On the Data Limitation of Small-State Stream Ciphers: Correlation Attacks on Fruit-80 and Plantlet

Yosuke Todo[1], Willi Meier[2], and Kazumaro Aoki[1]

[1] NTT Secure Platform Laboratories, Tokyo 180-8585, Japan
[2] FHNW, Windisch, Switzerland

Abstract. Many cryptographers have focused on lightweight cryptography, and a huge number of lightweight block ciphers have been proposed. On the other hand, designing lightweight stream ciphers is a challenging task due to the well-known security criteria, i.e., the state size of stream ciphers must be at least twice the key size. The designers of Sprout addressed this issue by involving the secret key not only in the initialization but also in the keystream generation, and the state size of such stream ciphers can be smaller than twice the key size. After the seminal work, some small-state stream ciphers have been proposed such as Fruit, Plantlet, and LIZARD. Unlike conventional stream ciphers, these small-state stream ciphers have the limitation of keystream bits that can be generated from the same key and IV pair. In this paper, our motivation is to show whether the data limitation claimed by the designers is proper or not. The correlation attack is one of the attack methods exploiting many keystream bits generated from the same key and IV pair, and we apply it to Fruit-80 and Plantlet. As a result, we can break the full Fruit-80, i.e., the designers' data limitation is not sufficient. We can also recover the secret key of Plantlet if it allows about 2^{53} keystream bits from the same key and IV pair.

Keywords: Small-state stream cipher, Grain, Correlation attack

1 Introduction

Lightweight cryptography has been a hot topic in the past few years. The availability of low-area implementation is one of the most common metrics for the "lightweight," and many such block ciphers have been proposed [1,2,3,4]. On the other hand, designing lightweight stream ciphers is a challenging topic. A time-memory-data trade-off (TMDTO) attack is a powerful generic attack against stream ciphers, and the state size of stream ciphers must be at least twice of the key length to avoid the TMDTO attack [5,6,7]. It implies that designing stream ciphers whose state size is small is impossible.

In FSE 2015, Armknecht and Mikhalev tackled this issue and designed a small-state stream cipher Sprout based on the Grain structure [8]. The claimed security level is 80 bits, although the state size of Sprout is 80 bits, which is

Table 1. State size, security level, and data limitation of Sprout, Plantlet, and Fruit.

Cipher	Size of NFSR	Size of LFSR	Security level	Data Limitation
Sprout	40 bits	40 bits	80 bits	2^{40}
Plantlet	40 bits	61 bits	80 bits	2^{30}
Fruit-80	37 bits	43 bits	80 bits	2^{43}
Fruit-128	63 bits	65 bits	128 bits	2^{65}

not enough to be secure against the TMDTO attack. However, the designers of Sprout introduced a new idea, where the secret key is involved not only in the initialization but also in the keystream generation. Then, the immunity against the TMDTO attack is higher and small-state stream ciphers become possible.

Unfortunately, full Sprout was exposed to many attacks soon after its proposal [9,10,11,12]. On the other hand, the idea that the secret key is involved in the keystream generation is promising, and two new small-state stream ciphers were proposed by taking these attacks into account. Fruit is a series of new small-state stream ciphers, and the initial version denoted as Fruit-v1 was proposed in [13]. However, Fruit-v1 was also broken by the divide-and-conquer attack [14] and correlation attack [15]. The designers of Fruit then updated the version of Fruit to be secure against these attacks [16] and proposed a 128-bit security version called Fruit-128 [17]. Recently, the designers proposed Fruit-80 as the formal journal publication [18]. Plantlet is another new small-state stream cipher [19] and is conservatively designed compared with Sprout and Fruit. State sizes of Sprout and Fruit are the same as their key lengths, while the state size of Plantlet is 101 bits to achieve 80-bit security. On the other hand, Plantlet is more carefully designed such that it has high performance under the condition that the secret key is stored in non-volatile memory.

On the Data Limitation of Small-State Stream Ciphers. In this paper, our focus is the data limitation, and this part is significantly different from the original Grain ciphers. For example, Grain-v1 does not have such a data limitation, i.e., 2^{80}-bit keystream can be generated. On the other hand, the designers of small-state stream ciphers establish a limitation of keystream generated from the same key and IV pair. Table 1 summarizes the state size, security level, and the limitation of Sprout, Plantlet, Fruit-80, and Fruit-128. The data limitations of Sprout and Fruit are derived from the size of LFSR, and such a limitation is plausible from the aspect of the security because the same internal state of the LFSR is repeated when the limitation is exceeded. On the other hand, Plantlet allows to output at most 2^{30}-bit keystream. This limitation is significantly smaller than the data limitation derived from the LFSR size.

The following question is naturally raised: If small-state stream ciphers output more keystream bits, can the secret key be recovered? The designers of Fruit said that *Fruit-80 is secure against all types of key recovery attacks without any limitation on the number of keystream bits* [18]. Moreover, the authors of Plantlet

Table 2. Summary of our key-recovery attacks.

Cipher	keystream	# IV	time	data	note
Fruit-80	2^{46}	1	$2^{54.5985}$	2^{46}	recovers 0.1501 bit of the weak key.
	2^{43}	2^{21}	$2^{77.8702}$	2^{64}	recovers the full key.
Plantlet	2^{55}	1	$2^{65.9362}$	2^{55}	recovers 1 bit of the key.
	2^{53}	2^{6}	$2^{75.0990}$	2^{59}	recovers the full key.

did not provide any plausible reason about the data limitation of Plantlet [19]. To show an answer for this question, we estimate a secure size of keystream against correlation attacks.

Our Contributions. A Grain-based structure is preferred to design lightweight stream ciphers because it is comparatively lightweight and had been believed to be secure. However, in CRYPTO 2018, Grain-v1 and the stream cipher mode of Grain-128a were broken by using the fast correlation attack [20], where the authors showed that there are too many linear approximations of the Grain-based structure. This is a potential vulnerability of the Grain-based structure, but the designers of small-state stream ciphers had not cared about security against correlation attacks seriously and it should be considered more carefully than the designers expected.

The goal of the correlation attack is to recover the initial state of the LFSR. Linear approximations are constructed, and many keystream bits generated from the same key and IV pair are used to distinguish the correct initial state of the LFSR. The more keystream is generated from the same key and IV pair, the easier the correlation attack. Therefore, the correlation attack is one of useful metrics to consider the impact of the data limitation.

A small-state stream cipher is a little different from the naive Grain-based structure, and this difference makes the correlation attack more difficult. The major difference is involving a round key during the keystream generation. Therefore, the constructed linear approximations also involve the round key like in linear cryptanalysis on block ciphers. We cannot exploit data where the involved round keys are different because the sign of the correlation depends on the involved round keys. This property surely enhances the security against correlation attacks. On the other hand, interestingly, involving the round key yields a new property that is useful for attackers. The conventional correlation attack does not recover the secret key directly because its goal is to recover the initial state of the LFSR. On the correlation attack on the small-state stream cipher, we can recover the secret key directly by observing the bias direction of its empirical correlation like Matsui's Algorithm 1 [21]. Moreover, since the bias direction does not depend on the IV, we show an extended correlation attack that uses keystream generated from the same key and different IVs.

We applied the correlation attacks to Fruit-80 and Plantlet, and Table 2 summarized the attack. The conventional correlation attack using the single IV

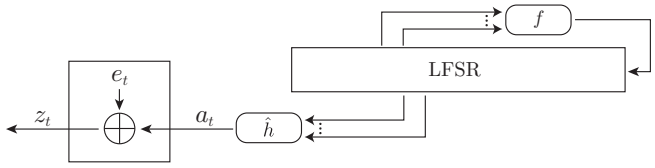

Fig. 1. Correlation attacks on Grain-based stream ciphers

can recover the secret key of Fruit-80 and Plantlet if they allow to output 2^{46}-bit and 2^{55}-bit keystream, respectively. The extended correlation attack using multiple IVs requires more data and time complexities, but it is useful to reduce the size of keystream generated from the same key and IV pair. The extended attack successfully breaks the full Fruit-80, and the secret key can be recovered with $2^{77.8702}$ time and 2^{43+21} data. Even if the extended attack is used, we cannot break the full Plantlet because it only allows to output at most 2^{30}-bit keystream. On the other hand, 2^{53}-bit keystream is enough to recover the secret key, and it is quite smaller than 2^{61} deduced by the size of the LFSR.

2 Correlation Attacks on Grain-Based Stream Ciphers

2.1 Notations

We first introduce some notations used in this paper. Let $B = \{\boldsymbol{b}_0, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{m-1}\}$ be a set of n-bit vectors. Then, $V(B) \subseteq \{0,1\}^n$ denotes a vector space spanned by B, i.e., $V(B) := \{\sum a_i \boldsymbol{b}_i : a_i \in \{0,1\}\}$.

Example 1. When B is given as $\{0100, 1101\}$, the vector space $V(B)$ is $\{0000, 0100, 1101, 1001\}$.

If all vectors in B are linearly independent, the cardinal number of $V(B)$ is 2^m, i.e., $|V(B)| = 2^m$.

2.2 Grain-Based Stream Ciphers

In this paper, we discuss the security of small-state stream ciphers, and many such ciphers adopt the so-called Grain structure. The Grain structure consists of an LFSR and NFSR, where the LFSR is updated independent of the NFSR and the NFSR is updated while involving the output of the LFSR. The keystream bit is generated as the output of a nonlinear filter function, and the domain of the filter function is made of tapping some bits from the LFSR and NFSR states.

