On the Pseudorandomness of Top-Level Schemes of Block Ciphers

Shiho Moriai[1] and Serge Vaudenay[2]*

[1] NTT Laboratories
1-1 Hikarinooka, Yokosuka, 239-0847 Japan
shiho@isl.ntt.co.jp

[2] Swiss Federal Institute of Technology (EPFL)
1015 Lausanne, Switzerland
serge.vaudenay@epfl.ch

Abstract. Block ciphers are usually based on one top-level scheme into which we plug "round functions". To analyze security, it is important to study the intrinsic security provided by the top-level scheme from the viewpoint of randomness: given a block cipher in which we replaced the lower-level schemes by idealized oracles, we measure the security (in terms of best advantage for a distinguisher) depending on the number of rounds and the number of chosen plaintexts. We then extrapolate a sufficient number of secure rounds given the regular bounds provided by decorrelation theory.
This approach allows the comparison of several generalizations of the Feistel schemes and others. In particular, we compare the randomness provided by the schemes used by the AES candidates.
In addition we provide a general paradigm for analyzing the security provided by the interaction between the different levels of the block cipher structure.

1 Introduction

From the attacker's viewpoint, the block cipher used by a given user can be considered as an instance of a random permutation over a message block space: since he only knows how the secret key has been chosen, he only has probabilistic information (in a Shannon sense) on the key and the permutation. In this setting, security can be formalized by pseudorandomness: if there is no way to distinguish the block cipher from an ideal random permutation, then we cannot attack it. Pseudorandomness more precisely means that no oracle circuit with polynomially many oracle gates can distinguish between the encryption function and a truly random permutation.

A block cipher usually made from a top-level oracle circuit that we call "scheme" (for instance the circuit of the Feistel scheme [4]) into which we plug lower-level circuits that we call "primitives" like round functions, S-boxes, and so

* Part of this work was done while the author was visiting NTT Laboratories.

on. An attack may succeed if it "bypasses" some of the primitives by using some intrinsic weaknesses of the scheme. For instance, differential cryptanalysis [1] can investigate the differentials in which some S-boxes play no role at all. This idea motivated this paper: we consider ideal models of the block ciphers by replacing the primitives by truly random functions and study the pseudorandomness provided by the scheme.

In this paper we investigate the randomness of several of the schemes used in many block ciphers. The target schemes are the Feistel scheme, variants of the Feistel scheme (the CAST256-like Feistel scheme, the MARS-like Feistel scheme, and the RC6-like Feistel scheme), and the SQUARE-like scheme used in SQUARE, Rijndael and Crypton.

The pseudorandomness of some general schemes were discussed in previous papers e.g. [9, 17]. In this paper we show how we can reach these kind of results and extensions in an easier and more systematic way by using the decorrelation theory introduced in [13–15].

In order to compare the schemes we study the threshold number of rounds needed to achieve randomness, a theoretically sufficient number of secure rounds against attacks that are limited to two chosen plaintexts or ciphertexts (which plays a crucial role in the security against differential and linear cryptanalysis), and the sufficient number of secure rounds, in practice, when we use a practical decorrelation module (as in DFC [5]) for primitives instead of an ideal primitive.

2 Decorrelation Theory and Randomness of Iterated Ciphers

2.1 Definitions and Basic Properties

The goal of decorrelation theory is to provide some kind of formal proof of security on block ciphers. This section describes the essential definitions and lemmas in decorrelation theory to prove the randomness of iterated ciphers.

Definition 1 (d-wise distribution matrix). *Given a random function F[1] from a set $\mathcal{M}_1$ to a set $\mathcal{M}_2$ and an integer d, we define the "d-wise distribution matrix" of F as the following $\mathcal{M}_1^d \times \mathcal{M}_2^d$-matrix.*

$$[F]^d_{(x_1,\ldots,x_d),(y_1,\ldots,y_d)} = \Pr[F(x_1) = y_1, \ldots, F(x_d) = y_d],$$

where $x_i \in \mathcal{M}_1$ and $y_i \in \mathcal{M}_2$ for $i = 1, \ldots, d$

Definition 2 (d-wise decorrelation bias). *Given a random function F from a set $\mathcal{M}_1$ to a set $\mathcal{M}_2$, a canonical idealized version F^* of F, an integer d, and a*

[1] Throughout this paper, "a random function F" means a random variable F which takes values in a set of functions, following regular probability theory. The same holds for "a random permutation C".

distance D over the matrix space $\boldsymbol{R}^{\mathcal{M}_1^d \times \mathcal{M}_2^d}$, we define the "d-wise decorrelation bias of F" as being the distance

$$\mathrm{Dec}_D^d(F) = D([F]^d, [F^*]^d).$$

In cases where the canonical idealized version F^* is not explicit, we will use the notation DecF in order to make implicit that F^* is a uniformly distributed random function, and DecP in order to make implicit that F^* is a uniformly distributed random permutation.

For instance, when talking about a block cipher as a random permutation C, the canonical idealized version C^* is a random permutation with uniform distribution. This canonical idealized version should be clear from the context.

