Cryptanalysis of a Generalized Unbalanced Feistel Network Structure*

Ruilin Li[1], Bing Sun[1], Chao Li[1,2], and Longjiang Qu[1,3]

[1]Department of Mathematics and System Science, Science College,
National University of Defense Technology, Changsha, 410073, China
securitylrl@gmail.com, happy_come@163.com
[2]State Key Laboratory of Information Security, Institute of Software,
Chinese Academy of Sciences, Beijing, 100190, China
lichao_nudt@sina.com
[3]National Mobile Communications Research Laboratory,
Southeast University, Nanjing, 210096, China
ljqu_happy@hotmail.com

Abstract. This paper reevaluates the security of GF-NLFSR, a new kind of generalized unbalanced Feistel network structure that was proposed at ACISP 2009. We show that GF-NLFSR itself reveals a very slow diffusion rate, which could lead to several distinguishing attacks. For GF-NLFSR containing n sub-blocks, we find an n^2-round integral distinguisher by algebraic methods and further use this integral to construct an $(n^2 + n - 2)$-round impossible differential distinguisher. Compared with the original $(3n - 1)$-round integral and $(2n - 1)$-round impossible differential, ours are significantly better.

Another contribution of this paper is to introduce a kind of non-surjective attack by analyzing a variant structure of GF-NLFSR, whose provable security against differential and linear cryptanalysis can also be provided. The advantage of the proposed non-surjective attack is that traditional non-surjective attack is only applicable to Feistel ciphers with non-surjective (non-uniform) round functions, while ours could be applied to block ciphers with bijective ones. Moreover, its data complexity is $\mathcal{O}(l)$ with l the block length.

Keywords: block ciphers, generalized unbalanced Feistel network, integral attack, impossible differential attack, non-surjective attack

1 Introduction

Differential cryptanalysis (DC) [6] and linear cryptanalysis (LC) [23] are the two most powerful known attacks on block ciphers since 1990s. For a new block

* The work in this paper is supported by the Natural Science Foundation of China (No: 60803156), the open research fund of State Key Laboratory of Information Security (No: 01-07) and the open research fund of National Mobile Communications Research Laboratory of Southeast University (No: W200807).

cipher algorithm, designers must guarantee that it can resist these two attacks. However, even the security against DC and LC can be proved, the algorithm may suffer other attacks, such as truncated differential attack [13], higher-order differential attack [13, 18], impossible differential attack [4, 14], boomerang attack [27], amplified boomerang attack [16], rectangle attack [5], integral attack [15], interpolation attack [12], non-surjective attack [24], algebraic attack [8], related-key attack [3], slide attack [1] and so on. Among these methods, integral attack and impossible differential attack are of special importance. Take the well-known 128-bit version block cipher Rijndael as an example, six rounds is sufficient for resisting DC and LC. However, by integral attack or impossible differential attack, one can break six, seven, even eight rounds [9, 11, 20, 29].

Integral cryptanalysis [15], which is especially well-suited for analyzing ciphers with primarily bijective components, was proposed by Knudsen *et al.*. In fact, it is a more generalization of Square attack [9], Saturation attack [19] and Multiset attack [2] proposed by Daemen *et al.*, Lucks, and Biryukov *et al.*, respectively. These methods exploit the simultaneous relationship between many encryptions, in contrast to differential cryptanalysis, where only pairs of encryptions are considered. Consequently, integral cryptanalysis applies to a lot of ciphers which are not vulnerable to DC and LC. These features have made integral an increasingly popular tool in recent cryptanalysis work.

The concept of using impossible differentials (differentials with probability 0) to retrieve the secret key of block ciphers was firstly introduced by Knudsen [14] against the DEAL cipher and further by Biham *et al.* [4] to attack Skipjack. Unlike differential cryptanalysis which recovers the right key through the obvious advantage of a high probability differential (differential characteristic), impossible differential cryptanalysis is a sieving attack that excludes all the wrong candidate keys using impossible differentials. Since its emergence, impossible differential cryptanalysis has been applied to attack many well-known block ciphers [20, 21, 28, 29].

Non-surjective attack [24] was introduced by Rijmen *et al.* and it is applicable to Feistel ciphers with non-surjective, or more generally, non-uniform round functions such as CAST and LOKI 91. If the round function of Feistel ciphers is non-surjective (non-uniform), then by analyzing the statistical bias of some expression derived from the round function, one can apply a key recovery attack. However, if the round function is a surjective (uniform) one, it is impossible to apply this kind of non-surjective attack.

At ACISP 2009, Choy *et al.* proposed a new block cipher structure called n-cell GF-NLFSR [7], which is a kind of generalized unbalanced Feistel network [26] containing n sub-blocks. The advantages of this structure are that it allows parallel computations for encryption and that it can provide provable security against DC and LC, given that the round function is bijective. Meanwhile, the designers show the existence of a $(3n-1)$-round integral distinguisher and a $(2n-1)$-round impossible differential distinguisher. In the same paper, a new block cipher Four-Cell is designed as an application of the theoretical model of 4-cell GF-NLFSR.

Main Contribution. (1) We demonstrate that GF-NLFSR itself reveals a very slow diffusion rate, which could lead to several distinguishing attacks. We especially apply *algebraic methods* to find integral distinguishers in n-cell GF-NLFSR. In this method, plaintexts of special forms as well as their indeterminate states are treated as polynomial functions over finite fields, and in many cases, more precise information among these states could be obtained, which would lead to a better distinguisher.

Our cryptanalytic results show that, for n-cell GF-NLFSR, there exists an n^2-round integral distinguisher, which could be extended to an (n^2+n-2)-round higher-order integral distinguisher. Furthermore, by studying the relationship between integral and truncated differential, an (n^2+n-2)-round impossible differential distinguisher could be constructed. These distinguishers are significantly better than the original ones.

(2) We introduce a kind of *non-surjective attack* by analyzing a variant structure of GF-NLFSR, whose provable security against DC and LC can also be provided. The advantage of the proposed attack is that traditional non-surjective attack is only applicable to Feistel ciphers with non-surjective (non-uniform) round functions, while ours could be applied to block ciphers with bijective ones. Moreover, its data complexity is $\mathcal{O}(l)$ with l the block length.

Outline. We begin with a brief description of n-cell GF-NLFSR in Section 2. Encryption properties of n-cell GF-NLFSR by every n rounds are studied in Section 3. The existence of n^2-round integral distinguisher and (n^2+n-2)-round impossible differential distinguisher are shown in Section 4 and Section 5, respectively. Section 6 presents a kind of non-surjective attack by analyzing a variant structure of GF-NLFSR. Section 7 contains results of the experiment with the proposed non-surjective attack on a toy cipher, and finally Section 8 is the conclusion.