We focus on the correlation attack [22,23] against the Grain structure. The correlation attack exploits high correlation between the initial state of the LFSR and corresponding keystream, and the goal is to recover the state of the LFSR. When we apply the correlation attack to the Grain structure, we simply regard the structure as the model described in Fig. 1. The difference from the

classical LFSR-based stream ciphers is the existence of a linear function $\hat{h}$, which is generated by linearly approximating the nonlinear filter function. Let $\{a_0, a_1, \ldots, a_{N-1}\}$ be an N-bit output sequence of $\hat{h}$. Then, an N-bit keystream $\{z_0, z_1, \ldots, z_{N-1}\}$ is computed as $z_t = a_t \oplus e_t$, where e_t is a binary noise. Let

$$f(x) = c_0 + c_1 x^1 + c_2 x^2 + \cdots + c_{n-1} x^{n-1} + x^n$$

be the feedback polynomial of the LFSR and $L^{(t)} = (\ell_t, \ell_{t+1}, \ldots, \ell_{t+n-1})$ be an n-bit internal state of the LFSR in round t. Then, the state is updated as

$$L^{(t+1)} = L^{(t)} \times F = L^{(t)} \times \begin{pmatrix} 0 \cdots 0\ 0 & c_0 \\ 1 \cdots 0\ 0 & c_1 \\ \vdots\ \ddots\ \vdots\ \vdots & \vdots \\ 0 \cdots 1\ 0 & c_{n-2} \\ 0 \cdots 0\ 1 & c_{n-1} \end{pmatrix},$$

where F is an $n \times n$ binary matrix that represents the feedback polynomial $f(x)$. In concrete Grain-based stream ciphers, the binary noise e_t is nonlinearly generated from the internal state in the LFSR and NFSR and the secret key.

2.3 Linear Approximations for Correlation Attacks

To understand the correlation attack, we first assume the simplest case, where there is $\hat{h}$ such that e_t itself is highly biased. Let p be the probability of $e_t = 1$, and the correlation c is defined as $c = 1 - 2p$. We guess the initial internal state $L^{(0)}$, calculate $\{a_0, a_1, \ldots, a_{N-1}\}$ from the guessed $L^{(0)}$ and $\hat{h}$, and evaluate $\sum_{t=0}^{N-1}(-1)^{a_t \oplus z_t}$, where the sum is computed over the set of integers. If the correct initial state is guessed, the sum is equal to $\sum_{t=0}^{N-1}(-1)^{e_t}$ and follows a normal distribution $\mathcal{N}(Nc, N)$ [3]. On the other hand, assuming that the sum behaves at random when an incorrect initial state is guessed, it follows $\mathcal{N}(0, N)$. To distinguish their distributions, we need to collect $N \approx O(1/c^2)$ bits of keystream.

Since the $\hat{h}$ function is linear, there is a corresponding linear mask Λ_h satisfying $\hat{h}(L^{(t)}) = \langle L^{(t)}, \Lambda_h \rangle$. Then, the output a_t is linearly computed as

$$a_t = \hat{h}(L^{(0)} \times F^t) = \langle L^{(0)} \times F^t, \Lambda_h \rangle = \langle L^{(0)}, \Lambda_h \times {}^{\mathrm{T}}F^t \rangle.$$

Once a high-biased $\hat{h}$ is found, the aim of attackers is to find $L^{(0)}$ such that $\sum_{t=0}^{N-1}(-1)^{z_t \oplus \langle L^{(0)}, \Lambda_h \times {}^{\mathrm{T}}F^t \rangle} = \sum_{t=0}^{N-1}(-1)^{e_t}$ is far from 0.

Modern stream ciphers are usually designed such that the binary noise e_t is balanced, but we may be able to observe a high bias by summing optimally

[3] If the correct initial state is guessed, it follows $\mathcal{N}(Nc, N - Nc^2)$. However, since N is huge and Nc^2 is small, $\mathcal{N}(Nc, N)$ is enough to approximate the distribution.

chosen binary noises. In other words, the following value

$$\begin{aligned}\bigoplus_{q\in\mathbb{T}_z} e_{t+q} &= \bigoplus_{q\in\mathbb{T}_z} a_{t+q} \oplus \bigoplus_{q\in\mathbb{T}_z} z_{t+q} \\ &= \bigoplus_{q\in\mathbb{T}_z} \langle L^{(0)}, \Lambda_{h,q} \times {}^{\mathrm{T}}F^{t+q}\rangle \oplus \bigoplus_{q\in\mathbb{T}_z} z_{t+q} \\ &= \left\langle L^{(0)}, \left(\bigoplus_{q\in\mathbb{T}_z} \left(\Lambda_{h,q} \times {}^{\mathrm{T}}F^{q}\right)\right) \times {}^{\mathrm{T}}F^{t} \right\rangle \oplus \bigoplus_{q\in\mathbb{T}_z} z_{t+q}\end{aligned}$$

could be biased. Note that the $\hat{h}$ function is generated by linearly approximating the filter function, and we do not need to use a common $\hat{h}$ function in all $q \in \mathbb{T}_z$. If a different $\hat{h}$ function is used, the corresponding linear mask is also different. Therefore, different linear masks $\Lambda_{h,q}$ can be used for each q in $\mathbb{T}_z$ in the equation above. For simplicity, we introduce Γ denoted by $\Gamma = \bigoplus_{q\in\mathbb{T}_z}(\Lambda_{h,q} \times {}^{\mathrm{T}}F^{q})$. Then, we can introduce the following parity-check equations

$$e'_t(\Gamma) = \left\langle L^{(0)}, \Gamma \times {}^{\mathrm{T}}F^{t} \right\rangle \oplus \bigoplus_{q\in\mathbb{T}_z} z_{t+q}. \tag{1}$$

We redefine p as the probability satisfying $e'_t(\Gamma) = 1$ for all possible t, and the correlation c is also redefined from the corresponding p.

2.4 Key-Recovery Algorithm Based on FWHT

The most straightforward algorithm requires the time complexity of $O(N2^n)$ to recover $L^{(0)}$. Chose et al. showed that the guess and evaluation procedure can be regarded as a Walsh-Hadamard transform [24]. The fast Walsh-Hadamard transform (FWHT) can be successfully applied to accelerate the algorithm, and it reduces the time complexity to $O(N + n2^n)$.

Definition 1 (Walsh-Hadamard Transform (WHT)). *Given a function $w : \{0,1\}^n \to \mathbb{Z}$, the WHT of w is defined as $\hat{w}(s) = \sum_{x\in\{0,1\}^n} w(x)(-1)^{\langle s,x\rangle}$.*

When $s \in \{0,1\}^n$ is guessed, the empirical correlation $\sum_{t=0}^{N-1}(-1)^{e'_t}$ is rewritten as

$$\begin{aligned}\sum_{t=0}^{N-1}(-1)^{e'_t} &= \sum_{t=0}^{N-1}(-1)^{\langle s,\Gamma\times{}^{\mathrm{T}}F^t\rangle\oplus\bigoplus_{q\in\mathbb{T}_z} z_{t+q}} \\ &= \sum_{x\in\{0,1\}^n}\left(\sum_{t\in\{0,1,\ldots,N-1|\Gamma\times{}^{\mathrm{T}}F^t=x\}}(-1)^{\langle s,x\rangle\oplus\bigoplus_{q\in\mathbb{T}_z} z_{t+q}}\right) \\ &= \sum_{x\in\{0,1\}^n}\left(\sum_{t\in\{0,1,\ldots,N-1|\Gamma\times{}^{\mathrm{T}}F^t=x\}}(-1)^{\bigoplus_{q\in\mathbb{T}_z} z_{t+q}}\right)(-1)^{\langle s,x\rangle}.\end{aligned}$$

Therefore, from the following public function w given as

$$w(x) := \sum_{t\in\{0,1,\ldots,N-1 | \Gamma\times {}^{\mathrm{T}}F^t=x\}} (-1)^{\bigoplus_{q\in\mathbb{T}_z} z_{t+q}},$$

we get $\hat{w}$ by using the FWHT, where $\hat{w}(s)$ is the empirical correlation when s is guessed.

2.5 Use of Multiple Linear Masks

In [20], Todo et al. showed that Grain-based stream ciphers have a huge number of high-biased linear masks.[4] The $\hat{h}$ function is generated by linearly approximating the filter function, and the filter function tends to have many linear approximate representations. For example, let us consider the following function

$$h(x) = x_0x_1 + x_2x_3 + x_4x_5 + x_6x_7 + x_0x_4x_8,$$

which is used in the filter function of Grain-128a [25] and Plantlet. Then, there are 2^8 linear masks Λ such that the correlation of $h(x) \oplus \langle x, \Lambda\rangle$ is $\pm 2^{-4}$. In other words, we can construct 2^8 high-biased linear masks, and each one generates a different linear mask Γ.

Assuming that there are m high-biased linear masks $(\Gamma_0, \Gamma_1, \ldots, \Gamma_{m-1})$ and letting c_i be the correlation when Γ_i is used, we compute

$$\sum_{i\in\{\{0,1,\ldots,m-1\}|c_i>0\}} (-1)^{e'_t(\Gamma_i)} - \sum_{i\in\{\{0,1,\ldots,m-1\}|c_i<0\}} (-1)^{e'_t(\Gamma_i)},$$

where $e'(\Gamma)$ is defined in Eq. (1). When we guess the initial state $L^{(0)}$, the value above follows a normal distribution $\mathcal{N}(mN\bar{c}, mN)$, where $\bar{c}$ is the average value of absolute values of c_i, i.e.,

$$\bar{c} = \frac{\sum_{i,c_i>0} c_i - \sum_{i,c_i<0} c_i}{m} = \frac{\sum_i |c_i|}{m}.$$

The key recovery algorithm based on the FWHT also works. Assuming that the data complexity (size of keystream) is N, the time complexity $O(N)$ is required to collect data. Then, we apply m high-biased linear masks for N data, and the time complexity is $O(mN)$. Finally, the FWHT is applied, and the time complexity is $O(n2^n)$. In total, the time complexity is $O(N + mN + n2^n)$.

3 Correlation Attacks on Small-State Stream Ciphers

Almost all small-state stream ciphers are based on the Grain structure, but it is modified from the original structure to avoid the time-memory-data trade-off (TMDTO) attack. Figure 2 shows the overview of Grain-based small-state

[4] Another contribution of [20] is to show the link between the parity-check equation and the multiplication over a finite field. This link is used to execute the correlation attack without guessing the whole of the initial state of the LFSR, but we do not use this technique because the size of the LFSR is small enough.

