

MultLA: A Framework of Fast Correlation Attack Exploiting Multiple Linear Approximations and Its Application to SNOW 5G

Jiang Wan, Lin Ding*, Jiekun Sun, Xinhai Wang, and Honglei Wang

Information Engineering University, Zhengzhou, China

Email: dinglin_cipher@163.com

Abstract. The complexity of fast correlation attack is fundamentally determined by the correlations of the linear approximations used. Conventional fast correlation attacks either rely on a single dominant linear approximation or use the lower bound to evaluate the resulting correlation of multiple linear approximations, which have the same output mask and distinct input masks. How to exploit multiple linear approximations with distinct output masks remains an open problem. In this paper, we propose a framework of fast correlation attack exploiting multiple linear approximations called MultLA to solve this problem. MultLA utilizes the arithmetic mean of the absolute correlations of linear approximations, rather than the lower bound of that. For SNOW 5G, our attack has time/memory/data complexities of $2^{279.25}/2^{265.03}/2^{263.62}$, which are all better than those of the existing work. When the time and memory complexities are slightly better than those of the existing work, the data complexity is significantly reduced from $2^{265.4}$ to $2^{261.21}$. To the best of our knowledge, our attacks achieve the best known cryptanalytic results of SNOW 5G up to now.

Keywords: Stream cipher · Fast correlation attack · Multiple linear approximations · SNOW 5G

1 Introduction

With the evolution of mobile communications from 3G to 5G, stream ciphers deployed in the system must not only support efficient software implementation in generic environments but also provide strong guarantees against cryptanalysis. The SNOW family of stream ciphers remains central to 3GPP standards. SNOW 3G [1] provided 128-bit security for 4G under hardware acceleration, but its limited software performance and high biases in nonlinear components made it unsuitable for 5G. This motivated the design of SNOW 5G [2], a systematic redesign optimized for high-speed software execution and 256-bit security. The SNOW-V stream cipher was introduced in [5] as a candidate for the 5G confidentiality and integrity algorithms, followed by an improved variant, SNOW-Vi,

* Corresponding author.

in [6]. In March 2024, 3GPP standardized a refined version of this design under the name SNOW 5G.

The cryptanalysis of the standardized SNOW 5G remains limited. The first cryptanalysis of SNOW 5G was presented in [13], which applied the automatic search technique from [17] to find a linear approximation with correlation $2^{-75.82}$ and mounted a correlation attack with complexity $2^{312.24}$. The best linear approximation found to date was presented by [11] in 2024. By characterizing high-correlation mask structures and applying a divide-and-conquer search, they obtained a linear approximation with correlation $-2^{-67.67}$. The resulting fast correlation attack achieves a complexity of $2^{279.8}$, improving the existing correlation attack by a factor of $2^{32.44}$. It is important to note that their attack is based on a single linear approximation and does not utilize multiple linear approximations.

Meier and Staffelbach [14] introduced the fast correlation attack in 1989. Since then, it has become a cornerstone in the cryptanalysis of LFSR-based stream ciphers and has been widely applied in subsequent works, such as [7,20,22]. Conventional fast correlation attacks are often based on a single linear approximation, and their success is fundamentally determined by the correlation of the single linear approximation used. Recent research has explored the usage of multiple linear approximations to enhance the attack. Murphy [15] derived the covariance between counting variables corresponding to different masks for multiple linear cryptanalysis in block ciphers, and proved that under appropriate conditions these counters remain statistically independent even if the underlying masks are linearly dependent. At CRYPTO 2018, Todo et al. [18] revisited the fast correlation attack from a finite field perspective and established a multiple linear approximations model. By proposing a new wrong-key distribution hypothesis, they proposed a novel attack algorithm that reduces both time and data complexities. Based on two methods to reduce the attacked bits, recently Ma et al. [12] proposed an improved fast correlation attack for stream ciphers with LFSR size exceeding the key size. The attack constructs parity-check equations which are constructed by XORing identical linear approximations across different clocks in the preprocessing phase. It then uses multiple linear approximations that further reduce attacked bits with expected probabilities in the online phase, ultimately recovers the full initial state by solving systems of linear equations.

Conventional fast correlation attacks either rely on a single dominant linear approximation or use the lower bound to evaluate the resulting correlation of multiple linear approximations, which have the same output mask and distinct input masks. As a result, many high-correlation linear approximations with distinct input and output masks cannot be utilized. Thus, how to exploit multiple linear approximations with distinct output masks remains an open problem. In this paper, we will bridge this gap by proposing a framework to solve this problem.

Technical contributions. In this paper, we propose a framework of fast correlation attack exploiting multiple linear approximations called **MultLA**. In this

framework, we construct a statistic for fast correlation attack that uses multiple linear approximations with distinct output masks. We give a definition of *pair-wise uncorrelated* linear approximations and prove that the introduced statistic is asymptotically normal when the linear approximations used are pairwise uncorrelated. This allows us to apply the Lindeberg-Lévy central limit theorem to estimate the complexities of the attack directly from the arithmetic mean of the individual absolute correlations. This results in improved attack complexity compared with the previous method that uses the lower bound to evaluate the resulting correlation of multiple linear approximations.

Applications. We find 19 distinct linear approximations for SNOW 5G with absolute correlation greater than 2^{-68} . By combining these pairwise uncorrelated linear approximations with the constructed statistic, we achieve the best known fast correlation attacks on the cipher. A comprehensive comparison with previous results is provided in Table 1. As shown in Table 1, our attacks achieve the best known cryptanalytic results of SNOW 5G up to now.

Table 1. Fast correlation attacks against SNOW 5G

Cipher	Number of LA	Type	Time/Memory/Data	Ref.
SNOW 5G	Single LA	/	$2^{312.24}/2^{297.12}/2^{297.12}$	[13]
	Single LA	/	$2^{279.8}/2^{265.5}/2^{265.4}$	[11]
	Multiple LAs	All-optimal	$2^{279.25}/2^{265.03}/2^{263.62}$	this paper
	Multiple LAs	Data-optimal	$2^{279.78}/2^{265.21}/2^{261.21}$	this paper

Organization. This paper is organized as follows. Section 2 presents the notation used throughout this paper and provides a brief overview of SNOW 5G. Section 3 briefly reviews the existing fast correlation attack based on multiple and multidimensional linear approximations. Section 4 proposes a general framework of fast correlation attack exploiting multiple linear approximations called MultLA. Section 5 apply the proposed framework to SNOW 5G. Section 6 concludes the paper.

2 Preliminaries

2.1 Notations and Definitions

We introduce the following notation and definitions.

- Let $\mathbb{F}_2$ denote the binary field, $\mathbb{F}_{2^m}$ its extension field of degree m , and $\mathbb{F}_2^m$ the m -dimensional vector space over $\mathbb{F}_2$.
- Let $\boxplus_{32}$ denote the modulo 2^{32} addition, $\boxplus_{16}$ denote the modulo 2^{16} addition, and $\oplus$ denote the bitwise XOR operation.

