Cryptographic Properties and Application of a Generalized Unbalanced Feistel Network Structure

Jiali Choy, Guanhan Chew, Khoongming Khoo and Huihui Yap

DSO National Laboratories
20 Science Park Drive, Singapore 118230
Email: cjiali,cguanhan,kkhoongm,yhuihui@dso.org.sg

Abstract. In this paper, we study GF-NLFSR, a Generalized Unbalanced Feistel Network (GUFN) which can be considered as an extension of the outer function FO of the KASUMI block cipher. We prove upper bounds for the differential and linear hull probabilities for any $n+1$ rounds of an n-cell GF-NLFSR. Besides analyzing security against differential and linear cryptanalysis, we provide a frequency distribution for upper bounds on the true differential and linear hull probabilities. We also demonstrate a $(2n-1)$-round impossible differential distinguisher and a $(3n-1)$-round integral attack distinguisher on the n-cell GF-NLFSR. As an application, we design a new block cipher Four-Cell based on a 4-cell GF-NLFSR. We prove the security of Four-Cell against differential, linear, and boomerang attack. Based on the 7-round impossible differential and 11-round integral attack distinguisher, we set the number of rounds of Four-Cell to be 25 for protection against these attacks. Furthermore, Four-Cell can be shown to be secure against other attacks such as higher order differential attack, cube attack, interpolation attack, XSL attack and slide attack.

Keywords: Block Ciphers, Generalized Unbalanced Feistel Network, Differential Probability, Linear Hull Probability.

1 Introduction

In this paper, we examine a family of block ciphers whose structure is modelled after that of a Generalized Unbalanced Feistel Network (GUFN). The GUFN was first suggested by Schneier et al. in [20]. Similar to conventional Feistel networks, unbalanced ones comprise of a concatenation of rounds. In each round, one part of the block controls the encryption of another part of the block. However, the two parts need not be of equal sizes.

The particular GUFN we shall be analyzing is an n-cell extension of the outer function FO of the KASUMI block cipher [24], which is a 2-cell structure. Besides being a GUFN, our structure can also be viewed as an n-cell *NonLinear Feedback Shift Register* (NLFSR). Thus, we call our structure a *Generalized Feistel-NonLinear Feedback Shift Register* (GF-NLFSR). In Section 3, we shall give a detailed description of the GF-NLFSR.

Many GUFN-based block ciphers have been constructed; some examples include the ciphers SMS4 [13] and CLEFIA [21]. While the true differential and linear hull

probabilities of these ciphers are not known in the open literature, they have been calculated for other GUFN-like constructions. In [24], these were derived for KASUMI's FO function, which is equivalent to a 2-cell GUFN. Similar analyses have been done in [25] for another GUFN-like round function, and also in [15]. To the best of our knowledge, bounds for the true differential and linear hull probabilities have not been proven for GUFN-based ciphers with n input cells. Analysis of true differentials and linear hulls is required in assessing vulnerability to attacks such as boomerang attack. In light of this, the study in our paper is both novel and useful.

In Sections 4 and 5, we prove that the true differential and linear hull probability of any $n+1$ rounds of the n-cell GF-NLFSR is bounded by p^2 where p is the maximal probability of the nonlinear function. In Section 6, we investigate the frequency distribution of the differential and linear hull probability of any $n+1$ rounds based on different input-output differentials/linear masks. From the frequency distribution, we see that the maximal probability p^2 only holds for a very tiny portion of all differentials/linear hulls. There are also other differentials/linear hulls having probability bounds $p^3, p^4, \ldots, p^n$, but we prove that almost all differentials/linear hulls have probability bound p^n. Furthermore, we compute the expected differential/linear hull probability bound and find this value to be close to $(2^{-B}+p)^n$ where B is the size of each cell in GF-NLFSR. These differential and linear hull probability bounds are achieved when the input differences and mask values are randomly chosen, which is likely when $n+1$ rounds of the n-cell GF-NLFSR is prepended and appended by additional cipher structures. In this case, the security of $n+1$ rounds of n-cell GF-NLFSR, in the sense of differential and linear hull probability bounds, is therefore much better than is typically believed. This motivates our study of the expected bounds in the Section 6.

Other than differential and linear cryptanalysis, in Sections 7 and 8, we also consider the security of GF-NLFSR against impossible differential and integral cryptanalysis. For the former this is done by finding impossible differential characteristics which play the role of a sieve, methodically rejecting the wrong key guesses and leaving the correct key. For GF-NLFSR, the maximum number of rounds for impossible differential characteristics was found to be $2n-1$. On the other hand, in an integral attack, the attacker looks at larger carefully chosen sets of encryptions, in which parts of the input text form a multiset. We studied the propagation of multisets through the cipher and unveiled a $(3n-1)$-round distinguisher for GF-NLSR.

As an application of our results on GF-NLFSR, we design a GUFN-based block cipher Four-Cell in Section 9. It is a 128-bit block cipher based on a 4-cell GF-NLFSR where each cell is 32-bit long. Besides proving practical security against differential and linear cryptanalysis, we are able to bound its true differential probability by $2^{-55.39}$ and linear hull probability by $2^{-52.96}$. Moreover, we show that with 99.9999% frequency, the differential and linear hull probability bounds are much lower at $2^{-110.78}$ and $2^{-105.91}$ respectively. These facts also allow us to prove its security against boomerang attack. Based on the results in Sections 7 and 8, we can deduce a 7-round impossible differential and an 11-round integral attack distinguisher on Four-Cell. To protect against these attacks, we set the number of rounds of Four-Cell to be 25. Furthermore, we explain

why Four-Cell is secure against other cryptanalysis like higher-order differential attack, cube attack, interpolation attack, XSL attack and slide attack.

Like the AES cipher, our Four-Cell block cipher can be proven secure against known block cipher attacks. In principle, it can use the same S-box (SubBytes) and MDS transform (MixColumn) as AES. However, it is more efficient (in hardware) in the sense that it uses less MDS transforms (25 compared to 40) than AES while keeping the number of S-boxes unchanged. Another advantage of the n-cell GF-NLFSR structure is that the nonlinear function in any n rounds can be computed in parallel. Therefore, any four rounds of the nonlinear transforms in our block cipher Four-Cell can be computed in parallel. This is not true for a general GUFN-based block cipher like SMS4 [13].

2 Definitions and Preliminaries

In this paper, we shall study the GF-NLFSR which can be considered as a particular instantiation of the Generalized Unbalanced Feistel Network defined in [20]. In what follows, the "+" symbol is used to denote finite field addition (XOR) over $GF(2)^n$ or ordinary addition, depending on the operands and context.

2.1 Differential Cryptanalysis

As is widely known, differential cryptanalysis [1] is a chosen-plaintext attack in which statistical key information is deduced from ciphertext blocks obtained by encrypting pairs of plaintext blocks with a specific bitwise difference under the target key. It studies the propagation of input differences to output differences in iterated transformations.

Let $f : GF(2)^m \mapsto GF(2)^m$ be a Boolean mapping composed of a number of rounds. The concept of *characteristic* was introduced: a sequence of difference patterns such that the output difference from one round corresponds to the input difference in the next round. On the other hand, in [10, 11], the concept of a *differential*, denoted by $\alpha \xrightarrow{f} \beta$, was presented, where the XORs in the inputs and outputs of the intermediate rounds are not fixed. We denote $DP(\alpha \xrightarrow{f} \beta) = Pr(f(x) + f(x + \alpha) = \beta)$, where α, β are fixed input and output differences.

Differential cryptanalysis exploits differential characteristics with high probability. However, even if the maximal differential characteristic probability is low, one cannot conclude that the cipher is secure against differential attack. Instead, one must show that the maximal differential probability of all differentials is low enough [11]. This property ensures *provable security* against differential cryptanalysis as opposed to *practical security* which simply considers the maximal differential characteristic probability.

