Parallelizing the Camellia and SMS4 Block Ciphers - Extended Version

Huihui Yap[1,2], Khoongming Khoo[1,2] and Axel Poschmann[2⋆]

[1] DSO National Laboratories, 20 Science Park Drive, Singapore 118230
[2] Division of Mathematical Sciences, School of Physical and Mathematical Sciences
Nanyang Technological University, Singapore
{yhuihui,kkhoongm}@dso.org.sg,aposchmann@ntu.edu.sg

Abstract. The n-cell GF-NLFSR (Generalized Feistel-NonLinear Feedback Shift Register) structure [8] is a generalized unbalanced Feistel network that can be considered as a generalization of the outer function FO of the KASUMI block cipher. An advantage of this cipher over other n-cell generalized Feistel networks, e.g. SMS4 [11] and Camellia [5], is that it is parallelizable for up to n rounds. In hardware implementations, the benefits translate to speeding up encryption by up to n times while consuming similar area and significantly less power. At the same time n-cell GF-NLFSR structures offer similar proofs of security against differential cryptanalysis as conventional n-cell Feistel structures. We also ensure that parallelized versions of Camellia and SMS4 are resistant against other block cipher attacks such as linear, boomerang, integral, impossible differential, higher order differential, interpolation, slide, XSL and related-key differential attacks.

Keywords: Generalized Unbalanced Feistel Network, GF-NLFSR, Camellia, SMS4

1 Introduction

1.1 Background and Motivation

Two very important security properties of block cipher structures are low differential and linear probability bounds for protection against differential and linear cryptanalysis. Choy et al. [8] had proven that the "true" differential/linear probabilities of any n rounds of the n-cell GF-NLFSR structure is p^2 if the differential/linear probability of the nonlinear function of each round is p. However, this result is applicable only if we use a nonlinear function with good provable differential/linear probability. One option is to use an S-box. However if the nonlinear function takes in 32-bit input, an S-box of this size would be infeasible to implement in terms of logic gates in hardware or as a look-up-table in memory. Other options would be to build a SDS (Substitution-Diffusion-Substitution) structure [21], use a Feistel structre [2] or even a nested Feistel structure for the nonlinear function [3], because there are provable bounds for the differential and linear probabilities of these structures.

However these nonlinear functions are too complex, and not suitable for space and speed efficient implementation. Therefore, the Substitution-then-Diffusion structure is usually implemented for the nonlinear functions. These structures are commonly called Substitution Permutation Networks (SPN) in the literature. Numerous examples of implementations where the SPN structure is used for the nonlinear functions of Feistel and Generalized Feistel Structures exist. They include DES [1], Camellia [5], SMS4 [11] and Clefia [22], to name a few. Motivated by these considerations, we would like to investigate the practical differential and linear probability bounds of the n-cell GF-NLFSR structure when the nonlinear function is a SPN structure.

⋆ The research was supported in part by the Singapore National Research Foundation under Research Grant NRF-CRP2-2007-03.

As applications, we would like to parallelize some of the abovementioned ciphers, where we replace the (Generalized) Feistel structures by the parallelizable GF-NLFSR structures, while keeping the internal components like S-boxes and linear diffusion to be the same. This would make encryption speed faster by up to n times. Two candidates which we find promising for parallelizing are the Camellia and SMS4 ciphers.

1.2 Related Works

In order to analyze the resistance of a block cipher against differential and linear cryptanalysis, we would like to lower bound the number of active S-boxes (S-boxes which contribute to the differential/linear probability) in any differential/linear characteristic path over a fixed number of rounds. Using such bounds, the cipher designer can choose a large enough number of rounds so that there are too many active S-boxes for differential/linear cryptanalysis to be successful.

Kanda [16] has proven that for a Feistel cipher with an SPN round function having branch number $\mathcal{B}$ (a measure of dispersion, please refer to Section 2 for the exact definition), the number of active S-boxes in any differential and linear characteristic path over every $4r$ rounds is at least $r\mathcal{B} + \lfloor \frac{r}{2} \rfloor$. Based on this lower bound, the authors of [5] designed the block cipher Camellia, which has practical provable security against differential and linear cryptanalysis.

1.3 Our Contribution

In Section 3, we provide a neat and concise proof of the result that for a $2nr$-round parallelizable n-cell GF-NLFSR structure with an SPN round function having branch number $\mathcal{B}$, the number of active S-boxes in any differential characteristic path is at least $r\mathcal{B} + \lfloor \frac{r}{2} \rfloor$. The result holds for any $n \geq 2$ in general, and we expect the result to be useful in the design and analysis of block cipher structures. For the case of a 2-cell GF-NLFSR structure, we have $r\mathcal{B} + \lfloor \frac{r}{2} \rfloor$ active S-boxes over every $4r$ rounds, which is the same as Kanda's result [16] for a conventional 2-cell Feistel structure. Motivated by this observation, we propose in Section 4 a parallelizable version of Camellia, p-Camellia, where we change the conventional Feistel structure to a 2-cell GF-NLFSR structure but keep all other components such as S-boxes and linear diffusion maps to be the same. We also ensure p-Camellia is secure against other cryptanalysis such as linear, boomerang, integral, impossible differential, higher order differential, interpolation and slide attacks. In addition, we assess the advantages of hardware implementations. For this reason we briefly introduce design strategies for hardware implementations. We then show that especially for applications with high throughput requirements, a 2-cell GF-NLFSR such as p-Camellia offers significant advantages over a conventional 2-cell Feistel structure such as Camellia. In particular, we show that an implementation of p-Camellia that processes two rounds in parallel has a maximum frequency that is nearly twice as high as it would be for Camellia while having similar area demands and significantly less power demands. We also show that for fully pipelined implementations a conventional 2-cell Feistel structure requires twice as many pipeline stages, and hence twice as many clock cycles delay, to achieve the same frequency as it is the case for a 2-cell GF-NLFSR.

In Section 6, we also apply a 4-cell GF-NLFSR structure to form a parallelizable version of SMS4 called p-SMS4. We change the generalized Feistel structure in both the main cipher and key schedule of SMS4 to a 4-cell GF-NLFSR structure but keep all other components such as S-boxes and linear diffusion maps to be the same. We first prove that p-SMS4 is secure against differential and linear cryptanalysis. In [7], Biryukov et al. showed a powerful related-key differential attack on AES-256 which can recover the secret key with complexity 2^{131} using 2^{35} related keys. We give a proof through the p-SMS4 key schedule that p-SMS4 is resistant against this attack. We also ensure p-SMS4 is secure against other block cipher cryptanalysis such as boomerang, integral, impossible differential, higher order differential, interpolation, slide and XSL attacks. A 4-cell GF-NLFSR structure offers also implementation advantages for round-based and parallelized hardware architectures. We show that a 4-cell GF-NLFSR structure, implemented in an architecture that processes four rounds in one clock cycle, has a significantly shorter critical path, and hence a higher maximum frequency, than a conventional 4-cell Feistel structure. In parallelized implementations

this advantage increases to a nearly four times higher maximum frequency while having similar area demands and significantly less power demands. In general the advantage is dependent on the number of branches, hence an n-cell GF-NLFSR has an advantage of a nearly n times higher maximum frequency.

2 Definitions and Preliminaries

In this section, we will list some definitions and summarize the results of Kanda in [16]. He has proven the upper bounds of the maximum differential and linear characteristic probabilities of Feistel ciphers with bijective SPN round function. More explicitly, the round function F-function comprises the key addition layer, the S-function and the P-function. Here we neglect the effect of the round key since by assumption, the round key, which is used within one round, consists of independent and uniformly random bits, and is bitwise XORed with data. The S-function is a non-linear transformation layer with m parallel d-bit bijective S-boxes whereas the P-function is a linear transformation layer. In particular, we have

$$S : (GF(2^d)^m \to (GF(2^d)^m \ , \ X = (x_1, \cdots, x_m) \mapsto Z = S(X) = (s_1(x_1), \cdots, s_n(x_n)),$$
$$P : (GF(2^d)^m \to (GF(2^d)^m \ , \ Z = (z_1, \cdots, z_m) \mapsto Y = P(Z) = (y_1, \cdots, y_n),$$
$$F : (GF(2^d)^m \to (GF(2^d)^m \ , \ X \mapsto Y = F(X) = P(S(X)).$$

Definition 1. *Let $x, z \in GF(2^d)$. Denote the differences and the mask values of x and z by Δx, Δz, and, Γx, Γz respectively. The differential and linear probabilities of each S-box s_i are defined as:*

$$DP^{s_i}(\Delta x \to \Delta z) = \frac{\#\{x \in GF(2^d) | s_i(x) \oplus s_i(x \oplus \Delta x) = \Delta z\}}{2^d},$$
$$LP^{s_i}(\Gamma z \to \Gamma x) = (2 \times \frac{\#\{x \in GF(2^d) | x \cdot \Gamma x = s_i(x) \cdot \Gamma z}{2^d} - 1)^2.$$

Definition 2. *The maximum differential and linear probabilities of S-boxes are defined as:*

$$p_s = \max_i \max_{\Delta x \neq 0, \Delta z} DP^{s_i}(\Delta x \to \Delta z),$$
$$q_s = \max_i \max_{\Gamma x, \Gamma z \neq 0} LP^{s_i}(\Gamma z \to \Gamma x).$$

This means that p_s, q_s are the upper bounds of the maximum differential and linear probabilities for all S-boxes.

Definition 3. *Let $X = (x_1, x_2, \cdots, x_m) \in GF(2^d)^m$. Then the Hamming weight of X is denoted by $H_w(X) = \#\{i | x_i \neq 0\}$.*

Definition 4. *[25] The branch number $\mathcal{B}$ of linear transformation θ is defined as follows:*

$$\mathcal{B} = \min_{x \neq 0}(H_w(x) + H_w(\theta(x)).$$

Consider Feistel ciphers with bijective SPN round functions as described previously. As mentioned in [16], for the differential case, $\mathcal{B}$ is taken to be the *differential* branch number, i.e. $\mathcal{B} = \min_{\Delta X \neq 0}(H_w(\Delta X) + H_w(\Delta Y))$, where ΔX is an input difference into the S-function and ΔY is an output difference of the P-function. On the other hand, for the linear case, $\mathcal{B}$ is taken to be the *linear* branch number, i.e. $\mathcal{B} = \min_{\Gamma Y \neq 0}(H_w(P^*(\Gamma Y)) + H_w(\Gamma Y))$, where ΓY is an output mask value of the P-function and P^* is a diffusion function of mask values concerning the P-function. Throughout this paper, $\mathcal{B}$ is used to denote differential or linear branch number, depending on the context.

