fig. 1). A query to Ψ_{2n}^{3r} results in one query to every of the $3r$ URF's. Let ${}_jV_\ell$ denote the input to the jth URF when queried with x_ℓ and let ${}_jW_\ell$ denote the input to the jth URF when queried with $\widetilde{X}_\ell$.

After k queries we say that an input ${}_jV_\ell$ resp. ${}_jW_\ell$ is unique (and we denote this event by ${}_j\tau_\ell$ resp. ${}_j\widetilde{\tau}_\ell$) if the jth URF was invoked only once with ${}_jV_\ell$ resp. ${}_jW_\ell$, i.e., for $1 \le i \le k$ and $1 \le j \le 3r$ we define

$$
\begin{aligned}
{}_j\tau_\ell &\iff (\forall i, i \neq \ell : {}_jV_\ell \neq {}_jV_i) \wedge (\forall i : {}_jV_\ell \neq {}_jW_i) \\
{}_j\widetilde{\tau}_\ell &\iff (\forall i, i \neq \ell : {}_jW_\ell \neq {}_jW_i) \wedge (\forall i : {}_jW_\ell \neq {}_jV_i).
\end{aligned}
$$

By ${}_q\xi_\ell$ we denote the event that ${}_{3q+2}V_\ell$, ${}_{3q+3}V_\ell$, ${}_{3q+2}W_\ell$ and ${}_{3q+3}W_\ell$ are all unique, i.e, for $0 \leq q \leq r-1$:

$$ {}_q\xi_\ell \iff {}_{3q+2}\tau_\ell \wedge {}_{3q+3}\tau_\ell \wedge {}_{3q+2}\widetilde{\tau}_\ell \wedge {}_{3q+3}\widetilde{\tau}_\ell . $$

By λ_ℓ we denote the event that ${}_q\xi_\ell$ holds for some $q, 0 \leq q \leq r-1$:

$$ \lambda_\ell \iff \bigvee_{q=0}^{r-1} {}_q\xi_\ell . $$

We can now define our monotone condition $\mathcal{M}$ for $\mathbf{H}$.

Definition 13 The condition $\mathcal{M}$ for $\mathbf{H}$ is defined as (Recall that m_k denotes the event that $M_k = 1$):

$$ m_k \iff \bigwedge_{\ell=1}^{k} \lambda_\ell \iff \bigwedge_{\ell=1}^{k} \bigvee_{q=0}^{r-1} {}_{3q+2}\tau_\ell \wedge {}_{3q+3}\tau_\ell \wedge {}_{3q+2}\widetilde{\tau}_\ell \wedge {}_{3q+3}\widetilde{\tau}_\ell . \tag{6} $$

So after $\mathbf{H}^{\mathcal{M}}$ has been queried with k inputs $x_1, \ldots, x_k$, the condition M_k holds if the following holds for $i = 1, \ldots, k$: There is an index q such that the four values that appear as the input to the URF's in the consecutive rounds $3q+2$ and $3q+3$, when Ψ_{2n}^{3r} is queried with x_i and $\tilde{X}_i$, are unique. In the proof of Lemma 8 we will use this fact to show that the right and left part of the output of $\mathbf{H}^{\mathcal{M}}$ has the same distribution whenever the condition holds, i.e., $\mathbf{H}_L^{\mathcal{M}} \stackrel{\circ}{=} \mathbf{H}_R^{\mathcal{M}}$. But first we give an upper bound on the probability of any non-adaptive strategy, making at most k queries, in provoking the event $\overline{m}_k$ (i.e. $M_k = 0$).