- Denote by $\Pr(A)$ the probability of event A .
- Let $E(\cdot)$ and $E^{-1}(\cdot)$ denote the AES encryption and decryption round functions, respectively, both with the round key set to zero.
- Denote by $|A|$ the number of elements in the set A .
- Denote by $\mathbb{E}[X]$ and $\text{Var}[X]$ the expectation and variance of a random variable X , respectively.
- Let n and m be positive integers such that $m \mid n$. Any element $x \in \mathbb{F}_2^n$ can be uniquely decomposed as

$$x = (x_0 \parallel x_1 \parallel \dots \parallel x_{d-1}),$$

where $d = n/m$, each $x_i \in \mathbb{F}_{2^m}$ for $0 \leq i \leq d-1$, and x_0 corresponds to the least significant m bits of x .

- Let $a = (a_{n-1}, a_{n-2}, \dots, a_0)$ and $b = (b_{n-1}, b_{n-2}, \dots, b_0)$ be two binary vectors of length n . The inner product of a and b is given by $a \cdot b = \bigoplus_{i=0}^{n-1} a_i \cdot b_i$.
- A function $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ is called a vectorial Boolean function. It can be expressed as $F(x) = (f_1(x), f_2(x), \dots, f_m(x))$, where each coordinate $f_i : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is a Boolean function of n variables.

Definition 1. Let X be a uniformly distributed binary random variable in $\mathbb{F}_2^n$. Given a Boolean function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ in n binary variables, the correlation of f to zero is defined as

$$c(f) = \Pr(f(X) = 0) - \Pr(f(X) = 1) = 2\Pr(f(X) = 0) - 1,$$

and the bias of f is defined as $\epsilon(f) = \Pr(f(X) = 0) - \frac{1}{2}$. It follows directly that $\epsilon(f) = \frac{1}{2}c(f)$.

Definition 2. Let X be uniformly distributed over $\mathbb{F}_2^n$. For a vectorial Boolean function $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ and linear masks $a \in \mathbb{F}_2^n$, $b \in \mathbb{F}_2^m$, the correlation of F is defined as

$$c(a, b) = \Pr(a \cdot X \oplus b \cdot F(X) = 0) - \Pr(a \cdot X \oplus b \cdot F(X) = 1),$$

and the bias of F is defined as

$$\epsilon(a, b) = \Pr(a \cdot X \oplus b \cdot F(X) = 0) - \frac{1}{2}.$$

2.2 A Brief Description of SNOW 5G

Continuing the evolution of the SNOW family beyond SNOW 3G, the designs of SNOW-V and its accelerated variant SNOW-Vi laid the foundations for modern 5G-oriented stream ciphers. SNOW 5G is derived from SNOW-Vi and retains the core LFSR/CFSM architecture of the SNOW family, which is further adapted through a structural modification that replaces the 32-bit modular addition with a 16-bit variant. The overall structure of SNOW 5G is illustrated in Figure 1.

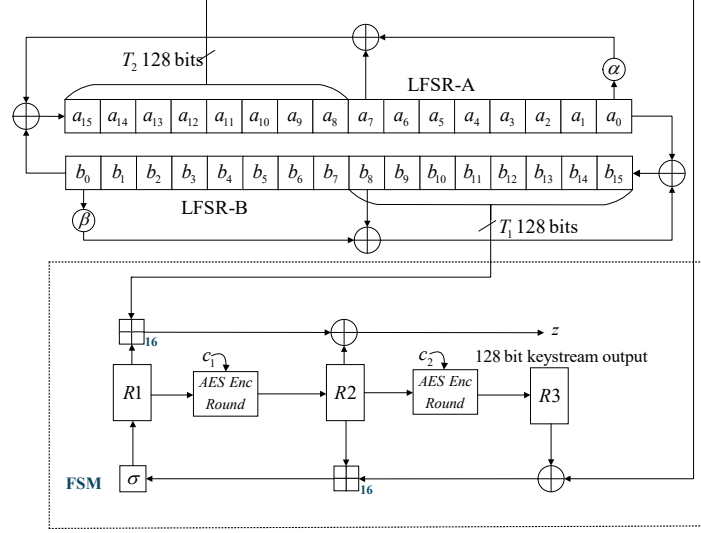


Fig. 1. The schematic of SNOW 5G

SNOW 5G adopts two parallel LFSRs named LFSR-A and LFSR-B over the field $F_{2^{16}}$, each consisting of 16 stages with 16 bits per stage. The FSM of SNOW 5G operates on three 128-bit registers, in contrast to the 32-bit registers used in SNOW 3G, while the sequences produced by LFSR-A and LFSR-B are generated according to irreducible polynomials $g^A(x)$ and $g^B(x)$ defined as

$$\begin{aligned} g^A(x) &= x^{16} + x^{14} + x^{11} + x^9 + x^6 + x^5 + x^3 + x^2 + 1 \in F_2[x], \\ g^B(x) &= x^{16} + x^{15} + x^{14} + x^{11} + x^{10} + x^7 + x^2 + x + 1 \in F_2[x]. \end{aligned}$$

Let $\alpha \in F_{2^{16}}^A$ be a root of $g^A(x)$ and $\beta \in F_{2^{16}}^B$ be a root of $g^B(x)$. At time t , the state of LFSR-A is denoted by $(a_{15}^{(t)}, a_{14}^{(t)}, \dots, a_0^{(t)})$, where each $a_i^{(t)} \in F_{2^{16}}^A$, and the state of LFSR-B is denoted by $(b_{15}^{(t)}, b_{14}^{(t)}, \dots, b_0^{(t)})$, where each $b_i^{(t)} \in F_{2^{16}}^B$. The state update equations are defined as

$$\begin{aligned} a_{15}^{(t+1)} &= b_0^{(t)} + \alpha a_0^{(t)} + a_7^{(t)} \bmod g^A(\alpha), \\ b_{15}^{(t+1)} &= a_0^{(t)} + \beta b_0^{(t)} + b_8^{(t)} \bmod g^B(\beta). \end{aligned}$$

In each update step, both LFSR-A and LFSR-B will update eight times, resulting in an update of 256 bits out of the total 512-bit LFSR state. Two 128-bit tap values T_1 and T_2 are then extracted from the LFSR and input to the FSM:

$$\begin{aligned} T_1^{(t)} &= (b_{15}^{(8t)}, b_{14}^{(8t)}, \dots, b_8^{(8t)}), \\ T_2^{(t)} &= (a_{15}^{(8t)}, a_{14}^{(8t)}, \dots, a_8^{(8t)}). \end{aligned}$$

The FSM takes two 128-bit words T_1 and T_2 from the LFSR as input and produces a 128-bit keystream word z_t as output. Let R_1, R_2, R_3 denote the three internal 128-bit registers of the FSM, the keystream output at time t is defined as

$$z^{(t)} = (R_1^{(t)} \boxplus_{16} T_1^{(t)}) \oplus R_2^{(t)}.$$

The registers are then updated as follows:

$$\begin{aligned} R_1^{(t+1)} &= \sigma \left(R_2^{(t)} \boxplus_{16} \left(R_3^{(t)} \oplus T_2^{(t)} \right) \right), \\ R_2^{(t+1)} &= E \left(R_1^{(t)} \right), \\ R_3^{(t+1)} &= E \left(R_2^{(t)} \right), \end{aligned}$$

where $\sigma = [0, 4, 8, 12, 1, 5, 9, 13, 2, 6, 10, 14, 3, 7, 11, 15]$ is a byte-oriented permutation that rearranges the 16 bytes of its input according to the permutation mapping σ . Let the 128-bit state be denoted as $S = (w_{15}, w_{14}, \dots, w_0)$, where each w_i ($0 \leq i \leq 15$) is a byte, can be arranged as

$$\begin{bmatrix} w_0 & w_4 & w_8 & w_{12} \\ w_1 & w_5 & w_9 & w_{13} \\ w_2 & w_6 & w_{10} & w_{14} \\ w_3 & w_7 & w_{11} & w_{15} \end{bmatrix}.$$