Definition 5. *A differential active S-box is defined as an S-box given a non-zero input difference. Similarly, a linear active S-box is defined as an S-box given a non-zero output mask value.*

Theorem 1. *Let $\mathcal{D}^{(r)}$ and $\mathcal{L}^{(r)}$ be the minimum number of all differential and linear active S-boxes for a r-round Feistel cipher respectively. Then the maximum differential and linear characteristic probabilities of the r-round cipher are bounded by $p_s^{D^{(r)}}$ and $q_s^{L^{(r)}}$ respectively.*

Note that Theorem 1 applies to any block cipher in general.

Theorem 2. *[16] The minimum number of differential (and linear) active S-boxes $\mathcal{D}^{(4r)}$ for $4r$-round Feistel ciphers with SPN round function is at least $r\mathcal{B} + \lfloor \frac{r}{2} \rfloor$.*

3 Practical Security Evaluation of GF-NLFSR against Differential and Linear Cryptanalysis

GF-NLFSR was proposed by Choy et al. in [8]. It is an n-cell extension of the outer function FO of the KASUMI block cipher which is a 2-cell structure [8].

Throughout this paper, we consider GF-NLFSR block ciphers with SPN (S-P) round function, as described in Section 1.2. In this paper, we assume that both the S-function and P-function are bijective.

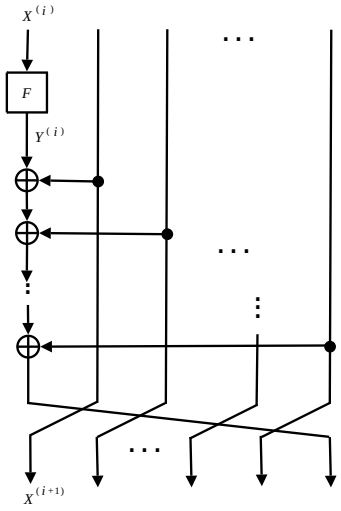

Fig. 1. i-th round of GF-NLFSR

With reference to Figure 1 , let $X^{(i)}$ and $Y^{(i)}$ be the input and output data to the i-th round function respectively. Then the GF-NLFSR block cipher can be defined as

$$X^{(i+n)} = Y^{(i)} \oplus X^{(i+1)} \oplus X^{(i+2)} \oplus \cdots \oplus X^{(i+n-1)}, \text{ for } i = 1, 2, \cdots . \tag{1}$$

3.1 Differential Cryptanalysis

We now investigate the minimum number of differental active S-boxes for GF-NLFSR block cipher. From equation (1), it can be shown almost immediately that there must be at least 2 differential active S-boxes over $(n+1)$-round of n-cell GF-NLFSR cipher.

Proposition 1 *The minimum number of differential active S-boxes for $(n+1)$-round n-cell GF-NLFSR cipher with bijective SPN round function satisfies $\mathcal{D}^{(n+1)} \geq 2$.*

Proof. Without loss of generality, we assume that the $n+1$ consecutive rounds run from the first round to the $(n+1)$-th round. Since the SPN round function is bijective, $\Delta Y^{(1)} = 0$ if and only if $\Delta X^{(1)} = 0$. From equation (1), we have

$$\Delta X^{(n+1)} = \Delta Y^{(1)} \oplus \Delta X^{(2)} \oplus \Delta X^{(3)} \oplus \cdots \oplus \Delta X^{(n)}, \tag{2}$$

from which it follows that there must exist at least two non-zero terms in equation (2) in order for equation (2) to hold. Therefore

$$\mathcal{D}^{(n+1)} = Hw(\Delta X^{(1)}) + \cdots + Hw(\Delta X^{(n+1)}) \geq 2.$$

□

Lemma 1. *Let $X = (x_1, x_2, \cdots, x_m)$ and $X' = (x'_1, x'_2, \cdots, x'_m) \in GF(2^d)^m$. Then*

$$H_w(X \oplus X') \leq H_w(X) + H_w(X').$$

Proof.

$$\begin{aligned}
&H_w(X \oplus X') \\
&= \#\{s | x_s \neq 0 \text{ and } x'_s = 0\} + \#\{t | x_t = 0 \text{ and } x'_t \neq 0\} + \#\{u | x_u \neq 0 \text{ and } x'_u \neq 0 \text{ and } x_u \neq x'_u\} \\
&\leq H_w(X) + \#\{t | x_t = 0 \text{ and } x'_t \neq 0\} \\
&\leq H_w(X) + H_w(X')
\end{aligned}$$

□

Lemma 2 is a straightforward generalization of Lemma 1.

Lemma 2. *Let X_1, X_2, $\cdots$, $X_k \in GF(2^d)^m$. Then*

$$H_w(X_1 \oplus X_2 \oplus \cdots \oplus X_k) \leq H_w(X_1) + H_w(X_2) + \cdots + H_w(X_k).$$

As stated in Theorem 1, to investigate the upper bound of the maximum differential characteristic probability of the GF-NLFSR cipher, we need to find a lower bound for $\mathcal{D}^{(r)}$, the number of differential active S-boxes for r consecutive rounds of the cipher. Then the differential characteristic probability of the r-round GF-NLFSR cipher is at most $p_s^{\mathcal{D}^{(r)}}$.

Lemma 3. *For n-cell GF-NLFSR cipher, the minimum number of differential active S-boxes in any $2n$ consecutive rounds satisfies $\mathcal{D}^{(2n)} \geq \mathcal{B}$.*

Proof. Without loss of generality, we assume that the $2n$ consecutive rounds run from the first round to the $2n$-th round. For $j = 1, \cdots, n$, note that at least one of $\Delta X^{(j)} \neq 0$. Let i be the smallest integer such that $\Delta X^{(i)} \neq 0$, where $1 \leq i \leq n$. Then

$$\begin{aligned}
\mathcal{D}^{(2n)} &= H_w(\Delta X^{(1)}) + H_w(\Delta X^{(2)}) + \cdots + H_w(\Delta X^{(2n)}) \\
&\geq H_w(\Delta X^{(i)}) + H_w(\Delta X^{(i+1)}) \cdots + H_w(\Delta X^{(i+n)}) \\
&\geq H_w(\Delta X^{(i)}) + H_w(\Delta X^{(i+1)} \oplus \cdots \oplus \Delta X^{(i+n)}), \text{ by Lemma 2,} \\
&= H_w(\Delta X^{(i)}) + H_w(\Delta Y^{(i)}) \\
&\geq \mathcal{B}.
\end{aligned}$$

□

Remark 1. From the above proof, we see that with probability $1 - \frac{1}{M}$, where M is the size of each cell, i.e. most of the time, we have $\Delta X^{(1)} \neq 0$. In that case, we are able to achieve at least $\mathcal{B}$ number of differential active S-boxes over $(n+1)$-round of n-cell GF-NLFSR cipher.

As a consequence of Lemma 3 and using a similar approach as [16], we have the following result.

Theorem 3. *The minimum number of differential active S-boxes for $2nr$-round n-cell GF-NLFSR cipher with bijective SPN round function satisfies*

$$\mathcal{D}^{(2nr)} \geq r\mathcal{B} + \lfloor \frac{r}{2} \rfloor.$$

In particular, when $n = 2$, the minimum number of differential active S-boxes for $4r$-round 2-cell GF-NLFSR cipher with bijective SPN round function is at least $r\mathcal{B} + \lfloor \frac{r}{2} \rfloor$. Hence we see that 2-cell GF-NLFSR cipher with bijective SPN round function has similar practical security against differential cryptanalysis as Feistel ciphers with bijective SPN round functions. Moreover, 2-cell GF-NLFSR has an added advantage that it realizes parallel computation of round functions, thus providing strong motivation for parallelizing ciphers with SPN round functions, as described in Section 4.

3.2 Linear Cryptanalysis

For the purpose of parallelizing Camellia and SMS4, we shall investigate the practical security of 2-cell and 4-cell GF-NLFSR cipher against linear cryptanalysis. Again from Theorem 1, to investigate the upper bound of the maximum linear characteristic probability of the GF-NLFSR cipher, we need to find a lower bound for $\mathcal{L}^{(r)}$, the number of linear active S-boxes for r consecutive rounds of the cipher. Then the linear characteristic probability of the r-round cipher is at most $q_s^{\mathcal{L}^{(r)}}$. We first consider the 2-cell GF-NLFSR cipher, followed by the 4-cell GF-NLFSR cipher.

Duality between Differential Characteristic and Linear Approximation As discussed in Section 3 of [18], when analyzing mask values in linear cryptanalysis, we need to consider the duality between differential characteristic and linear approximation, where each XOR is replaced by a joint and each joint is replaced by an XOR. Hence, in the case of 2-cell GF-NLFSR cipher, with reference to Figure 2, we have

$$\Gamma X^{(i+2)} = \Gamma Y^{(i)} \oplus \Gamma Y^{(i+1)}, \text{ for } i \geq 1, \tag{3}$$

where the input and output mask values to the i-th round F function are denoted by $\Gamma X^{(i)}$ and $\Gamma Y^{(i)}$ respectively.

Similarly, for 4-cell GF-NLFSR cipher, with reference to Figure 3, we have

$$\Gamma X^{(i+4)} = \Gamma Y^{(i)} \oplus \Gamma Y^{(i+1)} \oplus \Gamma Y^{(i+2)} \oplus \Gamma Y^{(i+3)}. \tag{4}$$

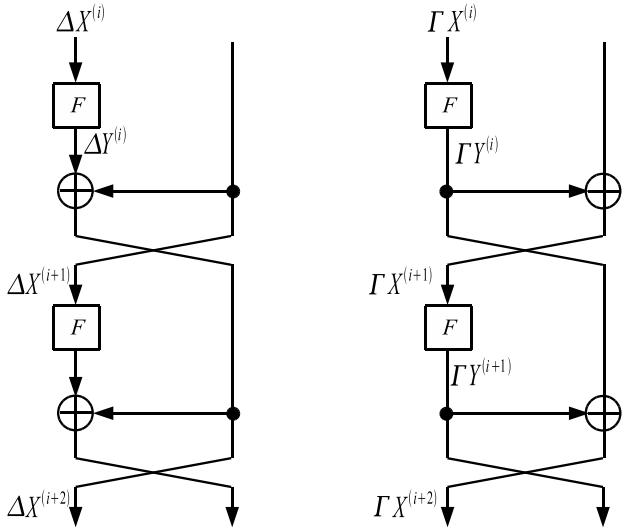

Fig. 2. (Left)2-cell GF-NLFSR cipher; (Right) Dual of 2-cell GF-NLFSR cipher

Lemma 4. *For 2-cell GF-NLFSR cipher with bijective SPN round function and linear branch number* $\mathcal{B} = 5$*, the minimum number of linear active S-boxes in any* 4 *consecutive rounds satisfies* $\mathcal{L}^{(4)} \geq 3$.