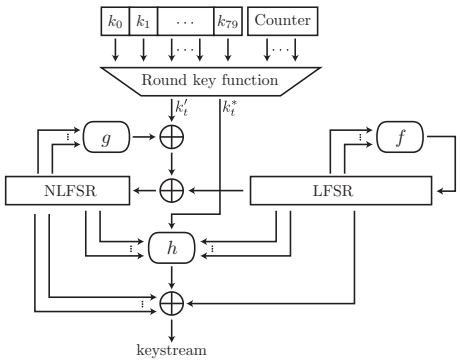

Fig. 2. Overview of Grain-based small-state stream ciphers

stream ciphers. The major difference is involving the round key in the state update function g and filter function h.

In this paper, we apply the correlation attack to Grain-based small-state stream ciphers. The basic attack strategy is the same as the correlation attack described in Sect. 2, but some different strategy is used to be optimized for the small-state stream ciphers. In this section, we summarize three major differences from the original correlation attack against the Grain structure.

3.1 Involving Round Keys

Generally, involved round keys makes correlation attacks difficult because constructed linear approximations also involve the round key. In other words, involved round keys must be constant when we collect data used in the correlation attack. The sequence of round keys usually has a small cycle to avoid degradation in efficiency. Assuming that the cycle length is ϕ, the available data decreases to N/ϕ when N-bit keystream is used. For example, in Fruit-80, both k'_t and k^*_t are generated from the secret key and a 7-bit counter. Therefore, the same pair of round keys is used every 2^7 rounds, i.e., $\phi = 2^7$. In Plantlet, only k'_t is used and $k'_t = k_{t \bmod 80} + c_t$. Since the counter c_t is public and linear, we can remove c_t efficiently. Thus, $\phi = 80$.

On the attack procedure of the correlation attack, the only difference is the interval of data sampling. Therefore, we can use the same attack strategy described in Sect. 2. Then, the empirical correlation follows $\mathcal{N}(mN\bar{c}/\phi, mN/\phi)$ when the correct initial state is guessed. Otherwise, it follows $\mathcal{N}(0, mN/\phi)$. Since m linear masks are used every ϕ rounds, the data and time complexities are N and $O(N + mN/\phi + n2^n)$.

3.2 Finding Multiple Linear Approximations with High Correlation

The reason why the Grain structure has many high-biased linear masks comes from the fact that there are many linear approximations of the filter function. Therefore, Grain-based small-state stream ciphers also inherit the property. On the other hand, finding such concrete linear masks is still difficult. A systematic method was used to find high-biased linear masks of Grain-128 and Grain-128a in [20], but the correlation found by the method is too small to attack small-state stream ciphers. Therefore, we use a more heuristic method. Recalling Eq. (1), the linear approximation exploits the sum $\bigoplus_{q\in\mathbb{T}_z} z_{t+q}$, where we exhaustively evaluate preferable $\mathbb{T}_z$. Unfortunately, only considering $\mathbb{T}_z$ is not enough because the sum $\bigoplus_{q\in\mathbb{T}_z} z_{t+q}$ always involves new bits, which are computed from the g function. Let $N_t = (n_t, n_{t+1}, \ldots, n_{t+m-1})$ be an internal state in the NFSR. Then, the new bit n_{t+m} is computed by $n_{t+m} = k'_t \oplus g_t \oplus \ell_t$, where g_t denotes the output of the g function and ℓ_t denotes the output of the LFSR in the tth round. We introduce a value b_t such that

$$b_t = n_{t+m} \oplus k'_t \oplus \ell_t \oplus g_t = 0.$$

Then, our linear approximations are constructed from $\bigoplus_{q\in\mathbb{T}_z} z_{t+q} \oplus \bigoplus_{i\in\mathbb{T}_b} b_{t+i}$.

Our task exhaustively evaluates preferable $\mathbb{T}_z$ and $\mathbb{T}_b$. To reduce the search space, we evaluate $\mathbb{T}_z$ and $\mathbb{T}_b$ such that the exploited number of rounds is small, i.e., the maximum number of values in $\mathbb{T}_z$ and $\mathbb{T}_b$ are not large.

3.3 Exploiting Keystream Generated from Different IVs

As shown in Sect. 3.1, the data must be sampled such that involved round keys are constant. While involved round keys are constant, they are still involved in the linear approximations. It implies that the bias direction depends on the involved round keys like in linear cryptanalysis on block ciphers.

We construct our linear approximations from $\bigoplus_{q\in\mathbb{T}_z} z_{t+q} \oplus \bigoplus_{i\in\mathbb{T}_b} b_{t+i}$, and the term $\bigoplus_{i\in\mathbb{T}_b} k'_{t+i}$ is included in $\bigoplus_{i\in\mathbb{T}_b} b_{t+i}$. Then, parity-check equations that we eventually construct change from Eq. (1) to

$$e'_t(\Gamma) \oplus \bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = \left\langle L^{(0)}, \Gamma \times {}^{\mathrm{T}}F^t \right\rangle \oplus \bigoplus_{q\in\mathbb{T}_z} z_{t+q}. \tag{2}$$

In other words, the bias direction is inverted if $\bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = 1$. It implies that we can easily recover the involved round keys by observing the bias direction.

Another important observation is that the bias direction is preserved unless the secret key changes. This property is useful when we consider an attack in which the size of keystream generated from the same key and IV pair is limited. Let N be the available keystream size, m be the number of linear masks, $\bar{c}$ be the average number of absolute values of correlations, and ϕ be the cycle length. Once we execute the correlation attack, the empirical correlation follows

$$\begin{cases} \mathcal{N}(mN\bar{c}/\phi, mN/\phi) & \text{for correct initial state and } \bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = 0, \\ \mathcal{N}(0, mN/\phi) & \text{for incorrect initial state,} \\ \mathcal{N}(-mN\bar{c}/\phi, mN/\phi) & \text{for correct initial state and } \bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = 1. \end{cases}$$

We introduce a threshold th such that

$$\Pr[|X| > th \mid X \sim \mathcal{N}(0, mN/\phi)] \leq 2^{-n},$$

where 2^n denotes the number of candidates of the initial state of the LFSR. We pick initial states of the LFSR whose absolute value of empirical correlation is larger than the threshold th and store only the information whether its bias direction is positive or negative. Then, one incorrect initial state remains in average, where we assume that it behaves randomly, i.e., the probability that the bias direction is positive is $1/2$. Similarly, let ϵ be the probability that the correct initial state survives, i.e.,

$$\epsilon = \Pr[X > th \mid X \sim \mathcal{N}(mN\bar{c}/\phi, mN/\phi)].$$

In other words, the bias direction leans toward positive with probability $1/2 + \epsilon$ when $\bigoplus_{i \in \mathbb{T}_b} k'_{t+i} = 0$. Similarly, it leans toward positive with probability $1/2 - \epsilon$ when $\bigoplus_{i \in \mathbb{T}_b} k'_{t+i} = 1$. Assuming that we repeat the attack procedure above over N_{iv} IVs under the fixed key, the number that the bias direction is positive follows a binomial distribution $\mathcal{B}(N_{iv}, 1/2 + \epsilon)$. Since it follows $\mathcal{B}(N_{iv}, 1/2)$ in the case of random behavior, we can distinguish the correct bias direction by using $N_{iv} = O(1/\epsilon^2)$ and recover $\bigoplus_{i \in \mathbb{T}_b} k'_{t+i}$. Since we repeat the correlation attack N_{iv} times, the data and time complexities are $N \times N_{iv}$ and $O(N_{iv} \times (N + mN/\phi + n2^n))$, respectively.

Note that this technique does not improve both time and data complexities. In other words, if we can collect enough keystream such that ϵ is almost 1, we do not need to use this technique because the naive correlation attack is always more efficient than this technique. This technique is useful only when there are data limitations about the keystream generated from the same key and IV pair.

4 Cryptanalysis on Full Fruit-80

In this section, we apply the correlation attack to Fruit-80. As shown in Sects. 2 and 3, we can estimate the data and time complexities by enumerating linear masks with high correlation and estimating the average value of correlations.

4.1 Specification of Fruit-80

The keystream generation of Fruit-80 is depicted in Fig. 2, where the sizes of NFSR and LFSR are 37 and 43 bits, respectively. Let $L^{(t)} = (\ell_t, \ell_{t+1}, \ell_{t+2}, \ldots, \ell_{t+42})$ and $N^{(t)} = (n_t, n_{t+1}, n_{t+2}, \ldots, n_{t+36})$ be the internal state of t rounds

after the initialization. Then, the state update function is defined as

$$\begin{aligned}
\ell_{t+43} &= \ell_{t+37} \oplus \ell_{t+28} \oplus \ell_{t+23} \oplus \ell_{t+18} \oplus \ell_{t+8} \oplus \ell_t,\\
n_{t+37} &= k'_t \oplus \ell_t \oplus g_t,\\
g_t &= n_t \oplus n_{t+10} \oplus n_{t+20} \oplus n_{t+12}n_{t+3} \oplus n_{t+14}n_{t+25}\\
&\quad \oplus n_{t+5}n_{t+23}n_{t+31} \oplus n_{t+8}n_{t+18} \oplus n_{t+28}n_{t+30}n_{t+32}n_{t+34},\\
z_t &= h_t \oplus \bigoplus_{j\in\mathbb{A}} n_{t+j} \oplus \ell_{t+38},\\
h_t &= k^*_t \cdot (n_{t+36} \oplus \ell_{t+19})\\
&\quad \oplus \ell_{t+6}\ell_{t+15} \oplus \ell_{t+1}\ell_{t+22} \oplus n_{t+35}\ell_{t+27} \oplus n_{t+1}n_{t+24} \oplus n_{t+1}n_{t+33}\ell_{t+42},
\end{aligned}$$

where $\mathbb{A} = \{0, 7, 19, 29, 36\}$.