Applying the σ operation to S transposes this matrix, yielding

$$\sigma(S) = \begin{bmatrix} w_0 & w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 & w_7 \\ w_8 & w_9 & w_{10} & w_{11} \\ w_{12} & w_{13} & w_{14} & w_{15} \end{bmatrix}.$$

3 A Brief Overview of Fast Correlation Attacks with Multiple and Multidimensional Linear Approximations

Multiple Linear Cryptanalysis. Ågren et al. [3] observed that when there exist m linear approximations that have high correlations with the same keystream output mask, while the correlations of all others are zero, the problem of recovering the initial state of the LFSR can be transformed into recovering the correct one from among m related solutions under a single linear approximation. Todo et al. [18] proposed a fast correlation attack method using multiple linear approximations at CRYPTO 2018. They established a multiple linear approximations model from the perspective of finite fields. Their key finding is that when m linear approximations with high correlations are found for the same output mask, the traditional statistical assumption concerning wrong keys no longer holds in the multiple approximation scenario. Consequently, they proposed a new wrong-key distribution hypothesis. Based on this theory, Todo et al. designed a novel attack algorithm which consists of the following three steps. First

they construct parity-check equations. Second, they bypass unknown variables by fixing part of the initial state bits and then use the Fast Walsh-Hadamard Transform (FWHT). Finally, they incorporate a crucial *removing* γ step and utilizing a Poisson distribution to uniquely determine the initial state of the LFSR. However, the success of this method critically depends on the precondition that the size n of the target cipher’s LFSR is approximately equal to its key size κ . To break this limitation, Ma et al. [12] significantly extended Todo’s method to apply to stream ciphers whose LFSR size is substantially larger than the key size (typically $n \geq 2\kappa$). The attack framework consists of three steps, i.e., first, they reduce the number of unknown variable bits involved in the parity-check equations using a collision reduction technique; second, they leverage multiple linear approximations to probabilistically bypass further bits, thereby reducing the number of bits to be guessed in the online phase to an even lower level; finally, they recover the full n -bit LFSR initial state by solving systems of linear equations and determining the solution that appears most frequently among the candidates.

However, all the works above focus on utilizing sets of linear approximations with the same keystream output mask. How to construct an efficient fast correlation attack by exploiting multiple linear approximations with distinct output masks remains an open problem.

Multidimensional Linear Cryptanalysis. Multidimensional linear approximations were introduced by Hermelin et al. in [8,9,10], based on a vector space of linear approximations that are not necessarily independent. Zhang et al. [21] proposed a vectorial framework for fast correlation attacks that utilizes multiple linear approximations through the novel concept of a correlation profile, combined with vectorized code reduction and general decoding algorithm to improve key recovery. Multidimensional linear cryptanalysis requires the linear approximations to form a vector space. In contrast, our new framework has no such requirement.

4 MultLA: A Framework of Fast Correlation Attack Exploiting Multiple Linear Approximations

This section introduces a framework called MultLA for computing the correlation in fast correlation attack that employs multiple linear approximations with distinct output masks. The proposed framework introduces a statistic designed for the set of high-correlation linear approximations. Combined with an evaluation framework based on the Lindeberg-Lévy central limit theorem, the framework enables a more effective usage of high-correlation linear approximations, leading to a more accurate estimation of the attack complexity.

4.1 Construction of a Statistic

Todo et al. [18] constructed parity-check equations by XORing different input masks with the same output mask, while Ma et al. [12] proposed the concept

of *correlated solutions* under the same output mask to handle equations. Unlike these existing works, our framework handles the set of high-correlation linear approximations with distinct output masks. Our focus is on complex stream ciphers, where high-correlation linear approximations are typically found via heuristic searches or automated tools. This framework explicitly requires the linear approximations to satisfy the pairwise uncorrelatedness condition. This condition not only ensures modeling the variance of the test statistic, but also enables more effective combination of multiple high-correlation linear approximations, leading to significantly reduced attack complexity.

Let us consider linear approximation of the following form

$$\beta \cdot Z_t = \alpha(F^T)^t \cdot S_0,$$

where F denotes the binary matrix that represents the feedback polynomial of the LFSR, S_0 denotes the initial state of the LFSR, Z_t is the keystream vector at time t which has the same dimension as S_0 , and α and β are masks. Assume there are w classes of linear approximations with distinct output masks within the class. By XORing the same linear approximations at v different clocks to eliminate the $B' = n - B$ bits, the attacker can construct M parity-check equations of the form

$$e_j(\alpha^{(i)}, \beta^{(i)}, x) = (\beta^{(i)})^B \cdot \left(\bigoplus_{u=1}^v Z_{t_{j,u}} \right)^B \oplus \left(\bigoplus_{u=1}^v \alpha^{(i)}(F^T)^{t_{j,u}} \right)^B \cdot S_0^B,$$

where $1 \leq i \leq w$ and $1 \leq j \leq Mw^{-1}$, and each equation has correlation ρ_i^v . Following the approach of Todo, Meier, and Aoki [19], we separate linear approximations according to the sign of their correlation. However, our focus is on the more general case with distinct output masks. Accordingly, we define the statistic $T(x)$ as

$$T(x) = \sum_{j=1}^{Mw^{-1}} \left(\sum_{\substack{i=1 \\ \rho_i^v > 0}}^w (-1)^{e_j(\alpha^{(i)}, \beta^{(i)}, x)} - \sum_{\substack{i=1 \\ \rho_i^v < 0}}^w (-1)^{e_j(\alpha^{(i)}, \beta^{(i)}, x)} \right).$$

This statistic combines parity-check equations that are individually constructed from distinct input-output mask pairs, based on the noise variables.

In [18], Todo et al. revisited the FCA from a perspective of finite field, and introduced a new property for the FCA using multiple linear approximations. Based on this, they proposed a wrong-key hypothesis from a new perspective. Let $A_t \in \{0, 1\}^n$ denote the first row vector of $(F^T)^t$, where the i -th row vector of $(F^T)^t$ is given by A_{t+i-1} . Let $\theta \in F_{2^n}$ be a root of an irreducible polynomial $f(x)$, θ^t becomes natural conversion of binary vector A_t . Furthermore, for any $\alpha \in \{0, 1\}^n$, its conversion of $\gamma \in F_{2^n}$ implies that the matrix multiplication $\alpha \times (F^T)^t$ corresponds to the field multiplication $\gamma \theta^t$. Since multiplication in the finite field $\mathbb{F}_{2^n}$ is commutative, it has $\gamma \cdot \theta^t = \theta^t \cdot \gamma$. Denote P_γ as an $n \times n$

binary matrix, where the i -th row vector of $(P_\gamma)^T$ is $\gamma\theta^{t-1}$, and because $\theta^t\gamma$ is the conversion of $A_t \times (P_\gamma)^T$, $\alpha \times (F^T)^t = A_t \times (P_\gamma)^T$ can be obtained. If $S_0 \times P_\gamma$ is guessed, parity-check equations can be generated from A_t instead of $\alpha \times (F^T)^t$. Therefore, if an incorrect initial state $S'_0 = S_0 \times P_{\gamma'}$ is guessed, we can get

$$\langle S'_0, \alpha^{(i)} \times (F^T)^t \rangle = \langle S_0 \times P_{\gamma'}, A_t \times P_{\gamma_i}^T \rangle = \langle S_0, A_t \times P_{\gamma_i \gamma'}^T \rangle = \langle S_0, \alpha^{(i')} \times (F^T)^t \rangle.$$