Lemma 7

$$ \mu(\mathbf{H}^{\mathcal{M}}, \overline{m}_k) \stackrel{\text{def}}{=} \max_{x^k \in \mathcal{X}^k} \mathsf{P}^{\mathbf{H}}_{\overline{m}_k | X^k}(x^k) \leq \frac{k^2}{2^{2n-1}} + \frac{k^{r+1}}{2^{r(n-3)}} . $$

Proof. Consider any $x^k \in \mathcal{X}^k$ (in particular the one maximising the second term in the lemma). Throughout the proof, all probabilities are in the random experiment where $\mathbf{H}$ is queried with x^k. So for example $\mathsf{P}[\overline{m}_k]$ denotes $\mathsf{P}^{\mathbf{H}}_{\overline{m}_k | X^k}(x^k)$. To establish the lemma, we must show that $\mathsf{P}[\overline{m}_k] \leq \frac{k^2}{2^{2n-1}} + \frac{k^{r+1}}{2^{r(n-3)}}$.

Let γ_k denote the event (defined on $\mathbf{H}$ after k queries) which holds if and only if x^k and $\widetilde{X}^k$ have no elements in common (i.e. all elements in $x^k \bigcup \widetilde{X}^k$ are distinct)[15]. The birthday bound gives us

$$ \mathsf{P}[\gamma_k] = \prod_{i=0}^{k-1} \left(1 - \frac{k+i}{2^{2n}}\right) \geq 1 - \frac{k^2}{2^{2n-1}} . \tag{7} $$

We will now upper bound $\mathsf{P}[\overline{m}_k | \gamma_k]$. In order to do so it helps to think of Ψ_{2n}^{3r} as a cascade of r blocks, each a Ψ_{2n}^{3} permutation. For $0 \leq q \leq r-1$ let

$$ {}_qS \stackrel{\text{def}}{=} (({}_{3q+1}V_1, {}_{3q}V_1), \ldots, ({}_{3q+1}V_k, {}_{3q}V_k), ({}_{3q+1}W_1, {}_{3q}W_1), \ldots, ({}_{3q+1}W_k, {}_{3q}W_k)) $$

[15] The x^k are all distinct, see comment after Definition 6.

denote the values appearing on the input to the $q+1$th block on input x^k and $\widetilde{X}^k$ (here ${}_0V_\ell$ and ${}_0W_\ell$ are as in figure 1). Let ${}_q\mathcal{S}$ denote the set of values ${}_qS$ may take on. Next we prove that, for any $\ell : 1 \le \ell \le k$, $h \in \{2,3\}$ and any possible ${}_qs \in {}_q\mathcal{S}$

$$\mathsf{P}[\neg_{3q+h}\tau_\ell | \gamma_k, {}_qS = {}_qs] \le \frac{2k}{2^n} \qquad \text{and} \qquad \mathsf{P}[\neg_{3q+h}\widetilde{\tau}_\ell | \gamma_k, {}_qS = {}_qs] \le \frac{2k}{2^n}. \tag{8}$$

Let ${}_qs_j$ denote the jth element in ${}_qs$ (e.g. ${}_0s_1 = \{{}_1v_1, {}_0v_1\}$) and let ${}_qs_j^L$ and ${}_qs_j^R$ denote the left and right part of ${}_qs_j$. We will only prove the first statement for $h=2$, the other cases are similar. The probability that $\neg_{3q+2}\tau_\ell$, conditioned on ${}_qS = {}_qs$, is the probability that there is a $t \ne \ell, 1 \le t \le 2k$, such that $f({}_qs_\ell^L) \oplus {}_qs_\ell^R = f({}_qs_t^L) \oplus {}_qs_t^R$, where f is a URF. This is at most $\frac{2k}{2^n}$ if all the $2k$ elements in ${}_qs$ are distinct, which is the case because we condition on γ_k.[16] Now for any ℓ, q and ${}_qs \in {}_q\mathcal{S}$ we get (using the union bound in the first and (8) in the second step)

$$\mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs] \le \sum_{\tau \in \{{}_{3q+2}\tau_\ell, {}_{3q+3}\tau_\ell, {}_{3q+2}\widetilde{\tau}_\ell, {}_{3q+3}\widetilde{\tau}_\ell\}} \mathsf{P}[\neg\tau | \gamma_k, {}_qS = {}_qs] \le \frac{8k}{2^n} = \frac{k}{2^{n-3}}. \tag{9}$$