Proposition 1 *[11] A block cipher with block length m is resistant against conventional differential attacks under an independent subkey assumption, if there does not exist any differential $\alpha \longrightarrow \beta$, $\alpha \neq 0$, ranging over all but a few rounds, such that $DP(\alpha \longrightarrow \beta) \gg 2^{-m}$.*

For key-dependent functions, we consider the average resistance against differential cryptanalysis, i.e. the average differential probability taken over the entire key set. More formally, let $F : GF(2)^m \times K \mapsto GF(2)^m$ be a key-dependent function. Denote $f_k = F(x, k)$ for each fixed $k \in K$. Let $\alpha, \beta \in GF(2)^m$ be constants. The *differential probability of the differential* $\alpha \xrightarrow{F} \beta$ is defined as $DP(\alpha \xrightarrow{F} \beta) = \frac{1}{|K|} \sum_{k \in K} DP(\alpha \xrightarrow{f_k} \beta)$. The *maximal differential probability* of F is defined as $DP(F_{max}) = \max_{\alpha \neq 0, \beta} DP(\alpha \xrightarrow{F} \beta)$.

2.2 Linear Cryptanalysis

Linear cryptanalysis [14] is a known-plaintext attack that tries to utilize high probability occurrences of linear expressions involving plaintext bits, ciphertext bits, and subkey bits.

As with the differential case, we must also distinguish between a linear characteristic and a linear hull. A *linear characteristic over* f consists of a sequence of mask values such that the output mask values from one round corresponds to the input mask values to the next round. On the other hand, a *linear hull*, denoted by $u \xleftarrow{f} w$, is the set of all linear characteristics with the same initial and terminal mask values. We denote $LP(u \xleftarrow{f} w) = [2 \cdot Pr(u \cdot f(x) = w \cdot x) - 1]^2$, where w, u are fixed input and output mask values.

Linear cryptanalysis takes advantage of linear characteristics with high correlation probability to recover key bits. However, in the evaluation of the strength of a block cipher against linear cryptanalysis, one must consider the linear hulls instead. Having low linear hull probability for all linear hulls will guarantee provable security against linear attacks [17].

Proposition 2 *[17] A block cipher with block length m is resistant against conventional linear cryptanalysis under an independent subkey assumption, if there does not exist any linear hull* $u \longleftarrow w$, $u \neq 0$, *ranging over all but a few rounds, such that* $LP(u \longleftarrow w) \gg 2^{-m}$.

For key-dependent functions, we consider the average resistance against linear cryptanalysis. Explicitly, let $F : GF(2)^m \times K \mapsto GF(2)^m$ be a key-dependent function. Denote $f_k(x) = F(x, k)$ for each fixed $k \in K$. Let $u, w \in GF(2)^m$ be constants. The *linear hull probability of the linear hull* $u \xleftarrow{F} w$ is defined as $LP(u \xleftarrow{F} w) = \frac{1}{|K|} \sum_{k \in K} LP(u \xleftarrow{f_k} w)$. The *maximal linear hull probability* of F is defined as $LP(F_{max}) = \max_{w, u \neq 0} LP(u \xleftarrow{F} w)$.

It was proven in [11] and [17] the following result about differential and linear hull probabilities of compositions of key-dependent mappings.

Fact 1 *[11, 17] Let* $F : GF(2)^m \times GF(2)^m \times K_1$ *and* $G : GF(2)^m \times GF(2)^m \times K_2$ *be key-dependent functions of the type* $F(x, k, k') = f(x + k, k')$, $G(x, k, k') = g(x + k, k')$, *where* $f : GF(2)^m \times K_1 \mapsto GF(2)^m$ *and* $g : GF(2)^m \times K_2 \mapsto GF(2)^m$ *are bijective for all fixed* $k_1 \in K_1$, $k_2 \in K_2$. *Then* $DP(\alpha \xrightarrow{G \circ F} \beta) = \sum_{\xi \in GF(2)^m} DP(\alpha \xrightarrow{f} \xi) DP(\xi \xrightarrow{g} \beta)$ *and* $LP(u \xleftarrow{G \circ F} w) = \sum_{v \in GF(2)^m} LP(u \xleftarrow{g} v) LP(v \xleftarrow{f} w)$.

In Sections 4 and 5, we shall be demonstrating provable security of our design structure against differential and linear cryptanalysis by studying its differential and linear hull probabilities. Fact 1 will be required in the proofs of our results later.

3 Description of the Structure

In this section, we will give a description of our design structure, which we call GF-NLFSR. It is essentially a generalization of the outer function, FO, of the KASUMI cipher. The FO function was first suggested by Matsui in [15, Figure 7] as one of the new structures of block ciphers with provable security against differential and linear cryptanalysis. It was then adopted in the design of KASUMI [24]. The following result was proven in the same paper regarding the maximal differential and linear hull probabilities of this function.

Fact 2 *[24, Theorem 2] Let F be the 3-round function shown in Figure 1 of [24] (i.e. a 2-cell GF-NLFSR) where each $F_i : GF(2)^B \times GF(2)^B \times K_i' \mapsto GF(2)^B$ is of the form $F_i(x, k_i, k_i') = f_i(x + k_i, k_i')$ and each $f_i : GF(2)^B \times K_i' \mapsto GF(2)^B$ is bijective for all fixed $k_i' \in K_i'$, where K_i' is the key space for k_i'.*

(1) If $DP((f_i)_{max}) \leq p$ for each i, then $DP(F_{max}) \leq p^2$.
(2) If $LP((f_i)_{max}) \leq q$ for each i, then $LP(F_{max}) \leq q^2$.

This function splits the input block into 2 sub-blocks of equal size. Our block cipher structure generalizes this by splitting the input block into n sub-blocks of equal size. Figure 1 below displays one round of GF-NLFSR. Explicitly, suppose we have a m-bit

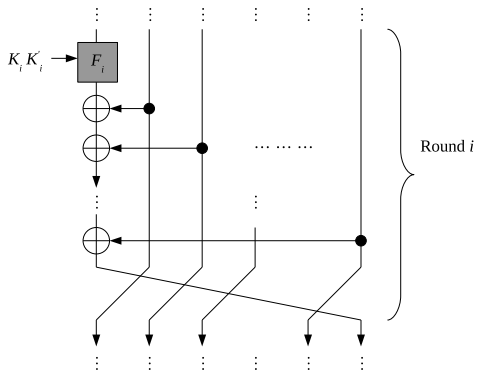

Fig. 1. One round of n-cell GF-NLFSR

block cipher, i.e. the input and output blocks are both of size $m = nB$ bits. Let the

internal state by denoted by $\mathcal{S} = (S_1, S_2, \ldots, S_n)$ where $S_i \in GF(2)^B$. Therefore the internal state consists of n sub-blocks of B bits each. The round keys of the cipher shall be denoted by k_i, k_i' $(i = 1, \ldots, n+1)$. Each F_i function is of the form

$$F_i : GF(2)^B \times GF(2)^B \times K_i' \mapsto GF(2)^B$$
$$F_i(x, k_i, k_i') = f_i(x + k_i, k_i')$$

where each $f_i : GF(2)^B \times K_i' \mapsto GF(2)^B$ is bijective for all fixed $k_i' \in K_i'$.

The round function R that maps $\mathcal{S}_i$ to $\mathcal{S}_{i+1}$ under the round keys k_i, k_i' is:

$$R : GF(2)^m \times GF(2)^B \times K_i' \mapsto GF(2)^m$$
$$((S_1, S_2, \ldots, S_n), k_i, k_i') \mapsto (S_2, S_3, \ldots, S_n, F_i(S_1, k_i, k_i') + S_2 + S_3 + \ldots + S_n)$$

4 Differential Probability

In this section, we present a result for the differential probability of an n-block GF-NLFSR over $n+1$ rounds which is similar to Fact 2.