This derivation reveals a relationship between the incorrect solution S'_0 and the linear mask $\gamma_i \gamma'$. If $\gamma_i \gamma'$ is a highly biased linear mask, we can also observe high correlation. To avoid this wrong-key issue, it suffices that at least one output mask corresponds to a unique input mask. Under this condition, an incorrect solution $S'_0 = S_0 \times M_{\gamma'}$ would replace the input mask $\alpha^{(i)}$ by a different mask that no longer corresponds to the original output mask $\beta^{(i)}$. This implies that the incorrect solution cannot satisfy parity-check equations and is thus rejected. Consequently, our framework is no longer constrained by the wrong key hypothesis of [18], which enables our framework to handle the set of high-correlation linear approximations with distinct output masks.

4.2 Theoretical Correlation Derivation

When combining multiple linear approximations, the statistical dependence among their corresponding masks must be considered. Following the framework of Murphy [15], we note that the covariance between two linear approximation statistics depends on the bias of their XOR-combined mask. Therefore, the linear approximations in our framework are required to satisfy that all pairwise XOR combinations yield balanced functions, i.e., zero bias. To facilitate the description, a definition is given as follows.

Definition 3 (Pairwise Uncorrelated Set of Linear Approximations). Let $\{(\alpha^{(i)}, \beta^{(i)})\}_{i=1}^w$ be a set of input-output mask pairs for a vectorial Boolean function $F : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$. We say that this set is **pairwise uncorrelated** if for every pair of distinct indices $i_1 \neq i_2$, the combined linear approximation defined by the XOR of the masks,

$$(\alpha^{(i_1)} \oplus \alpha^{(i_2)}) \cdot x \oplus (\beta^{(i_1)} \oplus \beta^{(i_2)}) \cdot F(x),$$

is a balanced Boolean function. Specifically, the bias ε_{i_1, i_2} satisfies

$$\varepsilon_{i_1, i_2} = \Pr_x \left[(\alpha^{(i_1)} \oplus \alpha^{(i_2)}) \cdot x \oplus (\beta^{(i_1)} \oplus \beta^{(i_2)}) \cdot F(x) = 0 \right] - \frac{1}{2} = 0.$$

Based on the definition above and the Lindeberg-Lévy central limit theorem, we now present the following theorem on the asymptotic distribution of the derived statistic. We assume that the noise variables $e_j(\alpha^{(i)}, \beta^{(i)}, x)$ corresponding to different time instances are independent, which is a standard assumption in correlation attacks and justified.

Theorem 1. Let $\{(\alpha^{(i)}, \beta^{(i)})\}_{i=1}^w$ be a pairwise uncorrelated set of linear approximations. Assume that at least one output mask corresponds to a unique input mask and that the correlations of these linear approximations are high. Define $\bar{\rho} = w^{-1} \sum_{i=1}^w |\rho_i^v|$, where ρ_i denotes the correlation of the i -th linear approximation and $v \geq 1$ is an integer parameter. Let M denote the number of parity-check equations, then the statistic $T(x)$ satisfies the following asymptotic normality:

- For the correct initial state, $T(x)$ converges in distribution to

$$\mathcal{N}(M\bar{\rho}, M),$$

- For the incorrect initial state, $T(x)$ converges in distribution to

$$\mathcal{N}(0, M).$$

Proof. Let x be a candidate initial state, and define $X_{i,j} = (-1)^{e_j(\alpha^{(i)}, \beta^{(i)}, x)}$, $G_j(x) = \sum_{\substack{i=1 \\ \rho_i^v > 0}}^w X_{i,j} - \sum_{\substack{i=1 \\ \rho_i^v < 0}}^w X_{i,j}$.

If x is the correct initial state, then $\Pr(X_{i,j} = 1) = (1 + \rho_i^v)/2$ and $\Pr(X_{i,j} = -1) = (1 - \rho_i^v)/2$, and we obtain $\mathbb{E}[X_{i,j}] = \rho_i^v$, $\text{Var}(X_{i,j}) = 1 - \rho_i^{2v}$. Considering the expectation and variance of $G_j(x)$ for $1 \leq j \leq Dw^{-1}$, we have

$$\mathbb{E}[G_j(x)] = \sum_{\substack{i=1 \\ \rho_i^v > 0}}^w \mathbb{E}[X_{i,j}] - \sum_{\substack{i=1 \\ \rho_i^v < 0}}^w \mathbb{E}[X_{i,j}] = \sum_{i=1}^w |\rho_i^v|.$$

and

$$\begin{aligned} \text{Var}(G_j(x)) &= \sum_{i=1}^w \text{Var}(X_{i,j}) + 2 \sum_{\substack{1 \leq i_1 < i_2 \leq w \\ \rho_{i_1}^v \rho_{i_2}^v > 0}} \text{Cov}(X_{i_1,j}, X_{i_2,j}) \\ &\quad - 2 \sum_{\substack{1 \leq i_1 < i_2 \leq w \\ \rho_{i_1}^v \rho_{i_2}^v < 0}} \text{Cov}(X_{i_1,j}, X_{i_2,j}), \end{aligned}$$

We recall that the covariance is given by

$$\text{Cov}(X_{i_1,j}, X_{i_2,j}) = \mathbb{E}[X_{i_1,j} X_{i_2,j}] - \mathbb{E}[X_{i_1,j}] \mathbb{E}[X_{i_2,j}].$$

For the term $\mathbb{E}[X_{i_1,j} X_{i_2,j}]$, we have

$$\mathbb{E}[X_{i_1,j} X_{i_2,j}] = (-1)^{e_j(\alpha^{(i_1)}, \beta^{(i_1)}, x) + e_j(\alpha^{(i_2)}, \beta^{(i_2)}, x)}.$$

We analyze the term $(-1)^{e_j(\alpha^{(i_1)}, \beta^{(i_1)}, x) + e_j(\alpha^{(i_2)}, \beta^{(i_2)}, x)}$, which corresponds to the equations of the XOR of two linear approximations. This is equivalent to a single linear approximation with combined masks $(\alpha^{(i_1)} \oplus \alpha^{(i_2)}, \beta^{(i_1)} \oplus \beta^{(i_2)})$. According to the condition of the theorem, any linear approximation constructed by XORing two distinct masks has zero correlation. Therefore, we can get

$$\mathbb{E}[X_{i_1,j} X_{i_2,j}] = \rho_{i_1, i_2}^v = 0.$$

Consequently, we obtain

$$\text{Var}(G_j(x)) = \sum_{i=1}^w (1 - \rho_i^{2v}) - 2 \sum_{\substack{1 \leq i_1 < i_2 \leq w \\ \rho_{i_1}^v \rho_{i_2}^v > 0}} \rho_{i_1}^v \rho_{i_2}^v + 2 \sum_{\substack{1 \leq i_1 < i_2 \leq w \\ \rho_{i_1}^v \rho_{i_2}^v < 0}} \rho_{i_1}^v \rho_{i_2}^v \approx w$$