The round keys k'_t and k^*_t are generated from the secret key and 7-bit counter $C^{(t)} = (c^0_t \| c^1_t \| \cdots \| c^6_t)$ as

$$\begin{aligned}
k'_t &= k_r \cdot k_{p+16} \cdot k_{q+48} \oplus k_r \cdot k_{p+16} \oplus k_{p+16} \cdot k_{q+48} \oplus k_r \cdot k_{q+48} \oplus k_{p+16},\\
k^*_t &= k_r \cdot k_{p+16} \oplus k_{p+16} \cdot k_{q+48} \oplus k_r \cdot k_{q+48} \oplus k_r \oplus k_{p+16} \oplus k_{q+48},
\end{aligned}$$

where $p = (c^1_t \| c^2_t \| c^3_t \| c^4_t \| c^5_t)$, $q = (c^2_t \| c^3_t \| c^4_t \| c^5_t \| c^6_t)$, and $r = (c^0_t \| c^1_t \| c^2_t \| c^3_t)$.

4.2 Enumerating Linear Masks with High Correlation

We exhaustively evaluated various $\mathbb{T}_z$ and $\mathbb{T}_b$ in the range that the maximum number of values in $\mathbb{T}_z$ and $\mathbb{T}_b$ are 8. As a result, $\mathbb{T}_z = \{0, 2, 3, 7\}$ and $\mathbb{T}_b = \{0, 1, 2, 6\}$ yielded the highest correlation.

Core Linear Approximate Representation. Let b_t be defined as $b_t = n_{t+37} \oplus k'_t \oplus \ell_t \oplus g_t = 0$, and let us consider the following sum of keystream.

$$\begin{aligned}
\bigoplus_{q\in\{0,2,3,7\}} z_{t+q} &= \bigoplus_{q\in\{0,2,3,7\}} z_{t+q} \oplus \bigoplus_{i\in\{0,1,2,6\}} b_{t+i}\\
&= \bigoplus_{i\in\{0,1,2,6\}} k'_{t+i} \oplus \bigoplus_{i\in\{0,1,2,6\}} \ell_{t+i} \oplus \bigoplus_{q\in\{0,2,3,7\}} \ell_{t+38+q}\\
&\quad \oplus \bigoplus_{q\in\{0,2,3,7\}} \left(h_{t+q} \oplus \bigoplus_{j\in\mathbb{A}} n_{t+q+j} \right) \oplus \bigoplus_{i\in\{0,1,2,6\}} \left(n_{t+37+i} \oplus g_{t+i} \right).
\end{aligned}$$

Since the internal state of the LFSR can be guessed in the correlation attack, $\bigoplus_{i\in\{0,1,2,6\}} \ell_{t+i} \oplus \bigoplus_{q\in\{0,2,3,7\}} \ell_{t+38+q}$ is computed. Therefore, assuming that the following Boolean function

$$g'_t = \bigoplus_{q\in\{0,2,3,7\}} \left(h_{t+q} \oplus \bigoplus_{j\in\mathbb{A}} n_{t+q+j} \right) \oplus \bigoplus_{i\in\{0,1,2,6\}} \left(n_{t+37+i} \oplus g_{t+i} \right) \tag{3}$$

is highly biased and the correlation of g'_t is c, the following approximation

$$\bigoplus_{q\in\{0,2,3,7\}} z_{t+q} \oplus \langle L^{(t)}, \Gamma_{base}\rangle = g'_t \oplus \bigoplus_{i\in\{0,1,2,6\}} k'_t \approx \bigoplus_{i\in\{0,1,2,6\}} k'_t$$

holds with the correlation c, where the linear mask Γ_{base} is defined as

$$\langle L^{(t)}, \Gamma_{base}\rangle = \bigoplus_{i\in\{0,1,2,6\}} \ell_{t+i} \oplus \bigoplus_{q\in\{0,2,3,7\}} \ell_{t+38+q}.$$

When we use the formula of Eq. (2), $e'_t(\Gamma_{base}) = g'$.

Generating Multiple Linear Approximations. Before we evaluate the correlation of g'_t, we first focus on the linear approximation of h_{t+q}, i.e., we focus on the correlation of the following function

$$\begin{aligned} h_{t+q} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle = {} & k^*_{t+q} \cdot (n_{t+q+36} \oplus \ell_{t+q+19}) \oplus \ell_{t+q+6}\ell_{t+q+15} \\ & \oplus \ell_{t+q+1}\ell_{t+q+22} \oplus n_{t+q+35}\ell_{t+q+27} \oplus n_{t+q+1}n_{t+q+24} \\ & \oplus n_{t+q+1}n_{t+q+33}\ell_{t+q+42} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle. \end{aligned}$$

When $k^*_{t+q} = 0$, six bits listed as ℓ_{t+q+1}, ℓ_{t+q+6}, ℓ_{t+q+15}, ℓ_{t+q+22}, ℓ_{t+q+27}, and ℓ_{t+q+42} are involved in h_{t+q}. Therefore, if other bits except for the six bits above are involved in $\langle L_{t+q}, \Lambda_{h,q}\rangle$, the correlation of $h_{t+q} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle$ is always 0. Therefore, $\Lambda_{h,q}$ must be chosen from the vector space $V(\boldsymbol{u}_1, \boldsymbol{u}_6, \boldsymbol{u}_{15}, \boldsymbol{u}_{22}, \boldsymbol{u}_{27}, \boldsymbol{u}_{42})$, where $\boldsymbol{u}_i$ denotes a unit vector whose $(i+1)$th element is 1 and the vector space $V(B)$ is defined in Sect. 2. When $k^*_t = 1$, $\Lambda_{h,q} \in \boldsymbol{u}_{19} + V(\boldsymbol{u}_1, \boldsymbol{u}_6, \boldsymbol{u}_{15}, \boldsymbol{u}_{22}, \boldsymbol{u}_{27}, \boldsymbol{u}_{42})$ because ℓ_{t+q+19} is linearly involved.

Recall Eq. (3), where $\bigoplus_{q\in\{0,2,3,7\}} h_{t+q}$ is used. Therefore, we introduce a linear mask Λ such that the following equation

$$\bigoplus_{q\in\{0,2,3,7\}} h_{t+q} \oplus \langle L_t, \Lambda\rangle = \bigoplus_{q\in\{0,2,3,7\}} \Big(h_{t+q} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle\Big)$$

holds. Then, Λ can take a value from the set $k^*_t\boldsymbol{u}_{19} + k^*_{t+2}\boldsymbol{u}_{21} + k^*_{t+3}\boldsymbol{u}_{22} + k^*_{t+7}\boldsymbol{u}_{26} + V(B)$, where

$$\begin{aligned} B = \{ & \boldsymbol{u}_1, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{u}_6, \boldsymbol{u}_8, \boldsymbol{u}_9, \boldsymbol{u}_{13}, \boldsymbol{u}_{15}, \boldsymbol{u}_{17}, \boldsymbol{u}_{18}, \boldsymbol{u}_{22}, \boldsymbol{u}_{24}, \boldsymbol{u}_{25}, \boldsymbol{u}_{27}, \boldsymbol{u}_{29}, \boldsymbol{u}_{30}, \boldsymbol{u}_{34}, \boldsymbol{u}_{42}, \\ & \boldsymbol{u}_{44} = \boldsymbol{u}_{38} + \boldsymbol{u}_{29} + \boldsymbol{u}_{24} + \boldsymbol{u}_{19} + \boldsymbol{u}_9 + \boldsymbol{u}_1, \\ & \boldsymbol{u}_{45} = \boldsymbol{u}_{39} + \boldsymbol{u}_{30} + \boldsymbol{u}_{25} + \boldsymbol{u}_{20} + \boldsymbol{u}_{10} + \boldsymbol{u}_2, \\ & \boldsymbol{u}_{49} = (\boldsymbol{u}_{37} + \boldsymbol{u}_{28} + \boldsymbol{u}_{23} + \boldsymbol{u}_{18} + \boldsymbol{u}_8 + \boldsymbol{u}_0) + \boldsymbol{u}_{34} + \boldsymbol{u}_{29} + \boldsymbol{u}_{24} + \boldsymbol{u}_{14} + \boldsymbol{u}_6\}. \end{aligned}$$

Since all vectors in B are linearly independent, $|V(B)| = 2^{21}$. Since the linear approximation involves k^*_t nonlinearly, the correlation depends on k^*_t. Therefore, we first assume weak keys as $(k^*_t, k^*_{t+2}, k^*_{t+3}, k^*_{t+7}) = (0, 1, 0, 0)$ because the use of this weak key yielded the highest correlation eventually. Note that this weak-key assumption can be removed in the attack procedure of the correlation attack.

We substitute $(0,1,0,0)$ for $(k_t^*, k_{t+2}^*, k_{t+3}^*, k_{t+7}^*)$, and then, Λ takes a value from the set $\boldsymbol{u}_{21} + V(B)$.