Proof. Let the input and output mask values to the i-th round F function be $\Gamma X^{(i)}$ and $\Gamma Y^{(i)}$ respectively. Note that since the F function is bijective, $\Gamma X^{(i)} = 0$ if and only if $\Gamma Y^{(i)} = 0$. Without loss of generality, we assume that the 4 consecutive rounds run from the first round to the fourth round. Thus the minimum number of linear active S-boxes over 4 consecutive rounds is given by

$$\mathcal{L}^{(4)} = H_w(\Gamma Y^{(1)}) + H_w(\Gamma Y^{(2)}) + H_w(\Gamma Y^{(3)}) + H_w(\Gamma Y^{(4)}).$$

As discussed in the previous section, we have, from Equation (3),

$$\Gamma X^{(i+1)} = \Gamma Y^{(i-1)} \oplus \Gamma Y^{(i)},$$

for $i = 2$ and 3. We consider all cases as follows, where $\mathcal{L}_i^{(r)}$ denotes the number of linear active S-boxes over r rounds for case i:

Case 1: $\Gamma X^{(1)} = 0$
This implies that $\Gamma X^{(2)} \neq 0$ and $\Gamma X^{(3)} = \Gamma Y^{(2)}$. Hence $\mathcal{L}_1^{(3)} \geq H_w(\Gamma X^{(2)}) + H_w(\Gamma X^{(3)}) = H_w(\Gamma X^{(2)}) + H_w(\Gamma Y^{(2)}) \geq \mathcal{B} = 5 \geq 3$. Thus $\mathcal{L}_1^{(4)} \geq \mathcal{L}_1^{(3)} \geq 3$.

Case 2: $\Gamma X^{(1)} \neq 0$ and $\Gamma X^{(2)} = 0$
This implies that $\Gamma X^{(3)} = \Gamma Y^{(1)}$. Hence $\mathcal{L}_2^{(3)} \geq H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(3)}) = H_w(\Gamma X^{(1)}) + H_w(\Gamma Y^{(1)}) \geq \mathcal{B} = 5 \geq 3$. Thus $\mathcal{L}_2^{(4)} \geq \mathcal{L}_2^{(3)} \geq 3$.

Case 3: $\Gamma X^{(1)} \neq 0$, $\Gamma X^{(2)} \neq 0$ and $\Gamma X^{(3)} = 0$
This implies that $\Gamma X^{(4)} = \Gamma Y^{(2)}$. Hence $\mathcal{L}_3^{(4)} \geq H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(2)}) + H_w(\Gamma X^{(4)}) = H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(2)}) + H_w(\Gamma Y^{(2)}) \geq 1 + \mathcal{B} = 6 \geq 3$.

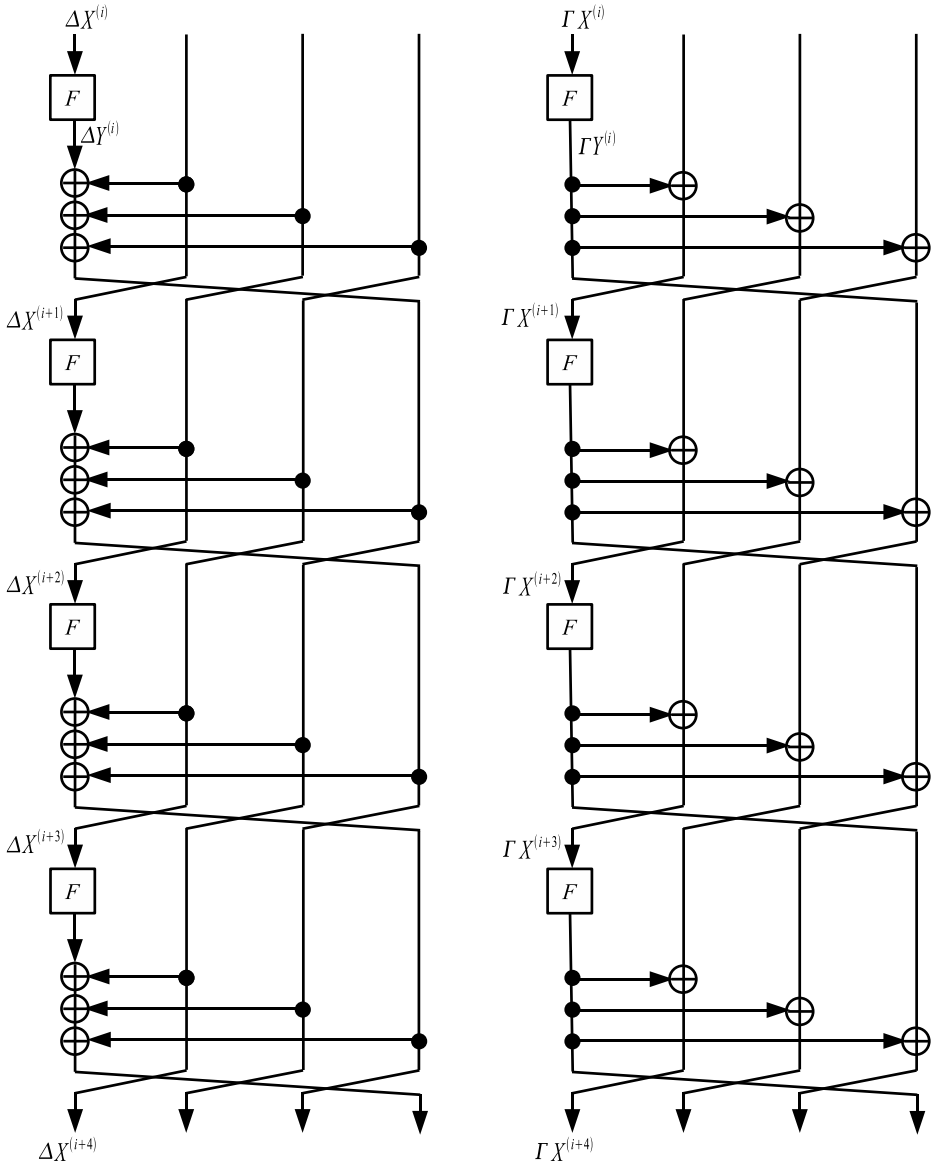

Fig. 3. (Left)4-cell GF-NLFSR cipher; (Right) Dual of 4-cell GF-NLFSR cipher

Case 4: $\Gamma X^{(1)} \neq 0$, $\Gamma X^{(2)} \neq 0$, $\Gamma X^{(3)} \neq 0$ and $\Gamma X^{(4)} = 0$
Then we obtain $\mathcal{L}_4^{(4)} \geq H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(2)}) + H_w(\Gamma X^{(3)}) \geq 1 + 1 + 1 = 3$.

Case 5: $\Gamma X^{(1)} \neq 0$, $\Gamma X^{(2)} \neq 0$, $\Gamma X^{(2)} \neq 0$ and $\Gamma X^{(4)} \neq 0$
Then we obtain $\mathcal{L}_5^{(4)} = H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(2)}) + H_w(\Gamma X^{(3)}) + H_w(\Gamma X^{(3)}) \geq 1 + 1 + 1 + 1 = 4 \geq 3$.

Therefore $\mathcal{L}^{(4)} \geq 3$.

□

Theorem 4. *For 2-cell GF-NLFSR cipher with bijective SPN round function and linear branch number $\mathcal{B} = 5$, we have*

(1) $\mathcal{L}^{(8)} \geq 7$,
(2) $\mathcal{L}^{(12)} \geq 11$,
(3) $\mathcal{L}^{(16)} \geq 15$.

Proof. Without loss of generality, we begin from the first round.

(1) From the proof of Lemma 4, over 8 rounds, we only need to check the case for $\Gamma X^{(1)} \neq 0$, $\Gamma X^{(2)} \neq 0$, $\Gamma X^{(3)} \neq 0$ and $\Gamma X^{(4)} = 0$. (In all remaining cases, there will be at least 7 linear active S-boxes over 8 rounds.) However $\Gamma X^{(3)} \neq 0$ and $\Gamma X^{(4)} = 0$ correspond to Case 1 of Lemma 4 for the four consecutive rounds that begin from the 4th round and end after the 7th round. Hence there will be at least $3+5 = 8$ linear active S-boxes. Therefore $\mathcal{L}^{(8)} \geq 7$.

(2) From (i), over 12 rounds, we only need to consider the case for $\Gamma X^{(i)} \neq 0$ for $i = 1, \cdots, 7$ and $\Gamma X^{(8)} = 0$. Following a similar argument to (i), we are definitely ensured of at least $7+5 = 12$ linear active S-boxes. Hence $\mathcal{L}^{(12)} \geq 11$.

(3) The proof is similar to that of (i) and (ii).

□

We conclude this section with the study of minimum number of active S-boxes for 4-cell GF-NLFSR.