Otherwise, if x is the incorrect initial state, then $\Pr(X_{i,j} = 1) = \Pr(X_{i,j} = -1) = \frac{1}{2}$, and we obtain $\mathbb{E}[X_{i,j}] = 0, \text{Var}(X_{i,j}) = 1$. In this case, we have

$$\mathbb{E}[G_j(x)] = \sum_{\substack{i=1 \\ \rho_i^v > 0}}^w \mathbb{E}[X_{i,j}] - \sum_{\substack{i=1 \\ \rho_i^v < 0}}^w \mathbb{E}[X_{i,j}] = 0,$$

$$\text{Cov}(X_{i_1,j}, X_{i_2,j}) = \mathbb{E}[X_{i_1,j} X_{i_2,j}] - \mathbb{E}[X_{i_1,j}] \mathbb{E}[X_{i_2,j}] = 0.$$

Thus, we obtain

$$\text{Var}(G_j(x)) = \sum_{i=1}^w \text{Var}(X_{i,j}) + 2 \sum_{1 \leq i_1 < i_2 \leq w} \text{Cov}(X_{i_1,j}, X_{i_2,j}) = w.$$

Under the above assumption, each $G_j(x)$ depends only on the noise variables at time instance j . For a fixed j , $G_j(x)$ combines the w linear approximations. By the independence assumption stated above, the variables $\{G_j(x)\}_{j=1}^{Mw^{-1}}$ are independent. Since each $G_j(x)$ is defined from the same set of linear approximations, their distributions are identical. Moreover, the random variables $\{G_j(x)\}_{j=1}^{Mw^{-1}}$ are independent across different j . Therefore, the sequence $\{G_j(x)\}_{j=1}^{Mw^{-1}}$ is independent and identically distributed. Consequently, by the Lindeberg-Lévy central limit theorem, the standardized statistic

$$\frac{T(x) - \mathbb{E}[T(x)]}{\sqrt{\text{Var}(T(x))}}$$

converges in distribution to a normal random variable as the number of terms increases. Specifically, we have

$$T(x) \sim \begin{cases} \mathcal{N}(M\bar{\rho}, M), & \text{correct initial state,} \\ \mathcal{N}(0, M) & \text{, incorrect initial state.} \end{cases}$$

□

According to the theorem, when a fast correlation attack employs parity-check equations from multiple linear approximations that are pairwise uncorrelated, the effective correlation reduces to the arithmetic mean of the individual absolute correlations. Combined with the condition that at least one output mask corresponds to a unique input mask, this allows us to accurately estimate the attack complexity.

4.3 Decoding with Multiple Output Masks via FWHT

Recall that each linear approximation used in our attack can be expressed in the form

$$\beta \cdot Z_t = \alpha(F^T)^t \cdot S_0,$$

where β is the output mask, α is the input mask, Z_t is the known keystream word at time t , and $S_0 \in \mathbb{F}_2^B$ is the unknown initial state to be recovered.

Following the approach of Berbain et al. [4], we employ the FWHT to efficiently evaluate all candidate states. For a fixed output mask β , the quantity $\beta \cdot Z_t$ is entirely determined by the known keystream and can thus be computed directly. Hence each linear approximation yields a known binary value

$$b_j = \beta \cdot Z_t \in \{0, 1\}.$$

Consequently, every equation can be rewritten in the unified form

$$\alpha^{(j)} \cdot S_0 = b_j, \quad j = 1, \dots, M,$$

where $\alpha^{(j)} \in \mathbb{F}_2^B$ is the input mask corresponding to the j -th equation. For any output mask β , the term $\beta \cdot Z_t$ is computable directly from the keystream. Each equation therefore provides a known bit $b_j = \beta \cdot Z_t$, reducing to $\alpha^{(j)} \cdot S_0 = b_j$. The output mask β serves only to produce b_j . Once b_j is obtained, the output mask β no longer appears in the decoding process.

We now describe how to efficiently evaluate all candidate states $y \in \mathbb{F}_2^B$. Define the function

$$S(y) = \sum_{j=1}^M (-1)^{\alpha^{(j)} \cdot y \oplus b_j}.$$

The correct state is expected to maximize $S(y)$. To accelerate the computation of $S(y)$ for all 2^B candidates, we construct a function $s(\alpha)$ by:

$$s(\alpha) = \sum_{\substack{j=1 \\ \alpha^{(j)}=\alpha}}^M (-1)^{b_j} = \#\{j : \alpha^{(j)} = \alpha, b_j = 0\} - \#\{j : \alpha^{(j)} = \alpha, b_j = 1\}.$$

That is, $s(\alpha)$ the difference between the number of equations with input mask α satisfying $b_j = 0$ and those satisfying $b_j = 1$. This function is computed by the M equations, requiring $O(M)$ time complexity.

Applying the Walsh transform to the function $s(\alpha)$ constructed above, we have

$$\widehat{s}(y) = \sum_{\alpha \in \mathbb{F}_2^B} s(\alpha) (-1)^{\alpha \cdot y}.$$

Substituting the definition of $s(\alpha)$ yields

$$\widehat{s}(y) = \sum_{j=1}^M (-1)^{b_j} (-1)^{\alpha^{(j)} \cdot y} = \sum_{j=1}^M (-1)^{\alpha^{(j)} \cdot y \oplus b_j} = S(y).$$

The time complexity for computing $\hat{s}(y)$ is $O(B \cdot 2^B)$. Hence the total time complexity of FWHT required for the decoding phase is $O(M + B \cdot 2^B)$, with memory $O(2^B)$.

5 Application to the SNOW 5G Stream Cipher

In this section, we apply the framework MultLA to launch a fast correlation attack to SNOW 5G. We find 19 linear approximations of SNOW 5G with absolute correlation greater than 2^{-68} . Based on these, we derive the corresponding attack complexity under the proposed framework. When the time and memory complexities are close to those of the existing work, the data complexity is reduced from $2^{265.4}$ to $2^{261.21}$ by using 9 linear approximations.

5.1 Multiple High-Correlation Linear Approximations of SNOW 5G

In [11], the authors proposed a theoretical framework to find high-correlation linear approximations for SNOW 5G, which reduces the search complexity by decomposing complex linear approximations into Type-I, Type-II, and Type-III. By exploiting constraints on active $\boxplus_{16}$ -blocks in the addition modulo 2^{16} and applying the divide-and-conquer method, they found the best approximation with correlation $-2^{-67.674}$.

Their analysis of SNOW 5G is based on the new theoretical framework they introduced. For the three-input modular addition involved in the N_3 linear approximation, an intermediate linear mask Θ is introduced. The correlation $c(N_3)$ is computed by summing over all intermediate linear masks Θ , i.e., $c(N_3) = \sum_{\Theta} c(P_3)c(P_4)$. As in Figure 2, this case of Θ gives the highest correlation for $c(N)$.

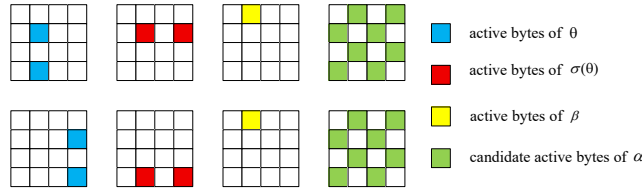


Fig. 2. The mask position that achieves the highest correlation for $c(N)$.