2 Description of n-cell GF-NLFSR

As shown in Fig. 1, assume the input, output and round key to the i-th round of n-cell GF-NLFSR are $(x_0^{(i)}, x_1^{(i)}, \ldots, x_{n-1}^{(i)}) \in \mathbb{F}_{2^b}^n$, $(x_0^{(i+1)}, x_1^{(i+1)}, \ldots, x_{n-1}^{(i+1)}) \in \mathbb{F}_{2^b}^n$, and $K_i = (k_i, k_i')$, then the round transformation can be described as follow:

$$(x_0^{(i)}, x_1^{(i)}, \ldots, x_{n-2}^{(i)}, x_{n-1}^{(i)}) \mapsto (x_0^{(i+1)}, x_1^{(i+1)}, \ldots, x_{n-2}^{(i+1)}, x_{n-1}^{(i+1)}),$$

where

$$\begin{cases} x_l^{(i+1)} = x_{l+1}^{(i)}, \quad \text{if } l = 0, 1, \ldots, n-2 \\ x_{n-1}^{(i+1)} = F(x_0^{(i)}, K_i) \oplus x_1^{(i)} \oplus x_2^{(i)} \oplus \ldots \oplus x_{n-1}^{(i)} \end{cases}$$

and $F(\cdot, K_i) \triangleq F_{K_i}(\cdot)$ is a permutation on $\mathbb{F}_{2^b}$.

From [7], this kind of generalized unbalanced Feistel network can provide its provable security against DC and LC, which is summarized in the following proposition.

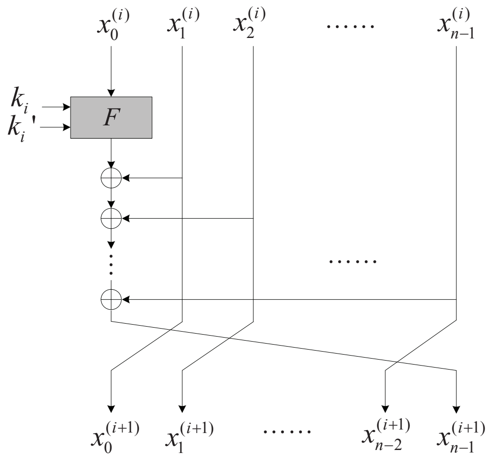

Fig. 1. The i-th round transformation of n-cell GF-NLFSR

Proposition 1. [7] *Let the round function of* n-cell GF-NLFSR $F : \mathbb{F}_{2^b} \times \mathbb{F}_{2^b} \times \Omega \to \mathbb{F}_{2^b}$ *be of the form* $F(x, k_i, k_i') = f(x \oplus k_i, k_i')$*, where* $f : \mathbb{F}_{2^b} \times \Omega \to \mathbb{F}_{2^b}$ *is bijective for all fixed* $k_i' \in \Omega$*. If the maximum differential (linear hull) probability of* f *satisfies* $DP(LP)_{max}(f) \leq p(q)$*, then the differential (linear hull) probability of the* $(n+1)$*-round encryption is upper bounded by* $p^2(q^2)$*.*

3 Encryption Property of n-cell GF-NLFSR

In this section, we study the encryption property of n-cell GF-NLFSR by every n rounds. From now on, the round function $F_{K_i}(x)$ is treated as a permutation polynomial over $\mathbb{F}_{2^b}$.

Firstly, according to the definition of n-cell GF-NLFSR, the following result could be obtained.

Proposition 2. *Let* $(x_0, x_1, \ldots, x_{n-1})$ *be the input of the* i*-th round of* n-cell GF-NLFSR*, and* $(y_0, y_1, \ldots, y_{n-1})$ *be the output of the* $(i+n-1)$*-th round, then*

$$\begin{cases} y_0 = F_{K_i}(x_0) \oplus x_1 \oplus x_2 \oplus \ldots \oplus x_{n-1} \\ y_m = F_{K_{i+m-1}}(x_{m-1}) \oplus F_{K_{i+m}}(x_m) \oplus x_m, \quad \textit{if } 1 \leq m \leq n-1 \end{cases}$$

and

$$\bigoplus_{j=0}^{n-1} y_j = F_{K_{i+n-1}}(x_{n-1}).$$

Proposition 2 can be verified directly by the encryption procedure of n-cell GF-NLFSR, based on which we could deduce the following proposition.

Proposition 3. *Let the input of* n-cell GF-NLFSR *be* $(x, c_1, \ldots, c_{n-1})$*, where* x *is a variable and each* c_i *is some constant with* $1 \le i \le n-1$*, let the output of the* r*-th round be* $\left(y_0^{(r)}(x), y_1^{(r)}(x), \ldots, y_{n-1}^{(r)}(x)\right)$*, and* $1 \le m \le n-1$*, then*

(1) $y_i^{(m \times n)}(x)$ *is a permutation polynomial over* $\mathbb{F}_{2^b}$ *if* $i = m$*,*

(2) $y_i^{(m \times n)}(x)$ *is a constant if* $i > m$*.*

Table 1 is the encryption results of every n rounds of n-cell GF-NLFSR when plaintexts are of the form $(x, c_1, \ldots, c_{n-1})$ as described in Proposition 3. Note that the first column denotes the round number, and each of the other columns represents the corresponding output sub-block. The letter C denotes some constant which could be different from each other. $P_m(x)$ is some permutation polynomial over $\mathbb{F}_{2^b}$ with $1 \le m \le n-1$, and those blank cells (elements under the diagonal) indicate that their behaviors are unknown.

An immediate conclusion, from Proposition 3 and Table 1, is that the diffusion rate of n-cell GF-NLFSR is very slow, since the input variable x needs at least $(n-1) \times n$ rounds to influence the last (rightmost) sub-block of the output.

Table 1. Output of every n rounds of n-cell GF-NLFSR

0	x	C	C	$\ldots$	C	$\ldots$	C	C
n		$P_1(x)$	C	$\ldots$	C	$\ldots$	C	C
$\vdots$			$\ddots$		$\vdots$		$\vdots$	$\vdots$
$(m-1) \times n$				$P_{m-1}(x)$	C	$\ldots$	C	C
$m \times n$					$P_m(x)$	$\ldots$	C	C
$\vdots$						$\ddots$	$\vdots$	$\vdots$
$(n-2) \times n$							$P_{n-2}(x)$	C
$(n-1) \times n$								$P_{n-1}(x)$

4 Integral Distinguisher of GF-NLFSR

4.1 Preliminaries

To apply integral cryptanalysis, one should first find an integral distinguisher of the reduced-round cipher, then apply the key recovery attack. In this section, we show how to construct an n^2-round integral distinguisher of n-cell GF-NLFSR by using algebraic techniques.

Firstly, recall that most traditional methods in finding integral distinguishers are based on the so-called *empirical methods*. They firstly treat each part of plaintexts with special forms as active or passive state (see definitions below), then study the property (active, passive or balanced) of its corresponding intermediate state after passing through several encryption rounds.