The reason we introduced ${}_qS$ is that, conditioned on ${}_qS$, the probability of any event defined on the last $3r-q$ rounds of Ψ_{2n}^{3r} does not depend on any event defined on the $3q$ first rounds of Ψ_{2n}^{3r}. The reason is that all rounds (random functions) are independent, and every interaction between the first $3q$ and the last $3q-r$ rounds is captured by ${}_qS$ (which specifies all values exchanged between round $3q$ and round $3q+1$). Let ${}_q\kappa$ be any event defined on the $3q$ first rounds of Ψ_{2n}^{3r}. By the above argument for any ${}_qs \in {}_q\mathcal{S}$,

$$\mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs, {}_q\kappa] = \mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs]. \tag{10}$$

For $0 \le q \le r-1$ let ${}_q\kappa \stackrel{\text{def}}{=} \bigwedge_{t=0}^{q-1} \neg_t\xi_\ell$. Now

$$\mathsf{P}[\neg\lambda_\ell | \gamma_k] = \mathsf{P}[\bigwedge_{q=0}^{r-1} \neg_q\xi_\ell | \gamma_k] = \prod_{q=0}^{r-1} \mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_q\kappa] = \tag{11}$$

$$\prod_{q=0}^{r-1} \sum_{{}_qs \in {}_q\mathcal{S}} \mathsf{P}[{}_qS = {}_qs | \gamma_k, {}_q\kappa] \mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs, {}_q\kappa] \le \left(\frac{k}{2^{n-3}}\right)^r = \frac{k^r}{2^{r(n-3)}}.$$

In the fourth step above we used $\mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs, {}_q\kappa] = \mathsf{P}[\neg_q\xi_\ell | \gamma_k, {}_qS = {}_qs] \le \frac{k}{2^{n-3}}$, which follows from (10) and (5). Now using the union bound in the third

[16] Consider any $t \ne \ell$. If ${}_qs_t^L = {}_qs_\ell^L$, then ${}_qs_t^R \ne {}_qs_\ell^R$ (because $\{{}_qs_t^L, {}_qs_t^R\} \ne \{{}_qs_\ell^L, {}_qs_\ell^R\}$) and thus $f({}_qs_\ell^L) \oplus {}_qs_\ell^R \ne f({}_qs_t^L) \oplus {}_qs_t^R$. If ${}_qs_t^L \ne {}_qs_\ell^L$, then $\mathsf{P}[f({}_qs_\ell^L) \oplus {}_qs_\ell^R = f({}_qs_t^L) \oplus {}_qs_t^R] = 1/2^n$. With the union bound we now get that the probability that $f({}_qs_\ell^L) \oplus {}_qs_\ell^R = f({}_qs_t^L) \oplus {}_qs_t^R$ for any of the $2k-1$ possible $t \ne \ell$, is at most $(2k-1)/2^n \le 2k/2^n$.

and (11) in the fourth step below, we obtain

$$\mathsf{P}[\overline{m}_k|\gamma_k] = \mathsf{P}[\neg \bigwedge_{\ell=1}^{k} \lambda_\ell|\gamma_k] = \mathsf{P}[\bigvee_{\ell=1}^{k} \neg\lambda_\ell|\gamma_k] \leq \sum_{\ell=1}^{k} \mathsf{P}[\neg\lambda_\ell|\gamma_k] \leq k\frac{k^r}{2^{r(n-3)}} \quad (12)$$

And finally, using (7) and (12) in the last step, we get

$$\mathsf{P}[\overline{m}_k] = \mathsf{P}[\overline{m}_k, \neg\gamma_k] + \underbrace{\mathsf{P}[\gamma_k]\mathsf{P}[\overline{m}_k|\gamma_k]}_{\mathsf{P}[\overline{m}_k,\gamma_k]} \leq \mathsf{P}[\neg\gamma_k] + \mathsf{P}[\overline{m}_k|\gamma_k] \leq \frac{k^2}{2^{2n-1}} + \frac{k^{r+1}}{2^{r(n-3)}}.$$