Given two random functions F and G from $\mathcal{M}_1$ to $\mathcal{M}_2$ we call "a distinguisher between F and G" any oracle Turing machine $\mathcal{A}^{\mathcal{O}}$ that can send $\mathcal{M}_1$-element queries to the oracle O and receive $\mathcal{M}_2$-element responses, and which finally outputs 0 or 1. In particular, the Turing machine can be probabilistic. In the following, the number of queries to the oracle will be limited to d. The distributions of F and G induce a distribution of $\mathcal{A}^F$ and $\mathcal{A}^G$, thus we can compute the probability that these probabilistic Turing machines output 1. We call the function

$$\mathrm{Adv}_{\mathcal{A}}(F, G) = \Pr[\mathcal{A}^F = 1] - \Pr[\mathcal{A}^G = 1].$$

the advantage $\mathcal{A}^{\mathcal{O}}$ achieves in distinguishing F from G.

We consider the classes Cl_{na}^d (resp. Cl_a^d) of non adaptive (resp. adaptive) distinguishers limited to d queries. Similarly, when F and G are permutations, we also consider the extension Cl_s^d of distinguishers that are limited to d queries but who can query either the function F/G or its inverse F^{-1}/G^{-1}. For any class of distinguishers Cl we will denote

$$\underset{\mathrm{Cl}}{\mathrm{BestAdv}}(F, G) = \max_{\mathcal{A} \in \mathrm{Cl}} \mathrm{Adv}_{\mathcal{A}}(F, G).$$

Lemma 1 (Equivalence between best advantage and decorrelation distance [13, 15]). *For any random functions F and G and any integer d, we have*

$$|||[F]^d - [G]^d|||_\infty = 2 \cdot \underset{\mathrm{Cl}_{na}^d}{\mathrm{BestAdv}}(F, G)$$
$$||[F]^d - [G]^d||_a = 2 \cdot \underset{\mathrm{Cl}_a^d}{\mathrm{BestAdv}}(F, G)$$
$$||[F]^d - [G]^d||_s = 2 \cdot \underset{\mathrm{Cl}_s^d}{\mathrm{BestAdv}}(F, G)$$

where $||.||_a$ and $||.||_s$ are special matrix norms defined in [15] and $|||.|||_\infty$ is the regular infinity associated matrix norm (the maximum of row sums).

Lemma 2 (Multiplicativity). *For any f and g, we denote by $f \circ g$ their composition. For any independent random functions $F_1, \ldots, F_r$, any integer d and any matrix norm we have*

$$\mathrm{DecF}^d(F_1 \circ \cdots \circ F_r) \leq \mathrm{DecF}^d(F_1) \cdots \mathrm{DecF}^d(F_r).$$

For any independent random permutations $C_1, \ldots, C_r$ we have

$$\mathrm{DecP}^d(C_1 \circ \cdots \circ C_r) \leq \mathrm{DecP}^d(C_1) \cdots \mathrm{DecP}^d(C_r).$$

Some known functions have quite small decorrelation biases called decorrelation modules. An example of decorrelation module is the NUT-IV decorrelation module.

Lemma 3 (NUT-IV decorrelation module with $d = 2$ [15]). *For an injection r from $\{0,1\}^m$ to $\mathrm{GF}(q)$ and a surjection π from $\mathrm{GF}(q)$ to $\{0,1\}^m$, it has been shown that the random function F, defined on $\{0,1\}^m$ by*

$$F(x) = \pi(r(K_0) + r(K_1)x)$$

for (K_0, K_1) uniformly distributed in $\{0,1\}^{2m}$, provides quite good decorrelation. Namely,

$$\mathrm{DecF}^2_{||.||_a}(F) \leq 2(q^2.2^{-2m} - 1).$$

For better implementation efficiency, we will only consider prime integers q in this paper. The reader can refer to Noilhan [11] for implementation issues. For instance, DFC uses $q = 2^{64} + 13$ for which we obtain $\mathrm{DecF}^2_{||.||_a}(F) \leq 2^{-58.3}$ (see [7]).

2.2 Basic Tools

The randomness of a cipher constructed using random primitives such as decorrelation modules can be proven using decorrelation theory. In order to deduce an upper bound on the decorrelation bias of the cipher from an upper bound on the decorrelation bias of these primitives, we use the following lemma.

Lemma 4 (Reduction to the randomness of ideal constructions [15]). *Let d be an integer, $F_1, \ldots, F_r, C_1, \ldots, C_s$ be $r + s$ independent random function oracles which are idealized by $F_1^*, \ldots, F_r^*, C_1^*, \ldots, C_s^*$ respectively, where the C_j and C_j^* are permutations. We let $\Omega^{F_1,\ldots,F_r,C_1,\ldots,C_s}$ be an oracle that can access the previous oracles and from each query x define an output $G(x)$. We assume that Ω is such that the number of queries to F_i is limited to some integer a_i, and the number of queries to C_j or C_j^{-1} is limited to b_j in total for any $i = 1, \ldots, r$ and $j = 1, \ldots, s$. We let G^* be the function defined by $\Omega^{F_1^*,\ldots,F_r^*,C_1^*,\ldots,C_s^*}$. We have*

$$\mathrm{Dec}^d_{||.||_a}(G) \leq \sum_{i=1}^{r} \mathrm{Dec}^{a_i d}_{||.||_a}(F_i) + \sum_{j=1}^{s} \mathrm{Dec}^{b_j d}_{||.||_s}(C_j) + \mathrm{Dec}^d_{||.||_a}(G^*)$$