Proposition 1. *Assume that the linear branch number $\mathcal{B} = 5$. Then the minimum number of linear active S-boxes for 5-round 4-cell GF-NLFSR cipher with bijective SPN round function satisfies $\mathcal{L}^{(5)} \geq 2$.*

Proof. Let $\Gamma X^{(i)}$ and $\Gamma Y^{(i)}$ be the input and output mask to the ith round function respectively. Since the round function is bijective, $\Gamma X^{(i)} = 0$ if and only if $\Gamma Y^{(i)} = 0$. It is evident from equation (4) that there cannot exist exactly one non-zero input mask for five consecutive rounds. The result now follows easily. □

Theorem 5. *Assume that the linear branch number $\mathcal{B} = 5$. Then the minimum number of linear active S-boxes for 10-round 4-cell GF-NLFSR cipher with bijective SPN round function satisfies $\mathcal{L}^{(10)} \geq \mathcal{B} + 1$.*

Proof. With no loss of generality, assume that the 10 rounds run consecutively from the first round to the tenth round. Let $\Gamma X^{(i)}$ and $\Gamma Y^{(i)}$ be the input and output mask to the ith round function respectively. Recall that due to the duality between differential characteristic and linear approximation, equation (4) holds. Let $\mathcal{M} = \{\Gamma X^{(1)}, \Gamma X^{(2)}, \Gamma X^{(3)}, \Gamma X^{(4)}\}$. We consider all the following cases, where $\mathcal{L}_j^{(r)}$ denotes the number of linear active S-boxes for r rounds for case j.

Case 1 : There is exactly one non-zero input mask in set $\mathcal{M}$, i.e. $\Gamma X^{(i)} \neq 0$ for some $i = 1, 2, 3$ or 4.

Then $\Gamma X^{(5)} = \Gamma Y^{(i)} \neq 0$. Since for four consecutive rounds, the input masks cannot be zero at the same time, we obtain

$$\begin{aligned}\mathcal{L}_1^{(9)} &= H_w(\Gamma X^{(i)}) + H_w(\Gamma X^{(5)}) + H_w(\Gamma X^{(6)}) + \cdots + H_w(\Gamma X^{(9)}) \\ &\geq H_w(\Gamma X^{(i)}) + H_w(\Gamma Y^{(i)}) + 1 \\ &\geq \mathcal{B} + 1.\end{aligned}$$

Case 2 : All input masks in $\mathcal{M}$ are non-zero.
By Proposition 1, we obtain

$$\begin{aligned}\mathcal{L}_2^{(9)} &= H_w(\Gamma X^{(1)}) + \cdots + H_w(\Gamma X^{(4)}) + H_w(\Gamma X^{(5)}) + \cdots + H_w(\Gamma X^{(9)}) \\ &\geq 4 + 2 \\ &= 6 \\ &\geq \mathcal{B} + 1.\end{aligned}$$

Case 3 : There are exactly three non-zero input masks in $\mathcal{M}$.
Let $\mathcal{S} = \{\Gamma X^{(5)}, \Gamma X^{(6)}, \Gamma X^{(7)}, \Gamma X^{(8)}\}$. If there are at least three non-zero input masks in $\mathcal{S}$, then we are done. Also, since the input masks for four consecutive rounds cannot be zero at the same time, at least one input mask in $\mathcal{S}$ is non-zero. This implies that we only need to check the following:
(i) There is exactly one non-zero input difference in $\mathcal{S}$.
Then $\Gamma X^{(9)} = \Gamma Y^{(j)}$ for $j = 5$, 6, 7 or 8. Hence

$$\mathcal{L}_3^{(9)} \geq 3 + H_w(\Gamma X^{(j)}) + H_w(\Gamma X^{(9)}) \geq \mathcal{B} + 3.$$

(ii) There are exactly two non-zero input masks in $\mathcal{S}$.

- Suppose $\Gamma X^{(5)} = 0$ and $\Gamma X^{(6)} \neq 0$. Then $\Gamma Y^{(1)} \oplus \Gamma Y^{(2)} \oplus \Gamma Y^{(3)} \oplus \Gamma Y^{(4)} = 0$ and it follows that $\Gamma X^{(6)} = \Gamma Y^{(1)} \neq 0$. Hence we are ensured of at least $\mathcal{B} + 2$ active S-boxes.
- Suppose $\Gamma X^{(6)} = 0$ and $\Gamma X^{(7)} \neq 0$. Then $\Gamma Y^{(2)} \oplus \Gamma Y^{(3)} \oplus \Gamma Y^{(4)} \oplus \Gamma Y^{(5)} = 0$ and it follows that $\Gamma X^{(7)} = \Gamma Y^{(2)} \neq 0$. Hence we are ensured of at least $\mathcal{B} + 2$ active S-boxes.
- Suppose $\Gamma X^{(6)} = \Gamma X^{(7)} = 0$, $\Gamma X^{(5)} \neq 0$ and $\Gamma X^{(8)} \neq 0$. Then it can be deduced easily that $\Gamma X^{(8)} = \Gamma Y^{(3)} \neq 0$, and there must be at least $\mathcal{B} + 2$ active S-boxes.
- Suppose $\Gamma X^{(7)} = \Gamma X^{(8)} = 0$, $\Gamma X^{(5)} \neq 0$ and $\Gamma X^{(6)} \neq 0$. It follows directly that $\Gamma X^{(9)} = \Gamma Y^{(4)}$. If $\Gamma X^{(9)} =\neq 0$, then we are done. Otherwise $\Gamma X^{(4)} = 0$ which implies that $\Gamma X^{(3)} \neq 0$. However, $0 = \Gamma X^{(8)} = \Gamma Y^{(3)} \neq 0$, which is a contradiction.

Case 4 : There are exactly two non-zero input masks in $\mathcal{M}$.
(i) Suppose $\Gamma X^{(5)} = 0$. Then $\Gamma X^{(6)} = \Gamma Y^{(1)}$.

- If $\Gamma X^{(1)} \neq 0$, then $\mathcal{L}_4^{(6)} \geq H_w(\Gamma X^{(1)}) + H_w(\Gamma Y^{(1)}) + 1 \geq \mathcal{B} + 1$.
- If $\Gamma X^{(1)} = 0$ and $\Gamma X^{(2)} \neq 0$, then $\Gamma X^{(7)} = \Gamma Y^{(2)}$ and so we obtain,

$$\mathcal{L}_4^{(7)} \geq H_w(\Gamma X^{(2)}) + H_w(\Gamma Y^{(2)}) + 1 \geq \mathcal{B} + 1.$$

- If $\Gamma X^{(1)} = 0$ and $\Gamma X^{(2)} = 0$, then $\Gamma X^{(3)} \neq 0$ and $\Gamma X^{(4)} \neq 0$. Hence $\Gamma X^{(8)} = \Gamma Y^{(3)}$, from which

$$\mathcal{L}_4^{(8)} \geq H_w(\Gamma X^{(3)}) + H_w(\Gamma X^{(4)}) + H_w(\Gamma X^{(8)}) \geq \mathcal{B} + 1,$$

  follows.

(ii) Suppose $\Gamma X^{(5)} \neq 0$ and $\Gamma X^{(6)} = 0$. It follows that $\Gamma X^{(7)} = \Gamma Y^{(2)}$.

- If $\Gamma X^{(2)} \neq 0$, then $\mathcal{L}_4^{(7)} \geq H_w(\Gamma X^{(2)}) + H_w(\Gamma Y^{(2)}) + 1 \geq \mathcal{B} + 1$.

– If $\Gamma X^{(2)} = 0$ and $\Gamma X^{(3)} \neq 0$, then $\Gamma X^{(8)} = \Gamma Y^{(3)}$. This implies that

$$\mathcal{L}_4^{(8)} \geq H_w(\Gamma X^{(3)}) + H_w(\Gamma Y^{(3)}) + 1 \geq \mathcal{B} + 1.$$

– If $\Gamma X^{(2)} = 0$ and $\Gamma X^{(3)} = 0$, then $\Gamma X^{(1)} \neq 0$ and $\Gamma X^{(4)} \neq 0$. This implies that $\Gamma X^{(9)} = \Gamma Y^{(4)}$, and so

$$\mathcal{L}_4^{(9)} \geq H_w(\Gamma X^{(4)}) + H_w(\Gamma X^{(9)}) + H_w(\Gamma X^{(1)}) \geq \mathcal{B} + 1.$$

(iii) Suppose $\Gamma X^{(5)} \neq 0$, $\Gamma X^{(6)} \neq 0$ and $\Gamma X^{(7)} = 0$. Then $\Gamma X^{(8)} = \Gamma Y^{(3)}$.

– If $\Gamma X^{(3)} \neq 0$, then $\mathcal{L}_4^{(8)} \geq H_w(\Gamma X^{(3)}) + H_w(\Gamma Y^{(3)}) + 1 \geq \mathcal{B} + 1$.
– If $\Gamma X^{(3)} = 0$ and $\Gamma X^{(4)} \neq 0$, then $\Gamma X^{(9)} = \Gamma Y^{(4)}$. This implies that

$$\mathcal{L}_4^{(9)} \geq H_w(\Gamma X^{(4)}) + H_w(\Gamma Y^{(4)}) + 1 \geq \mathcal{B} + 1.$$

– If $\Gamma X^{(3)} = 0$ and $\Gamma X^{(4)} = 0$, then $\Gamma X^{(1)} \neq 0$ and $\Gamma X^{(2)} \neq 0$. This implies that $\Gamma X^{(10)} = \Gamma Y^{(5)} \neq 0$. So

$$\mathcal{L}_4^{(10)} \geq H_w(\Gamma X^{(5)}) + H_w(\Gamma X^{(10)}) + H_w(\Gamma X^{(1)}) + H_w(\Gamma X^{(2)}) + H_w(\Gamma X^{(6)}) \geq \mathcal{B} + 3.$$

(iv) Suppose $\Gamma X^{(5)} \neq 0$, $\Gamma X^{(6)} \neq 0$ and $\Gamma X^{(7)} \neq 0$. If $\Gamma X^{(8)} \neq 0$ or $\Gamma X^{(9)} \neq 0$, then there will be at least 6 linear active S-boxes and we are done. Otherwise $\Gamma X^{(8)} = \Gamma X^{(9)} = 0$ and $\Gamma X^{(10)} = \Gamma Y^{(5)} \neq 0$ and we obtain

$$\mathcal{L}_4^{(10)} \geq H_w(\Gamma X^{(5)}) + H_w(\Gamma X^{(10)}) + H_w(\Gamma X^{(6)}) + H_w(\Gamma X^{(7)}) \geq \mathcal{B} + 2.$$

Hence considering all cases, we conclude that $\mathcal{L}^{(10)} \geq \mathcal{B} + 1$. □

Corollary 1. *The minimum number of linear active S-boxes for* 9*-round* 4*-cell GF-NLFSR cipher with bijective SPN round function satisfies* $\mathcal{L}^{(9)} \geq 4$.