□

Lemma 8

$$\mathbf{H}_L^{\mathcal{M}} \mathrel{\mathring{=}} \mathbf{H}_R^{\mathcal{M}}.$$

Proof. We will show that for all k and all $x^k, g', g'' \in I_{2n}^k$ we have

$$\mathsf{P}^{\mathbf{H}^{\mathcal{M}}}_{m_k Y^k Z^k|X^k}(g', g'', x^k) = \mathsf{P}^{\mathbf{H}^{\mathcal{M}}}_{m_k Y^k Z^k|X^k}(g'', g', x^k). \quad (13)$$

If we sum the two terms above over all g'', we see that (13) implies that for all $x^k, g' \in I_{2n}^k$

$$\mathsf{P}^{\mathbf{H}_L^{\mathcal{M}}}_{m_k Y^k|X^k}(g', x^k) = \mathsf{P}^{\mathbf{H}_R^{\mathcal{M}}}_{m_k Z^k|X^k}(g', x^k) \quad (14)$$

holds. Note that this is exactly the statement of the lemma.

The space of internal randomness for $\mathbf{H}$ (see Definition 1) consists of the function tables for the $3r$ $\mathbf{R}_n$'s which build Ψ_{2n}^{3r} (each uniform random in $\{0,1\}^{n2^n}$) and a number uniform random between 1 and $2^{2n}!$ defining one possible permutation on $\{0,1\}^{2n}$. Thus the internal randomness of $\mathbf{H}$ is an element chosen uniformly random in $\mathcal{R} \stackrel{\text{def}}{=} \{0,1\}^{3rn2^n} \times [1, 2^{2n}!]$.

Let $\mathcal{R}(x^k, g', g'') \subset \mathcal{R}$ be such that iff the internal randomness of $\mathbf{H}$ is $\rho \in \mathcal{R}(x^k, g', g'')$, then the system will output (g', g'') on input x^k *and* M_k will hold (note that M_k is determined by x^k and ρ). With this we can write (13) as

$$\frac{|\mathcal{R}(x^k, g', g'')|}{|\mathcal{R}|} = \frac{|\mathcal{R}(x^k, g'', g')|}{|\mathcal{R}|}. \quad (15)$$

We will prove (15) by showing a bijection between the sets $\mathcal{R}(x^k, g', g'')$ and $\mathcal{R}(x^k, g'', g')$, which implies that they have the same cardinality.

Consider $\mathbf{H}$ with internal randomness $\rho \in \mathcal{R}(x^k, g', g'')$ was queried with x^k. Let V, W and $\widetilde{X}$ be as defined before Definition 13. Note that $V, W, \widetilde{X}$ are determined by ρ and x^k, we use the corresponding small letters to denote the values taken by V, W and $\widetilde{X}$. Also all $3r$ URF's are deterministic functions when ρ is fixed, we denote the function in the jth round by f_j.

Let $\alpha_1, \ldots, \alpha_k$ be such that ${}_{3\alpha_\ell+2}\tau_\ell \wedge_{3\alpha_\ell+3} \tau_\ell \wedge_{3\alpha_\ell+2} \widetilde{\tau}_\ell \wedge_{3\alpha_\ell+3} \widetilde{\tau}_\ell$ for $\ell = 1, \ldots, k$. By (6) and the fact that M_k holds, such $\alpha_1, \ldots, \alpha_k$ exist. If there are several possibilities for α_ℓ then let it be, say, the smallest possible value. Note that for $1 \le \ell \le k$, ρ defines the following relations:

$$
\begin{aligned}
f_{3\alpha_\ell+2}({}_{3\alpha_\ell+2}v_\ell) &= {}_{3\alpha_\ell+1}v_\ell \oplus {}_{3\alpha_\ell+3}v_\ell & \quad f_{3\alpha_\ell+2}({}_{3\alpha_\ell+2}w_\ell) &= {}_{3\alpha_\ell+1}w_\ell \oplus {}_{3\alpha_\ell+3}w_\ell \\
f_{3\alpha_\ell+3}({}_{3\alpha_\ell+3}v_\ell) &= {}_{3\alpha_\ell+2}v_\ell \oplus {}_{3\alpha_\ell+4}v_\ell & \quad f_{3\alpha_\ell+3}({}_{3\alpha_\ell+3}w_\ell) &= {}_{3\alpha_\ell+2}w_\ell \oplus {}_{3\alpha_\ell+4}w_\ell
\end{aligned}
$$

Let $\phi_{x^k}(\rho)$ be a transformation on ρ which for $1 \le \ell \le k$ changes the function table of $f_{3\alpha_\ell+2}$ resp. $f_{3\alpha_\ell+3}$ on inputs ${}_{3\alpha_\ell+2}v_\ell$ and ${}_{3\alpha_\ell+2}w_\ell$ resp. ${}_{3\alpha_\ell+3}v_\ell$ and ${}_{3\alpha_\ell+3}w_\ell$ to

$$
\begin{aligned}
f_{3\alpha_\ell+2}({}_{3\alpha_\ell+2}v_\ell) &= {}_{3\alpha_\ell+1}v_\ell \oplus {}_{3\alpha_\ell+3}w_\ell & \quad f_{3\alpha_\ell+2}({}_{3\alpha_\ell+2}w_\ell) &= {}_{3\alpha_\ell+1}w_\ell \oplus {}_{3\alpha_\ell+3}v_\ell \\
f_{3\alpha_\ell+3}({}_{3\alpha_\ell+3}w_\ell) &= {}_{3\alpha_\ell+2}v_\ell \oplus {}_{3\alpha_\ell+4}w_\ell & \quad f_{3\alpha_\ell+3}({}_{3\alpha_\ell+3}v_\ell) &= {}_{3\alpha_\ell+2}w_\ell \oplus {}_{3\alpha_\ell+4}v_\ell
\end{aligned}
$$

Then $\phi_{x^k}(\rho)$ is in $\mathcal{R}(x^k, g'', g')$. To see this, first note that ϕ_{x^k} only changes $f_{3\alpha_\ell+2}$ and $f_{3\alpha_\ell+3}$ on inputs that are unique. Consider the two cases where the internal randomness of $\mathbf{H}$ is ρ and $\phi_{x^k}(\rho)$ respectively. On input x_ℓ, ${}_jv_\ell$ and ${}_jw_\ell$ are equal for $j \le 3\alpha_\ell + 2$ in both cases, this is because for $j \le 3\alpha_\ell + 1$ the input/output behaviour of the internal functions f_j on inputs ${}_jv_\ell$ and ${}_jw_\ell$ is not affected by ϕ_{x^k}. With $({}_{3\alpha_\ell+1}v_\ell, {}_{3\alpha_\ell+2}v_\ell)$ and $({}_{3\alpha_\ell+1}w_\ell, {}_{3\alpha_\ell+2}w_\ell)$ being equal in both cases, we see by the definition of $\phi_{x^k}(\rho)$ that the values $({}_{3\alpha_\ell+3}v_\ell, {}_{3\alpha_\ell+4}v_\ell)$ and $({}_{3\alpha_\ell+3}w_\ell, {}_{3\alpha_\ell+4}w_\ell)$ are exchanged in both cases. With this also all ${}_jv_\ell$ and ${}_jw_\ell$ are exchanged for all $j \ge 3\alpha_\ell+4$ (ϕ_{x^k} does not affect the input/output behaviour of the internal functions f_j on inputs ${}_jv_\ell$ and ${}_jw_\ell$ for $j \ge 3\alpha_\ell + 4$), so the outputs g'_ℓ and g''_ℓ are also exchanged for all $1 \le \ell \le k$, and thus $\phi_{x^k}(\rho) \in \mathcal{R}(x^k, g'', g')$.