In addition, if the Ω construction defines a permutation G, assuming that computing G^{-1} leads to the same a_i, b_j and c_k limits, we have

$$\mathrm{Dec}^d_{||.||_s}(G) \leq \sum_{i=1}^{r} \mathrm{Dec}^{a_i d}_{||.||_a}(F_i) + \sum_{j=1}^{s} \mathrm{Dec}^{b_j d}_{||.||_s}(C_j) + \mathrm{Dec}^d_{||.||_s}(G^*).$$

Lemma 5 ([16]). *Let d be an integer. Let F be a random function from a set $\mathcal{M}_1$ to a set $\mathcal{M}_2$. We let $\mathcal{X}$ be the subset of $\mathcal{M}_1^d$ of all $(x_1, \ldots, x_d)$ with pairwise different entries. We let F^* be a uniformly distributed random function from $\mathcal{M}_1$ to $\mathcal{M}_2$. We know that for all $x \in \mathcal{X}$ and $y \in \mathcal{M}_2^d$ the value $[F^*]^d_{x,y}$ is the constant $p_0 = (\#\mathcal{M}_2)^{-d}$. We assume there exists a subset $\mathcal{Y} \subseteq \mathcal{M}_2^d$ and two positive real values ϵ_1 and ϵ_2 such that*

- $(\#\mathcal{Y})p_0 \geq 1 - \epsilon_1$
- $\forall x \in \mathcal{X} \quad \forall y \in \mathcal{Y} \quad [F]^d_{x,y} \geq p_0(1 - \epsilon_2)$.

This yields $\mathrm{DecF}^d_{||.||_a}(F) \leq 2\epsilon_1 + 2\epsilon_2$.

This lemma intuitively means that if $[F]^d_{x,y}$ is close to $[F^*]^d_{x,y}$ for all x and almost all y, then the decorrelation bias of F is small. We have a twin lemma for the $||.||_s$ norm. Here, since we can query y as well, the approximation must hold for all x and y.

Lemma 6 ([16]). *Let d be an integer. Let C be a random permutation on a set $\mathcal{M}$. We let $\mathcal{X}$ be the subset of $\mathcal{M}^d$ of all $(x_1, \ldots, x_d)$ with pairwise different entries. We let F^* be a uniformly distributed random function on $\mathcal{M}$. We let C^* be a uniformly distributed random permutation on $\mathcal{M}$. We have*

- *if* $[C]^d_{x,y} \geq [C^*]^d_{x,y}(1 - \epsilon)$ *for all x and y in $\mathcal{X}$ then* $\mathrm{DecP}^d_{||.||_s}(F) \leq 2\epsilon$
- *if* $[C]^d_{x,y} \geq [F^*]^d_{x,y}(1 - \epsilon)$ *for all x and y in $\mathcal{X}$ then* $\mathrm{DecP}^d_{||.||_s}(F) \leq 2\epsilon + 2d^2(\#\mathcal{M})^{-1}$.

2.3 Examples

First this section studies how many rounds are required for Luby-Rackoff's randomness assuming round functions to be random ones. This is related to the "lack of randomness" provided by the upper-level design. The required numbers of rounds for the Feistel scheme and some generalized Feistel schemes are shown in [17, Section 3.2].

Hereafter we use the following notations. I_n denotes the set of all n-bit strings, $\{0,1\}^n$. H_n denotes the set of all $I_n \mapsto I_n$ functions and P_n denotes the set of all such permutations. By $x \in_U X$ we mean that x is drawn randomly and uniformly from a finite set X.

Lemma 7 (Luby-Rackoff 1986 [9]). *Let $(F_1^*, F_2^*, F_3^*, F_4^*) \in_U \left(H_{\frac{m}{2}}\right)^4$ be four independent random functions. We have*

$$\mathrm{DecF}^d_{||.||_a}(\Psi(F_1^*, F_2^*, F_3^*)) \leq 2d^2 \cdot 2^{-\frac{m}{2}}$$
$$\mathrm{DecP}^d_{||.||_a}(\Psi(F_1^*, F_2^*, F_3^*)) \leq 2d^2 \cdot 2^{-\frac{m}{2}}$$
$$\mathrm{DecP}^d_{||.||_s}(\Psi(F_1^*, F_2^*, F_3^*, F_4^*)) \leq 2d^2 \cdot 2^{-\frac{m}{2}}$$

Here $\Psi(F_1,\ldots,F_r)$ is the notation introduced by Luby and Rackoff in order to denote a Feistel scheme where the i-th round function is F_i.[2]

This lemma is tight in the sense that 2 rounds are not enough for pseudorandomness and 3 rounds are not enough for super-pseudorandomness. Indeed, we can make a simple distinguisher against a 2-round Feistel scheme with $d=2$ queries with an advantage equal to $1-2^{-\frac{m}{2}}$ by querying random (a,b) and (a,c) plaintexts and checking that the right half difference is equal to $b\oplus c$. The same holds for super-pseudorandomness with 3 rounds (see Patarin [12]): we can query for the encryption of (a,b) and $(a,b\oplus\delta)$, obtain (x,y) and (x',y') respectively, query for the decryption of $(x,y\oplus\delta)$ and $(x',y'\oplus\delta)$, and check that the obtained left halves are equal.