Proof. The result follows easily from the proof of Theorem 5. □

4 Application 1: Parallelizing Camellia

4.1 Brief Description of Camellia

Camellia was jointly developed by NTT and Mitsubishi Electric Corporation. According to [5], Camellia uses an 18-round Feistel structure for 128-bit key, and a 24-round Feistel structure for 192-bit and 256-bit keys, with additional input/output whitenings and logical functions called the FL-function and FL^{-1}-function inserted every 6 rounds. Its F-function uses the SPN (Substitution-Permutation Network) structure, whereby the non-linear layer comprises eight S-boxes in parallel while the linear layer can be represented using only bytewise exclusive-ORs. Note that the F-function is bijective.

For security against differential and linear cryptanalysis, the branch number of the linear layer should be optimal, i.e. branch number = 5. In addition, the S-boxes adopt functions affine equivalent to the inversion function in $GF(2^8)$ which achieves the best known of the maximum differential and linear probabilities 2^{-6} [5].

The key schedule of Camellia is slightly different for the 128-bit key version and the 192-bit/256-bit key version. Despite the slight differences, the key schedule is relatively simple and consists of two main steps. One (or two) 128-bit subkey materials are first derived from the secret key via some Feistel network. The round keys are then generated by rotating the secret key itself and the derived subkeys by various amounts.

For more details of the structure of Camellia, readers are referred to [4].

4.2 Parallelizing Camellia : p-Camellia

In this section, we propose another version of the existing Camellia block cipher, which we call p-Camellia (*"parallelizable" Camellia*). As described previously, Camellia uses a Feistel network structure. For the encryption procedure of p-Camellia, we shall replace the Feistel network with the 2-cell GF-NLFSR block cipher structure instead, as depicted in Figure 4 of Appendix. Other components such as number of rounds, S-function, P-function and the key schedule for the different key versions etc remain unchanged. In addition, similar to Camellia, there are input/output whitenings which are represented by the XOR symbols at the beginning/end of p-Camellia cipher in Figure 4.

5 Figure of p-Camellia

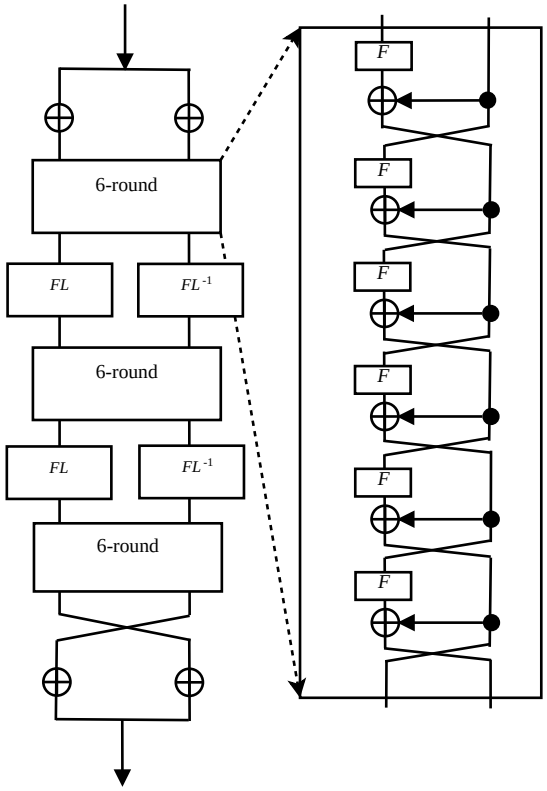

Fig. 4. p-Camellia block cipher structure

5.1 Differential and Linear Cryptanalysis of p-Camellia

Following the same approach in [4], denote the maximum differential and linear characteristic probabilities of p-Camellia reduced to 16-round by p and q respectively. Recall that since both p-Camellia and Camellia use the same F-function, in the case of p-Camellia, the maximum differential and linear probability of the S-boxes are 2^{-6}. From [4], the differential branch numbers is equal to 5. By considering the P^*-function of Camellia as in [16], the linear branch number is verified to be 5.

Over 16 rounds, there are four 4-round blocks. By virtue of Theorem 3, where $n = 2$ and $r = 4$, we have

$$p \leq (2^{-6})^{4\times 5+2} = 2^{-132} < 2^{-128}.$$

By Theorem 4, we obtain $q \leq (2^{-6})^{15} = 2^{-90}$. This implies that an attacker needs to collect at least 2^{90} chosen/known plaintexts to mount an attack, which is not feasible in practice.

This implies that there is no effective differential or linear characteristic for p-Camellia reduced to more than 15 rounds. In other words, p-Camellia offers sufficient security against differential and linear attack.

5.2 Other Attacks on p-Camellia

In this section, we briefly examine the protection of p-Camellia against various known attacks. Since p-Camellia uses the same components as Camellia, we expect that p-Camellia offers similar level of protection against most of the attacks, as compared to Camellia.

Boomerang Attack To perform boomerang attack, the cipher is split into two shorter ciphers E_0 and E_1 such that the differential probability of each part is known to be large. Suppose an adversary split 16 rounds into E_0 and E_1 with r and $16 - r$ rounds respectively. By Theorem 3, the characteristic differential probability of each sub-ciphers would be bounded by $p_0 \leq (2^{-30})^{\lfloor r/4 \rfloor}$ and $p_1 \leq (2^{-30})^{\lfloor (16-r)/4 \rfloor}$. (Note that we ignore the last term in the upper bound of Theorem 3 for ease of calculation.) It can be easily verified that $\lfloor r/4 \rfloor + \lfloor (16 - r)/4 \rfloor \geq 3$ for $r = 1, \ldots, 15$. Consequently,

$$p_0^2 \times p_1^2 \leq 2^{-60\times 3} = 2^{-180} < 2^{-128},$$

and thus p-Camellia is secure against boomerang attack.

Impossible Differential Attack Impossible differential attack is a chosen plaintext attack and is an extension of differential cryptanalysis. The main idea of this attack is to construct an impossible differential characteristic which is then used to filter wrong key guesses. Employing similar techniques as [23], we can prove the following result.

Proposition 2. *Let e_1 denote a subblock which is non-zero in the first byte position and zero in the remaining byte positions. For 2-cell GF-NLFSR cipher with bijective SPN round function and differential branch number $\mathcal{B} \geq 3$, there is at least one 5-round impossible differential, namely of the form $(e_1, 0) \nrightarrow_5 (\beta, \beta)$, where β is a non-zero fixed difference.*
(Note that here we only consider $\mathcal{B} \geq 3$ since linear transformation layers with $\mathcal{B} = 2$ are unlikely to be used as they do not aid in the protection of the cipher against differential attack.)

Proof. Suppose for a contradiction that $(e_1, 0) \rightarrow_5 (\beta, \beta)$ is possible. In the direction of encryption, after 3 rounds, we have $(e_1, 0) \rightarrow (PS(e_1),\ PS(e_1) \oplus PSPS(e_1))$. On the other hand, decrypting two rounds, we obtain $(S^{-1}P^{-1}(\beta), 0) \leftarrow (\beta, \beta)$. Hence

$$\begin{aligned} PS(e_1) \oplus PSPS(e_1) &= 0, \\ P(S(e_1) \oplus SPS(e_1)) &= 0, \\ S(e_1) \oplus SPS(e_1) &= 0. \end{aligned} \tag{5}$$

However,

$$H_w(SPS(e_1) \oplus S(e_1)) \geq (\mathcal{B} - 1) - 1 = \mathcal{B} - 2 \geq 3 - 2 = 1,$$

which is a contradiction with equation 5. □

Since for p-Camellia, $\mathcal{B} = 5$, by Proposition 2, there is at least a 5-round impossible differential in p-Camellia. We have not found impossible differentials with more than 5 rounds. As explained in [5], we expect that the presence of the FL- and FL^{-1} functions will greatly increase the difficulty of performing impossible differential attack on p-Camellia since the functions change the differential paths depending on key values.

Integral Attack In an integral attack, the attacker studies the propagation of multisets of chosen plaintexts of which part is held constant, and another part varies through all possibilities (also said to be *active*) through the cipher. There is a 4-round integral distinguisher of 2-cell GF-NLFSR [8], namely $(A, C) \rightarrow (S_0, S_1)$, where C is constant, A is active and $S_0 \oplus S_1$ is active. We have not found integral distinguishers with more than 4 rounds. An adversary can extend an integral attack distingusher by at most three rounds. That means he would need to extend the integral attack distinguisher from 4 to $18 - 3 = 15$ rounds which seems unlikely.

Slide Attack The slide attack works on ciphers with cyclical structures over a few rounds. According to [5], the FL- and FL^{-1}- functions are inserted between every 6 rounds to provide non-regularity across rounds. In addition, different subkeys are used for every round, making slide attack unlikely.

We now proceed to examine the protection of p-Camellia against higher order differential attack and interpolation attack. We will adopt a similar approach as [5], which is somewhat heuristic but adequate for us to have a comprehensive and insightful discussion.

Higher Order Differential Attack Higher order differential attack was introduced by Knudsen in [17]. This attack works especially well on block ciphers with components of low algebraic degree such as the KN-Cipher [13], whereby the ciphers can be represented as Boolean polynomials of low degree in terms of the plaintext. The attack requires $O(2^{t+1})$ chosen plaintext when the cipher has degree t.

p-Camellia uses exactly the same S-boxes as Camellia and it was confirmed in [5] that the degree of the Boolean polynomial of every output bit of the S-boxes is 7 by finding Boolean polynomial for every outpit bit of the S-boxes. Hence, similar to Camellia, the degree of an intermediate bit in the encryption process should increase as the data passes through many S-boxes. Indeed, let (α_i, β_i) be the input to the $(i+1)$-th round of p-Camellia. Suppose $\deg(\alpha_0) = \deg(\beta_0) = 1$. After the first round, $\deg(\alpha_1) = \deg(\beta_0) = 1$ while $\deg(\beta_1) = \deg(F(\alpha_0) \oplus \beta_0) = 7$. Continuing this process, we see that the degrees of α_i and β_i for $i = 0, 1, 2, \cdots$, increases as follows: $(1, 1), (1, 7), (7, 7), (7, 49), (49, 49), (49, 127), (127, 127), \cdots$

That is, the degrees increase exponentially as the number of rounds increase and reach the maximum degree of 127 after the 6th round, implying that it is highly unlikely that higher order differential attack will work.