Finally note that $\phi_{x^k}(\phi_{x^k}(\rho)) = \rho$, thus ϕ_{x^k} is a bijection (actually even an involution) between $\mathcal{R}(x^k, g', g'')$ and $\mathcal{R}(x^k, g'', g')$, and hence $|\mathcal{R}(x^k, g', g'')| = |\mathcal{R}(x^k, g'', g')|$. □

Proof (of Theorem 1). There is a random automaton $\mathbf{H}$ (cf. Figure 1) with two outputs such that $\mathbf{H}_L \equiv \Psi_{2n}^{3r}$ and $\mathbf{H}_R \equiv \mathbf{P}_{2n}$ (Lemma 6). Here $\mathbf{H}_L$ and $\mathbf{H}_R$ denote $\mathbf{H}$, but where only the right and left half, respectively, is seen at the output. There is a condition $\mathcal{M}$ defined for $\mathbf{H}$ (Definition 13) such that $\mathbf{H}_L^{\mathcal{M}} \mathrel{\hat{=}} \mathbf{H}_R^{\mathcal{M}}$ (Lemma 8) and $\mu(\mathbf{H}^{\mathcal{M}}, \overline{m}_k) \le \frac{k^2}{2^{2n-1}} + \frac{k^{r+1}}{2^{r(n-3)}}$ (Lemma 7). With Corollary 1 and the observation that[17] $\mu(\mathbf{H}_R^{\mathcal{M}}, \overline{m}_k) = \mu(\mathbf{H}^{\mathcal{M}}, \overline{m}_k)$ we now get

$$\Delta_k\left(\langle \Psi_{2n}^r (\Psi_{2n}^r)^{-1} \rangle, \langle \mathbf{P}_{2n} \rangle\right) \le \frac{k^2}{2^{2n-2}} + \frac{k^{r+1}}{2^{r(n-3)-1}}$$

and the theorem follows with Lemma 5. □

[17] In the non-adaptive case it does not matter how much of the output the distinguisher can see.

6 Conclusions

In this paper we showed that the number of queries needed to distinguish a uniform random permutation (URP) from Ψ_{2n}^r (the r-round Feistel-permutation with independent uniform random functions) by any computationally unbounded adaptive distinguisher making combined plaintext/ciphertext queries, approaches the information theoretic upper-bound $O(2^n)$, as r is increased.

The proof of our main theorem is based on the framework of [7]. In this framework, for our case, one must define two random automata with the input/output behaviour of Ψ_{2n}^r and a URP, respectively. Then one must give a condition for each automaton, such that both have the same input/output behaviour as long as the condition holds. The expected probability of any distinguisher (as above) making k queries, in making this condition fail, is now an upper bound for the advantage of any distinguisher (as above) in distinguishing Ψ_{2n}^r from a URP.

We proposed a new result (Lemma 4, see also Corollary 1) which reduces the arising problem of upper bounding the probability of any *adaptive* distinguisher making *combined chosen plaintext/ciphertext queries* in making the condition fail, to the case where one only has to consider a distinguisher making *non-adaptive chosen plaintext* queries. This lemma is generic and can be applied to any random permutation, but it comes at a price: The bound now only holds for a cascade of two of the originally considered random permutations.

We took a new approach in defining the two random automata as discussed above. Only one random automaton **H** with two outputs was defined, such that **H** has the same input/output behaviour as Ψ_{2n}^r or a URP when only the right or left part of the output is considered. One now must only find a single condition for **H** such that the input/output behaviour, when only the left or the right half of the output is considered, is identical whenever the condition holds. We do not know how to prove our result without this trick, and think that it could be useful for the analysis of other systems as well.

Patarin conjectured that the information theoretic upper bound is already reached if the number of rounds is *constant* (5 or maybe 6 or 7), this question is still open. If the conjecture is true, then collision arguments (like "as long as there is some input that has not appeared yet ... we cannot distinguish"), as used here and in many other papers, will be too weak as to prove it.[18] Maybe adopting ideas from [1], and arguing about linear independence (like "as long as some internal inputs are linearly independent...") would be more successful.