Theorem 1 *Let F be the $(n+1)$-round function in Figure 2 (left) of Appendix B where each $F_i : GF(2)^B \times GF(2)^B \times K_i' \mapsto GF(2)^B$ is of the form $F_i(x, k_i, k_i') = f_i(x + k_i, k_i')$ and each $f_i : GF(2)^B \times K_i' \mapsto GF(2)^B$ is bijective for all fixed $k_i' \in K_i'$. If $DP((f_i)_{max}) \leq p$ for each i, then $DP(F_{max}) \leq p^2$.*

Proof. Let the input difference of F be $\alpha = (\alpha_1, \ldots, \alpha_n) \neq 0$ and the output difference be $\beta = (\beta_1, \ldots, \beta_n) \neq 0$, where $\alpha_i, \beta_i \in GF(2)^B$ for $i = 1, 2, \ldots, n$. Also let the output difference of F_1 be ϵ.

In general, the input-output differences for all F_i's in the n-cell GF-NLFSR can be summarized as follows:

$$\begin{array}{rcl}
\alpha_1 & \xrightarrow{F_1} & \epsilon \\
\alpha_2 & \xrightarrow{F_2} & \epsilon + \alpha_2 \qquad\qquad +\beta_1 \\
\alpha_3 & \xrightarrow{F_3} & \epsilon + \alpha_2 + \alpha_3 \qquad\qquad +\beta_1 + \beta_2 \\
\vdots & & \vdots \qquad\qquad\qquad\qquad \vdots \\
\alpha_n & \xrightarrow{F_n} & \epsilon + \alpha_2 + \alpha_3 + \ldots + \alpha_n \; +\beta_1 + \beta_2 + \ldots + \beta_{n-1} \\
\epsilon + \alpha_2 + \alpha_3 + \ldots + \alpha_n & \xrightarrow{F_{n+1}} & \beta_1 + \beta_2 + \ldots + \beta_{n-1} + \beta_n
\end{array} \tag{1}$$

From Fact 1, we have the following:

$$DP(\alpha \xrightarrow{F} \beta) = \sum_{\epsilon \in GF(2)^B} DP(\alpha_1 \xrightarrow{F_1} \epsilon) DP(\alpha_2 \xrightarrow{F_2} \epsilon + \alpha_2 + \beta_1) DP(\alpha_3 \xrightarrow{F_3} \epsilon + \alpha_2 + \alpha_3 + \beta_1 + \beta_2) \ldots DP(\epsilon + \alpha_2 + \ldots + \alpha_n \xrightarrow{F_{n+1}} \beta_1 + \ldots + \beta_n). \quad (2)$$

We shall show that at least 2 input differences in Equation 2 are non-zero when $\alpha \neq 0$. This implies that $DP(\alpha \xrightarrow{F} \beta) \leq p^2$. It suffices to prove this fact for the cases where only one of $\alpha_1, \alpha_2, \ldots, \alpha_n$ is non-zero.

(1) Suppose that only $\alpha_1 \neq 0$, then $\epsilon \neq 0$ (otherwise, $DP(\alpha_1 \xrightarrow{F_1} \epsilon = 0)$). Therefore, the input difference of F_{n+1}, i.e. $\epsilon + \alpha_2 + \ldots + \alpha_n = \epsilon$, is non-zero.
(2) Suppose that only $\alpha_2 \neq 0$, then the input difference of F_{n+1}, i.e. $\epsilon + \alpha_2 + \ldots + \alpha_n = \alpha_2$, is non-zero.

$\vdots$

(n) Suppose that only $\alpha_n \neq 0$, then the input difference of F_{n+1}, i.e. $\epsilon + \alpha_2 + \ldots + \alpha_n = \alpha_n$, is non-zero.

Therefore, at least 2 of the input differences are non-zero and $DP(\alpha \xrightarrow{F} \beta) \leq p^2$. □

5 Linear Hull Probability

We also have a result similar to Fact 2 for the linear hull probability of GF-NLFSR over $n+1$ rounds where the internal state is split into n equally sized blocks.

Theorem 2 *Let F be the $(n+1)$-round function in Figure 2 (right) of Appendix B where each $F_i : GF(2)^B \times GF(2)^B \times K_i' \mapsto GF(2)^B$ is of the form $F_i(x, k_i, k_i') = f_i(x + k_i, k_i')$ and each $f_i : GF(2)^B \times K_i' \mapsto GF(2)^B$ is bijective for all fixed $k_i' \in K_i'$. If $LP((f_i)_{max}) \leq q$ for each i, then $LP(F_{max}) \leq q^2$.*

Proof. Let the output mask value of F be $u = (u_1, \ldots, u_n) \neq 0$ and the input mask value be $w = (w_1, \ldots, w_n) \neq 0$. If the output mask value of F_1 is ϵ, it can be easily derived that we have the following individual round approximations:

$$\begin{array}{lllllll}
\epsilon & \xleftarrow{F_1} & w_1 & & & & \\
u_1 + u_2 & \xleftarrow{F_2} & \epsilon & + w_2 & & & \\
u_2 + u_3 & \xleftarrow{F_3} & \epsilon & & + w_3 & + u_1 + u_2 & \\
\vdots & & & \vdots & & \vdots & \\
u_{n-1} + u_n & \xleftarrow{F_n} & \epsilon & & + w_n + u_1 & & + u_{n-1} \\
u_n & \xleftarrow{F_{n+1}} & \epsilon & & & + u_1 & + u_n
\end{array}$$

Then Fact 1 gives

$$LP(u \xleftarrow{F} w) = \sum_{\epsilon \in GF(2)^B} LP(\epsilon \xleftarrow{F_1} w_1) LP(u_1 + u_2 \xleftarrow{F_2} \epsilon + w_2) LP(u_2 + u_3 \xleftarrow{F_3} \epsilon + w_3 + u_1 + u_2) \ldots LP(u_{n-1} + u_n \xleftarrow{F_n} \epsilon + w_n + u_1 + u_{n-1}) LP(u_n \xleftarrow{F_{n+1}} \epsilon + u_1 + u_n). \quad (3)$$

We shall show that at least 2 output mask values in Equation 3 are non-zero when $u, w \neq 0$. This will then imply that $LP(u \xleftarrow{F} w) \leq q^2$. If all the output mask values are equal to 0, i.e.

$$\epsilon = u_1 + u_2 = u_2 + u_3 = \ldots = u_{n-1} + u_n = u_n = 0,$$

then

$$\begin{aligned} & u_1 = u_2 = \ldots = u_n = 0 \\ \Rightarrow\ & u = 0 \end{aligned}$$

which gives a contradiction. Therefore, at least 1 output mask value is non-zero. Now we show that if only one of them is non-zero, then we will arrive at a contradiction.

(1) Suppose that only $\epsilon \neq 0$. Then $u_1 = u_2 = \ldots = u_n = 0$ which is a contradiction since $u \neq 0$.

(2) Suppose that only $u_1 + u_2 \neq 0$. Note that if $\epsilon = 0$, then $w_1 = 0$; otherwise, $LP(\epsilon \xleftarrow{F_1} w_1) = 0$. If $w_1 = 0$, then for other non-zero values of ϵ, $LP(\epsilon \xleftarrow{F_1} w_1) = 0$.

$$\begin{aligned} & \epsilon = u_2 + u_3 = u_3 + u_4 = \ldots = u_{n-1} + u_n = u_n = 0 \\ \Rightarrow\ & \epsilon + u_1 + u_n = u_1 = 0 \ \text{ (otherwise, } LP(u \xleftarrow{F} w) = 0) \\ & \text{and } \ u_2 = u_3 = \ldots = u_n = 0 \\ \Rightarrow\ & u = 0 \end{aligned}$$

which gives a contradiction.