Interpolation Attack The interpolation attack [14] works on block ciphers that can be expressed as an equation in $GF(2^d)$ with few monomials. p-Camellia uses the same components as Camellia and it was shown in [5] that as the data passes through many S-boxes and the P-function, the cipher became a complex function which is a sum of many multi-variate monomials over $GF(2^8)$. Hence we also expect p-Camellia to be secure against interpolation attack.

5.3 Implementation Advantages

Before we discuss the implementation advantages of p-Camellia we briefly introduce hardware implementation strategies for block ciphers that consist of a round-function that is iterated several times. While software implementations have to process single operations in a serial manner, hardware implementations offer more flexibility for parallelization. Generally speaking there exist three major architecture strategies for the implementation of block ciphers: *serialized*, *round-based*, and *parallelized*. In a *serialized* architecture only a fraction of a single round is processed in one clock cycle. These lightweight implementations allow reduction in area and power consumption at the cost of a rather long processing time. If a complete round is performed in one clock cycle, we have a *round-based* architecture. This implementation strategy usually offers the best time-area product and throughput per area ratio. A *parallelized* architecture processes more than one round per clock cycle, leading to a rather long critical path. A longer critical path leads to a lower

maximum frequency but also requires the gates to drive a higher load (fanout), which results in larger gates with a higher power consumption. By inserting intermediate registers (a technique called *pipelining*), it is possible to split the critical path into fractions, thus increasing the maximum frequency. Once the pipeline is filled, a complete encryption can be performed in one clock cycle with such an architecture. Consequently, this implementation strategy yields the highest throughput at the cost of high area demands. Furthermore, since the pipeline has to be filled, each pipelining stage introduces a delay of one clock cycle.

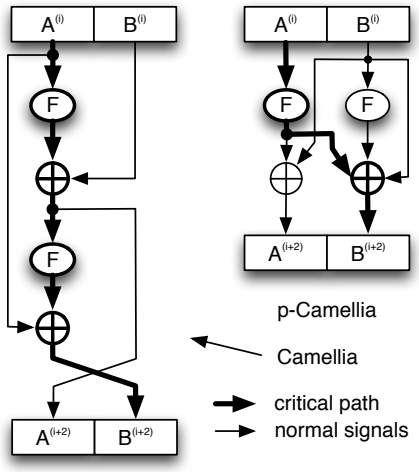

Fig. 5. Possible hardware architecture of two rounds of Camellia (left) and p-Camellia (right).

From a lightweight perspective, *i.e.* if we consider serialized architectures, it is no wonder that area, power and timing demands stay the same for Camellia and p-Camellia, since no operation was introduced or removed. Also a round-based p-Camellia implementation is as efficient as a round-based Camellia implementation. However, if we consider applications that require high throughput, p-Camellia has significant advantages. If we consider an architecture that implements two rounds in one clock cycle (see Figure 5), Camellia's critical path involves two F-functions and two 2-input XOR gates, compared to only one F-function and one 3-input XOR gate for p-Camellia. Since Camellia inserts every six rounds the FL and FL^{-1} functions, it is advantageous to parallelize this fraction of Camellia/p-Camellia. In this case the critical path of Camellia consists of six F-functions, six 2-input XOR gates and the delay of FL/FL^{-1} while p-Camellia's critical path only consists of three F-functions, three 3-input XOR gates, and the delay of FL/FL^{-1}. Given the fact that the F-function consists of a 2-input XOR gate (key addition), several combinatorial gates (S-box) and an extensive XOR network (P-function), the delay difference between a 2-input and a 3-input XOR gate is negligible. Hence p-Camellia can achieve a maximum frequency that is nearly twice as high as it would be for Camellia while having similar or lower area and power demands. In case pipelining is applied, Camellia requires twice as much pipelining stages as p-Camellia to achieve the same maximum frequency, resulting in a delay that is twice as high.

To substantiate our claims we have implemented the round function of Camellia and p-Camellia each with a 128-bit key in VHDL. We obtained area, timing and power figures figures for a 180 nm ASIC technology from UMC using Synopsys Design Vision for synthesis. Table 1 depicts a comparison of the hardware implementation results of the round function of Camellia and p-Camellia. This is a typical setup in a co-processor or instruction set extension scenario. As expected, the area requirements of 4877 GE for one instance of the round function are the same for Camellia and p-Camellia and double to 9754 GE for two instances. Also the maximum frequency of 229.4 MHz is the same for Camellia and p-Camellia in the one

Table 1. Comparison of the implementation results of the round function of Camellia and p-Camellia on UMC *180 nm* ASIC technology.

	Camellia				p-Camellia			
	1 round		2 rounds		1 round		2 rounds	
	abs.	%	abs.	%	abs.	%	abs.	%
Area (GE)	4877	100	9754	200	4877	100	9754	200
power* (mW)	2.65	100	8.38	316.5	2.65	100	5.2	196.2
max Freq. (MHz)	229.4	100	117.8	51.4	229.4	100	221.2	96.5
max T'put (Gbps)	29.4	100	30.2	103	29.4	100	56.6	192.9

*at a frequency of 100 MHz and a supply voltage of 1.8V.

round implementation. However, as depicted in Figur 5 the critical path for two consecutive instances of the round function of Camellia is nearly twice as long as for p-Camellia. Consequently, the maximum frequency achievable for Camellia drops to 51.4% while it only slightly decreases to 96.5% for p-Camellia. p-Camellia cannot achieve exactly twice the maximum frequency, because it XORs three summands, while Camellia only XORs two summands. The maximum throughput of a 1 round implementation is the same for Camellia and p-Camellia and achieves 29.4 Gbps (Giga bits per second). A two round Camellia implementation slightly increases the maximum throughput by a mere 2.7% to 30.2 Gbps, while p-Camellia boosts the maximum throughput to 56.6 Gbps - an increment of 92.9% compared to the 1 round Camellia implementation and still 87.8% higher than the 2 round Camellia implementation.

For all architectures we simulated the power consumption at a frequency of 100 MHz and a supply voltage of 1.8 Volt. 1 round of Camellia and p-Camellia require both 2.65 mW. While the power consumption for the 2 rounds implementation of Camellia increases more than 3 times (+216%) to 8.38 mW, it less than doubles for the 2 rounds implementation of p-Camellia (+96%) to 5.2 mW compared to the 1 round implementations. These figures highlight the advantages of p-Camellia over Camellia from a power perspective.

6 Application 2: Parallelizing SMS4

6.1 Brief Description of SMS4

According to [11], SMS4 takes in a 128-bit key and uses a 32-round generalized Feistel structure to transform the plaintext to the ciphertext. Each round of the generalized Feistel transformation transforms four 32-bit words X_i, $i = 0, 1, 2, 3$, as follows:

$$(X_0, X_1, X_2, X_3, rk) \mapsto (X_1, X_2, X_3, X_0 \oplus T(X_1 \oplus X_2 \oplus X_3 \oplus rk)), \tag{6}$$

where rk denotes the round key. In each round, the nonlinear function T does the following operations in sequence: 32-bit subkey addition, S-box Subsitution (layer of four 8-bit S-boxes) and lastly, a 32-bit linear transformation L.

It is well-known that the S-boxes adopt functions affine equivalent to the inversion function in $GF(2^8)$ [15, 10], which achieves the best known maximum differential and linear probabilities of 2^{-6}. Furthermore, it can be verified that the branch number of the linear transformation L is $\mathcal{L}_d = 5$. This gives optimal spreading effect which increases the number of active S-boxes for protection against differential and linear cryptanalysis.

The key schedule of SMS4 XORs the secret key MK with a constant FK and passes it through a nearly-identical 32-round structure as the main SMS4 cipher. The only difference is that the 32-bit linear transformation L is replaced by a simpler linear transformation L', which can be verified to have branch number $\mathcal{L'}_d = 4$. The 32-bit nonlinear output of the i-th round of the key schedule is taken to be the i-th round subkey of the main cipher. For more details, please refer to [11].

6.2 Parallelizing SMS4: p-SMS4

In this section, we propose another version of the existing SMS4 block cipher, which we call p-SMS4 (*"parallelizable" SMS4*). As described previously, SMS4 uses a generalized Feistel network structure described by equation (6). For the encryption procedure of p-SMS4, we shall replace the generalized Feistel network with the 4-cell GF-NLFSR block cipher structure described by:

$$(X_0, X_1, X_2, X_3, rk) \mapsto (X_1, X_2, X_3, X_1 \oplus X_2 \oplus X_3 \oplus T(X_0 \oplus rk)). \tag{7}$$

Other components such as number of rounds and the T-function, which consists of four S-boxes and a L-function, remain the same as SMS4. One round of p-SMS4 corresponds to a 4-cell version of the structure in Figure 1, where the nonlinear function $F(\cdot)$ is the T-function used in SMS4.

The key schedule of p-SMS4 XORs the secret key MK with a constant FK and passes it through an identical 32-round structure as the main cipher of p-SMS4 described by equation (7). The constant FK, S-box and the linear transformation L' of the key schedule remain the same as SMS4. We need the key schedule to have the same structure as the main cipher so that it is also parallelizable in hardware, and thus can be made "on-the-fly".

6.3 Differential and Linear Cryptanalysis of p-SMS4

Su et al. proved bounds for the differential characteristic probability of the SMS4 cipher in [19]. One of the results they proved was that in every 7 rounds of the SMS4 cipher, there are at least 5 active S-boxes. However, there are currently no known bounds on the linear characteristic probability of SMS4 to the best of our knowledge.

Similarly for the p-SMS4 cipher, we can easily compute the differential characteristic bound by Theorem 3. Denote the maximum differential probability of p-SMS4 reduced to 29-round by p (we assume a minus-3 round attack where the attacker guesses three subkeys with complexity 2^{96}).

Recall that both p-SMS4 and SMS4 use the same T-function. In the case of p-SMS4, the maximum differential probability of the S-boxes is 2^{-6} and $\mathcal{L}_d = 5$. By virtue of Theorem 3 with $n = 4$ and $r = 5$, the first 24 rounds has $5 \times 3 + \lfloor 3/2 \rfloor = 16$ active S-boxes. Over the next 5 rounds, we have 2 active S-boxes by Proposition 1. Therefore the differential characteristic probability over 29 rounds satisfies:

$$p \leq (2^{-6})^{16} \times (2^{-6})^2 = 2^{-108}.$$

This implies that an attacker needs to collect at least 2^{108} chosen plaintext-ciphertext pairs to launch an attack. This is not feasible in practice. Moreover by Remark 1, for random input differences, we have at least 5 active S-boxes every 5 rounds with probability $1 - 2^{-32}$. Only 2^{-32} of the time do we need 8 rounds to ensure at least 5 active S-boxes. Thus we expect the bound for the differential characteristic probability to be even lower. In summary, we have shown that p-SMS4 offers sufficient security against differential cryptanalysis.