This lemma can be formally proven by using Lemma 5 and 6. From Lemma 2 and 4 this is generalized for a permutation on $\{0,1\}^m$ consisting of r rounds of Feistel transformations:

$$\mathrm{DecP}^d_{||.||_a}(\Psi(F_1,\ldots,F_r)) \leq \left(2d^2\cdot 2^{-\frac{m}{2}} + 3\max_i \mathrm{DecF}^d_{||.||_a}(F_i)\right)^{\left\lfloor \frac{r}{3}\right\rfloor}$$

$$\mathrm{DecP}^d_{||.||_s}(\Psi(F_1,\ldots,F_r)) \leq \left(2d^2\cdot 2^{-\frac{m}{2}} + 4\max_i \mathrm{DecF}^d_{||.||_a}(F_i)\right)^{\left\lfloor \frac{r}{4}\right\rfloor}$$

for any independent functions $F_1,\ldots,F_r \in H_{\frac{m}{2}}$. This leads to the following conclusions about the regular Feistel scheme with $m=128$.

- The *threshold number of rounds* for achieving the security result is 3 for pseudorandomness and 4 for super-pseudorandomness, when $d \ll 2^{32}$.
- The *theoretical sufficient number of secure rounds* for achieving the decorrelation bias of 2^{-m} is $\frac{\alpha m}{\frac{m}{2}-1-2\log_2 d}$ with $\alpha=3$ for pseudorandomness and $\alpha=4$ for super-pseudorandomness, when $d\ll 2^{32}$. This leads to 9 and 12 rounds, respectively, for $d=2$.
- When using the NUT-IV decorrelation module with $d=2$, $m=128$ and $q=2^{64}+13$ in each round (as for instance DFC), these numbers of rounds provide decorrelation biases less than 2^{-m} for the corresponding norms.

Here we used an arbitrary threshold of 2^{-m} for the decorrelation bias used in order to compare different schemes. Since 2^{-m} yields a level of security given by exhaustive search on m bits, we believe it is a relevant objective criterion for comparing schemes. We also focused on $d=2$ which leads to security against differential and linear cryptanalysis.

3 Several Cases

3.1 CAST256-like Feistel Scheme

CAST-256 is an AES candidate based on a generalized Feistel scheme called "Type-1 transformation" by Zheng-Matsumoto-Imai [17] and denoted by Ψ_1.

[2] In order to be consistent with further schemes, the first round here maps the left half through F_1 and add to the right half.

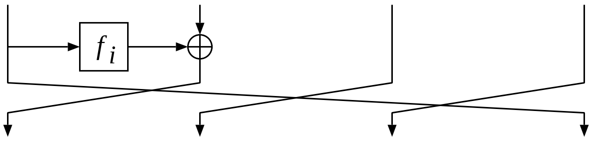

Fig. 1. CAST256-like Feistel Scheme

Formally, we define $\Psi_1 \in H_m$ as $\Psi_1()(x) = x$ and

$$\begin{aligned}&\Psi_1(f_1, \ldots, f_r)(x_1, \ldots, x_k) = \\ &\Psi_1(f_2, \ldots, f_r)(f_1(x_1) + x_2, x_3, x_4, \ldots, x_k, x_1)\end{aligned}$$

for any primitive set $f_1, \ldots, f_r \in H_{\frac{m}{k}}$. Here k is the number of branches and r is the number of rounds.

Lemma 8 (Zheng-Matsumoto-Imai 1989 [17]). *For independent and uniformly distributed random functions $F_1^*, \ldots, F_{2k-1}^* \in_U H_{\frac{m}{k}}$ and an integer d, we have*

$$\mathrm{DecP}^d_{||.||_a}(\Psi_1(F_1^*, \ldots, F_{2k-1}^*)) \leq 2(k-1)d^2 \cdot 2^{-\frac{m}{k}}$$

It can easily be shown that the number of rounds of $2k-1$ for pseudorandomness is actually minimal. For instance, if we take $2k-2$ rounds and $d = 2$, we can submit two chosen plaintexts for which only the input of the rightmost branch has changed. The input difference in this branch will always be equal to the output difference in the second branch, which leads to a distinguisher of advantage $1 - 2^{-k}$.

We however notice that a number of rounds of $k^2 - k$ is not enough for super-pseudorandomness. With $k(k-1)$ rounds, we can decrypt $(y_1, y_2, \ldots, y_k)$ and $(y_1', y_2, \ldots, y_k)$, obtain $(x_1, \ldots, x_k)$ and $(x_1', \ldots, x_k')$ respectively, and check that $x_1 \oplus x_1' = y_1 \oplus y_1'$. This actually shows that the inverse of the Ψ_1 scheme is not pseudorandom unless the number of rounds is very large. Actually, the CAST256 cipher is a construction like

$$\left(\Psi_1(f_r, \ldots, f_{\frac{r}{2}+1})\right)^{-1} \circ \Psi_1(f_1, \ldots, f_{\frac{r}{2}}).$$

We can show that the above attack generalizes to this scheme for $r \leq 4k-6$, that $r = 4k-4$ is enough for pseudorandomness, and that $r = 4k-2$ is enough for super-pseudorandomness.