We briefly describe the mask search procedure below. First, according to the active column of Θ , we determine the active columns of γ . We get the 255 candidates of γ that activate only one S-box. Second, for each candidate γ , we test all possible values of n and compute $c(P_3)$ via the modular-addition automaton, retaining the pairs (γ, n) that yield large $|c(P_3)|$. Third, based on the active bytes of $\sigma(\Theta)$, we fix the active positions of h and $\beta \oplus \xi$, consider

Table 2. Linear approximations of SNOW 5G with absolute correlation greater than 2^{-68} .

α_1	α_2	α_3	α_4	l_1	l_2	l_3	l_4	$ \rho $	Ref.
0xb0000000	0x20041	0x80006000	0x2	0xf0000000	0x20061	0xc0006000	0x2	$2^{-67.674}$	[11]
0xb0000000	0x20041	0x8000b000	0x2	0xf0000000	0x20061	0xc000f000	0x2	$2^{-67.695}$	Ours
0xb0000000	0x20060	0x80006000	0x2	0xf0000000	0x20060	0xc0006000	0x2	$2^{-67.821}$	Ours
0xb0000000	0x200b0	0x80006000	0x2	0xf0000000	0x200f0	0xc0006000	0x2	$2^{-67.826}$	Ours
0xb0000000	0x20041	0x80006000	0x2	0xf0000000	0x20061	0xc0004000	0x2	$2^{-67.833}$	Ours
0xb0000000	0x20060	0x8000b000	0x2	0xf0000000	0x20060	0xc000f000	0x2	$2^{-67.842}$	Ours
0xb0000000	0x200b0	0x8000b000	0x2	0xf0000000	0x200f0	0xc000f000	0x2	$2^{-67.847}$	Ours
0xb0000000	0x20041	0x8000d000	0x2	0xf0000000	0x20061	0xc0009000	0x2	$2^{-67.902}$	Ours
0x03000000	0x20041	0x80006000	0x2	0x03000000	0x20061	0xc0006000	0x2	$2^{-67.923}$	Ours
0x03000000	0x20041	0x8000b000	0x2	0x03000000	0x20061	0xc000f000	0x2	$2^{-67.944}$	Ours
0xb0000000	0x20008	0x80006000	0x2	0xf0000000	0x2000c	0xc0006000	0x2	$2^{-67.957}$	Ours
0xb0000000	0x20041	0x8000f000	0x2	0xf0000000	0x20061	0xc000b000	0x2	$2^{-67.958}$	Ours
0xb0000000	0x20041	0x80006000	0x2	0xf0000000	0x20041	0xc0006000	0x2	$2^{-67.972}$	Ours
0xb0000000	0x20008	0x8000b000	0x2	0xf0000000	0x2000c	0xc000f000	0x2	$2^{-67.978}$	Ours
0xb0000000	0x20041	0x33006000	0x2	0xf0000000	0x20061	0x23006000	0x2	$2^{-67.980}$	Ours
0xb0000000	0x20060	0x80006000	0x2	0xf0000000	0x20060	0xc0004000	0x2	$2^{-67.980}$	Ours
0xb0000000	0x200b0	0x80006000	0x2	0xf0000000	0x200f0	0xc0004000	0x2	$2^{-67.985}$	Ours
0xb0000000	0x20041	0x8000b000	0x2	0xf0000000	0x20041	0xc000f000	0x2	$2^{-67.992}$	Ours

all corresponding candidates of h and ξ , and for each combination exhaustively search Θ to compute $c(N_3) = \sum_{\Theta} c(P_3)c(P_4)$, retaining the candidates with large $|c(N_3)|$. Fourth, for the selected β and γ , we select candidates of m that yield large $|c(N_1)|$, again using the modular-addition automaton. Fifth, for each retained pair (h, ξ) , we search over the active bytes of α and l and compute $c(N_2) = c(P_1)c(P_2) = c(P_1) \sum_{\zeta_2} c(M_{3,\zeta_2})c(M_{4,\zeta_2})$, retaining candidates of α and l with large $|c(N_2)|$. Finally, we combine all components and compute the overall correlation $c(N) = c(N_1)c(N_2)c(N_3)$, and retain the masks $(\gamma, n, h, \xi, \beta, m, \alpha, l)$ whose absolute correlation $|c(N)|$ is sufficiently large. Following this method, we experimentally found 18 linear approximations with absolute correlation greater than 2^{-68} , as listed in Table 2. Here, $\alpha_1, \dots, \alpha_4$ and $l_1, \dots, l_4$ are the values of α and l in their respective columns 1–4. Notably, all these approximations have the same values for the mask $\gamma, n, h, \xi, \beta, m$. Specifically, we have:

$$\begin{aligned}
\gamma &= 0, 0x81ec5a80, 0, 0, \\
n &= 0, 0x81e85a00, 0, 0, \\
h &= 0, 0x7f40, 0, 0xe1a0, \\
\xi &= 0, 0x7f40, 0, 0x8180, \\
\beta &= 0, 0x40, 0, 0, \\
m &= 0, 0x40, 0, 0.
\end{aligned}$$

It can be observed that the output mask γ takes the same value across all considered masks in this table. Consequently, for any pair of linear masks, the correlation of XOR of them is identically zero, i.e., they satisfy the condition of Theorem 1.

In addition to the above 18 linear approximations whose masks $\gamma, n, h, \xi, \beta, m$ have identical values, we also find an additional linear approximation with correlation $2^{-67.996}$, specified as follows:

$$\begin{aligned}\gamma &= 0, 0\text{x}81\text{ec}5\text{a}80, 0, 0, \\ n &= 0, 0\text{x}81\text{e}85\text{f}00, 0, 0, \\ h &= 0, 0\text{x}7\text{f}40, 0, 0\text{x}\text{e}1\text{a}0, \\ \xi &= 0, 0\text{x}7\text{f}40, 0, 0\text{x}8180, \\ \beta &= 0, 0\text{x}40, 0, 0, \\ m &= 0, 0\text{x}40, 0, 0, \\ \alpha &= 0\text{x}\text{b}0000000, 0\text{x}20041, 0\text{x}80006000, 0\text{x}2, \\ l &= 0\text{x}\text{f}0000000, 0\text{x}20061, 0\text{x}\text{c}0006000, 0\text{x}2.\end{aligned}$$

5.2 Complexity analysis

We recover the initial state after initialization by following the correlation attack procedure described in [17]. The fast correlation attack of SNOW 5G consists of two stages, i.e., a preprocessing stage where parity-check equations are generated, and a processing stage in which decoding is performed using the collected equations.

The linear approximation can be written as

$$\alpha \cdot z^{(t-1)} \oplus \beta \cdot z^{(t)} \oplus \gamma \cdot z^{(t+1)} \oplus l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)} = 0,$$

which holds with correlation ρ^2 if the most significant B bits of the initial state are guessed correctly.