(3) Suppose that only $u_2 + u_3 \neq 0$. Then

$$\begin{aligned} & \epsilon = u_1 + u_2 = u_3 + u_4 = \ldots = u_{n-1} + u_n = u_n = 0 \\ \Rightarrow\ & \epsilon + u_1 + u_n = u_1 = 0 \ \text{ (otherwise, } LP(u \xleftarrow{F} w) = 0) \\ \Rightarrow\ & u_2 = u_1 = 0 \ \text{ and } \ u_3 = u_4 = \ldots = u_n = 0 \\ \Rightarrow\ & u = 0 \end{aligned}$$

which gives a contradiction.

$$\vdots$$

(n) Suppose that only $u_{n-1} + u_n \neq 0$. Then

$$\begin{aligned} & \epsilon = u_1 + u_2 = u_2 + u_3 = \ldots = u_{n-2} + u_{n-1} = u_n = 0 \\ \Rightarrow\ & \epsilon + u_1 + u_n = u_1 = 0 \ \text{ (otherwise, } LP(u \xleftarrow{F} w) = 0) \\ \Rightarrow\ & u_1 = u_2 = \ldots = u_{n-1} = 0 \ \text{ and } \ u_n = 0 \\ \Rightarrow\ & u = 0 \end{aligned}$$

which gives a contradiction.

$(n+1)$ Suppose that only $u_n \neq 0$. Then

$$\begin{aligned}
&\epsilon = u_1 + u_2 = u_2 + u_3 = \ldots = u_{n-1} + u_n = 0 \\
\Rightarrow\ &u_1 = u_2 = \ldots = u_{n-1} = u_n \\
\Rightarrow\ &w_1 = 0,\ \epsilon + w_2 = w_2 = 0,\ \epsilon + w_3 + u_1 + u_2 = w_3 = 0, \ldots, \\
&\epsilon + w_n + u_1 + u_{n-1} = w_n = 0 \ \text{ (otherwise, } LP(u \xleftarrow{F} w) = 0) \\
\Rightarrow\ &w = 0
\end{aligned}$$

which gives a contradiction.

Therefore, at least 2 of the output mask values must be non-zero and $LP(u \xleftarrow{F} w) \leq q^2$. □

6 Frequencies of Differential and Linear Hull Probabilities and Expected Value

Here we calculate the approximate number of input-output differences $(\alpha \longrightarrow \beta)$ or mask values $(u \longleftarrow w)$ with $DP(\alpha \xrightarrow{F} \beta) \leq p^x$ or $LP(u \xleftarrow{F} w) \leq q^x$ respectively $(x = 2, \ldots, n)$. With reference to the sequence of differences and mask values stated in Sections 4 and 5, let $\Delta = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $\Omega = \{u_1 + u_2, u_2 + u_3, \ldots, u_{n-1} + u_n, u_n\}$.

Define $N_d(x)$ (respectively $N_l(x)$) as the number of input-output differences (α, β) (respectively input-output masks (w, u)) when there are x non-zero entries in Δ (respectively Ω). From the structure of n-cell, having x non-zero entries in Δ or Ω will ensure $DP(\alpha \xrightarrow{F} \beta) \leq p^x$ or $LP(u \xleftarrow{F} w) \leq q^x$ respectively. The only exception is when $x = 1$, where we still have $DP(\alpha \xrightarrow{F} \beta) \leq p^2$ or $LP(u \xleftarrow{F} w) \leq q^2$ by Theorems 1 and 2.

Various cases for the input-output pairs and their corresponding bounds are shown in Table 1 in Appendix A. When there are x non-zero entries in Δ, the number of possible input-output differences is given by $N_d(x) = \binom{n}{x}(2^B - 1)^x(2^{nB} - 1)$. This is because there are $\binom{n}{x}$ possible input differences with x non-zero entries where each non-zero entry has $2^B - 1$ possibilities, and there are $2^{nB} - 1$ possibilities for the non-zero output difference. We have an identical formula for $N_l(x)$ by a similar reason.

Based on the values $N_d(x)$ and $N_l(x)$, we see that when an attacker uses plaintexts such that the input differences α (output mask values u resp.) are randomly chosen, he is more likely to obtain a bound much lower than p^2 (q^2 resp.) since most of the input differences α (output mask values u resp.) give rise to differential probabilities $DP(\alpha \xrightarrow{F} \beta)$ (linear hull probabilities $LP(u \xleftarrow{F} w)$ resp.) whose bounds are much smaller than p^2 (q^2 resp.). Such a scenario may occur when, for example, the $(n+1)$-round structure is an intermediate portion of a cipher so that the attacker does not have much control over the input differences (output mask values resp.). This motivates our desire to have more practically useful differential and linear hull probability bounds. For this purpose, we make the following definitions:

Definition 1 *The* expected differential probability *is defined as* $E_d = \frac{\sum_{\alpha,\beta\neq 0} DP(\alpha \xrightarrow{F} \beta)}{\#\{(\alpha,\beta)|\alpha,\beta\neq 0\}}$ *and the* expected linear probability *is defined as* $E_l = \frac{\sum_{w,u\neq 0} LP(u \xleftarrow{F} w)}{\#\{(w,u)|w,u\neq 0\}}$

Note that $\sum_{x=2}^{n} N_d(x) = (2^{nB}-1)^2$ which is the total number of differences with both input and output non-zero. We may make a similar observation for the linear case. From this table, we may directly calculate the proportion of input-output differences (mask values resp.) with differential (linear hull resp.) probability $\leq p^x$ (q^x resp.). Denote the approximate proportion of input-output differences with differential probability $\leq p^x$ by $P_d(x) = \frac{N_d(x)}{\#\{(\alpha,\beta)|\alpha,\beta\neq 0\}}$. Likewise, denote the approximate proportion of input-output mask values with linear hull probability $\leq q^x$ by $P_l(x) = \frac{N_l(x)}{\#\{(w,u)|w,u\neq 0\}}$. It can be computed that the statistics are heavily skewed towards the lowest probabilities instead of p^2 or q^2. For example, when $n = 4$, $B = 8$, and when $n = 4$, $B = 16$, we have the following proportions shown in Table 2 in Appendix A.

Using the frequency values in Table 1, we can derive that

$$\begin{aligned}
E_d &\leq \frac{1}{(2^{nB}-1)^2}\left[\binom{n}{1}(2^B-1)\cdot(2^{nB}-1)p^2 + \sum_{x=2}^{n}\binom{n}{x}(2^B-1)^x\cdot(2^{nB}-1)p^x\right] \\
&< \frac{1}{(2^{nB}-1)}\left[\binom{n}{1}(2^B-1)p + \sum_{x=2}^{n}\binom{n}{x}(2^B-1)^x p^x\right] \\
&< \frac{1}{(2^{nB}-1)}\left[\sum_{x=0}^{n}\binom{n}{x}(2^B-1)^x p^x\right] \\
&= \frac{1}{(2^{nB}-1)}(1+(2^B-1)p)^n \\
&\approx (2^{-B}+p)^n, \qquad (4)
\end{aligned}$$

where we have approximated $2^B - 1$ and $2^{nB} - 1$ by 2^B and 2^{nB} respectively because B is usually much larger than 1. Similarly, we have $E_l \leq (2^{-B} + q)^n$.

For example, when $n = 4$, $B = 8$ and $p = 2^{-6}$, the bound in (4) is approximately $2^{-22.7}$, which is much better than the 2^{-12} bound obtained from Theorem 1.

7 Impossible Differential Characteristics

Impossible differential cryptanalysis is a variant of differential cryptanalysis against block ciphers. It was applied against Skipjack to reject wrong key candidates by using input and output difference pairs whose probabilities are zero. It can also be used to attack a 5-round Feistel structure even though the 3-round Feistel structure with bijective round functions are provably secure against differential and linear cryptanalysis.

In impossible differential cryptanalysis, impossible differential characteristics are used to retrieve a subkey material for the first or last several rounds of block ciphers. Thus the security of a block cipher against impossible differential cryptanalysis can be evaluated by impossible differential characteristics [26].