Proof (sketch). We use Lemma 5 for evaluating $\mathrm{DecF}^d_{||.||_a}$.

For $\mathrm{DecP}_{||.||_a}$ we let $\mathcal{Y}$ be the set of all $y = (y_1, \ldots, y_d)$ where $y_i = (y_i^1, \ldots, y_i^k)$ such that $y_i^j \neq y_{i'}^j$ for $j > 1$ and $i < i'$. We get $\epsilon_1 = (k-1)\frac{d(d-1)}{2}2^{-\frac{m}{k}}$. We then consider the event in which the first entry after the $(k-1)$th round takes pairwise different values for $x_1, \ldots, x_d$. Upper bounding the probability when this event occurs we get $\epsilon_2 = (k-1)\frac{d(d-1)}{2}2^{-\frac{m}{k}}$. Thus $\mathrm{DecF}^d_{||.||_a}(F) \leq 2(k-1)d(d-1)2^{-\frac{m}{k}}$.

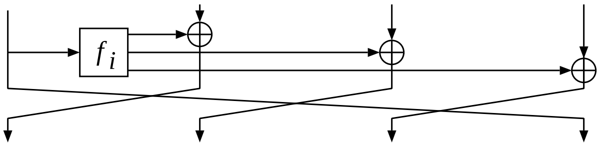

Fig. 2. MARS-like Feistel Scheme

Here, ϵ_2 is evaluated as the number of unexpected equalities between two outputs from a single circuit of depth $k-1$ with k inputs and internal F_j^* and additions times the probability it occurs, which is at most the depth $k-1$ times $2^{-\frac{m}{k}}$.

Now to get DecP from DecF, from $\mathrm{DecF}^d_{||.||_a}(C^*) \leq d(d-1)2^{-m}$ and the triangular inequality we have

$$\mathrm{DecP}^d_{||.||_a}(F) \leq \mathrm{DecF}^d_{||.||_a}(F) + \mathrm{DecP}^d_{||.||_a}(F^*) \leq \mathrm{DecF}^d_{||.||_a}(F) + d^2 2^{-m}.$$

We then notice that the obtained upper bound for $\mathrm{DecF}^d_{||.||_a}$ can be written $\mathrm{DecF}^d_{||.||_a}(F) \leq Ad(d-1)2^{-\frac{m}{k}}$ for some $A \geq 2$. For $d \leq A2^{m-\frac{m}{k}}$ we thus obtain $\mathrm{DecP}^d_{||.||_a}(F) \leq Ad^2 2^{-\frac{m}{k}}$. For larger d, this bound is greater than $A^3 2^{m\left(2-\frac{3}{k}\right)}$ which is greater than 8 since $m \geq k \geq 2$. Since $\mathrm{DecP}^d_{||.||_a}(F)$ is always less than 2, the bound is thus still valid. □

Thus the required number of rounds for the CAST256-like scheme is proven to be $2k-1$, where k is the number of branches. That is, the required numbers of rounds for the Feistel scheme and the CAST256-like scheme are 3 and 7, respectively.

This leads to the following conclusions about the CAST256-like scheme with $k = 4$ branches and $m = 128$.

- The *threshold number of rounds* is 7 for pseudorandomness when $d \ll 2^{16}$. For super-pseudorandomness, this threshold is larger than 13.
- For $d = 2$, the *theoretical sufficient number of secure rounds* is 35 for pseudorandomness.
- For the NUT-IV decorrelation module with $d = 2$, $m = 128$ and $q = 2^{32}+15$, the sufficient number of rounds is 42 pseudorandomness.

3.2 MARS-like Feistel Scheme

Similarly, we define the MARS-like generalized Feistel scheme denoted by $\Psi_1' \in H_m$ as $\Psi_1'()(x) = x$ and

$$\begin{aligned}&\Psi_1'(f_1, \ldots, f_r)(x_1, \ldots, x_k) = \\ &\Psi_1'(f_2, \ldots, f_r)(f_1^2(x_1) + x_2, f_1^3(x_1) + x_3, \ldots, f_1^k(x_1) + x_k, x_1)\end{aligned}$$

where $f_i = (f_i^2, \ldots, f_i^k)$, $f_i^2, \ldots, f_i^k \in H_{\frac{m}{k}}$.

Lemma 9. *For independent uniformly distributed random functions $F_i^* \in_U H_{\frac{m}{k}}$ for $i = 1, \ldots, 2k$ and $j = 2, \ldots, k$ and an integer d, we have*

$$\mathrm{DecP}^d_{||\cdot||_a}(\Psi_1'(F_1^*, \ldots, F_{k+1}^*)) \leq 2d^2 \cdot 2^{-\frac{m}{k}}$$
$$\mathrm{DecP}^d_{||\cdot||_s}(\Psi_1'(F_1^*, \ldots, F_{2k}^*)) \leq 2d^2 \cdot 2^{-\frac{m}{k}}$$

It can easily be shown that the number of rounds of $k+1$ for pseudorandomness is actually minimal since a difference in the last input branch only remains unchanged after k rounds. Similarly, for $2k-1$ rounds, we can merge the first $k-1$ branches and consider that we have a regular 3-round Feistel scheme, and we can apply the same attack for proving it is not super-pseudorandom.