The internal state of the LFSR is guessed in the correlation attack. Namely, if $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ is biased for multiple Λ, we can construct multiple linear approximations. The following

$$\sum_{q\in\{0,2,3,7\}} z_{t+q} \oplus \langle L^{(t)}, \Gamma_{base} \rangle \oplus \langle L^{(t)}, \Lambda \rangle = \sum_{i\in\{0,1,2,6\}} k'_{t+i} \oplus g'_t \oplus \langle L^{(t)}, \Lambda \rangle \approx \sum_{i\in\{0,1,2,6\}} k'_{t+i}$$

represents linear approximations for our correlation attack, and the correlation that this approximation holds coincides with the correlation of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$. We want to evaluate correlations of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ for $\Lambda \in \boldsymbol{u}_{21} + V(B)$. To evaluate them simply, we extract independent terms from $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ as

$$\begin{aligned}
&g'_t \oplus \langle L^{(t)}, \Lambda \rangle \\
&= \ell_{t+6}\ell_{t+15} \oplus \ell_{t+6} \cdot \Lambda[6] \oplus \ell_{t+15} \cdot \Lambda[15] && (4)\\
&\quad \oplus \ell_{t+9}\ell_{t+18} \oplus \ell_{t+9} \cdot \Lambda[9] \oplus \ell_{t+18} \cdot \Lambda[18] && (5)\\
&\quad \oplus \ell_{t+3}\ell_{t+24} \oplus \ell_{t+3} \cdot \Lambda[3] \oplus \ell_{t+24} \cdot \Lambda[24] && (6)\\
&\quad \oplus \ell_{t+4}\ell_{t+25} \oplus \ell_{t+4} \cdot \Lambda[4] \oplus \ell_{t+25} \cdot \Lambda[25] && (7)\\
&\quad \oplus \ell_{t+1}\ell_{t+22} \oplus \ell_{t+13}\ell_{t+22} \oplus \ell_{t+1} \cdot \Lambda[1] \oplus \ell_{t+13} \cdot \Lambda[13] \oplus \ell_{t+22} \cdot \Lambda[22] && (8)\\
&\quad \oplus n_{t+42}\ell_{t+34} \oplus \ell_{t+34} \cdot \Lambda[34]. && (9)\\
&\quad \oplus g''_t \oplus \langle L^{(t)}, \Lambda' \rangle,
\end{aligned}$$

where $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ is the remaining term after extracting six lines. Equation (4) is independent of other terms, and each correlation is $\pm 2^{-1}$ for 2^2 linear masks $\Lambda[6,15] \in \{00, 01, 10, 11\}$. Similarly, the correlations of Eqs. (5), (6), and (7) are also $\pm 2^{-1}$ for 2^2 linear masks. In Eq. (8), the correlation is $\pm 2^{-1}$ for $\Lambda[1,13,22] \in \{000, 001, 110, 111\}$. In Eq. (9), the correlation is 2^{-1} for any $\Lambda[34]$. In total, the correlation of the above six lines is $\pm 2^{-6}$, and their signs are determined by $\Lambda[1,3,4,6,9,13,15,18,22,24,25,34]$, and the number of linear masks is $2^{2+2+2+2+2+1}$. In other words, there are 2^{11} linear masks $\Lambda[1,3,4,6,9,13,15,18,22,24,25,34]$ satisfying $g'_t \oplus \langle \Lambda, L \rangle \approx g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ with correlation $\pm 2^{-6}$.

Finally, we want to evaluate the correlation of $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$, and it is calculated by using the brute force method. Eventually, we can find 12 Λ' whose absolute values of correlations are $2^{-17.415}$ and $2^{-17.8301}$, and please refer to Appendix A in Supplementary Material in detail. Since there are 2^{11} linear masks Λ satisfying $g'_t \oplus \langle L^{(0)}, \Lambda \rangle \approx g''_t + \langle L^{(0)}, \Lambda' \rangle$ with correlation $\pm 2^{-6}$, there are 12×2^{11} linear masks Λ such that the correlations of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ are $\pm 2^{-23.415}$ and $\pm 2^{-23.8301}$.

Table 3. Success probability of correlation attack on Fruit-80.

keystream	2^{40}	2^{41}	2^{42}	2^{43}	2^{44}	2^{45}	2^{46}	2^{47}	threshold
probability	0 %	0 %	0 %	0.27 %	18.36 %	96.09 %	100.00 %	100.00 %	$th_{2^{-42}}$
	0 %	0 %	0 %	0.01 %	3.76 %	81.15 %	100.00 %	100.00 %	$th_{2^{-52}}$

4.3 Correlation Attack against Fruit-80

There are 12×2^{11} linear masks whose correlations are $\pm 2^{-23.415}$ and $\pm 2^{-23.8301}$, respectively. Thus, the attack parameter is

$$m = 24 \times 2^{11}, \quad \bar{c} = \frac{2^{-23.415} + 2^{-23.8301}}{2} = 2^{-23.6077}.$$

We assume that N keystream bits are observed. The linear approximation depends on k'_t and k^*_t, and the same (k'_t, k^*_t) is used every 2^7 rounds. Therefore, $\phi = 2^7$. As we already showed in Sect. 3, the empirical correlation follows $\mathcal{N}(mN\bar{c}/\phi, mN/\phi)$ if we guess the initial state correctly and $\sum_{i\in\{0,1,2,6\}} k'_t = 0$. When $\sum_{i\in\{0,1,2,6\}} k'_t = 1$, the bias direction is inverted, i.e., $\mathcal{N}(-mN\bar{c}/\phi, mN/\phi)$. Otherwise, we assume that the empirical correlation behaves randomly, i.e., $\mathcal{N}(0, mN/\phi)$.

We introduce a threshold th_p satisfying $\Pr[|X| > th_p \mid X \sim \mathcal{N}(0, mN/\phi)] = p$, and pick initial states whose absolute value of the empirical correlation is greater than th_p. Table 3 summarizes the probability that the correct initial state survives. To avoid all-zero initial state of the LFSR, the leftmost bit of the initial state of the LFSR is forced to 1. Therefore, the number of candidates of the initial state of the LFSR is 2^{42}. Therefore, using 2^{46} keystream with $th_{2^{-52}}$ is enough to recover the initial state of the LFSR uniquely. Then, the data and time complexities are $N = 2^{46}$ and $N + mN/\phi + n2^n = 2^{54.5985}$.

Reducing Data Complexity and Removing Weak-Key Assumption. As explained in Table 3, using 2^{43} keystream is not enough to recover the initial state of the LFSR. Even if $th_{2^{-42}}$ is used, the survival probability is $\epsilon = 0.27\%$. Besides, it assumes the use of the weak key. To enhance the success probability and remove the weak-key assumption, we exploit the technique described in Sect. 3.3.

We use 2^{43} keystream generated from the same key and IV pair. We pick initial states of the LFSR whose absolute value of empirical correlation is larger than the threshold $th_{2^{-42}}$ and store only the information whether its bias direction is positive or negative. Then, we repeat this procedure while changing IVs, and let N_{iv} be the number of repetitions. If the secret key belongs to the weak key, the number that the bias direction is positive follows a binomial distribution $\mathcal{B}(N_{iv}, 1/2 + \epsilon)$. If the secret key is not weak key, it follows $\mathcal{B}(N_{iv}, 1/2)$.

Figure 3 shows the comparison of the binomial distributions $\mathcal{B}(N_{iv}, 1/2)$ and $\mathcal{B}(N_{iv}, 1/2 \pm \epsilon)$, where $N_{iv} = 2^{21}$ and $\epsilon = 0.27\% = 2^{-1.8890}$. We can distinguish

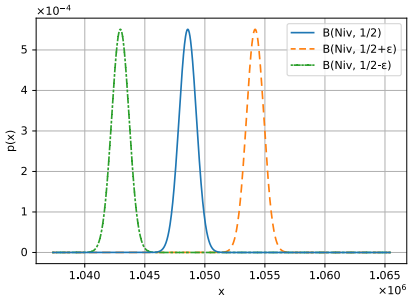

Fig. 3. Comparison of binomial distributions on Fruit-80

three binomial distributions enough. As a result, the data and time complexities are

$$N \times N_{iv} = 2^{43} \times 2^{21} = 2^{64},$$
$$N_{iv} \times (N + mN/\phi + n2^n) = 2^{21} \times (2^{43} + 24 \times 2^{11+43-7} + 42 \times 2^{42}) \approx 2^{72.67},$$

respectively.

Finally, we analyze the round-key functions, which are not balanced. The probabilities satisfying $k_t^* = 1$ and $k_t^* = 0$ are $3/4$ and $1/4$, respectively. Moreover, the probabilities satisfying $k_t' = 1$ and $k_t' = 0$ are $3/8$ and $5/8$, respectively. Therefore, the probability satisfying weak keys is $(1/4)^3 \times (3/4) = 2^{-6.4150}$. In other words, we can recover $-\log_2(15/16)$, $4-\log_2(3/8)$, $4-\log_2(5/8)$ bits of information with probabilities $(1-2^{-6.4150})$, $2^{-6.4150} \times (3/8)$, and $2^{-6.4150} \times (5/8)$, respectively. Therefore, only 0.1501 bits of information is recovered. On the other hand, exploited rounds t are restricted as $t \in \mathbb{S}_i$, where $\mathbb{S}_i := \{2^7 \times j + i \mid j = \{0, 1, \ldots, N/2^7 - 1\}\}$. We can repeat this attack procedure for $\mathbb{S}_1, \mathbb{S}_2, \ldots, \mathbb{S}_i$. By taking the trade-off with the brute-force search into account, the time complexity is optimal when 27 sets are used, i.e., $27 \times 2^{72.67} + 2^{80-0.1501\times 27} \approx 2^{77.8702}$. Note that we assume that the exhaustive search of the secret key can be immediately filtered by using the recovered round keys, and we believe that it is possible because the round key function is very simple.

5 Plantlet

5.1 Specification

Plantlet is another Grain-based small-state stream cipher and consists of a 61-bit LFSR and 40-bit NFSR. Let $L^{(t)}$ and $N^{(t)}$ be the internal state in round t after the initialization, and they are represented as

$$L^{(t)} = (\ell_t, \ell_{t+1}, \ell_{t+2}, \ldots, \ell_{t+60}),$$
$$N^{(t)} = (n_t, n_{t+1}, n_{t+2}, \ldots, n_{t+39}).$$

Then, the state update function is defined as

$$\begin{aligned}
\ell_{t+61} &= \ell_{t+54} \oplus \ell_{t+43} \oplus \ell_{t+34} \oplus \ell_{t+20} \oplus \ell_{t+14} \oplus \ell_t, \\
n_{t+40} &= k'_t \oplus \ell_t \oplus g_t, \\
g_t &= n_t \oplus n_{t+13} \oplus n_{t+19} \oplus n_{t+35} \oplus n_{t+39} \oplus n_{t+2}n_{t+25} \oplus n_{t+3}n_{t+5} \\
&\quad \oplus n_{t+7}n_{t+8} \oplus n_{t+14}n_{t+21} \oplus n_{t+16}n_{t+18} \oplus n_{t+22}n_{t+24} \oplus n_{t+26}n_{t+32} \\
&\quad \oplus n_{t+33}n_{t+36}n_{t+37}n_{t+38} \oplus n_{t+10}n_{t+11}n_{t+12} \oplus n_{t+27}n_{t+30}n_{t+31}, \\
z_t &= h_t \oplus \ell_{t+30} \oplus \bigoplus_{j \in \mathbb{A}} n_{t+j}, \\
h_t &= n_{t+4}\ell_{t+6} \oplus \ell_{t+8}\ell_{t+10} \oplus \ell_{t+32}\ell_{t+17} \oplus \ell_{t+19}\ell_{t+23} \oplus n_{t+4}\ell_{t+32}n_{t+38},
\end{aligned}$$

where $\mathbb{A} = \{1, 6, 15, 17, 23, 28, 34\}$. Moreover, the round key k'_t is defined as

$$k'_t = k_{t \bmod 80} \oplus c_t,$$

where c_t is 0 and 1 when $0 \leq (t \bmod 8) \leq 3$ and $4 \leq (t \bmod 8) \leq 7$, respectively.