A general tool, called U-method, was introduced by [26] to find the maximum number of rounds for impossible differential characteristics. An algorithm, Algorithm 1, was also provided to compute the maximum length of impossible differential characteristics that can be found by the U-method. Interested readers may refer to [26] for the technicalities. By modifying Algorithm 1, we can determine the impossible differential characteristics of the block cipher structures. The following result for our block cipher n-cell GF-NLFSR is based on the simulation. Here, a r-round impossible differential characteristic is denoted by $\alpha \nrightarrow_r \beta$ where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$.

Proposition 3 *The maximum number of rounds for impossible differential characteristics that can be found by the U-method for n-cell GF-NLFSR is $2n-1$. Generalized impossible differential characteristics are*

$$(0, \cdots, 0, \alpha_n) \nrightarrow_{2n-1} (\beta_1, \beta_2, 0, \cdots, 0), \text{ where } \alpha_n \neq 0, \beta_1 = \beta_2 \neq 0,$$

and,

$$(0, \cdots, 0, \alpha_n) \nrightarrow_{2n-1} (\beta_1, 0, \cdots, 0, \beta_n), \text{ where } \alpha_n \neq 0, \beta_1 = \beta_n \neq 0.$$

In particular, when $n = 4$, a 7-round impossible differential characteristic is $(0, 0, 0, \gamma) \nrightarrow_7 (\gamma, \gamma, 0, 0)$, with the input and output differences to and after each round as follows:

$$(0,0,0,\gamma) \rightarrow (0,0,\gamma,\gamma) \rightarrow (0,\gamma,\gamma,0) \rightarrow (\gamma,\gamma,0,0) \rightarrow (\gamma,0,0,\gamma+\delta) \rightarrow (0,0,\gamma+\delta,?)$$

$$\rightarrow (0,\gamma+\delta,?,?) \neq (0,\gamma,\gamma,0) \leftarrow (\gamma,\gamma,0,0),$$

where γ, δ and ? denote nonzero nonfixed, nonzero fixed, and, nonfixed differences respectively. We can thus use a 7-round impossible differential to conduct an impossible differential attack.

8 Integral Attack

The integral attack is a cryptanalytic technique on block ciphers. It was originally proposed by Knudsen and Wagner in 2002 [9] and has since been adapted to cryptanalyse various ciphers. In this attack, a set of chosen plaintexts is encrypted and the corresponding ciphertexts are decrypted a certain number of rounds using all possible subkey guesses. The plaintext set is chosen such that one part is held constant while another part varies over all possibilities. Using this plaintext set, a distinguisher is produced after a certain number of rounds, which enables the attacker to determine the correct subkey which was used for partial decryption. We shall show that for $n \geq 2$, GUFN has a $(3n-1)$-round distinguisher, where n is the number of blocks.

In this section, we let the n-tuple $(A, c, \ldots, c)$ denote a set of 2^B plaintexts where the leftmost block of B bits vary over all 2^B possibilities, while the other blocks, represented by c, are constants. We shall let uppercase letters represent a set that varies over all values in $GF(2)^B$. Thus, we obtain the set $(c, c, \ldots, c, D + c)$ after a

round of encryption. Consequently, we have the following sequence of round-by-round output sets after n rounds:

$$\begin{aligned}(A, c, \ldots, c) &\rightarrow (c, c, \ldots, c, D + c)\\ &\vdots\\ &\rightarrow (c, D + c, D + c, c, \ldots, c)\\ &\rightarrow (D + c, D + c, c, c, \ldots, c).\end{aligned}$$

After another n rounds we have the following:

$$\begin{aligned}(D + c, D + c, c, \ldots, c) &\rightarrow (D + c, c, \ldots, c, D + E + c)\\ &\rightarrow (c, \ldots, c, D + E + c, D + E + G + c)\\ &\rightarrow (c, \ldots, c, D + E + c, D + E + G + c, G + c)\\ &\rightarrow (c, \ldots, c, D + E + c, D + E + G + c, G + c, c)\\ &\vdots\\ &\rightarrow (D + E + c, D + E + G + c, G + c, c, \ldots, c).\end{aligned}$$

The distinguishing property of certain values in the set is subsequently destroyed block by block. We obtain this, after another $n - 1$ rounds:

$$\begin{aligned}(D + E + c, D + E + G + c, G + c, c, \ldots, c) &\rightarrow (D + E + G + c, G + c, c, \ldots, c, ?)\\ &\vdots\\ &\rightarrow (c, ?, ?, \ldots, ?).\end{aligned}$$

The set of ciphertexts after $3n - 1$ rounds of encryption will be constant in the leftmost block. This property can be exploited as a distinguisher if the cipher has only slightly more than $3n - 1$ rounds. Although some minor detail of the above proof does not apply for the cases $n = 2$ and $n = 3$, it can be easily verified, using a similar approach, that the $(3n - 1)$-round result still holds for these values of n.

9 Application : New Block Cipher Four-Cell

As an application, we design a new 128-bit block cipher, Four-Cell, with 128-bit key size. It uses the block cipher structure described in Section 3 with four cells where each cell is a 32-bit word. The block cipher has 25 rounds and uses two types of nonlinear functions for round i, defined as follows:

$$f_i(x_i, k_i, 0) = MDS(S(x_i + k_i)), \text{ for rounds } i = 1, 2, \ldots, 5 \text{ and } i = 21, 22, \ldots, 25.$$

$$f_i(x_i, k_i, k_i') = S(MDS(S(x_i + k_i)) + k_i'), \text{ for rounds } i = 6, 7, \ldots, 20.$$

Here, $S : GF(2^8)^4 \rightarrow GF(2^8)^4$ is defined as

$$S(x_1, x_2, x_3, x_4) = (Inv(x_1), Inv(x_2), Inv(x_3), Inv(x_4)),$$

where $Inv : GF(2^8) \rightarrow GF(2^8)$ is affine equivalent to $x \mapsto x^{254}$ on $GF(2^8)$ (e.g., the AES S-box). $MDS : GF(2^8)^4 \rightarrow GF(2^8)^4$ is a 4-byte to 4-byte maximal distance separable transform with optimal branch number 5 (e.g., the MixColumn operation in AES). Note that one subkey and one layer of S-box is used for rounds $1, 2, \ldots, 5$ and $21, 22, \ldots, 25$ while two subkeys and two layers of S-boxes are used for rounds $6, 7, \ldots, 20$. Moreover, we XOR a 128-bit post-whitening key K_{26} to the output after 25 rounds.

We leave the implementation of a secure key schedule open to the reader. One possibility would be to use a similar cipher with 26 rounds as the key schedule. The only difference is that the nonlinear function for all rounds is defined as $f_i(x_i, c_i, 0) = MDS(S(x_i + c_i))$ where the c_i's are distinct randomly generated constants. The input to the key-schedule is the secret key K. The least significant 32 output bits (i.e., nonlinear output) of round i of the key schedule can be used as the i^{th}-round cipher subkey k_i. For rounds $i = 6, 7, \ldots, 20$, we can take the next 32 least significant output bits of round i of the key schedule to be k'_i. The post-whitening key K_{26} is the 128-bit output of round 26 of the key schedule.

In the following section, we demonstrate the security of Four-Cell against a slew of cryptanalytic attacks, in addition to differential and linear cryptanalysis.