Let N be the number of keystream words generated under a given key-IV pair. Since the feedback polynomials determine each LFSR state bit at any clock as a linear combination of the initial state $u \in \mathbb{F}_2^{512}$, there exists, for each $1 \leq t \leq N - 1$, a vector $\Gamma^{(t)} \in \mathbb{F}_2^{512}$ such that

$$\Gamma^{(t)} \cdot u = l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)}.$$

In the preprocessing stage, we search for two clocks t_1 and t_2 such that the lower $(512 - B)$ bits of $\Gamma^{(t_1)}$ and $\Gamma^{(t_2)}$ are equal, i.e., $\Gamma_l^{(t_1)} = \Gamma_l^{(t_2)}$, where $\Gamma_l^{(t)}$ denotes the lower $(512 - B)$ bits of $\Gamma^{(t)}$. Then, XORing the two equations yields

$$(\Gamma^{(t_1)} \oplus \Gamma^{(t_2)}) \cdot u = (\Gamma_h^{(t_1)} \oplus \Gamma_h^{(t_2)}) \cdot u_h,$$

where u_h denotes the most significant B bits of u , and the lower $512 - B$ bits cancel out. To achieve this, we compute and store all $\Gamma^{(t)} \in \mathbb{F}_2^{512}$ for $1 \leq t \leq N - 1$. This step requires $O(N)$ time and $O(N)$ memory. We then sort the vectors $\Gamma^{(t)}$ by their lower $(512 - B)$ -bit parts using the fast sorting algorithm, which takes $O(N \log N)$ time. We obtain pairs (t_1, t_2) with $\Gamma_l^{(t_1)} = \Gamma_l^{(t_2)}$, each yielding a parity-check equation that involves only the high part u_h . This results

in approximately $M = \binom{N}{2} 2^{n-B} \approx 2^{-(n-B+1)} N^2$ parity-check equations. The preprocessing stage has total time complexity $O(N + N \log N)$, and both its data and memory complexities are $O(N)$.

In the processing stage, once enough parity-check equations are collected, the highest B bits of the initial state are recovered. We can then collect parity-check equations involving only the highest B bits of the initial state as below:

$$\begin{aligned} T(\hat{u}_h) = & \alpha \cdot (z^{(t_1-1)} \oplus z^{(t_2-1)}) \oplus \beta \cdot (z^{(t_1)} \oplus z^{(t_2)}) \\ & \oplus \gamma \cdot (z^{(t_1+1)} \oplus z^{(t_2+1)}) \oplus (\Gamma_h^{(t_1)} \oplus \Gamma_h^{(t_2)}) \cdot \hat{u}_h. \end{aligned}$$

To successfully decode, according to Shannon's Theorem, $n/M < (\bar{\rho}^2)^2 / (2 \ln 2)$ should be satisfied, i.e., $N > \sqrt{2 \ln 2} \cdot B \cdot \bar{\rho}^2 2^{\frac{n-B+1}{2}}$. Based on these M parity-check equations, we compute the statistic

$$T(u_h) = \sum_{j=1}^{Mw^{-1}} \sum_{i=1}^w (-1)^{e_j(\alpha^{(i)}, \beta^{(i)}, u_h)}$$

for each candidate $u_h \in \mathbb{F}_2^B$. The value of u_h that maximizes $T(u_h)$ is taken as the correct key, denoted by $\hat{u}_h$. The statistic $T(u_h)$ can be computed using the FWT as done in [17]. We analyze the time complexity by dividing it into two cases according to the relationship between B and M . If $B > \log_2 M$, the time complexity can be calculated as $O(N + N \log_2 N + 2^B (\lceil \log_2 M \rceil + 1))$; otherwise, the time complexity is $O(N + N \log_2 N + M + B2^B)$. The memory complexity is $O(N + 2^B)$, and the data complexity is $O(N/n)$, where n denotes the number of linear approximations satisfying Theorem 1.

Table 3. Time-Memory-Data trade-offs in fast correlation attack on SNOW 5G

	$n = 2$	$n = 4$	$n = 5$	$n = 7$	$n = 9$
$\bar{\rho}^2$	$2^{-135.369}$	$2^{-135.501}$	$2^{-135.533}$	$2^{-135.576}$	$2^{-135.627}$
B	263	264	264	264	264
Time	$2^{279.25}$	$2^{279.53}$	$2^{279.59}$	$2^{279.67}$	$2^{279.78}$
Memory	$2^{265.03}$	$2^{265.14}$	$2^{265.15}$	$2^{265.18}$	$2^{265.21}$
Data	$2^{263.62}$	$2^{262.26}$	$2^{261.97}$	$2^{261.53}$	$2^{261.21}$

Considering the trade-offs among time, memory and data complexities, we evaluate various choices of n and their corresponding linear approximations, as summarized in Table 3. The time complexity achieves the lowest $2^{279.25}$ when $n = 2$. In this case, all complexities are all better than those of the existing work. The data complexity achieves the lowest $2^{261.21}$ when $n = 9$. In this case, time and memory complexities are slightly better than those of the existing work.

6 Conclusion

In this paper, we introduce a new framework of exploiting multiple linear approximations with distinct output masks called MultLA. To achieve this, we define a statistic and analyze the distribution of this statistic theoretically. By applying the MultLA framework to SNOW 5G, our attack achieves time, memory, and data complexities of $2^{279.25}$, $2^{265.03}$, and $2^{263.62}$, respectively, which are all better than the previous best results. When the time and memory complexities are slightly better than those of the existing work, the data complexity is significantly reduced from $2^{265.4}$ to $2^{261.21}$. To the best of our knowledge, our attacks achieve the best known cryptanalytic results of SNOW 5G up to now.

References

1. 3GPP. Specification of the 3GPP confidentiality and integrity algorithms UEA2 & UIA2: Document 2: SNOW 3G specification. ETSI TR 135 922 V7.0.0, 2006. Version 1.1.
2. 3GPP. Specification of the SNOW 5G based 256-bit algorithm set: Specification of the 256-NEA4 encryption, the 256-NIA4 integrity, and the 256-NCA4 authenticated encryption algorithm for 5g; document 1: Algorithm specification. Technical Report 3GPP TS 35.244 V18.0.0, 3rd Generation Partnership Project (3GPP), March 2024. Release 18.
3. Martin Ågren, Martin Hell, Thomas Johansson, and Willi Meier. Grain-128a: a new version of grain-128 with optional authentication. *Int. J. Wirel. Mob. Comput.*, 5(1):48–59, 2011.
4. Côme Berbain, Henri Gilbert, and Alexander Maximov. Cryptanalysis of grain. In Matthew J. B. Robshaw, editor, *Fast Software Encryption, 13th International Workshop, FSE 2006, Graz, Austria, March 15-17, 2006, Revised Selected Papers*, volume 4047 of *Lecture Notes in Computer Science*, pages 15–29. Springer, 2006.
5. Patrik Ekdahl, Thomas Johansson, Alexander Maximov, and Jing Yang. A new SNOW stream cipher called SNOW-V. *IACR Trans. Symmetric Cryptol.*, 2019(3):1–42, 2019.
6. Patrik Ekdahl, Thomas Johansson, Alexander Maximov, and Jing Yang. SNOW-Vi: an extreme performance variant of SNOW-V for low-end cpus. *IACR Cryptol. ePrint Arch.*, page 236, 2021.
7. Xinxin Gong and Bin Zhang. Resistance of SNOW-V against fast correlation attacks. *IACR Trans. Symmetric Cryptol.*, 2021(1):378–410, 2021.
8. Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. Multidimensional linear cryptanalysis of reduced round serpent. In Yi Mu, Willy Susilo, and Jennifer Seberry, editors, *Information Security and Privacy, 13th Australasian Conference, ACISP 2008, Wollongong, Australia, July 7-9, 2008, Proceedings*, volume 5107 of *Lecture Notes in Computer Science*, pages 203–215. Springer, 2008.
9. Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. Multidimensional extension of matsui’s algorithm 2. In Orr Dunkelman, editor, *Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers*, volume 5665 of *Lecture Notes in Computer Science*, pages 209–227. Springer, 2009.
10. Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. Multidimensional linear cryptanalysis. *J. Cryptol.*, 32(1):1–34, 2019.