5.2 Enumerating Linear Masks with High Correlation

We heuristically searched for various $\mathbb{T}_z$ and $\mathbb{T}_b$, where we restricted the number of elements in $\mathbb{T}_z$ and the maximum number of values in $\mathbb{T}_z$ and $\mathbb{T}_b$ to 2 and 13, respectively. As a result, $\mathbb{T}_z = \{0, 12\}$ and $\mathbb{T}_b = \{1, 3, 5, 7, 8, 9, 10\}$ yielded the highest correlation.

Core Linear Approximate Representation. Let b_t be defined as $b_t = n_{t+40} \oplus k'_t \oplus \ell_t \oplus g_t = 0$, and let us consider the following sum of keystream bits.

$$\begin{aligned}
\bigoplus_{q \in \{0,12\}} z_{t+q} &= \bigoplus_{q \in \{0,12\}} z_{t+q} \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} b_{t+i} \\
&= \bigoplus_{i \in \{1,3,5,7,8,9,10\}} k'_{t+i} \oplus \bigoplus_{q \in \{0,12\}} \ell_{t+30+q} \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} \ell_{t+i} \\
&\quad \oplus \bigoplus_{q \in \{0,12\}} \left(h_{t+q} \oplus \bigoplus_{j \in \mathbb{A}} n_{t+q+j} \right) \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} \left(n_{t+40+i} \oplus g_{t+i} \right).
\end{aligned}$$

Since the internal state of the LFSR can be guessed in the correlation attack, $\bigoplus_{q \in \{0,12\}} \ell_{t+30+q} \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} \ell_{t+i}$ is computed. Therefore, assuming that the following Boolean function

$$g'_t = \bigoplus_{q \in \{0,12\}} \left(h_{t+q} \oplus \bigoplus_{j \in \mathbb{A}} n_{t+q+j} \right) \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} \left(n_{t+40+i} \oplus g_{t+i} \right) \quad (10)$$

is highly biased and the correlation of g'_t is c, the following linear approximation

$$\bigoplus_{q\in\{0,12\}} z_{t+q} \oplus \langle L^{(t)}, \Gamma_{base}\rangle = \bigoplus_{i\in\{1,3,5,7,8,9,10\}} k'_{t+i} \oplus g'_t \approx \bigoplus_{i\in\{1,3,5,7,8,9,10\}} k'_{t+i}$$

holds with the correlation c, where Γ_{base} is defined as

$$\langle L^{(t)}, \Gamma_{base}\rangle = \bigoplus_{i\in\{1,3,5,7,8,9,10\}} \ell_{t+i} \oplus \bigoplus_{q\in\{0,12\}} \ell_{t+30+q}.$$

Generating Multiple Linear Approximations. We first focus on the linear approximation of h_{t+q}, i.e., we focus on the correlation of the following function

$$\begin{aligned} h_{t+q} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle = {} & n_{t+q+4}\ell_{t+q+6} \oplus \ell_{t+q+8}\ell_{t+q+10} \oplus \ell_{t+q+32}\ell_{t+q+17} \\ & \oplus \ell_{t+q+19}\ell_{t+q+23} \oplus n_{t+q+4}\ell_{t+q+32}n_{t+q+38} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle. \end{aligned}$$

Seven bits listed as ℓ_{t+q+6}, ℓ_{t+q+8}, ℓ_{t+q+10}, ℓ_{t+q+17}, ℓ_{t+q+19}, ℓ_{t+q+23}, and ℓ_{t+q+32} are involved in h_{t+q}. Therefore, $\Lambda_{h,q}$ must be chosen from the vector space $V(\boldsymbol{u}_6, \boldsymbol{u}_8, \boldsymbol{u}_{10}, \boldsymbol{u}_{17}, \boldsymbol{u}_{19}, \boldsymbol{u}_{23}, \boldsymbol{u}_{32})$, where $\boldsymbol{u}_i$ denotes a unit vector whose $(i+1)$th element is 1 and the vector space $V(B)$ is defined in Sect. 2.

Recall Eq. (10), where $\bigoplus_{q\in\{0,12\}} h_{t+q}$ is used. Therefore, we introduce a linear mask Λ such that the following equation

$$\bigoplus_{q\in\{0,12\}} h_{t+q} \oplus \langle L_t, \Lambda\rangle = \bigoplus_{q\in\{0,12\}} \Big(h_{t+q} \oplus \langle L_{t+q}, \Lambda_{h,q}\rangle\Big)$$

holds. Then, Λ can take a value from the set $V(B)$, where

$$B = \{\boldsymbol{u}_6, \boldsymbol{u}_8, \boldsymbol{u}_{10}, \boldsymbol{u}_{17}, \boldsymbol{u}_{18}, \boldsymbol{u}_{19}, \boldsymbol{u}_{20}, \boldsymbol{u}_{22}, \boldsymbol{u}_{23}, \boldsymbol{u}_{29}, \boldsymbol{u}_{31}, \boldsymbol{u}_{32}, \boldsymbol{u}_{35}, \boldsymbol{u}_{44}\}.$$

Since all vectors in B are linearly independent, $|V(B)| = 2^{14}$.

The internal state of the LFSR is guessed in the correlation attack. Namely, if $g'_t \oplus \langle L^{(t)}, \Lambda\rangle$ is biased for multiple Λ, we can construct multiple linear approximations. The following

$$\bigoplus_{q\in\{0,12\}} z_{t+q} \oplus \langle L^{(t)}, \Gamma_{base}\rangle \oplus \langle L^{(t)}, \Lambda\rangle \approx \bigoplus_{i\in\{1,3,5,7,8,9,10\}} k'_{t+i}$$

represents linear approximations for our correlation attack, and the probability that this approximation holds coincides with the correlation of $g'_t \oplus \langle L^{(t)}, \Lambda\rangle$. We want to evaluate correlations of $g'_t \oplus \langle L^{(t)}, \Lambda\rangle$ for $\Lambda \in V(B)$. To evaluate them simply, we extract independent terms from $g'_t \oplus \langle L^{(t)}, \Lambda\rangle$ as

$$\begin{aligned} & g'_t \oplus \langle L^{(t)}, \Lambda\rangle \\ & = \ell_{t+8}\ell_{t+10} \oplus \ell_{t+8} \cdot \Lambda[8] \oplus \ell_{t+10} \cdot \Lambda[10] && (11) \\ & \quad \oplus \ell_{t+19}\ell_{t+23} \oplus \ell_{t+19} \cdot \Lambda[19] \oplus \ell_{t+23} \cdot \Lambda[23] && (12) \\ & \quad \oplus \ell_{t+20}\ell_{t+22} \oplus \ell_{t+20} \cdot \Lambda[20] \oplus \ell_{t+22} \cdot \Lambda[22] && (13) \\ & \quad \oplus \ell_{t+31}\ell_{t+35} \oplus \ell_{t+31} \cdot \Lambda[31] \oplus \ell_{t+35} \cdot \Lambda[35] && (14) \\ & \quad \oplus g''_t \oplus \langle L^{(t)}, \Lambda'\rangle, \end{aligned}$$

Table 4. Success probability of correlation attack on Plantlet.

keystream	2^{50}	2^{51}	2^{52}	2^{53}	2^{54}	2^{55}	2^{56}	2^{57}	threshold
probability	0 %	0 %	0.06 %	18.17 %	99.14 %	100.00 %	100.00 %	100.00 %	$th_{2^{-60}}$
	0 %	0 %	0 %	4.93 %	94.93 %	100.00 %	100.00 %	100.00 %	$th_{2^{-70}}$

where $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ is the remaining term after extracting four lines. The correlations of Eqs. (11), (12), (13), and (14) are $\pm 2^{-1}$ for 2^2 linear masks. In total, the correlation of the above four lines is $\pm 2^{-4}$, and their signs are determined by $\Lambda[8, 10, 19, 20, 22, 23, 31, 35]$, and the number of linear masks is $2^{2+2+2+2} = 2^8$. In other words, there are 2^8 linear masks $\Lambda[8, 10, 19, 20, 22, 23, 31, 35]$ satisfying $g'_t \oplus \langle \Lambda, L \rangle \approx g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ with correlation $\pm 2^{-4}$.

Finally, we want to evaluate the correlation of $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$, and it is calculated by using the brute force method. Eventually, we can find 12 Λ' whose absolute values of correlations are $2^{-22.142}$, and please refer to Appendix B in Supplementary Material in detail. Since there are 2^8 linear masks Λ satisfying $g'_t \oplus \langle L^{(0)}, \Lambda \rangle \approx g''_t + \langle L^{(0)}, \Lambda' \rangle$ with correlation $\pm 2^{-4}$, there are 12×2^8 linear masks Λ such that the correlations of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ are $\pm 2^{-26.142}$.

5.3 Correlation Attack against Plantlet

There are 12×2^8 linear masks whose correlations are $\pm 2^{-26.142}$. Thus, the attack parameter is $m = 12 \times 2^8$ and $\bar{c} = 2^{-26.142}$. We assume that N keystream bits are observed. The linear approximation depends on $k'_t = k_{t \bmod 80} \oplus c_t$. Since the same $k_{t \bmod 80}$ is used every 80 rounds and c_t is public, $\phi = 80$.