9.1 Security of Four-Cell

In Sections C.1 and C.2 in Appendix, we show that the differential and linear characteristic probabilities of Four-Cell are at most $2^{-192} < 2^{-128}$. Therefore it is practically secure against differential and linear cryptanalysis. In Section C.3 in Appendix, we show that the true differential and linear probabilities of Four-Cell are at most $2^{-55.39}$ and $2^{-52.96}$ respectively. However, this bound is tight only for a negligible number of input-output differences and masks. The expected differential and linear probabilities are actually $2^{-110.5}$ and $2^{-105.79}$ respectively. Based on the true differential probability, we show in Section C.4 in Appendix that if we split Four-Cell into two sub-ciphers with true differential probabilities p and q, then $(pq)^2 \leq 2^{-221.57} \ll 2^{-128}$. This will ensure Four-Cell is secure against boomerang attack. In Section C.5 in Appendix, we show that there is an 10-round attack based on a 7-round impossible differential distinguisher. But it is unlikely that it will work against the full cipher which needs a 21-round distinguisher. In Section C.6 in Appendix, we show that there is a 14-round attack based on an 11-round integral attack distinguisher. But it is unlikely that it will work against the full cipher which needs a 21-round distinguisher. In Section C.7 in Appendix, we show that Four-Cell is secure against higher order differential and cube attacks after 9 rounds, because the algebraic degree of the cipher attains the maximum degree 127. We also explain that interpolation attack might not work as the cipher will be a complex multivariable equation over $GF(2^8)$. In Section C.8, we give some background on the XSL attacks and explain why it might not work on our cipher. Finally in Section C.9, we explain that Four-Cell is secure against slide attack because of its distinct round structures and distinct round subkeys.

9.2 Implementation Considerations

The Four-Cell cipher uses 160 S-boxes based on the inversion function on $GF(2^8)$. This is the same as the number of S-boxes used in AES. However only 25 MDS transform are used when compared to AES, which uses 40 MDS transforms. This might make the cipher faster in hardware implementations where the S-box and MDS are not combined into a T-table. Moreover, note that the computation of the nonlinear function in any 4 consecutive rounds of the cipher can be performed in parallel for faster encryption speed, giving it an added advantage over other GUFNs such as SMS4. Thus the Four-Cell cipher which (like the AES cipher) has provable security against existing block cipher attacks can be viewed as a viable alternative.

Also note that although the inverse cipher of Four-cell is distinct from Four-cell itself and therefore coding might potentially take up more space in hardware, it is still useful for modes of operation such as counter mode, output feedback (OFB) mode, and cipher feedback (CFB) mode, where no inverse cipher is required.

10 Acknowledgement

The authors would like to thank the anonymous reviewer of CT-RSA who pointed out the integral attack on Four-Cell.

References

1. E. Biham and A. Shamir, "Differential Cryptanalysis of the Data Encryption Standard", Springer-Verlag, New York, 1993.
2. A. Biryukov and D. Wagner, "Slide Attack", LNCS 1636, *FSE'99*, pp. 245-259, Springer-Verlag, 1999.
3. C. Cid and G. Leurent, "An Analysis of the XSL Algorithm", LNCS 3788, *Asiacrypt 2005*, pp. 333-352, Springer-Verlag, 2005.
4. N. Courtois and J. Pieprzyk, "Cryptanalysis of Block Ciphers with Overdefined Systems of Equations", IACR eprint server, `http://www.iacr.org`, 2002/044, March 2002.
5. N. Courtois and J. Pieprzyk, "Cryptanalysis of Block Ciphers with Overdefined Systems of Equations", LNCS 2501, *Asiacrypt 2002*, pp. 267-287, Springer-Verlag, 2002.
6. J. Daemen and V. Rijmen, *The Design of Rijndael: AES, The Advanced Encryption Standard*, Springer, 2002.
7. I. Dinur and A. Shamir, "Cube Attacks on Tweakable Black Box Polynomials", *Cryptology Eprint Archive*, Report 2008/385.
8. T. Jakobsen and L. R. Knudsen, "Attacks on Block ciphers of Low Algebraic Degree", *Journal of Cryptology*, vol.14, pp. 197-210, Springer, 2001.
9. L.R. Knudsen and D. Wagner, "Integral Cryptanalysis", LNCS 2365, *Fast Encryption Software 2002*, pp. 112-127, Springer, 2002.
10. X. Lai, "On the Design and Security of Block Ciphers", Thesis, 1992.
11. X. Lai, J.L. Massey, and S. Murphy, "Markov Ciphers and Differential Cryptanalysis", *Eurocrypt 1991*, LNCS 547, pp. 17-38, Springer-Verlag, 1991.
12. C.W. Lim and K. Khoo, "An Analysis of XSL Applied on BES", LNCS 4593, *Fast Software Encryption 2007*, pp. 242-253, Springer-Verlag, 2007.

13. F. Liu, W. Ji, L. Hu, J. Ding, S. Lv, A. Pyshkin, and R. Weinmann, "Analysis of the SMS4 Block Cipher", *ACISP 2007*, LNCS 4586, pp. 158-170, Springer-Verlag, 2007.
14. M. Matsui, "Linear Cryptanalysis Method for DES Cipher.", *Eurocrypt 1993*, LNCS 765, Springer-Verlag, 1994.
15. M. Matsui, "New Structure of Block Ciphers with Provable Security Against Differential and Linear Cryptanalysis", *Fast Software Encryption 1996*, LNCS 1039, pp. 205-218, Springer-Verlag, 1996.
16. S. Murphy and M. Robshaw, "Essential Algebraic Structure within the AES", LNCS 2442, *Crypto 2002*, pp. 1-16, Springer-Verlag, 2002.
17. K. Nyberg, "Linear Approximation of Block Ciphers", *Eurocrypt 1994*, LNCS 950, pp. 439-444, Springer-Verlag, 1994.
18. K. Nyberg, "Generalized Feistel Networks", *Asiacrypt 1996*, LNCS 1163, pp. 91-104, Springer, 1996.
19. S. Park, S.H. Sang, S. Lee, and J. Lim, "Improving the Upper Bound on the Maximum Differential and the Maximum Linear Hull Probability for SPN Structures and AES", *Fast Software Encryption 2003*, LNCS 2887, pp. 247-260, Springer-Verlag, 2003.
20. B. Schneier and J. Kelsey, "Unbalanced Feistel Networks and Block-Cipher Design", *Fast Software Encryption 1996*, LNCS 1039, pp. 121-144, Springer, 1996.
21. T. Shirai, K. Shibutani, T. Akishita, S. Moriai and T. Iwata, "The 128-bit Blockcipher CLEFIA (Extended Abstract)", *Fast Software Encryption 2007*, LNCS 4593, pp. 181-195, Springer-Verlag, 2007.
22. J. Daemen, L. Knudsen, and V. Rijmen, "The Block Cipher Square", *Fast Software Encryption 1997*, LNCS 1267, pp. 149-165, Springer-Verlag, 1997.
23. D. Wagner, "The Boomerang Attack", *Fast Software Encryption 1999*, LNCS 1636, pp. 156-170, Springer-Verlag, 1999.
24. J. Wallen, "Design Principles of the KASUMI Block Cipher", http://www.tml.tkk.fi/Opinnot/Tik-110.501/2000/papers/wallen.pdf, June 2008.
25. W. Wu, W. Zhang, and D. Lin, "On the Security of Generalized Feistel Scheme with SP Round Function", International Journal of Network Security, Vol. 3, No. 3, pp. 215-224, 2006.
26. 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.

A Tables

Table 1. Frequencies of differential and linear hull probabilities

Differential probability	Linear hull probability	$N_d(x)/N_l(x)$	# of elements in Δ (or Ω resp.) which are non-zero
$\leq p^n$	$\leq q^n$	$(2^B-1)^n \cdot (2^{nB}-1)$	n
$\leq p^{n-1}$	$\leq q^{n-1}$	$\binom{n}{n-1}(2^B-1)^{n-1} \cdot (2^{nB}-1)$	$n-1$
$\leq p^{n-2}$	$\leq q^{n-2}$	$\binom{n}{n-2}(2^B-1)^{n-2} \cdot (2^{nB}-1)$	$n-2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\leq p^3$	$\leq q^3$	$\binom{n}{3}(2^B-1)^3 \cdot (2^{nB}-1)$	3
$\leq p^2$	$\leq q^2$	$\binom{n}{2}(2^B-1)^2 \cdot (2^{nB}-1)$	2
$\leq p^2$	$\leq q^2$	$\binom{n}{1}(2^B-1) \cdot (2^{nB}-1)$	1