Acknowledgements

We would like to thank Serge Vaudenay for his valuable comments which helped to improve the presentation of this paper.

[18] After $O(2^{nr/(r+1)})$ queries to Ψ_{2n}^r, we have a constant probability that there was a query where none of the r URF's was invoked with an input on which it was never invoked before.

References

1. W. Aiello and R. Venkatesan, Foiling birthday attacks in length-doubling transformations - Benes: A non-reversible alternative to Feistel, *Advances in Cryptology - EUROCRYPT '96*, Lecture Notes in Computer Science, vol. 1070, pp. 307–320, Springer-Verlag, 1996.
2. O. Goldreich, S. Goldwasser, and S. Micali, How to construct random functions, *Journal of the ACM*, vol. 33, no. 4, pp. 210–217, 1986.
3. L. R. Knudsen, The security of Feistel ciphers with six rounds or less, *Journal of Cryptology* vol. 15, no. 3, pp. 207–222, Springer-Verlag, 2002.
4. S. Lucks, Faster Luby-Rackoff ciphers, *Fast Software Encryption*, Lecture Notes in Computer Science, vol. 1039, pp. 189–203, Springer-Verlag, 1996.
5. M. Luby and C. Rackoff, How to construct pseudo-random permutations from pseudo-random functions, *SIAM J. on Computing*, vol. 17, no. 2, pp. 373–386, 1988.
6. U. M. Maurer, A simplified and generalized treatment of Luby-Rackoff pseudo-random permutation generators, *Advances in Cryptology - EUROCRYPT '92*, Lecture Notes in Computer Science, vol. 658, pp. 239–255, Springer-Verlag, 1992.
7. ——, Indistinguishability of random systems *Advances in Cryptology - EUROCRYPT '02*, Lecture Notes in Computer Science, vol. 2332, pp. 110–132, Springer-Verlag, 2002.
8. M. Naor and O. Reingold, On the construction of pseudorandom permutations: Luby-Rackoff revisited, *Journal of Cryptology*, vol. 12, no. 1, pp. 29–66, 1999.
9. J. Patarin, How to construct pseudorandom permutations from a single pseudo-random function, *Advances in Cryptology - EUROCRYPT '92*, Lecture Notes in Computer Science, vol. 658, pp. 256–266, Springer-Verlag, 1992.
10. ——, About Feistel schemes with six (or more) rounds, *Fast Software Encryption*, Lecture Notes in Computer Science, vol. 1372, pp. 103–121, Springer-Verlag, 1998.
11. S. Patel, Z. Ramzan, and G. Sundaram, Towards making Luby-Rackoff ciphers optimal and practical, *Fast Software Encryption*, Lecture Notes in Computer Science, vol. 1636, pp. 171–185, Springer-Verlag, 1999.
12. J. Pieprzyk, How to construct pseudorandom permutations from single pseudo-random functions, *Advances in Cryptology - EUROCRYPT '90*, Lecture Notes in Computer Science, vol. 473, pp. 140–150, Springer-Verlag, 1990.
13. Z. Ramzan and L. Reyzin, On the round security of symmetric-key cryptographic primitives, *Advances in Cryptology - CRYPTO '00*, Lecture Notes in Computer Science, vol. 1880, pp. 376–393, Springer-Verlag, 2000.
14. S. Vaudenay, Adaptive-attack norm for decorrelation and super-pseudorandomness, *Proc. of SAC'99*, Lecture Notes in Computer Science, vol. 1758, pp. 49–61, Springer-Verlag, 2000.
15. Y. Zheng, T. Matsumoto, and H. Imai, Impossibility and optimality results on constructing pseudorandom permutations, *Advances in Cryptology - EUROCRYPT '89*, Lecture Notes in Computer Science, vol. 434, pp. 412–422, Springer-Verlag, 1989.