Definition 1. *A set* $\{a_i | a_i \in \mathbb{F}_{2^b}, 0 \le i \le 2^b - 1\}$ *is* active*, if for any* $0 \le i < j \le 2^b - 1$*,* $a_i \ne a_j$*. We use* **A** *to denote the active set.*

Definition 2. *A set* $\{a_i | a_i \in \mathbb{F}_{2^b}, 0 \le i \le 2^b - 1\}$ *is* passive *or* constant*, if for any* $0 < i \le 2^b - 1$*,* $a_i = a_0$*. We use* **C** *to denote the passive set.*

Definition 3. *A set* $\{a_i | a_i \in \mathbb{F}_{2^b}, 0 \le i \le 2^b - 1\}$ *is* balanced*, if the* XOR*-sum of all element of the set is 0, that is* $\oplus_{i=0}^{2^b-1} a_i = 0$*. We use* **B** *to denote the balanced set.*

Moreover, three principles are widely used when applying empirical methods: (1) An active set remains active after passing a bijective transform. (2) The linear combination of several active/balanced sets is a balanced set. (3) The property of a balanced set after passing a nonlinear transformation is generally unknown.

Obviously, the third one is the bottleneck of empirical methods, thus if one could determine the property of a balanced set after it passes a nonlinear transformation, integral distinguisher with more rounds can be constructed.

4.2 n^2-Round Integral Distinguisher of n-cell GF-NLFSR

By using the empirical method, the designers presented the following $(3n-1)$-round integral distinguisher:

$$(A, C, C, \ldots, C) \to (C, ?, ?, \ldots, ?),$$

where A is active in $\mathbb{F}_{2^b}$, C is constant in $\mathbb{F}_{2^b}$, and ? is unknown.

Now we describe the newly constructed n^2-round integral in the following theorem, the proof is based on algebraic methods. See Appendix B for a 16-round integral distinguisher of 4-cell GF-NLFSR as an example.

Theorem 1. *There is an* n^2*-round integral distinguisher of* n-cell GF-NLFSR*:*

$$(A, C, \ldots, C) \to (S_0, S_1, \ldots, S_{n-1}),$$

where A *is active,* C *is constant and* $(S_0 \oplus S_1 \oplus \ldots \oplus S_{n-1})$ *is active.*

Proof. Let the input of n-cell GF-NLFSR be $(x, c_1, \ldots, c_{n-1})$ and the output of the $((n-1) \times n)$-th round be

$$\left(y_0^{((n-1)\times n)}(x), y_1^{((n-1)\times n)}(x), \ldots, y_{n-1}^{((n-1)\times n)}(x)\right),$$

then $y_{n-1}^{((n-1)\times n)}(x)$ is a permutation polynomial by Proposition 3.

Assume the output of the n^2-round is

$$\left(y_0^{(n^2)}(x), y_1^{(n^2)}(x), \ldots, y_{n-1}^{(n^2)}(x)\right),$$

according to Proposition 2,

$$y_0^{(n^2)}(x) \oplus y_1^{(n^2)}(x) \oplus \ldots \oplus y_{n-1}^{(n^2)}(x) = F_{K_{n^2}}\left(y_{n-1}^{((n-1)\times n)}(x)\right).$$

Since $y_{n-1}^{((n-1)\times n)}(x)$ is a permutation polynomial, so is $F_{K_{n^2}}\left(y_{n-1}^{((n-1)\times n)}(x)\right)$, which ends the proof. $\square$

From the idea of higher-order integral [15], the above n^2-round integral can be extended to an (n^2+n-2)-round higher-order one.

Theorem 2. *There is an (n^2+n-2)-round higher-order integral distinguisher of n-cell GF-NLFSR:*

$$(A_0, A_1, \ldots, A_{n-2}, C) \to (S_0, S_1, \ldots, S_{n-1}),$$

where $(A_0, A_1, \ldots, A_{n-2})$ is active in $\mathbb{F}_{2^b}^{n-1}$, C is constant and $(S_0 \oplus S_1 \oplus \ldots \oplus S_{n-1})$ is balanced.

Proof. First, according to bijective property of the encryption structure of n-cell NLFSR, if the input is $(x_0, x_1, \ldots, x_{n-2}, c)$, where $(x_0, x_1, \cdots, x_{n-1})$ is active in $\mathbb{F}_{2^b}^{n-1}$, $c \in \mathbb{F}_{2^b}$ is constant, after $n-2$ rounds encryption, the intermediate state must be $(y_0, c, y_2, \ldots, y_{n-1})$, where $(y_0, y_2, \ldots, y_{n-1})$ is active in $\mathbb{F}_{2^b}^{n-1}$.

Next, let's focus on the set containing these $2^{(n-1)b}$ intermediate states after $n-2$ rounds encryption. Fix $(y_2, y_3, \ldots, y_{n-1}) \in \mathbb{F}_{2^b}^{n-2}$, we thus get a structure with 2^b elements, which is the input of the n^2-round integral distinguisher as shown in Theorem 1(From now on, we call this structure a Λ set).

Now, these $2^{(n-1)b}$ intermediate states can be divided into $2^{(n-2)b}$ indistinguishable Λ sets. When each Λ set passes through the n^2 rounds encryption, the XOR sum of the n sub-blocks of outputs is active (thus balanced) in $\mathbb{F}_{2^b}$. Consequently, the XOR sum of the n sub-blocks of outputs for these $2^{(n-2)b}$ indistinguishable Λ sets is balanced. Let $E_j^{(i)}(\cdot)$ denote the j-th sub-block after i rounds encryption of the input, then we can explain the higher-order integral distinguisher as follows:

$$\begin{aligned}
&\bigoplus_{x_0,x_1,\ldots,x_{n-2}} \bigoplus_{j=0}^{n-1} E_j^{(n^2+n-2)}(x_0, x_1, \ldots, x_{n-2}, c) \\
&= \bigoplus_{y_0,y_2,\ldots,y_{n-1}} \bigoplus_{j=0}^{n-1} E_j^{(n^2)}(y_0, c, y_2, \ldots, y_{n-1}) \\
&= \bigoplus_{y_2,\ldots,y_{n-1}} \left(\bigoplus_{y_0} \bigoplus_{j=0}^{n-1} E_j^{(n^2)}(y_0, c, y_2, \ldots, y_{n-1}) \right) \\
&= \bigoplus_{y_2,\ldots,y_{n-1}} 0 \\
&= 0
\end{aligned}$$

□

5 Impossible Differential of GF-NLFSR

By using the $\mathcal{U}$-method [17], the designers of n-cell GF-NLFSR found a $(2n-1)$-round impossible differential: $(0, 0, 0, \ldots, \alpha) \nrightarrow (\psi, \psi, 0, \ldots, 0)$, where $\alpha \neq 0$, $\psi \neq 0$. In this section, we show how to construct an (n^2+n-2)-round impossible differential by studying the relationship between integral and truncated differential as described in the following theorem:

Theorem 3. *The n^2-round integral distinguisher of Theorem 1 corresponds to the following n^2-round truncated differential with probability* 1*:*

$$(\delta, 0, \ldots, 0) \rightarrow (\delta_0, \delta_1, \ldots, \delta_{n-1}),$$

where $\delta \neq 0$ and $\delta_0 \oplus \delta_1 \oplus \ldots \oplus \delta_{n-1} \neq 0$.