Proof (sketch). Using Lemma 5 we let $\mathcal{Y}$ be the set of all $(y_1, \ldots, y_d)$ such that $y_i^k \neq y_j^k$ for $i \neq j$. We get $\epsilon_1 = \frac{d(d-1)}{2} 2^{-\frac{m}{k}}$. We focus on the event that the first output after $k-1$ rounds leads to no collision. We get $\epsilon_2 = \frac{d(d-1)}{2} 2^{-\frac{m}{k}}$.

For $\mathrm{DecP}^d_{||\cdot||_s}$ we use the same event. □

This leads to the following conclusions about the MARS-like scheme with $k = 4$ branches and $m = 128$.

- The *threshold number of rounds* is 5 for pseudorandomness and 8 for super-pseudorandomness, when $d \ll 2^{16}$.
- For $d = 2$, the *theoretical sufficient number of secure rounds* is 25 for pseudorandomness and 40 for super-pseudorandomness.
- For the NUT-IV decorrelation module with $d = 2$, $m = 128$ and $q = 2^{32} + 15$, the sufficient number of rounds is as for the ideal case.

3.3 RC6-like Feistel Scheme

The RC6 block cipher is designed to be secure by mixing operations that are efficiently implemented on most modern processors.

One controversial additional operation is the data dependent rotation. Such a scheme cannot provide pseudorandomness nor super-pseudorandomness.[3] Indeed, the attack in Gilbert et al. [6] exhibits an efficient polynomial time distinguisher.

However, we can consider RC6′, a transformation of RC6 WITHOUT the data dependent rotations. The structure of RC6′ can be regarded as a generalized Feistel scheme, which is similar to "Type-2 transformation" named by Zheng-Matsumoto-Imai [17] assuming that primitives are independent random functions. Formally, as the RC6′-like Feistel scheme $\Psi_2 \in H_m$ is defined for k even and r a multiple of $\frac{k}{2}$, by $\Psi_2()(x) = x$ and

$$\Psi_2(f_1, \ldots, f_r)(x_1, \ldots, x_k) =$$
$$\Psi_2(f_{\frac{k}{2}+1}, \ldots, f_r)(x_2, f_2(x_4) + x_3, \ldots, x_{k-2}, f_{\frac{k}{2}}(x_k) + x_{k-1}, x_k, f_1(x_2) + x_1),$$

[3] As was mentioned by Joux during the third Advanced Encryption Standard workshop, although Iwata and Kurosawa had claimed the opposite two days before at the FSE00 workshop [8].

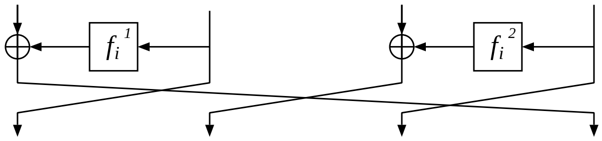

Fig. 3. RC6′-like Feistel Scheme

where $f_1, \ldots, f_r \in H_{\frac{m}{k}}$. We consider this as r rounds which are processed in bunch of $\frac{k}{2}$ parallel rounds.

Lemma 10. *For independent uniformly distributed random functions $F_1^*, \ldots, F_{k^2}^* \in_U H_{\frac{m}{k}}$ and an integer d, we have*

$$\mathrm{DecP}^d_{||\cdot||_a}(\Psi_2(F_1^*, \ldots, F^*_{\frac{k}{2}(k+1)})) \leq \frac{k^2}{2} d^2 \cdot 2^{-\frac{m}{k}}$$
$$\mathrm{DecP}^d_{||\cdot||_s}(\Psi_2(F_1^*, \ldots, F^*_{k^2})) \leq \frac{k^2}{2} d^2 \cdot 2^{-\frac{m}{k}}$$

It can easily be shown that the number of rounds of $\frac{k}{2}(k+1)$ for pseudorandomness is actually minimal. Tightness of the k^2 bound for super-pseudorandomness is still open. (We already know that it is tight for $k = 2$.)

Proof (sketch). Similarly, we use Lemma 9 for evaluating $\mathrm{DecP}^d_{||\cdot||_a}$. For Ψ_2 we let $\mathcal{Y}$ be the set of all y such that $y_i^j \neq y_{i'}^j$ for odd j and $i < i'$. We get $\epsilon_1 = \frac{k}{2} \times \frac{d(d-1)}{2} 2^{-\frac{m}{k}}$. We consider the event in which all even entries after the $(k-1)$th bunch of rounds takes pairwise different values for $x_1, \ldots, x_d$. We get $\epsilon_2 = \frac{k}{2}(k-1) \times \frac{d(d-1)}{2} 2^{-\frac{m}{k}}$. Thus $\mathrm{DecF}^d_{||\cdot||_a}(F) \leq \frac{k^2}{2} d(d-1) 2^{-\frac{m}{k}}$. For $\mathrm{DecP}^d_{||\cdot||_s}$, we add $k-1$ more bunch of rounds and study the probability that we get $\mathcal{Y}$ if we invert them on $y_1, \ldots, y_d$. The result comes from Lemma 6. □

This leads to the following conclusions about the RC6′-like scheme with $k = 4$ branches and $m = 128$.