11. YINUO LIU, JING YANG, and TIAN TIAN. Theoretical linear cryptanalysis of the 5G standard candidate SNOW 5G. *IACR Trans. Symmetric Cryptol.*, 2024(2):141–165, 2024.
12. SUDONG MA, CHENHUI JIN, JIE GUAN, TING CUI, and ZHEN SHI. Improved fast correlation attack using multiple linear approximations and its application on SOSEMANUK. *IEEE Trans. Inf. Theory*, 70(10):7484–7497, 2024.
13. SUDONG MA, CHENHUI JIN, ZHEN SHI, TING CUI, and JIE GUAN. Correlation attacks on SNOW-V-Like stream ciphers based on a heuristic MILP model. *IEEE Trans. Inf. Theory*, 70(6):4478–4491, 2024.
14. WILLI MEIER and OTHMAR STAFFELBACH. Fast correlation attacks on certain stream ciphers. *J. Cryptol.*, 1(3):159–176, 1989.
15. S. MURPHY. The independence of linear approximations in symmetric cryptanalysis. *IEEE Trans. Inf. Theory*, 52(12):5510–5518, 2006.
16. KAISA NYBERG and JOHAN WALLÉN. Improved linear distinguishers for SNOW 2.0. In Matthew J. B. Robshaw, editor, *Fast Software Encryption, 13th International Workshop, FSE 2006, Graz, Austria, March 15-17, 2006, Revised Selected Papers*, volume 4047 of *Lecture Notes in Computer Science*, pages 144–162. Springer, 2006.
17. ZHEN SHI, CHENHUI JIN, JIYAN ZHANG, TING CUI, LIN DING, and YU JIN. A correlation attack on full SNOW-V and SNOW-Vi. In Orr Dunkelman and Stefan Dziembowski, editors, *Advances in Cryptology - EUROCRYPT 2022 - 41st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Trondheim, Norway, May 30 - June 3, 2022, Proceedings, Part III*, volume 13277 of *Lecture Notes in Computer Science*, pages 34–56. Springer, 2022.
18. YOSUKE TODO, TAKANORI ISOBE, WILLI MEIER, KAZUMARO AOKI, and BIN ZHANG. Fast correlation attack revisited - cryptanalysis on full grain-128a, grain-128, and grain-v1. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part II*, volume 10992 of *Lecture Notes in Computer Science*, pages 129–159. Springer, 2018.
19. YOSUKE TODO, WILLI MEIER, and KAZUMARO AOKI. On the data limitation of small-state stream ciphers: Correlation attacks on fruit-80 and plantlet. In Kenneth G. Paterson and Douglas Stebila, editors, *Selected Areas in Cryptography - SAC 2019 - 26th International Conference, Waterloo, ON, Canada, August 12-16, 2019, Revised Selected Papers*, volume 11959 of *Lecture Notes in Computer Science*, pages 365–392. Springer, 2019.
20. BIN ZHANG, RUITAO LIU, XINXIN GONG, and LIN JIAO. Improved fast correlation attacks on the sosemanuk stream cipher. *IACR Trans. Symmetric Cryptol.*, 2023(4):83–111, 2023.
21. BIN ZHANG, RUITAO LIU, WILLI MEIER, SIWEI SUN, DENGGUO FENG, and WENLING WU. Vectorial fast correlation attacks. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 285–316. Springer, 2025.
22. BIN ZHANG, CHAO XU, and WILLI MEIER. Fast correlation attacks over extension fields, large-unit linear approximation and cryptanalysis of SNOW 2.0. In Rosario Gennaro and Matthew Robshaw, editors, *Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I*, volume 9215 of *Lecture Notes in Computer Science*, pages 643–662. Springer, 2015.

A Linear Approximation of Addition Modulo 2^n

The results in this appendix are based on the analysis of modular addition by Nyberg and Wallén [16]. Below, we present the theorem that is used in our paper.

Theorem 2. [16] *Let $k \geq 1$ be a fixed integer, and let $L, C \in \mathbb{F}_2^k$ be the constant vectors $L = (1, 1, \dots, 1)$ and $C = (1, 0, \dots, 0)$. Let $A_0, \dots, A_{2^{k+1}-1}$ be the $k \times k$ matrices defined by*

$$(A_r)_{d,c} = 2^{-k} \left| \left\{ x \in \mathbb{F}_2^k : u \cdot g(x, c) = x \cdot v, f(x, c) = d \right\} \right| \\ - 2^{-k} \left| \left\{ x \in \mathbb{F}_2^k : u \cdot g(x, c) \neq x \cdot v, f(x, c) = d \right\} \right|,$$

where

$$r = u2^k + \sum_{h=1}^k v_h 2^{h-1}, v = (v_1, \dots, v_k), x = (x_1, \dots, x_k), \\ c, d \in \{0, \dots, k-1\}, \\ f : \{0, \dots, k-1\}^2 \rightarrow \{0, \dots, k-1\}, f(x, c) = \lfloor (w_H(x) + c)/2 \rfloor, \\ g : \{0, \dots, k-1\}^2 \rightarrow \{0, 1\}, g(x, c) = (w_H(x) + c) \bmod 2.$$

The functions f and g are the carry and sum functions for the basic school-book method for adding k integers in binary. Let $n \geq 1$ be an integer, and let $(u; w^{(1)}, \dots, w^{(k)})$ be a linear approximation for the addition modulo 2^n with k inputs. Let $z = z_{n-1} \dots z_1 z_0$ be the binary word associated with this approximation. Then the correlation of the approximation is given by

$$\text{cor}(u; w^{(1)}, \dots, w^{(k)}) = LA_{z_{n-1}} \dots A_{z_1} A_{z_0} C.$$

For $k = 2$, the correlation can be computed as

$$\text{cor}(u; w^{(1)}, w^{(2)}) = L' A'_{z_{n-1}} \dots A'_{z_1} A'_{z_0} C'.$$

Let $H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ be the 2×2 Hadamard matrix. The vectors and matrices are expressed in the basis defined by the Hadamard matrix H_2 , i.e., $L = L' H_2^{-1}, C = H_2 C', A_{z_i} = H_2 A'_{z_i} H_2^{-1}$. This change of basis yields the following matrix $L = (1, 0), C = (1, 1)^T$, and

$$A_0 = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1 = A_2 = -A_4 = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ A_7 = \frac{1}{2} \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \quad -A_3 = A_5 = A_6 = \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Let $e_0 = (1 \ 0)$, $e_1 = (0 \ 1)$. The matrices A_i satisfy the following identities:

$$\begin{cases} e_0 A_i = 0, & i \neq 0, 7, \\ e_0 A_0 = e_0, & e_0 A_7 = e_1, \\ e_1 A_0 = e_1 A_5 = e_1 A_6 = \frac{1}{2} e_1, \\ e_1 A_3 = -\frac{1}{2} e_1, \\ e_1 A_1 = e_1 A_2 = e_1 A_7 = \frac{1}{2} e_0, \\ e_1 A_4 = -\frac{1}{2} e_0. \end{cases}$$