Proof. Let the input of the n-cell GF-NLFSR be $(x, c_1, c_2, \ldots, c_{n-1})$, after n^2 rounds, the output is $(q_0(x), q_1(x), \ldots, q_{n-1}(x))$, then according to Proposition 2, $q_0(x) \oplus q_1(x) \oplus \ldots \oplus q_{n-1}(x) \triangleq q(x) \in \mathbb{F}_{2^b}[x]$ is a permutation polynomial.

Assume two inputs are $(x_1, c_1, c_2, \ldots, c_{n-1})$ and $(x_2, c_1, c_2, \ldots, c_{n-1})$ with $x_1 \neq x_2$, thus $q(x_1) \neq q(x_2)$. Now the input difference is $(\delta, 0, \ldots, 0)$ with $\delta = x_1 \oplus x_2 \neq 0$, and the output difference is $(\delta_0, \delta_1, \ldots, \delta_{n-1})$, satisfying $\delta_0 \oplus \delta_1 \oplus \ldots \oplus \delta_{n-1} = q(x_1) \oplus q(x_2) \neq 0$. □

Theorem 4. *There exists an (n^2+n-2)-round impossible differential in n-*cell GF-NLFSR *of the following form:*

$$(\delta, 0, \ldots, 0) \not\rightarrow (\psi, \psi, 0, \ldots, 0),$$

where $\delta \neq 0$ and $\psi \neq 0$.

Proof. From encrypt direction, the n^2-round truncated differential $(\delta, 0, \ldots, 0) \rightarrow (\delta_0, \delta_1, \ldots, \delta_{n-1})$ is with probability 1, where $\delta \neq 0$ and $\delta_0 \oplus \delta_1 \oplus \ldots \oplus \delta_{n-1} \neq 0$. From decrypt direction, the $(n-2)$-round truncated differential $(\psi, \psi, 0, \ldots, 0) \rightarrow (0, \ldots, 0, \psi, \psi)$ is with probability 1. Since $\psi \oplus \psi = 0$, we find a contradiction. □

Remark. Wu *et al.* [30] independently found the same (n^2+n-2)-round impossible differential through a more direct approach. By using the 18-round impossible differential when $n = 4$, they presented a key recovery attack on the full round block cipher Four-Cell. Due to these new distinguishers and the full round attack, the designers have modified Four-Cell to Four-Cell$^+$ for better protection against the integral and impossible differential attacks.

6 A Kind of Non-surjective Attack

Our goal for introducing this kind of attack is that traditional non-surjective attack is only applicable to Feistel ciphers with non-surjective (non-uniform) round functions, while ours could be applied to block ciphers with bijective ones. Moreover, its data complexity is $\mathcal{O}(l)$ with l the block length.

To this end, we describe a variant structure of n-cell GF-NLFSR, denoted as n-cell VGF-NLFSR. As shown in Fig. 2, the main difference between these two structures is the round function. In n-cell VGF-NLFSR, the round function is $F(x \oplus K_i)$ with F bijective. One can easily demonstrate that the provable security against DC and LC for n-cell VGF-NLFSR can be provided using the same technique in [7]. Furthermore, Proposition 2 and 3 also suit for n-cell VGF-NLFSR, thus there exist the same n^2-round integral and $(n^2 + n - 2)$-round impossible differential as in n-cell GF-NLFSR.

Now, we introduce the non-surjective attack by analyzing VGF-NLFSR in the following two subsections.

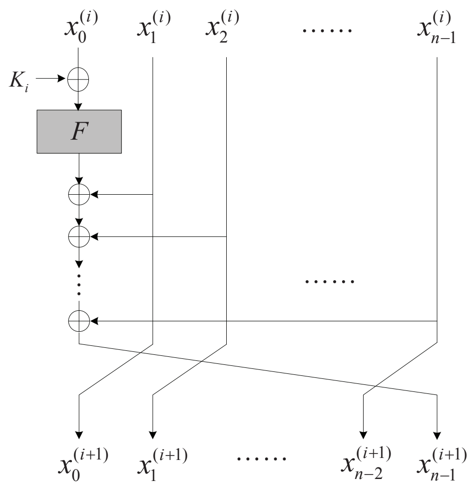

Fig. 2. The i-th round transformation of n-cell VGF-NLFSR

6.1 Description of the Non-surjective Distinguisher

Let the input of n-cell VGF-NLFSR be $(x, c_1, \dots, c_{n-1})$, according to Proposition 2 and Proposition 3, $y_{n-1}^{((n-2)\times n)}$ is a constant, say C, and

$$\bigoplus_{j=0}^{n-1} y_j^{(n^2-n)} = F(C \oplus K_{n^2-n}) \triangleq C'.$$

Thus

$$y_0^{(n^2-n)} = C' \oplus \bigoplus_{j=1}^{n-1} y_j^{(n^2-n)}.$$

Assume the output of the n^2-th round is $(q_0(x), q_1(x), \dots, q_{n-1}(x))$, from Proposition 2, we have

$$q_0(x) = F\left(y_0^{(n^2-n)} \oplus K_{n^2-n+1}\right) \oplus \bigoplus_{j=1}^{n-1} y_j^{(n^2-n)}.$$

Let $t = y_0^{(n^2-n)} \oplus K_{n^2-n+1}$, then

$$\begin{aligned} q_0(x) &= F(t) \oplus t \oplus K_{n^2+n-1} \oplus C' \\ &= F(t) \oplus t \oplus C^*, \end{aligned}$$

where $C^* = K_{n^2+n-1} \oplus C'$ represents some unknown constant.