Table 2. Distribution of proportions

Differential probability	Linear hull probability	x	$P_d(x)/P_l(x)$	
			$n=4, B=8$	$n=4, B=16$
p^4	q^4	4	0.9844663148	0.9999389662
p^3	q^3	3	0.1544260886	0.0000610323
p^2	q^2	2	0.0000910763	$0.1396955440 \times 10^{-8}$

Table 3. Distribution of Differential and Linear Hull Probabilities of the Four-Cell Cipher

Differential probability	Linear hull probability	x	Frequency	$P_d(x)/P_l(x)$
$2^{-110.78}$	$2^{-105.91}$	4	$2^{256.000000}$	0.999999
$2^{-83.09}$	$2^{-79.43}$	3	$2^{226.000000}$	9.31×10^{-10}
$2^{-55.39}$	$2^{-52.96}$	2	$2^{194.584962}$	3.25×10^{-18}

B Figures

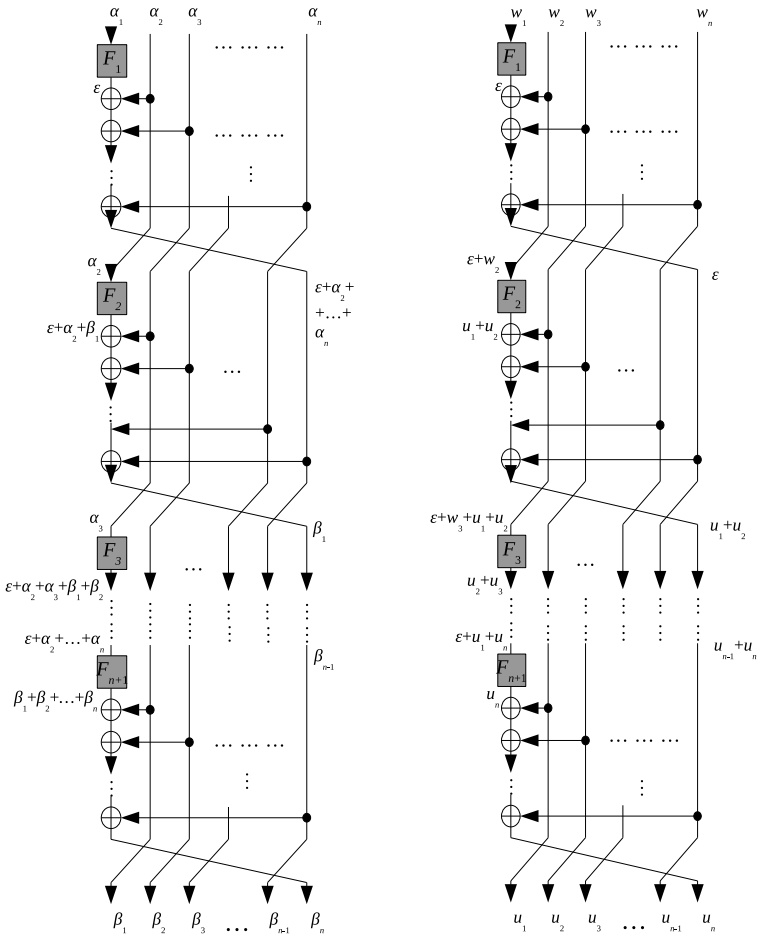

Fig. 2. Sequence of differences(left)/mask values(right) for $n+1$ rounds of GF-NLFSR

C Security of Four-Cell

C.1 Security of Four-Cell against Differential Cryptanalysis

Because of the structure of the cipher, an adversary can guess the post-whitening and some of the last four subkeys and perform the attack with a 21-round differential distinguisher. Each nonlinear function for rounds $i = 1, 2, \ldots, 5$ has the same maximal differential probability as an S-box, which is $4/256 = 2^{-6}$ from [6]. The differential probability of the first five rounds is at most $(2^{-6})^2 = 2^{-12}$ by Theorem 1. The nonlinear function for rounds $i = 6, 7, \ldots, 20$ has differential characteristic probability at most $(4/256)^5 = 2^{-30}$ because of the effect of the MDS transform which causes at least 5 S-boxes to be active. The differential characteristic probability for the next 15 rounds is at most $(2^{-30})^2 \times (2^{-30})^2 \times (2^{-30})^2 = 2^{-180}$ by Theorem 1. The probability of any 20-round (and thus 21-round) differential characteristic is at most:

$$2^{-180} \times 2^{-12} = 2^{-192} < 2^{-128} = 2^{-blocksize}.$$

Therefore Four-Cell is secure against differential cryptanalysis.

C.2 Security of Four-Cell against Linear Cryptanalysis

In a similar way, we will estimate the correlation of a 20-round linear characteristic. From [6], the correlation of an S-box has magnitude at most 32 which implies a linear probability of at most $(32/256)^2 = 2^{-6}$. Similar to our reasoning for differential cryptanalysis, we can split the first 20 rounds into four 5-round sub-ciphers and apply Theorem 2. We see that the linear characteristic probability for 20 rounds (and thus 21 rounds) of the cipher is at most:

$$(2^{-6})^2 \times ((2^{-6})^5)^2 \times ((2^{-6})^5)^2 \times ((2^{-6})^5)^2 = 2^{-192} < 2^{-128} = 2^{-blocksize}.$$

Thus the cipher is secure against linear cryptanalysis.

C.3 Actual Differential and Linear Hull Probability of Four-Cell

We shall need the following result from [19].

Proposition 4 *([19, Theorem 1 and 2]) Assume that the round keys, which are XORed to the input data at each round, are independent and uniformly random. If $Br(D) = k$, the probability of each differential of the SDS structure is bounded by:*

$$\max\left(\max_{1\le i\le n}\max_{1\le u\le 2^m-1}\sum_{j=1}^{2^m-1} DP^{S_i}(u\to j)^k, \max_{1\le i\le n}\max_{1\le u\le 2^m-1}\sum_{j=1}^{2^m-1} DP^{S_i}(j\to u)^k\right)$$

The linear hull probability of the SDS structure is bounded by:

$$\max\left(\max_{1\le i\le n}\max_{1\le u\le 2^m-1}\sum_{j=1}^{2^m-1} LP^{S_i}(u\to j)^k, \max_{1\le i\le n}\max_{1\le u\le 2^m-1}\sum_{j=1}^{2^m-1} LP^{S_i}(j\to u)^k\right)$$

Using Proposition 4, Park proved that the differential probability of the SDS structure corresponding to the nonlinear function F_i for $i = 6, 7, \ldots, 15$ of the Four-Cell cipher is $2^{-27.696}$ in [19, Section 4]. The linear correlation of each S-box takes the values

$$0, \pm 4, \pm 8, \pm 12, \pm 16, \pm 20, \pm 24, \pm 28, \pm 32$$

with frequencies

$$17, 48, 36, 40, 34, 24, 36, 16, 5,$$

respectively. By substituting these values for the linear probability ($= (correlation/256)^2$) in Proposition 4, the linear hull probability is at most $2^{-26.478}$.

By substituting the differential and linear hull probabilities of the SDS structure in Table 1 of Appendix A, we get in Table 3 the distribution of any 5 rounds of the Four-Cell cipher between rounds $i = 6, 7, \ldots, 20$. This will also give an upper bound for the differential and linear hull probabilities of the cipher. Table 3 shows that for 5 intermediate rounds of the cipher, the true differential and linear probabilities are at most $2^{-55.39}$ and $2^{-52.96}$ respectively. However, this happens only for a negligible number of input-output differences over 5 rounds. Over more rounds or when the input differences cannot be controlled, a more accurate measure is the expected differential and linear probability over 5 rounds, which is given by $(2^{-32} + 2^{-27.696})^4 \approx 2^{-110.5}$ and $(2^{-32} + 2^{-26.478})^4 \approx 2^{-105.79}$ respectively.