- The *threshold number of rounds* is 5 for pseudorandomness and between 5 and 8 for super-pseudorandomness, when $d \ll 2^{16}$.
- For $d = 2$, the *theoretical sufficient number of secure rounds* is 25 for pseudorandomness and between 25 and 40 for super-pseudorandomness.
- For the NUT-IV decorrelation module with $d = 2$, $m = 128$ and $q = 2^{32} + 15$, the sufficient number of rounds as the ideal case.

3.4 Square-like Scheme

In this paper we discuss only the Rijndael scheme. The pseudorandomness of other Square-like schemes will be described in the full paper. Let us formalize the Rijndael scheme on k^2 values by

$$\begin{aligned}&\Sigma(f_1,\ldots,f_r)(x_1,\ldots,x_{k^2}) = \\ &\Sigma(f_2,\ldots,f_r)(\mathrm{MixCol}(\mathrm{ShiftRow}(f_1^1(x_1),\ldots,f_1^{k^2}(x_{k^2}))))\end{aligned}$$

where $f_i = (f_i^1,\ldots,f_i^{k^2})$, $f_i^1,\ldots,f_i^{k^2} \in H_{\frac{m}{k^2}}$, the ShiftRow transformation is a fixed linear transformation on the rows of a $k \times k$ matrix which consists in mixing them, and the MixCol transformation is a fixed linear transformation on the columns [3].

Lemma 11. *For independent uniformly distributed random functions $F_1^*,\ldots,$ F_5^* and an integer d, we have*

$$\begin{aligned}\mathrm{DecP}^d_{||.||_a}(\Sigma(F_1^*,\ldots,F_3^*)) \le 2k^2d^2\cdot 2^{-\frac{m}{k^2}} \\ \mathrm{DecP}^d_{||.||_s}(\Sigma(F_1^*,\ldots,F_5^*)) \le 2k^2d^2\cdot 2^{-\frac{m}{k^2}}\end{aligned}$$

Thus achieving decorrelation to the order $d \geq \frac{1}{k\sqrt{2}}2^{\frac{m}{2k^2}}$ does not seem possible with this design. (For $m = 128$ and $k = 4$, this is $d = 2\sqrt{2}$.)

It can easily be shown that the number of rounds of 3 for pseudorandomness is actually minimal. The tightness of the 5 bound depends on the instance of the cipher.

Proof (sketch). We use Lemma 9 for evaluating $\mathrm{DecP}^d_{||.||_a}$. We let $\mathcal{Y}$ be the set of all $y = (y_1,\ldots,y_d)$ that take different values on all positions before the last MixCol and ShiftRow transformations. We have $\epsilon_1 = k^2\frac{d(d-1)}{2}2^{-\frac{m}{k^2}}$. We consider the event that after two rounds we obtain different values on all positions. Provided that the MixCol transformation has good diffusion properties we obtain $\epsilon_2 = k^2\frac{d(d-1)}{2}2^{-\frac{m}{k^2}}$. □

This leads to the following conclusions about the Rijndael scheme with $k^2 = 4^2$ branches and $m = 128$.

- The *threshold number of rounds* is 3 for pseudorandomness and between 3 and 5 for super-pseudorandomness, when $d < 3$.
- For $d = 2$, the *theoretical sufficient number of secure rounds* is 384 for pseudorandomness and between 384 and 640 for super-pseudorandomness.
- For the NUT-IV decorrelation module with $d = 2$, $m = 128$ and $q = 2^8 + 1$, the bounds of decorrelation theory cannot guaranty any low decorrelation bias for any number of rounds.

4 Conclusion

We studied the randomness provided by several schemes used in block ciphers. We focused on the schemes for AES candidates in particular (see Table 1). The randomness so discovered is a good measure for evaluating the security from a

Table 1. Randomness of several schemes (when $d = 2, k = 4, m = 128$)

Scheme	Feistel	CAST256-like	MARS-like	RC6′-like	Rijndael
Threshold number of rounds for p.r.	3	7	5	5	3
sufficient number of rounds for p.r. (ideal)	9	35	25	25	384
sufficient number of rounds for p.r. (NUT-IV)	9	42	25	25	∞
Threshold number of rounds for s.p.r.	4	≥ 13	8	5–8	3–5
sufficient number of rounds for s.p.r. (ideal)	12		40	25–40	384–640
sufficient number of rounds for s.p.r. (NUT-IV)	12		40	25–40	∞
Example	Twofish, DFC	CAST-256	MARS		Rijndael

Note: "p.r." and "s.p.r." mean pseudorandomness and super-pseudorandomness, respectively.

randomness viewpoint but the readers should take care to note that it doesn't show the actual security of a cipher based on one of the schemes. To study the intrinsic security provided by the general schemes, we decomposed the ciphers into a general scheme and internal primitives, ignoring the components that we considered do not affect its randomness. We also assumed that internal primitives are ideal random ones.