Let $f(t) = F(t) \oplus t$, and define $\mathcal{D}_f = \{y | y = f(t),\ t \in \mathbb{F}_{2^b}\}$. From the above fact, we have the following n^2-round distinguisher:

Theorem 5. *Let the input to* n-cell VGF-NLFSR *be* $(x, c_1, \ldots, c_{n-1})$*, where* c_i *is constant, and the output of the* n^2*-th round be* $(q_0(x), q_1(x), \ldots, q_{n-1}(x))$*, then there exists some constant* $C^* \in \mathbb{F}_{2^b}$*, such that for any* $x \in \mathbb{F}_{2^b}$*,* $q_0(x) \oplus C^* \in \mathcal{D}_f$*.*

Consider the distinguisher in Theorem 5, in this situation, the input to the (n^2+1)-th round function F is $q'(x) = q_0(x) \oplus K_{n^2+1}$, let $c^* = C^* \oplus K_{n^2+1}$, then $q'(x) \oplus c^* = q_0(x) \oplus C^*$. In other words, for all $x \in \mathbb{F}_{2^b}$, there exists some constant c^*, such that $q'(x) \oplus c^* \in \mathcal{D}_f$. Thus we could get the following theorem:

Theorem 6. *Let the input of* n-cell VGF-NLFSR *be* $(x, c_1, \ldots, c_{n-1})$*, where* c_i *is constant, and the input of the* (n^2+1)*-th round function* F *be* $q'(x)$*, then there exists some constant* $c^* \in \mathbb{F}_{2^b}$*, such that for any* $x \in \mathbb{F}_{2^b}$*,* $q'(x) \oplus c^* \in \mathcal{D}_f$*.*

One should note that if $\mathcal{D}_f = \mathbb{F}_{2^b}$, then both $F(x)$ and $F(x) \oplus x$ are permutations on $\mathbb{F}_{2^b}$, which indicates that $F(x)$ is an *orthormorphic permutation* [22]. Since the number of all orthormorphic permutations is small, in general, for a randomly chosen permutation $F(x)$, $f(x) = F(x) \oplus x$ can be seen as a random function (as the Davies-Meyer construction in hash function), thus $\mathcal{D}_f \subsetneq \mathbb{F}_{2^b}$. From now on, we will call the above distinguisher a *non-surjective distinguisher*, since the range of the function f is only a subset of $\mathbb{F}_{2^b}$.

6.2 Description of the Non-surjective Attack

By using the non-surjective distinguisher, one can attack (n^2+n')-round n-cell VGF-NLFSR by Algorithm 1, where $n' > 1$.

Algorithm 1: Non-surjective attack on n-cell VGF-NLFSR	
Step 1	Compute and store $\mathcal{D}_f$.
Step 2	Given t plaintexts $(x_i, c_1, \ldots, c_{n-1})$, obtain the corresponding (n^2+n')-round ciphertexts, $i = 1, \ldots, t$.
Step 3	Guess the last $(n'-1)$ round-keys $rk = (rk_1, rk_2, \ldots, rk_{n'-1})$, decrypt the ciphertext to get the input of the (n^2+1)-round function F, denoted by $q'_{rk}(x_i)$.
Step 4	For all x_i in Step 2, test whether there exists some constant c^* satisfying $q'_{rk}(x_i) \oplus c^* \in \mathcal{D}_f$. If not, the guessed round-keys rk must be wrong.
Step 5	If necessary, repeat Step 2 $\sim$ Step 5 to further filter the wrong round keys until only one left.

In order to estimate the complexity of the above attack, we need the following two lemmas and their proofs can be found in Appendix A.

Lemma 1. *Given* $A \subseteq \mathbb{F}_{2^b}$*,* $|A|$ *denotes the number of different elements in* A*. For a randomly chosen set* $X \subseteq \mathbb{F}_{2^b}(|X| \le |A|)$*, let* p *be the probability that there exists some constant* $c \in \mathbb{F}_{2^b}$*, such that* $X \oplus c = \{x \oplus c | x \in X\} \subseteq A$*, then*

$$p \le 2^b \times \frac{|A|}{2^b} \times \frac{|A|-1}{2^b-1} \times \ldots \times \frac{|A|-(|X|-1)}{2^b-(|X|-1)}.$$

Lemma 2. *Let $f(x)$ be a random function from $\mathbb{F}_q$ to $\mathbb{F}_q$, $\mathcal{D}_f = \{f(x)|x \in \mathbb{F}_q\}$, let $\epsilon = E(|\mathcal{D}_f|)$ and $\sigma^2 = V(|\mathcal{D}_f|)$ be the expectation and variance of $|\mathcal{D}_f|$, respectively, then*

(i) $\lim_{q\to\infty} \frac{\epsilon}{q} = 1 - \frac{1}{e} \approx 0.632,$

(ii) $\lim_{q\to\infty} \frac{\sigma^2}{q} = \frac{e-2}{e^2} \approx 0.097.$

From Lemma 1, for a randomly chosen $X \subseteq \mathbb{F}_{2^b}$, if $|X| \ll |A|$, the upper bound of p can be well approximated by $2^b \times \left(|A|/2^b\right)^{|X|}$.

From Lemma 2, when q is large, the *Chebyshev Inequality* [27] indicates

$$\mathbf{Pr}\left(||\mathcal{D}_f| - \epsilon| \leq l\sigma\ \right) \geq 1 - \frac{1}{l^2}.$$

If we choose $q = 2^b$ and $l = 10$, then for a randomly chosen f,

$$\mathbf{Pr}\left(0.63 \times 2^b - 3 \times 2^{b/2} \leq |\mathcal{D}_f| \leq 0.63 \times 2^b + 3 \times 2^{b/2}\right) \geq 0.99.$$

Thus we can estimate with high probability that $|\mathcal{D}_f|$ is less than $0.63 \times 2^b + 3 \times 2^{b/2}$. Moreover, when b is large, $|\mathcal{D}_f|$ can be approximated by 0.63×2^b.

Now, the data, time and space complexity of the proposed non-surjective attack can be analyzed as follows:

Data Complexity. Firstly, we note that when applying integral attack to n-cell VGF-NLFSR, one must choose at least a structure of all possible $(x, c_1, \ldots, c_{n-1})$, where c_i's are constants. While for the non-surjective attack, only a fraction of them are needed.

Assume the number of chosen plaintexts as $(x, c_1, \ldots, c_{n-1})$ is t, let $\mathcal{T}$ denote the set of their corresponding ciphertexts, $\mathcal{T}_{rk}$ denote the set of the input to the (n^2+1)-round F function from decrypting the ciphertexts in $\mathcal{T}$ by guessing the last $n'-1$ round keys rk.

The crucial step in Algorithm 1 is to check whether there exists a constant $c^* \in \mathbb{F}_{2^b}$ such that $\mathcal{T}_{rk} \oplus c^* \subseteq \mathcal{D}_f$. Assume wrong key values can pass such test with probability P_{err}, then from Lemma 1,

$$P_{err} \leq (2^{(n'-1)b} - 1) \times 2^b \times \binom{|\mathcal{D}_f|}{t} \Big/ \binom{2^b}{t} \triangleq P_t,$$

thus in order to identify the right keys for the last $n'-1$ rounds, P_{err} must be small enough. If b is large, and $t \ll |\mathcal{D}_f|$,

$$P_t \approx 2^{n'b} \times \left(|\mathcal{D}_f|/2^b\right)^t \approx 2^{n'b} \times 0.63^t.$$

Let $P_t = 2^{-\lambda}$, where the parameter λ is related to the success probability, and can be deduced by experiments, then $P_{err} \leq P_t = 2^{-\lambda}$, which indicates that the probability that wrong key values can pass the test in Step 4 is less than $2^{-\lambda}$.