Denote the maximum linear probability of p-SMS4 reduced to 28-round by q. Recall that the maximum linear probability of the S-boxes is 2^{-6} and the linear branch number is 5. By Theorem 5 and Corollary 1, we deduce that there must be at least 16 linear active S-boxes. Hence $q \leq (2^{-6})^{16} = 2^{-96}$. This implies that an attacker needs to collect at least 2^{96} chosen/known plaintexts to mount a linear attack, which is not feasible in practice.

This implies that there is no effective differential or linear characteristic for p-SMS4 reduced to more than 29 rounds. In other words, p-SMS4 offers sufficient security against differential and linear attack.

6.4 Related-Key Differential Attack on p-SMS4

Related-key differential attacks have been shown to have the devastating effect of recovering the secret key of AES-256 with a complexity of 2^{131} using 2^{35} related keys in [7]. In related-key differential attack, there

are non-zero differential inputs into both the cipher and the key schedule. The adversary tries to find a differential characteristic path in the key schedule with probability p_k and a differential characteristic path in the main cipher with probability $p_{c|k}$ that holds, on the condition that the key schedule differential path is true. The attacker can then launch the attack with complexity $O(1/(p_k \times p_{c|k}))$ where he can tweak the secret key $1/p_k$ times to get that many related keys. In AES-256, we have $p_k = 2^{-35}$ and $p_{c|k} = 2^{-93}$.

Because the p-SMS4 key schedule uses a 4-cell GF-NLFSR structure, we can try to bound the probability p_k of a differential characteristic path in the key schedule by Theorem 3. However, Theorem 3 cannot be directly applied to the main cipher to derive the differential characteristic probability $p_{c|k}$ because there are subkey differential input into every round.

We use the fact that the key schedule uses the inversion S-box with differential probability 2^{-6} and that the linear transform L' has branch number $\mathcal{L'}_d = 4$. By Theorem 3 with $n = 4$ and $r = 4$, every 24 rounds of the key schedule has $4 \times 3 + \lfloor 3/2 \rfloor = 13$ active S-boxes. With a computation similar to Section 6.3, we have another 2 active S-boxes over the next 5 rounds giving:

$$p_k \leq (2^{-6})^{13} \times (2^{-6})^2 = 2^{-90}.$$

over 29 rounds of the key schedule. That means the complexity of any minus-3 round related-key differential attack is at least $O(2^{90})$ and uses at least 2^{90} related keys, which is not feasible in practice. Again, by a similar explanation as in Section 6.3 based on Remark 1, most of the time we have 5 active S-boxes per 5 rounds and we expect p_k to be lower and the attack complexity to be higher.

In [6], a related-key boomerang attack on AES-256 with a complexity of 2^{119} using 4 related keys is presented but it assumes a more powerful adverserial model. In a similar way, we can show through the p-SMS4 key schedule differential structure that related-key boomerang attack is infeasible.

6.5 Other Attacks on p-SMS4

Boomerang Attack Suppose an adversary performs a minus-3 round attack on 29 rounds of p-SMS4. He would need to split 29 rounds into two sub-ciphers E_0, E_1 with r and $29-r$ rounds respectively, where $r = 1, \cdots, 28$. By Proposition 1 and Theorem 3, $p_0 \leq (2^{-6})^{5\times\lfloor\frac{r}{8}\rfloor+2\times\lfloor\frac{r \bmod 8}{5}\rfloor}$ and $p_1 \leq (2^{-6})^{5\times\lfloor\frac{29-r}{8}\rfloor+2\times\lfloor\frac{(29-r) \bmod 8}{5}\rfloor}$. (Note that we ignore the last term in the upper bound of Theorem 3 for ease of calculation.) For $r = 1, \cdots, 28$, let $n_8 = \lfloor\frac{r}{8}\rfloor + \lfloor\frac{29-r}{8}\rfloor$ and $n_5 = \lfloor\frac{r \bmod 8}{5}\rfloor + \lfloor\frac{(29-r) \bmod 8}{5}\rfloor$. It can be easily checked that there are only three combinations of values that n_8 and n_5 can take, as summarized in the Table 2.

Now $p_0 \times p_1 \leq (2^{-6})^{5n_8+2n_5}$. This implies that

$$p_0^2 \times p_1^2 \leq (2^{-12})^{5n_8+2n_5}.$$

The upper bounds of $p_0^2 \times p_1^2$ for each combination of n_8 and n_5 are also given in Table 2. From Table 2, we see that $p_0^2 \times p_1^2 < 2^{-128}$. Hence p-SMS4 is secure against boomerang attack.

n_8	n_5	r	$p_0^2 \times p_1^2$
3	0	$1, \cdots, 4, 9, \cdots, 12, 17, \cdots, 20, 25, \cdots, 28$	$\leq (2^{-12})^{15} = 2^{-180}$
3	1	5, 8, 13, 16, 21, 24	$\leq (2^{-12})^{15+2} = 2^{-204}$
2	2	6, 7, 14, 15, 22, 23	$\leq (2^{-12})^{10+4} = 2^{-168}$

Table 2. Values of n_8, n_5 and upper bounds of $p_0^2 \times p_1^2$ for $r = 1, \cdots, 28$

Impossible Differential Attack According to [8, 20, 24], there is at least one 18-round impossible differential distinguisher in the 4-cell GF-NLFSR, which results in a 25-round impossible differential attack with complexity 2^{123} and uses 2^{115} chosen plaintext encryptions. An identical attack is applicable to 25-round p-SMS4 with the same complexity. However, that attack is unlikely to work on the full p-SMS4 cipher, which has 32 rounds.

Integral Attack According to [8, 20], there is at least one 16-round integral attack distinguisher in the 4-cell GF-NLFSR starting with one active 32-bit word. A naive key guessing attack can extend this distinguisher by at most 3 rounds at the end (guessing more rounds of keys may make the complexity too close to 2^{128}). An adversary may extend the attack by 4 rounds in front, starting with 3 active words and using the method of [12]. Using these means, we expect a $4 + 16 + 3 = 23$ round attack on p-SMS4 and the full 32 rounds will be secure against integral attack.

Slide Attack The slide attack works on ciphers with cyclical structures over a few rounds. However the subkeys used in every round are nonlinearly derived from the previous subkey. Thus the subkeys are all distinct and there is no simple linear relation between them, making slide attack unlikely.

XSL Attack In [15], Ji and Hu showed that the eprint XSL attack on SMS4 embedded in $GF(2^8)$ can be applied with complexity 2^{77}. A similar analysis can be applied on p-SMS4 to show that the complexity of the eprint XSL attack on p-SMS4 embedded in $GF(2^8)$ is also 2^{77}. However, it was shown in [10] by Choy et al. that Ji and Hu's analysis might be too optimistic and the actual complexity of the compact XSL attack on embedded SMS4 is at least $2^{216.58}$. We can use an analysis identical to the ones used in [10] to show that the complexity of the compact XSL attack on p-SMS4 is also at least $2^{216.58}$.

Using a similar approach as [5], we discuss the protection of p-SMS4 against higher order differential attack and interpolation attack in the remaining of this section.

Higher Order Differential Attack As mentioned previously, higher order differential attack is generally applicable to ciphers that can be represented as Boolean polynomials of low degree in terms of the plaintext. The attack requires $O(2^{t+1})$ chosen plaintext when the cipher has degree t.

p-SMS4 uses exactly the same S-boxes as SMS4 where the degree of the Boolean polynomial of every output bit of the S-boxes is 7. Making the assumption that when we compose two *randomly chosen* S-boxes F, G of degree t_1, t_2, $F \circ G$ should have degree $t_1 t_2$. We expect the degree of an intermediate bit in the encryption process to increase exponentially as the data passes through many S-boxes.

Indeed, by the 4^{th} round, every output bit will have degree 7. By the 8^{th} round, every output bit will have degree $7^2 = 49$. By the 12^{th} round, every output bit will have degree $min(7^3, 127) = 127$ in terms of the plaintext bits. Therefore p-SMS4 is secure against higher order differential attack.

Interpolation Attack The interpolation attack works on block ciphers that can be expressed as an equation in $GF(2^d)$ with few monomials. p-SMS4 uses the same components as SMS4 and as the data passes through many S-boxes and L-functions, the cipher will became a complex function which is a sum of exponentially many multi-variate monomials over $GF(2^8)$. Hence we expect p-SMS4 to be secure against interpolation attack.

6.6 Implementation Advantages

Similar to p-Camellia we will assess the implementation advantages of p-SMS4 over SMS4 with respect to serialized, round-based and parallelized architectures. In case of SMS4 the XOR sum of three branches

forms the input to the F-function and its output is XORed to the last branch while p-SMS4 uses one branch as the input for the F-function and XORs its output to the remaining three branches. This difference allows more flexible implementations of p-SMS4 compared to SMS4, because the XOR sum of four signals can be achieved by either using three 2-input XOR gates or combining a 3-input XOR gate with a 2-input XOR gate. The first option is faster (0.33 ns vs. 0.45 ns) while the second option requires less area (256 GE vs. 235 GE), which is an advantage for lightweight implementations. Beside this flexibility, p-SMS4 has similar characteristics as SMS4 for a serialized implementation. The critical path of a round-based p-SMS4 implementation is shorter than that of SMS4, since it consists of the F-function and a 2-input XOR gate compared to a 3-input XOR gate, the F-function and a 2-input XOR gate for SMS4.