The results in Table 1 show that the regular Feistel scheme is the best in that it requires the fewest number of rounds for pseudorandomness and super-pseudorandomness. However, when comparing the randomness of several schemes we should take account of the computational cost of random primitives. For example, for the Feistel scheme we assume the random functions on $\{0,1\}^{64}$, and for the CAST256-like[4], MARS-like, and RC6-like schemes, we assume the random functions on $\{0,1\}^{32}$, whose computational cost is much cheaper than the former. Under the same assumption of the computational cost of random functions on $\{0,1\}^{32}$, the MARS-like scheme is the best. Table 1 separates the schemes according to the size of the internal random functions.

Our results show that the schemes that use random primitives with smaller input/output sizes are less secure, which is not surprising because the randomness bias is larger in these cases. We should interpret these conclusions with great care. Indeed, our results do not mean that Rijndael (or Serpent[5]) is not secure, or less secure than regular Feistel schemes. Rather they mean that the latter can benefit from stronger security arguments: we can prove that an efficient attack against — say Twofish — must use an unexpected property of the round function, whereas an attack against Serpent may hold for any set of (random) S-boxes.

[4] Table 1 considers the Ψ_1 structure only and not the $\Psi_1^{-1} \circ \Psi_1$ scheme on which CAST256 is based. This latter scheme increases the threshold number of rounds for p.r. to 12.

[5] A preliminary study suggested that the Serpent scheme requires too many rounds for randomness, because the size of primitives is too small (4 bits).

References

1. E. Biham, A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, 1993.
2. L. Carter, M. Wegman. Universal Classes of Hash Functions. *Journal of Computer and System Sciences*, vol.18, pp.143–154, 1979.
3. J. Daemen, V. Rijmen. AES Proposal: Rijndael.
URL: `http://www.esat.kuleuven.ac.be/~rijmen/rijndael/`
4. H. Feistel. Cryptography and Computer Privacy. *Scientific American*, vol. 228, pp. 15–23, 1973.
5. H. Gilbert, M. Girault, P. Hoogvorst, F. Noilhan, T. Pornin, G. Poupard, J. Stern, S. Vaudenay. Decorrelated Fast Cipher: an AES Candidate. (Extended Abstract.) In *Proceedings from the First Advanced Encryption Standard Candidate Conference*, National Institute of Standards and Technology (NIST), August 1998.
6. H. Gilbert, H. Handschuh, A. Joux, S. Vaudenay. A Statistical Attack on RC6. To appear in the proceedings of FSE00.
7. L. Granboulan, P. Nguyen, F. Noilhan, S. Vaudenay. DFCv2. To appear in the proceedings of SAC00.
8. T. Iwata, K. Kurosawa. On the Pseudorandomness of AES Finalists — RC6, Serpent, MARS and Twofish. To appear in the proceedings of FSE00.
9. M. Luby, C. Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM Journal on Computing*, vol. 17, pp. 373–386, 1988.
10. M. Matsui. The First Experimental Cryptanalysis of the Data Encryption Standard. In *Advances in Cryptology CRYPTO'94*, Santa Barbara, California, U.S.A., Lectures Notes in Computer Science 839, pp. 1–11, Springer-Verlag, 1994.
11. F. Noilhan. Software Optimization of Decorrelation Module. In *Selected Areas in Cryptography*, Kingston, Ontario, Canada, Lectures Notes in Computer Science 1758, pp. 175–183, Springer-Verlag, 2000.
12. J. Patarin. *Etude des Générateurs de Permutations Basés sur le Schéma du D.E.S.*, Thèse de Doctorat de l'Université de Paris 6, 1991.
13. S. Vaudenay. Provable Security for Block Ciphers by Decorrelation. In *STACS 98*, Paris, France, Lectures Notes in Computer Science 1373, pp. 249–275, Springer-Verlag, 1998.
14. S. Vaudenay. On the Lai-Massey Scheme. *Advances in Cryptology — ASIACRYPT'99*, Singapore, Lecture Notes in Computer Science 1716, pp.8–19, Springer-Verlag, 1999.
15. S. Vaudenay. Adaptive-Attack Norm for Decorrelation and Super-Pseudorandomness. Technical report LIENS-99-2, Ecole Normale Supérieure, 1999. In *Selected Areas in Cryptography*, Kingston, Ontario, Canada, Lectures Notes in Computer Science 1758, pp. 49–61, Springer-Verlag, 2000.
16. S. Vaudenay. On Provable Security for Conventional Cryptography. Invited talk. To appear in the proceedings of ICISC' 99, LNCS, Springer-Verlag.
17. Y. Zheng, T. Matsumoto, H. Imai. On the Construction of Block Ciphers Provably Secure and Not Relying on Any Unproved Hypotheses (Extended Abstract). *Advances in Cryptology — CRYPTO'89*, Santa Barbara, California, U.S.A., Lecture Notes in Computer Science 435, pp.461–480, Springer-Verlag, 1990.