From $2^{n'b} \times 0.63^t = 2^{-\lambda}$, we get $t \approx \frac{3}{2}n'b + \frac{3}{2}\lambda$. Thus the data complexity of the above non-surjective attack is $\mathcal{O}(b)$.

To sum up, for attacking $(n^2 + n')$-round n-cell VGF-NLFSR, the data complexity is about $\frac{3}{2}n'b + \frac{3}{2}\lambda$.

Time Complexity. As explained before, Step 4 of Algorithm 1 needs to verify whether there exists a constant $c^* \in \mathbb{F}_{2^b}$, s.t. $\mathcal{T}_{rk} \oplus c^* \subseteq \mathcal{D}_f$ for each possible r_k. Assume for each possible c^*, the time complexity for testing whether $\mathcal{T}_{rk} \oplus c^* \subseteq \mathcal{D}_f$ is equivalent to u encryptions, then the time complexity is about

$$\left(\frac{3}{2}n'b + \frac{3}{2}\lambda\right) \times (2^{(n'-1)b}) \times 0.63 \times 2^b \times u \approx (n'b + \lambda) \times 2^{n'b} \times u,$$

thus a good algorithm for testing whether one set is included in another is required.

Space Complexity. Since one must store $\mathcal{D}_f$ to apply the non-surjective attack, the space complexity is about 0.63×2^b.

7 Experiments with the Proposed Non-surjective Attack

This section describes a 32-bit toy cipher based on 4-cell VGF-NLFSR, where the round function is defined by $F(x, k) = S(x \oplus k)$ with S as the S-box of AES. It is well known that the differential (linear hull) probability of the S-box of AES is upper bounded by 2^{-6}, thus the differential (linear hull) probability for five rounds is upper bounded by $(2^{-6})^2 = 2^{-12}$. Now we can see that the differential (linear) characteristic probability for 15 rounds is at most $(2^{-12})^3 = 2^{36} \leq 2^{-32}$, that is to say such toy cipher with more than 15 rounds is practically secure against DC and LC.

As an example, we use the method in Section 6 to mount a non-surjective attack on the 18-round toy cipher. In this case, $b = 8$ and $|\mathcal{D}_f| = 163 \approx 0.63 \times 2^8$. Table 2 lists our experimental results. For each $\lambda = 2, 4, 6, 8, 10$, t_λ denotes the number of chosen plaintexts and p_λ denotes the success probability, where the "success" means the adversary can uniquely recover the right 18-th round key. For each chosen parameter λ, we do the non-surjective attack 1000 times, and in each time the plaintext as well as the encryption key are randomly generated. The success probabilities are $0.474, 0.758, 0.873, 0.965, 0.992$.

One could also apply the integral attack to the 18-round toy cipher, however, to get a high success probability, its data complexity is about $2 \times 2^8 = 2^9$.

Table 2. Experiments with the non-surjective attack on the 18-round toy cipher

parameter λ	chosen plaintexts $t_\lambda = 3b + 1.5\lambda$	success probability p_λ
2	27	0.474
4	30	0.758
6	33	0.873
8	36	0.965
10	39	0.992

8 Conclusion

This paper presents several security analysis on GF-NLFSR. Although such structure allows parallel computations for encryption and can even provide its provable security against DC and LC, the structure itself reveals a very slow diffusion rate, which could lead to several distinguishing attacks.

For n-cell GF-NLFSR, our cryptanalytic results show that there exists an n^2-round integral distinguisher, which could be extended to an (n^2+n-2)-round higher-order one. Based on this n^2-round integral distinguisher, an (n^2+n-2)-round impossible differential is constructed. These results are significantly better than the original ones and thus imply that the security of n-cell GF-NLFSR must be carefully reevaluated.

Besides, a kind of non-surjective attack is proposed, which is different in essence with the one introduced by Rijmen *et al.*, since traditional non-surjective attack is only applicable to Feistel ciphers with non-surjective (non-uniform) round functions while ours can be applied to block ciphers with round functions being bijective. To demonstrate this, we describe a variant structure of n-cell GF-NLFSR, whose round function is defined by $F(x \oplus K)$. The provable security against DC and LC can also be provided for this variant structure, however, by using the proposed non-surjective attack, an efficient key recovery attack with very low data complexity could be mounted. Some experimental results are given for this non-surjective attack on a toy cipher based on the S-box of AES.

It is interesting that whether this kind of non-surjective attack can be applied to other block ciphers.

Acknowledgments

The authors wish to thank the anonymous reviewers of ACISP 2010 for their valuable suggestions and comments.

References

1. A. Biryukov and D. Wagner. Slide Attack. FSE 1999, LNCS 1636, pp. 245–259, Springer-Verlag, 1999.
2. A. Biryukov, and A. Shamir. Structural Cryptanalysis of SASAS. EUROCRYPT 2001, LNCS 2045, pp. 394–405, Springer–Verlag, 2001.
3. E. Biham. New Types of Cryptanalytic Attacks Using Related Keys. EUROCRYPT 1993, LNCS 765, pp. 398–409, Springer-Verlag, 1994.
4. E. Biham, A. Biryukov, A. Shamir. Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. EUROCRYPT 1999, LNCS 1592, pp. 12–23, Springer-Verlag, 1999.
5. E. Biham, O. Dunkelman, and N. Keller. The Rectangle Attack- Rectangling the Serpent. EUROCRYPT 2001, LNCS 2045, pp. 340–357, Springer-Verlag, 2001.
6. E. Biham and A. Shamir. Differential Cryptanalysis of DES-like Cryptosystems. Journal of Cryptology, Vol 3, pp. 3–72, Springer-Verlag, 1991.