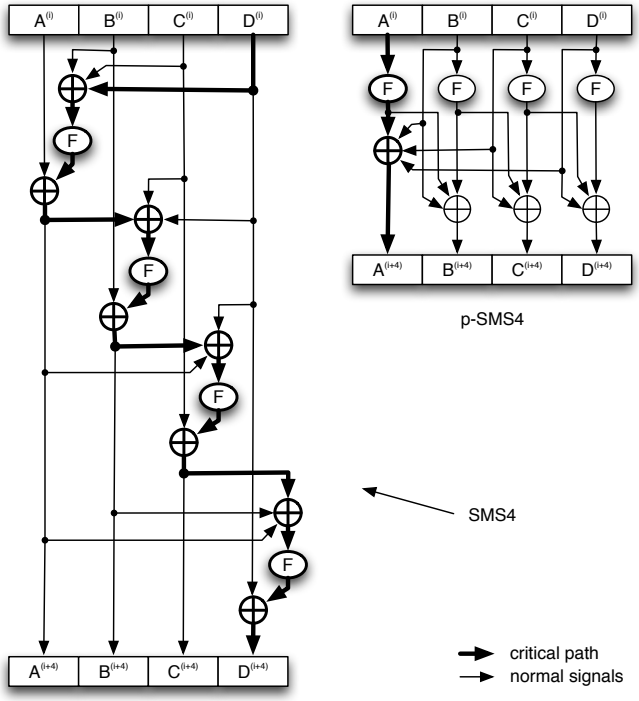

Fig. 6. Possible hardware architecture of four rounds of SMS4 (left) and p-SMS4 (right).

For parallelized implementations p-SMS4 offers even greater advantages. If we consider an implementation that processes four rounds in one clock cycle (see figure 6), the critical path of p-SMS consists only of the F-function and two 2-input XOR gates while SMS4's critical path consists of four F-functions, four 2-input XOR gates and four 3-input XOR gates. Hence, the maximum frequency and thus the maximum throughput that can be achieved with p-SMS4 using such an architecture is around four times higher while the area and power consumption are similar or lower compared to a corresponding SMS4 implementation. A similar frequency can be achieved for SMS4 by inserting three pipelining stages, which significantly increases the area and power consumption and introduces a delay of three clock cycles.

To substantiate our claims we have implemented the round function of SMS4 and p-SMS4 in VHDL. We obtained area, timing and power figures figures for a 180 nm ASIC technology from UMC using Synopsys Design Vision for synthesis. Table 3 depicts a comparison of the hardware implementation results of the round function of SMS4 and p-SMS4. This is a typical setup in a co-processor or instruction set extension scenario. As expected, the area requirements of 2924 GE for one instance of the round function are the

Table 3. Comparison of the implementation results of the round function of SMS4 and p-SMS4 on UMC *180 nm* ASIC technology.

	SMS4				p-SMS4			
	1 round		4 rounds		1 round		4 rounds	
	abs.	%	abs.	%	abs.	%	abs.	%
Area (GE)	2924	100	11546	394.9	2924	100	11574	395.9
power* (mW)	1.81	100	11.38	627.5	1.39	76.8	5.9	322.3
max Freq. (MHz)	288.2	100	73.1	25.4	290.7	100.9	267.4	92.8
max T'put (Gbps)	36.9	100	37.4	101.4	37.2	100.9	136.9	371.1

*at a frequency of 100 MHz and a supply voltage of 1.8V.

same for SMS3 and p-SMS4 and nearly quadruple to 11546 GE and 11574 GE for four instances. The 1 round implementation of p-SMS4 achieves a slightly higher maximum frequency of 290.7 MHz compared to SMS4 with 288.2 MHz. However, as depicted in Figur 6 the critical path for four consecutive instances of the round function of SMS4 is nearly four times as long as for p-SMS4. Consequently, the maximum frequency achievable for SMS4 drops to 25.4% while it only slightly decreases to 92.8% for p-SMS4. The maximum throughput of a 1 round implementation is the about same for SMS4 and p-SMS4 and achieves 36.9 Gbps and 37.2 Gbps, respectively. A four round SMS4 implementation slightly increases the maximum throughput by a mere 1.4% to 37.4 Gbps, while p-SMS4 boosts the maximum throughput to 136.9 Gbps - an increment of 271.1% compared to the 1 round SMS4 implementation and still 266% higher than the 4 round SMS4 implementation.

For all architectures we simulated the power consumption at a frequency of 100 MHz and a supply voltage of 1.8 Volt. 1 round of SMS4 requires 1.81 mW and a similar p-SMS4 implementation rewuires 1.39 mW. While the power consumption for the 4 rounds implementation of SMS4 increases more than 6 times (+528%) to 11.38 mW, it less than quadruples for the 4 rounds implementation of p-SMS4 (+222%) to 5.85 mW compared to the 1 round implementations. These figures highlight the advantages of p-SMS4 over SMS4 from a power perspective.

From these estimates it becomes clear that the implementation advantages of our newly proposed parallelizable Feistel-structure becomes even larger with a growing number of branches. In fact, an n-cell GF-NLFSR can be implemented using n rounds in parallel while having nearly the same critical path as for a single round implementation. This translates to an about n times higher maximum frequency while the area and power consumption are similar then for a conventional Feistel structure.

7 Conclusion

In this paper we proposed the use of n-cell GF-NLFSR structure to parallelize (Generalized) Feistel structures. We used two examples, p-Camellia and p-SMS4, and showed that they offer sufficient security against various known existing attacks. At the same time, as compared to their conventional Feistel structure counterparts Camellia and SMS4, their hardware implementations achieve a maximum frequency that is about n times higher, where n is the number of Feistel branches, while having similar area demands and significantly less power demands. These estimates indicate that of n-cell GF-NLFSRs are particularly well suited for applications that require a high throughput.

References

1. National Bureau of Standards, Data Encryption Standard, FIPS-Pub.46. National Bureau of Standards, U.S. Department of Commerce, Washington D.C., January 1977.
2. "SKIPJACK and KEA Algorithm Specifications" at `http://csrc.nist.gov/groups/ST/toolkit/documents/skipjack/skipjack.pdf`.

3. "Universal Mobile Telecommunications System (UMTS); Specification of the 3GPP confidentiality and integrity algorithms; Document 2: Kasumi specification" at `http://www.etsi.org/website/document/algorithms/ts_135202v070000p.pdf`.
4. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita, "Specification of Camellia - A 128-Bit Block Cipher", 2000. at `http://info.isl.ntt.co.jp/camellia/`.
5. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita, "*Camellia*: A 128-Bit Block Cipher Suitable for Multiple Platforms, Design and Analysis ", SAC 2000, LNCS 2012, pp. 39-56, Springer-Verlag, 2001.
6. A. Biryukov and D. Khovratovich, "Related-Key Cryptanalysis of the Full AES-192 and AES-256", IACR eprint server, 2009/317 (June 2009), `http://eprint.iacr.org/2009/317`.
7. A. Biryukov, D. Khovratovich, and I. Nikolic, "Distinguisher and Related-Key Attack on the Full AES-256 (Extended Version)", IACR eprint server, 2009/241 (June 2009), `http://eprint.iacr.org/2009/241`.
8. J. Choy, G. Chew, K. Khoo and H.Yap, "Cryptographic Properties and Application of a Generalized Unbalanced Feistel Network Structure (Revised Version)", *Cryptology Eprint Archive*, Report 2009/178 (July 2009). (Revision of [9])
9. 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.
10. J. Choy, H.Yap and K. Khoo, "An Analysis of the Compact XSL Attack on BES and Embedded SMS4", to appear in proceedings of CANS 2009, Springer-Verlag, 2009.
11. W. Diffe and G. Ledin, "SMS4 Encryption Algorithm for Wireless Networks", Cryptology ePrint Archive: Report 2008/329.
12. K. Hwang, W. Lee, S. Lee, S. Lee and J. Lim, "Saturation Attacks on Reduced Round Skipjack", LNCS 2365, FSE 2002, pp. 100-111, Springer-Verlag, 2002.
13. T. Jakobsen and L. R. Knudsen, "The Interpolation Attack on Block Ciphers", LNCS 1267, *FSE 1997*, pp. 28-40, Springer-Verlag, 1997.
14. T. Jakobsen and L. R. Knudsen, "Attacks on Block Ciphers of Low Algebraic Degree", Journal of Cryptology, Vol. 14, pp. 197-210, Springer, 2001.
15. W. Ji and L. Hu, "New Description of SMS4 by an Embedding over $GF(2^8)$", LNCS 4859, *Indocrypt 2007*, pp. 238-251, Springer-Verlag, 2007.
16. M. Kanda, "Practical Security Evaluation against Differential and Linear Cryptanalysis for Feistel Ciphers with SPN Round Function", SAC 2000, LNCS 2012, pp. 324-338, Springer-Verlag, 2001.
17. L. R. Knudsen, "Truncated and Higher Order Differentials", LNCS 1008, *FSE 1994*, pp. 196-211, Springer-Verlag, 1995.
18. M. Matsui, "On Correlation Between the Order of S-boxes and the Strength of DES", Eurocrypt 1994, LNCS 950, pp. 366-375, 1995.
19. B. Su, W. Wu and W. Zhang, "Differential Cryptanalysis of SMS4 Block Cipher", *Cryptology Eprint Archive*, Report 2010/062 (Feb 2010).
20. R. Li, B. Sun and C. Li, "Distinguishing Attack on a Kind of Generalized Unbalanced Feistel Network", *Cryptology Eprint Archive*, Report 2009/360 (July 2009).
21. S. Park, S. Sung, S. Lee, J. Lim, "Improving the Upper Bound on the Maximum Differential and the Maximum Linear Hull Probability for SPN Structures and AES", FSE 2003, LNCS 2887, pp. 247-260, Springer-Verlag, 2003.
22. T. Shirai, K. Shibutani, T. Akishita, S. Moriai, T. Iwata, "The 128-Bit Blockcipher CLEFIA (Extended Abstract)", FSE 2007, LNCS 4593, pp. 181-195, Springer-Verlag, 2007.
23. Y. Wei, P. Li, B. Sun and C. Li, "Impossible Differential Cryptanalysis on Feistel Ciphers with *SP* and *SPS* Round Functions", *ACNS 2010*, LNCS 6123, pp. 105-122, Springer, 2010.
24. W. Wu, L. Zhang, L. Zhang and W. Zhang, "Security Analysis of the GF-NLFSR Structure and Four-Cell Block Cipher", *Cryptology Eprint Archive*, Report 2009/346 (July 2009).
25. V. Rijmen, J. Daemon, B. Preneel, A. Bosselaers, and E. D. Win, "The cipher SHARK", *Fast Software Encryption - Third International Workshop*, LNCS 1039, pp. 99-111, 1996.