C.4 Protection against Boomerang Attacks

There is also a stronger form of differential attack called boomerang attack [23]. It splits $R - 4$ rounds of Four-Cell into 2 shorter ciphers such that the differential probability of each part is known to be large, say with probability p for the differential $\alpha \to \beta$ for the first part and probability q for the differential $\gamma \to \delta$ for the second part. The distinguisher is the following boomerang process:

(1) Ask for the encryption of a pair of plaintexts (P_1, P_2) such that $P_1 + P_2 = \alpha$ and denote the corresponding ciphertexts by (C_1, C_2).
(2) Calculate $C_3 = C_1 + \delta$ and $C_4 = C_2 + \delta$, and ask for the decryption of the pair (C_3, C_4). Denote the corresponding plaintexts by (P_3, P_4).
(3) Check whether $P_3 + P_4 = \alpha$.

For a random permutation, the probability that the last condition is satisfied is $2^{-blocksize}$. The probability that a quartet of plaintexts and ciphertexts satisfies the boomerang conditions is $(pq)^2$. Therefore, we have a distinguisher which distinguishes between the cipher being attacked and a random cipher if $(pq)^2 < 2^{-blocksize}$.

For Four-Cell, similar to our computation of the 20-round differential characteristic probability, 15 rounds of the cipher already has maximal differential characteristic probability $((2^{-6})^5)^2 \times ((2^{-6})^5)^2 \times (2^{-6})^2 = 2^{-132}$ which is less than 2^{-128}. Thus it is unlikely that an adversary can find a good differential over 15 rounds and any good differential is likely to involve 14 or less rounds. Thus when the adversary splits $25 - 4 = 21$ rounds into two sub-ciphers, they will each contain at least 5 rounds where the nonlinear function involves 2 layers of S-boxes. We just saw that the differential

probabilities p, q of the 2 sub-ciphers are at most $2^{-55.39}$. Thus $(pq)^2 \leq 2^{-221.570} < 2^{-128}$ and Four-Cell is secure against boomerang attack.

In another variant of the boomerang attack, intermediate differences, β and γ, are allowed to vary so that the adversary only needs to find several high probability differential paths of the same initial and terminal differences α and δ. However, to get a good distinguisher for the attack to succeed, the adversary would need to identify more than $2^{221.570-128} = 2^{93.57}$ such optimal paths, which is unlikely.

Remark 1. We have used the assumption that if the differential characteristic probability of R' rounds of a cipher is less than $2^{-blocksize}$, then it is not likely that a good differential over R' rounds can be found. This is in line with the common approach of practical provable security against differential cryptanalysis employed in the proofs of security of ciphers like AES [6]. Thus if our assumption is not true, then the approach is wrong because although we can prove that the differential characteristic probability is less than $2^{-blocksize}$, we can still find a differential with high probability to launch differential cryptanalysis.

C.5 Protection against Impossible Differential Attack

According to the attack in Section 7, there is a 7-round impossible characteristic that begins with the differential $(0, 0, 0, \gamma)$ and ends in the differential $(\gamma, \gamma, 0, 0)$. This can be used in a 8, 9 or 10-round attack with complexity 2^{32}, 2^{64} or 2^{96} by guessing up to three subkeys to verify if the 7^{th} round output has the required impossible differential. To launch an impossible differential attack on the full 25-round Four-Cell cipher, the adversary would need to guess 128 bits of the post-whitening key K_{26}. Even if the adversary can bypass this post-whitening key, he can extend an impossible differential characteristic by at most four rounds. That means he would need to extend the impossible differential characteristic from 7 to $25 - 4 = 21$ rounds which seems unlikely.

C.6 Protection against Integral Attack

According to the attack in Section 8, there is an 11-round integral attack distinguisher: the adversary starts with (A, c, c, c), where the first 32-bit word A ranges through all 2^{32} vectors and the other words are kept constant; and the 11-round output would be of the form $(c, ?, ?, ?)$, which is constant in the first word and unknown in the remaining three words. This can be exploited in a 12, 13 or 14-round attack with complexity 2^{32}, 2^{64} or 2^{96} by guessing up to three subkeys to verify if the 11^{th} round output (computed from the 2^{32} ciphertexts) is constant in the first word. To launch an integral attack on the full 25-round Four-Cell cipher, the adversary would need to guess, in addition to the subkeys used in the round functions, 128 bits of post-whitening key K_{26}. Even if the adversary can bypass the post-whitening key, he can extend an integral attack distinguisher by at most four rounds. That means he would need to extend the integral attack distinguisher from 11 to $25 - 4 = 21$ rounds which seems unlikely.

C.7 Protection against Higher Order Differential, Cube and Interpolation Attacks

The algebraic degree of any round with a single layer of S-box (rounds 1-5, 21-25) is 7 while that of any round with two layers of S-boxes (rounds 6-20) is 49. By the 4^{th} round, every output bit will have degree 7. By the 8^{th} round, 1 output word will have degree 49 (composition of two balanced functions both of degree 7) while 3 output words will have degree 127 (composition of two balanced function of degree 7 and 49). By the 9^{th} round, all output words will have degree 127 (maximal degree for balanced 128-bit function). There are two known attacks, higher order differential [8] and later, cube attacks [7], which exploits the algebraic degree d of a block cipher in terms of the plaintext. However, both has data and time complexity of magnitude $O(2^d)$. Thus, they will be ineffective against Four-Cell when there are 9 or more rounds.

The interpolation attack [8] works on block ciphers that can be expressed as an equation in $GF(2^n)$ with few monomials. In Four-Cell, if we use the AES S-box, each S-box is a sum of 8 monomials in $GF(2^8)$. However, if we compose the S-boxes with the MDS transforms (e.g. AES MixColumn) over several rounds, the block cipher will become a complex multi-variable function which is a sum of many monomials over $GF(2^8)$. Thus it will be secure against interpolation attack. Moreover, the cipher can be made more secure against interpolation attack by choosing each S-box to be a random pre- and post-affine transform of x^{-1}. Then the expression for each S-box will be a sum of 255 monomials over $GF(2^8)$.

C.8 Protection against XSL Attacks

The XSL attack [4, 5] tries to solve the sparse equations formed by the plaintext, ciphertexts and intermediate variables in a block cipher computation. It multiplies the existing equations with monomials to form new equations. The aim is to do this process intelligently so as to increase the number of equations more quickly than the number of new monomials, such that eventually we have more equations than monomials. Then the set of equations can be solved by linearization to reveal the secret key.

However, the second XSL attack (the better of the two attacks in [4]) has complexity 2^{203} and is ineffective against AES-128. Four-Cell and AES-128 are block ciphers with 128-bit block size, and have a comparable number of S-boxes (based on x^{-1} in $GF(2^8)$) in both the main cipher and key schedule. So we expect them to be described by a comparable number of equations and monomials which might lead to similarly ineffective XSL attack complexity.

There is also a more powerful attack where we embed AES in $GF(2^8)$ to form the BES (Big Encryption Standard) cipher [16]. In that case, each S-box can be expressed as 24 quadratic equations based on 41 monomials over $GF(2^8)$, instead of 24 quadratic equations based on 81 monomials over $GF(2)$. When we apply the second XSL attack [4] on BES, the reduction in the number of monomials causes the attack complexity to drop from 2^{203} to only 2^{87}. However, it has been shown in [12] that the number of linearly independent equations in the XSL attack on BES might be over-estimated and the actual complexity of this attack should be at least 2^{401}.

In a similar way to [16], Four-Cell can also be embedded in $GF(2^8)$ to form a Big Four-Cell cipher. We can also show that the XSL attack will not work on this embedded Big Four-Cell cipher by a method similar to that in [12].

C.9 Protection against Slide Attack

The slide attack [2] works on ciphers which have cyclical structures over a few rounds, i.e. the cipher structure and subkeys are repeated over every few rounds. It can usually be protected against if the subkeys of each rounds are different. In our cipher, besides having a different subkey for every round, the cipher rounds are different between rounds 6 to 20 (which has two layers of S-boxes) and the other rounds (which has one layer of S-box). Therefore, slide attack will not work.