7. J. Choy, G. Chew, K. Khoo and H. Yap. Cryptographic Properties and Application of a Generalized Unbalanced Feistel Network Structure. ACISP 2009, LNCS 5594, pp. 73–89, Springer-Verlag, 2009.
8. N. Courtois and J. Pieprzyk. Cryptanalysis of Block Ciphers with Overdefined Systems of Equations. ASIACRYPT 2002, LNCS 2501, pp. 267–287, Springer-Verlag, 2002.
9. J. Daemen , L. Knudsen, V. Rijmen. The Block Cipher Square. FSE 1997, LNCS 1267, pp. 149–165, Springer-Verlag, 1997.
10. W. Feller. An Introduction to Probability Theory and Its Applications, 3rd Edition. Wiley, New York, 1968.
11. N. Ferguson, J. Kelsey, S. Lucks, B. Schneier, M. Stay, D. Wagner, and D. Whiting. Improved Cryptanalysis of Rijndael. FSE 2000, LNCS 1978, pp. 213–230, Springer-Verlag, 2001.
12. T. Jackobsen, L. Knudsen. The Interpolation Attack on Block Cipher. FSE 1997, LNCS 1008, pp. 28–40, Springer-Verlag, 1997.
13. L. Knudsen. Truncated and High Order Differentials. FSE 1995, LNCS 1008, pp. 196–211, Springer-Verlag 1995.
14. L. Knudsen. DEAL – A 128-bit Block Cipher. Technical Report 151, Department of Informatics, University of Bergen, Bergen, Norway, Feb. 1998.
15. L. Knudsen, D. Wagner. Integral Cryptanalysis. FSE 2002, LNCS 2365, pp. 112–127, Springer-Verlag, 2002.
16. J. Kelsey, T. Kohno and B. Schneier. Amplified Boomerang Attacks against Reduced-round MARS and Serpent. FSE 2000, LNCS 1978, pp. 75–93, Springer-Verlag, 2001.
17. J. Kim, S. Hong, J. Sung, S. Lee, J. Lim, and S. Sung. Impossible Differential Cryptanalysis for Block Cipher Structures. Indocrypt 2003, LNCS 2904, pp. 82–96, Springer-Verlag 2003.
18. X. Lai. High Order Derivatives and Differential Cryptanalysis. Communications and Cryptography, pp. 227–233, 1994.
19. S. Lucks. The Saturation Attack — A Bait for Twofish. FSE 2001, LNCS 2355, pp. 1–15, Springer–Verlag, 2002.
20. J. Lu, O. Dunkelman, N. Keller and J. Kim. New Impossible Differential Attacks on AES. Indocrypt 2008, LNCS 5365, pp. 279–293, Springer-Verlag 2008.
21. J. Lu, J. Kim, N. Keller and O. Dunkelman. Improving the Efficiency of Impossible Differential Cryptanalysis of Reduced Camellia and MISTY1. CT-RSA 2008, LNCS 4904, pp. 370–386, Springer-Verlag 2008.
22. L. Mittenthal. Block Substitutions Using Orthomorphic Mappings. In Advances in Applied Mathematics, Vol 16(1), pp. 59–71, 1995.
23. M. Matsui. Linear Cryptanalysis Method for DES Cipher. EUROCRYPT 1993, LNCS 765, pp. 386–397, Springer-Verlag, 1993.
24. V. Rijmen, B. Preneel and E. De Win. On Weaknesses of Non-surjective Round Functions. Designs, Codes, and Cryptography, VOL 12, pp. 253–266, Springer-Verlag, 1997.
25. F. Roberts and B. Tesman. Applied Combinatorics, 2nd Edition. Pearson Education, 2005.
26. B. Schneier, J. Kelsey. Unbalanced Feistel Networks and Block Cipher Design. FSE 1996, LNCS 1039, pp.121–144, Springer-Verlag, 1996.
27. D. Wanger. The Boomerang Attack. FSE 1999, LNCS 1636, pp.156–170, Springer-Verlag, 1999.

28. W. Wu, W. Zhang and D. Feng. Impossible differential cryptanalysis of Reduced-Round ARIA and Camellia. Journal of Compute Science and Technology 22(3), pp. 449–456, Springer-Verlag, 2007.
29. W. Zhang, W. Wu, and D. Feng. New Results on Impossible Differential Cryptanalysis of Reduced AES. ICISC 2007, LNCS 4817, pp.239–250, Springer-Verlag, 2007.
30. W. Wu, L. Zhang, L. Zhang and W. Zhang. Security Analysis of the GF-NLFSR Structure and Four-Cell Block Cipher. ICICS 2009, LNCS 5927, pp.17–31, Springer-Verlag, 2009.

A Proofs of Lemma 1 and Lemma 2

1. Proof of Lemma 1

First note that the number of different sets chosen from $\mathbb{F}_{2^b}$ with $|X|$ elements is $\binom{2^b}{|X|}$. Consider the subset $A \subseteq \mathbb{F}_{2^b}$, the number of different sets chosen from A with $|X|$ elements is $\binom{|A|}{|X|}$. Now for every fixed $c \in \mathbb{F}_{2^b}$, the probability p_c that $X \oplus c \subseteq A$ is upper bound by $\binom{|A|}{|X|}/\binom{2^b}{|X|}$. Thus we have

$$p = \sum_{c\in\mathbb{F}_{2^b}} p_c \leq 2^b \times \binom{|A|}{|X|}/\binom{2^b}{|X|}.$$

□

2. Proof of Lemma 2

Lemma 2 can be extended to a more general situation, where $\mathbb{F}_q$ can be replaced by any set with n elements and we will prove this more general conclusion. Note that the result of (i) can also be found in [24], however, by using their technique, one could not get the result of (ii). So, we introduce a formal method and prove these two results in a unified approach.

Given a set S, $|S| = n$, let f be a random function from S to S and $\mathcal{D}_f = \{f(a)|a \in S\} \subseteq S$.

(i) By the definition of expectation,

$$\epsilon = \sum_f \frac{1}{n^n} \times |\mathcal{D}_f| = \frac{1}{n^n} \times \sum_f |\mathcal{D}_f|. \tag{1}$$

From the "Principle of Inclusive and Exclusive" [25], we have

$$\begin{aligned}\sum_f |\mathcal{D}_f| &= \sum_{t=1}^{n} t \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n \\ &= \sum_{t=1}^{n} t \cdot \binom{n}{t} \cdot \sum_{u=1}^{t} \binom{t}{u} \cdot (-1)^{t-u} \cdot u^n \text{ (where } u = t-i)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{u=1}^{n} u^n \cdot \sum_{t=u}^{n} t \cdot \binom{n}{t} \cdot \binom{t}{u} \cdot (-1)^{t-u} \\
&= \sum_{u=1}^{n} u^n \cdot \sum_{k=0}^{n-u} (k+u) \cdot \binom{n}{k+u} \cdot \binom{k+u}{u} \cdot (-1)^{k} \text{ (where } k = t-u\text{)} \\
&= \sum_{u=1}^{n} u^n \cdot \sum_{k=0}^{n-u} (k+u) \cdot \binom{n}{u} \cdot \binom{n-u}{k} \cdot (-1)^{k} \\
&= \sum_{u=1}^{n} u^n \cdot \binom{n}{u} \cdot \sum_{k=0}^{n-u} (k+u) \cdot \binom{n-u}{k} \cdot (-1)^{k} \\
&\triangleq A + B, \qquad (2)
\end{aligned}
$$

where

$$
\begin{aligned}
A &= \sum_{u=1}^{n} u^n \cdot \binom{n}{u} \cdot \sum_{k=0}^{n-u} k \cdot \binom{n-u}{k} \cdot (-1)^{k} \\
&= \sum_{u=1}^{n-1} u^n \cdot \binom{n}{u} \cdot \sum_{k=1}^{n-u} k \cdot \binom{n-u}{k} \cdot (-1)^{k} \\
&= \sum_{u=1}^{n-1} u^n \cdot \binom{n}{u} \cdot \sum_{k=1}^{n-u} (n-u) \cdot \binom{n-u-1}{k-1} \cdot (-1)^{k} \\
&= -\sum_{u=1}^{n-1} u^n \cdot \binom{n}{u} \cdot \sum_{k'=0}^{n-u-1} (n-u) \cdot \binom{n-u-1}{k'} \cdot (-1)^{k'} \\
&= -n \cdot (n-1)^n,
\end{aligned}
$$

and

$$
B = \sum_{u=1}^{n} u^{n+1} \cdot \binom{n}{u} \cdot \sum_{k=0}^{n-u} \binom{n-u}{k} \cdot (-1)^{k} = n^{n+1}.
$$