As shown in Sect. 3, the empirical correlation follows $\mathcal{N}(mN\bar{c}/\phi, mN/\phi)$ if we guess the initial state correctly and $\bigoplus_{i \in \{1,3,5,7,8,9,10\}} k'_t = 0$. On the other hand, when $\bigoplus_{i \in \{1,3,5,7,8,9,10\}} k'_t = 1$, the bias direction is inverted, i.e., $\mathcal{N}(-mN\bar{c}/\phi, mN/\phi)$. Otherwise, we assume that the empirical correlation behaves randomly, i.e., $\mathcal{N}(0, mN/\phi)$.

We introduce th_p, which was defined in Sect. 4, and pick initial states whose absolute value of the empirical correlation is greater than th_p. Table 4 summarizes the probability that the correct initial state survives. Similarly to Fruit-80, one bit in the initial state of the LFSR is forced to 1. Therefore, the number of candidates of the initial state of the LFSR is 2^{60}. Therefore, using 2^{55} keystream with $th_{2^{-70}}$ is enough to recover the initial state of the LFSR uniquely. Then, the data and time complexities are $N = 2^{55}$ and $N + mN/\phi + n2^n = 2^{65.9362}$.

Reducing Data Complexity. Similarly to the application to Fruit-80, we exploit the technique described in Sect. 3.3. We use 2^{53} keystream generated from the same key and IV pair, and pick initial states of the LFSR whose absolute value of empirical correlation is larger than the threshold $th_{2^{-60}}$ and store the information whether its bias direction is positive or negative. We repeat this

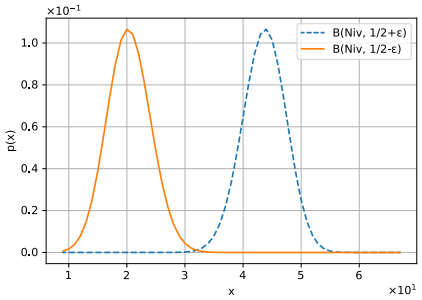

Fig. 4. Comparison of binomial distributions on Plantlet.

procedure while changing IVs, and let N_{iv} be the number of repetitions. When $\bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = 0$, the number that the bias direction is positive follows a binomial distribution $\mathcal{B}(N_{iv}, 1/2+\epsilon)$, where $\epsilon = 18.17\% = 2^{-2.4604}$. When $\bigoplus_{i\in\mathbb{T}_b} k'_{t+i} = 1$, it follows $\mathcal{B}(N_{iv}, 1/2-\epsilon)$.

Figure 4 shows the comparison of the binomial distributions when $N_{iv} = 2^6$, and we can distinguish two binomial distributions enough. As a result, the data and time complexities are

$$N \times N_{iv} = 2^{53} \times 2^6 = 2^{59},$$
$$N_{iv} \times (N + mN/\phi + n2^n) = 2^6 \times (2^{53} + 12 \times 2^{8+53}/80 + 60 \times 2^{60}) \approx 2^{71.92},$$

respectively. Unlike Fruit-80, it is very easy to analyze the time complexity to recover the secret key due to its simple round key function. Since all round keys are balanced, one procedure can recover 1 bit of information. Moreover, we can repeat this attack procedure for $\mathbb{S}_1, \mathbb{S}_2, \ldots, \mathbb{S}_i$, where $\mathbb{S}_i$ is defined in Sect. 4. By taking the trade-off with the brute-force search into account, the time complexity is optimal when 8 sets are used, i.e., $8 \times 2^{71.92} + 2^{80-8} \approx 2^{75.0990}$.

6 Conclusion

In this paper, we discussed the data limitation of keystream generated by stream ciphers using the same key and IV pair. We proposed correlation attacks for the small-state stream ciphers and applied them to two Grain-like small-state stream ciphers, Fruit-80 and Plantlet. The data limitation of Fruit-80 is derived by designers from the size of the component LFSR, but our correlation attack can successfully recover the secret key and break full Fruit-80. It implies that the claimed data limitation is not sufficient. On Plantlet, 2^{53}-bit keystream is required to recover the secret key. The data limitation is 2^{30} bits, which comes from the expectation that such a keystream length is sufficient for a current practical use. Thanks to this conservative claimed security, our correlation attack

cannot break full Plantlet, but 2^{53} is quite smaller than the data limitation derived from the size of the LFSR.

The round key is involved in the state update function or filter function in the small-state stream ciphers. When involved round keys are distinct, the absolute value of the observed correlation is the same but the bias direction could be reversed. Therefore, in this paper, we used keystream bits in which involved round keys are common. On the other hand, similar circumstances often happen in a multi-dimensional linear attack for block cipher, and a chi-squared method is successfully used to improve the attack. Thus, adopting the chi-squared method is one of the future works to improve our attacks.

Acknowledgments. The authors thank the anonymous SAC 2019 reviewers for careful reading and many helpful comments.

References

1. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: An ultra-lightweight block cipher. In Paillier, P., Verbauwhede, I., eds.: CHES 2007. Volume 4727 of LNCS., Springer, Heidelberg (September 2007) 450–466
2. Guo, J., Peyrin, T., Poschmann, A., Robshaw, M.J.B.: The LED block cipher. In Preneel, B., Takagi, T., eds.: CHES 2011. Volume 6917 of LNCS., Springer, Heidelberg (September / October 2011) 326–341
3. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY family of block ciphers and its low-latency variant MANTIS. In Robshaw, M., Katz, J., eds.: CRYPTO 2016, Part II. Volume 9815 of LNCS., Springer, Heidelberg (August 2016) 123–153
4. Banik, S., Pandey, S.K., Peyrin, T., Sasaki, Y., Sim, S.M., Todo, Y.: GIFT: A small present - towards reaching the limit of lightweight encryption. In Fischer, W., Homma, N., eds.: CHES 2017. Volume 10529 of LNCS., Springer, Heidelberg (September 2017) 321–345
5. Babbage, S.H.: Improved "exhaustive search" attacks on stream ciphers. In: European Convention on Security and Detection, 1995. (May 1995) 161–166
6. Golic, J.D.: Cryptanalysis of alleged A5 stream cipher. In Fumy, W., ed.: EUROCRYPT'97. Volume 1233 of LNCS., Springer, Heidelberg (May 1997) 239–255
7. Biryukov, A., Shamir, A.: Cryptanalytic time/memory/data tradeoffs for stream ciphers. In Okamoto, T., ed.: ASIACRYPT 2000. Volume 1976 of LNCS., Springer, Heidelberg (December 2000) 1–13
8. Armknecht, F., Mikhalev, V.: On lightweight stream ciphers with shorter internal states. In Leander, G., ed.: FSE 2015. Volume 9054 of LNCS., Springer, Heidelberg (March 2015) 451–470
9. Lallemand, V., Naya-Plasencia, M.: Cryptanalysis of full Sprout. In Gennaro, R., Robshaw, M.J.B., eds.: CRYPTO 2015, Part I. Volume 9215 of LNCS., Springer, Heidelberg (August 2015) 663–682
10. Esgin, M.F., Kara, O.: Practical cryptanalysis of full Sprout with TMD tradeoff attacks. In Dunkelman, O., Keliher, L., eds.: SAC 2015. Volume 9566 of LNCS., Springer, Heidelberg (August 2016) 67–85

11. Banik, S.: Some results on Sprout. In Biryukov, A., Goyal, V., eds.: INDOCRYPT 2015. Volume 9462 of LNCS., Springer, Heidelberg (December 2015) 124–139
12. Zhang, B., Gong, X.: Another tradeoff attack on Sprout-like stream ciphers. In Iwata, T., Cheon, J.H., eds.: ASIACRYPT 2015, Part II. Volume 9453 of LNCS., Springer, Heidelberg (November / December 2015) 561–585
13. Ghafari, V.A., Hu, H., Xie, C.: Fruit: Ultra-lightweight stream cipher with shorter internal state. Cryptology ePrint Archive, Report 2016/355 (2016) `http://eprint.iacr.org/2016/355`.
14. Dey, S., Sarkar, S.: Cryptanalysis of full round Fruit. Cryptology ePrint Archive, Report 2017/087 (2017) `http://eprint.iacr.org/2017/087`.
15. Zhang, B., Gong, X., Meier, W.: Fast correlation attacks on Grain-like small state stream ciphers. IACR Trans. Symm. Cryptol. **2017**(4) (2017) 58–81
16. Vahid Amin Ghafari, Honggang Hu, Y.C.: Fruit-v2: Ultra-lightweight stream cipher with shorter internal state. IACR Cryptology ePrint Archive **2016** (2016) 355
17. Ghafari, V.A., Hu, H., Alizadeh, M.: Necessary conditions for designing secure stream ciphers with the minimal internal states. Cryptology ePrint Archive, Report 2017/765 (2017) `http://eprint.iacr.org/2017/765`.
18. Ghafari, V.A., Hu, H.: Fruit-80: A secure ultra-lightweight stream cipher for constrained environments. Entropy **20**(3) (2018) 180
19. Mikhalev, V., Armknecht, F., Müller, C.: On ciphers that continuously access the non-volatile key. IACR Trans. Symm. Cryptol. **2016**(2) (2016) 52–79 `http://tosc.iacr.org/index.php/ToSC/article/view/565`.
20. Todo, Y., Isobe, T., Meier, W., Aoki, K., Zhang, B.: Fast correlation attack revisited - cryptanalysis on full Grain-128a, Grain-128, and Grain-v1. In Shacham, H., Boldyreva, A., eds.: CRYPTO 2018, Part II. Volume 10992 of LNCS., Springer, Heidelberg (August 2018) 129–159
21. Matsui, M.: On correlation between the order of S-boxes and the strength of DES. In Santis, A.D., ed.: EUROCRYPT'94. Volume 950 of LNCS., Springer, Heidelberg (May 1995) 366–375
22. Siegenthaler, T.: Correlation-immunity of nonlinear combining functions for cryptographic applications. IEEE Trans. Information Theory **30**(5) (1984) 776–780
23. Meier, W., Staffelbach, O.: Fast correlation attacks on certain stream ciphers. J. Cryptology **1**(3) (1989) 159–176
24. Chose, P., Joux, A., Mitton, M.: Fast correlation attacks: An algorithmic point of view. In Knudsen, L.R., ed.: EUROCRYPT 2002. Volume 2332 of LNCS., Springer (2002) 209–221
25. Ågren, M., Hell, M., Johansson, T., Meier, W.: Grain-128a: a new version of Grain-128 with optional authentication. IJWMC **5**(1) (2011) 48–59