From (1) and (2), we get

$$
\epsilon = \frac{A+B}{n^n} = \frac{1}{n^n} \times \left(n^{n+1} - n \cdot (n-1)^n\right) = n - n \cdot (1 - 1/n)^n.
$$

Thus

$$
\lim_{n\to\infty} \frac{\epsilon}{n} = \lim_{n\to\infty} \left(1 - \left(1 - \frac{1}{n}\right)^n\right) = 1 - \frac{1}{e}.
$$

(ii) By the definition of variance,

$$
\sigma^2 = \sum_{f} \frac{1}{n^n} \times (\,|\mathcal{D}_f| - \epsilon\,)^2 = \frac{1}{n^n} \times \sum_{f} (\,|\mathcal{D}_f| - \epsilon\,)^2. \qquad (3)
$$

From the result of (i),

$$
\begin{aligned}
&\sum_{f} (\,|\mathcal{D}_f| - \epsilon\,)^2 \\
&= \sum_{t=1}^{n} \left(t - n\left(1 - \left(1 - \frac{1}{n}\right)^n\right)\right)^2 \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n \\
&= \sum_{t=1}^{n} \left(t^2 - 2nt\left(1 - \left(1 - \frac{1}{n}\right)^n\right) + \left(1 - \left(1 - \frac{1}{n}\right)^n\right)^2 \cdot n^2\right) \\
&\qquad \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n \\
&\triangleq A + B + C, \qquad (4)
\end{aligned}
$$

where

$$
\begin{aligned}
A &= \sum_{t=1}^{n} t^2 \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n, \\
B &= -2n\left(1 - \left(1 - \frac{1}{n}\right)^n\right) \cdot \sum_{t=1}^{n} t \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n, \\
C &= \left(1 - \left(1 - \frac{1}{n}\right)^n\right)^2 \cdot n^2 \cdot \sum_{t=1}^{n} \cdot \binom{n}{t} \cdot \sum_{i=0}^{t-1} \binom{t}{t-i} \cdot (-1)^i \cdot (t-i)^n.
\end{aligned}
$$

Using the same technique as in the proof of (i), after careful calculation,

$$
\begin{aligned}
A &= n^{n+2} - 2n(n-1)^{n+1} + \left(2(n-2)^n \binom{n}{2} - n(n-1)^n\right), \\
B &= -2n\left(1 - \left(1 - \frac{1}{n}\right)^n\right) \cdot (n^{n+1} - n(n-1)^n), \\
C &= \left(1 - \left(1 - \frac{1}{n}\right)^n\right)^2 \cdot n^2 \cdot n^n.
\end{aligned}
$$

From (3) and (4), we get

$$\sigma^2 = \frac{A + B + C}{n^n}.$$

Thus

$$\lim_{n \to \infty} \frac{\sigma^2}{n} = \lim_{n \to \infty} \frac{A + B + C}{n^{n+1}} = \frac{e - 2}{e^2}.$$

□

B 16-Round Integral Distinguisher of 4-Cell GF-NLFSR

0	x	C_1	C_2	C_3
1	C_1	C_2	C_3	$y \oplus C_4$
2	C_2	C_3	$y \oplus C_4$	$y \oplus C_5$
3	C_3	$y \oplus C_4$	$y \oplus C_5$	C_6
4	$y \oplus C_4$	$y \oplus C_5$	C_6	C_7
5	$y \oplus C_5$	C_6	C_7	$y \oplus z \oplus C_8$
6	C_6	C_7	$y \oplus z \oplus C_8$	$y \oplus z \oplus w \oplus C_9$
7	C_7	$y \oplus z \oplus C_8$	$y \oplus z \oplus w \oplus C_9$	$w \oplus C_{10}$
8	$y \oplus z \oplus C_8$	$y \oplus z \oplus w \oplus C_9$	$w \oplus C_{10}$	C_{11}
9	$y \oplus z \oplus w \oplus C_9$	$w \oplus C_{10}$	C_{11}	$t_1 \oplus y \oplus z \oplus C_{12}$
10	$w \oplus C_{10}$	C_{11}	$t_1 \oplus y \oplus z \oplus C_{12}$	$t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{13}$
11	C_{11}	$t_1 \oplus y \oplus z \oplus C_{12}$	$t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{13}$	$t_2 \oplus w \oplus u \oplus C_{14}$
12	$t_1 \oplus y \oplus z \oplus C_{12}$	$t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{13}$	$t_2 \oplus w \oplus u \oplus C_{14}$	$u \oplus C_{15}$
13	$t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{13}$	$t_2 \oplus w \oplus u \oplus C_{14}$	$u \oplus C_{15}$	$t_3 \oplus t_1 \oplus y \oplus z \oplus C_{16}$
14	$t_2 \oplus w \oplus u \oplus C_{14}$	$u \oplus C_{15}$	$t_3 \oplus t_1 \oplus y \oplus z \oplus C_{16}$	$t_4 \oplus t_3 \oplus t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{17}$
15	$u \oplus C_{15}$	$t_3 \oplus t_1 \oplus y \oplus z \oplus C_{16}$	$t_4 \oplus t_3 \oplus t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{17}$	$t_5 \oplus t_4 \oplus t_2 \oplus w \oplus u \oplus C_{18}$
16	$t_3 \oplus t_1 \oplus y \oplus z \oplus C_{16}$	$t_4 \oplus t_3 \oplus t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{17}$	$t_5 \oplus t_4 \oplus t_2 \oplus w \oplus u \oplus C_{18}$	$t_5 \oplus u \oplus v \oplus C_{19}$

The parameters in the above 16-round integral distinguisher are as follows: C_i, $4 \leq i \leq 19$ is passive (constant) in $\mathbb{F}_{2^b}$, x, y, z, w, u, v are active in $\mathbb{F}_{2^b}$, and t_j, $1 \leq j \leq 5$ is some unknown intermediate value in $\mathbb{F}_{2^b}$. It can be easily verified that

$$\begin{aligned} &(t_3 \oplus t_1 \oplus y \oplus z \oplus C_{16}) \oplus (t_4 \oplus t_3 \oplus t_2 \oplus t_1 \oplus y \oplus z \oplus w \oplus C_{17}) \oplus (t_5 \oplus t_4 \oplus t_2 \oplus w \oplus u \oplus C_{18}) \oplus (t_5 \oplus u \oplus v \oplus C_{19}) \\ &= v \oplus C_{16} \oplus C_{17} \oplus C_{18} \oplus C_{19}. \end{aligned}$$