A Correlation of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ on Fruit-80

In this section, we show the detailed method to evaluate the correlation of $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$. As we already showed in Sect. 4, we first extract independent terms from $g'_t \oplus \langle L^{(t)}, \Lambda \rangle$ as

$$\begin{aligned}
&g'_t \oplus \langle L^{(t)}, \Lambda \rangle \\
&= \ell_{t+6}\ell_{t+15} \oplus \ell_{t+6} \cdot \Lambda[6] \oplus \ell_{t+15} \cdot \Lambda[15] \\
&\quad \oplus \ell_{t+9}\ell_{t+18} \oplus \ell_{t+9} \cdot \Lambda[9] \oplus \ell_{t+18} \cdot \Lambda[18] \\
&\quad \oplus \ell_{t+3}\ell_{t+24} \oplus \ell_{t+3} \cdot \Lambda[3] \oplus \ell_{t+24} \cdot \Lambda[24] \\
&\quad \oplus \ell_{t+4}\ell_{t+25} \oplus \ell_{t+4} \cdot \Lambda[4] \oplus \ell_{t+25} \cdot \Lambda[25] \\
&\quad \oplus \ell_{t+1}\ell_{t+22} \oplus \ell_{t+13}\ell_{t+22} \oplus \ell_{t+1} \cdot \Lambda[1] \oplus \ell_{t+13} \cdot \Lambda[13] \oplus \ell_{t+22} \cdot \Lambda[22] \\
&\quad \oplus n_{t+42}\ell_{t+34} \oplus \ell_{t+34} \cdot \Lambda[34]. \\
&\quad \oplus g''_t \oplus \langle L^{(t)}, \Lambda' \rangle,
\end{aligned}$$

where $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ is the remaining term after extracting the first six lines. Then, there are 2^{11} linear masks $\Lambda[1, 3, 4, 6, 9, 13, 15, 18, 22, 24, 25, 34]$ satisfying $g'_t \oplus \langle \Lambda, L \rangle \approx g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ with correlation $\pm 2^{-6}$.

Our next goal is to evaluate the correlation of $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$, which is described as

$$\begin{aligned}
&g''_t \oplus \langle L^{(t)}, \Lambda' \rangle \\
&= n_{t+38} \oplus \ell_{t+21} \\
&\quad \oplus n_{t+35}\ell_{t+27} \oplus n_{t+1}n_{t+24} \oplus n_{t+1}n_{t+33}\ell_{t+42} \\
&\quad \oplus \ell_{t+8}\ell_{t+17} \oplus n_{t+37}\ell_{t+29} \oplus n_{t+3}n_{t+26} \oplus n_{t+3}n_{t+35}\ell_{t+44} \\
&\quad \oplus n_{t+38}\ell_{t+30} \oplus n_{t+4}n_{t+27} \oplus n_{t+4}n_{t+36}\ell_{t+45} \\
&\quad \oplus \ell_{t+8}\ell_{t+29} \oplus n_{t+8}n_{t+31} \oplus n_{t+8}n_{t+40}\ell_{t+49} \\
&\quad \oplus \bigoplus_{q \in \{0,2,3,7\}} \left(\bigoplus_{j \in \mathbb{A}} n_{t+q+j} \right) \oplus \bigoplus_{i \in \{0,1,2,6\}} \left(n_{t+37+i} \oplus g_{t+i} \right) \oplus \langle L^{(t)}, \Lambda' \rangle,
\end{aligned}$$

where

$$\begin{aligned}
\langle L^{(t)}, \Lambda' \rangle = (&\ell_{t+21} \oplus \Lambda'[8] \cdot \ell_{t+8} \oplus \Lambda'[17] \cdot \ell_{t+17} \oplus \Lambda'[27] \cdot \ell_{t+27} \oplus \Lambda'[29] \cdot \ell_{t+29} \\
&\oplus \Lambda'[30] \cdot \ell_{t+30} \oplus \Lambda'[42] \cdot \ell_{t+42} \oplus \Lambda'[44] \cdot \ell_{t+44} \oplus \Lambda'[45] \cdot \ell_{t+45} \oplus \Lambda'[49] \cdot \ell_{t+49}).
\end{aligned}$$

Here, the indices 44, 45, and 49 exceeds the length of Λ, i.e., 43. Therefore, $\Lambda'[44, 45, 49]$ are computed by using the feedback function f as

$$\begin{aligned}
\Lambda'[44] &= \Lambda'[38] \oplus \Lambda'[29] \oplus \Lambda'[24] \oplus \Lambda'[19] \oplus \Lambda'[9] \oplus \Lambda'[1], \\
\Lambda'[45] &= \Lambda'[39] \oplus \Lambda'[30] \oplus \Lambda'[25] \oplus \Lambda'[20] \oplus \Lambda'[10] \oplus \Lambda'[2] \\
\Lambda'[49] &= (\Lambda'[37] \oplus \Lambda'[28] \oplus \Lambda'[23] \oplus \Lambda'[18] \oplus \Lambda'[8] \oplus \Lambda'[t]) \\
&\quad \oplus \Lambda'[34] \oplus \Lambda'[29] \oplus \Lambda'[24] \oplus \Lambda'[14] \oplus \Lambda'[6].
\end{aligned}$$

We expand all terms in $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle$ as

$$
\begin{aligned}
& g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle \\
&= n_{t+38} \oplus \cancel{\ell_{t+21}} \\
&\quad \oplus n_{t+35}\ell_{t+27} \oplus n_{t+1}n_{t+24} \oplus n_{t+1}n_{t+33}\ell_{t+42} \\
&\quad \oplus \ell_{t+8}\ell_{t+17} \oplus n_{t+37}\ell_{t+29} \oplus n_{t+3}n_{t+26} \oplus n_{t+3}n_{t+35}\ell_{t+44} \\
&\quad \oplus n_{t+38}\ell_{t+30} \oplus n_{t+4}n_{t+27} \oplus n_{t+4}n_{t+36}\ell_{t+45} \\
&\quad \oplus \ell_{t+8}\ell_{t+29} \oplus n_{t+8}n_{t+31} \oplus n_{t+8}n_{t+40}\ell_{t+49} \\
&\quad \oplus \cancel{n_t} \oplus \cancel{n_{t+7}} \oplus n_{t+19} \oplus n_{t+29} \oplus \cancel{n_{t+36}} \\
&\quad \oplus \cancel{n_{t+2}} \oplus n_{t+9} \oplus \cancel{n_{t+21}} \oplus n_{t+31} \oplus \cancel{n_{t+38}} \\
&\quad \oplus n_{t+3} \oplus \cancel{n_{t+10}} \oplus \cancel{n_{t+22}} \oplus n_{t+32} \oplus \cancel{n_{t+39}} \\
&\quad \oplus \cancel{n_{t+7}} \oplus n_{t+14} \oplus \cancel{n_{t+26}} \oplus \cancel{n_{t+36}} \oplus \cancel{n_{t+43}} \\
&\quad \oplus n_{t+37} \oplus (\cancel{n_t} \oplus \cancel{n_{t+10}} \oplus n_{t+20} \oplus n_{t+12}n_{t+3} \oplus n_{t+14}n_{t+25} \\
&\qquad \oplus n_{t+5}n_{t+23}n_{t+31} \oplus n_{t+8}n_{t+18} \oplus n_{t+28}n_{t+30}n_{t+32}n_{t+34}) \\
&\quad \oplus \cancel{n_{t+38}} \oplus (n_{t+1} \oplus n_{t+11} \oplus \cancel{n_{t+21}} \oplus n_{t+13}n_{t+4} \oplus n_{t+15}n_{t+26} \\
&\qquad \oplus n_{t+6}n_{t+24}n_{t+32} \oplus n_{t+9}n_{t+19} \oplus n_{t+29}n_{t+31}n_{t+33}n_{t+35}) \\
&\quad \oplus \cancel{n_{t+39}} \oplus (\cancel{n_{t+2}} \oplus n_{t+12} \oplus \cancel{n_{t+22}} \oplus n_{t+14}n_{t+5} \oplus n_{t+16}n_{t+27} \\
&\qquad \oplus n_{t+7}n_{t+25}n_{t+33} \oplus n_{t+10}n_{t+20} \oplus n_{t+30}n_{t+32}n_{t+34}n_{t+36}) \\
&\quad \oplus \cancel{n_{t+43}} \oplus (n_{t+6} \oplus n_{t+16} \oplus \cancel{n_{t+26}} \oplus n_{t+18}n_{t+9} \oplus n_{t+20}n_{t+31} \\
&\qquad \oplus n_{t+11}n_{t+29}n_{t+37} \oplus n_{t+14}n_{t+24} \oplus n_{t+34}n_{t+36}n_{t+38}n_{t+40}) \\
&\quad \oplus (\cancel{\ell_{t+21}} \oplus \Lambda'[8] \cdot \ell_{t+8} \oplus \Lambda'[17] \cdot \ell_{t+17} \oplus \Lambda'[27] \cdot \ell_{t+27} \oplus \Lambda'[29] \cdot \ell_{t+29} \\
&\qquad \oplus \Lambda'[30] \cdot \ell_{t+30} \oplus \Lambda'[42] \cdot \ell_{t+42} \oplus \Lambda'[44] \cdot \ell_{t+44} \oplus \Lambda'[45] \cdot \ell_{t+45} \oplus \Lambda'[49] \cdot \ell_{t+49}).
\end{aligned}
$$

There are 35 bits in the NFSR and 9 bits in the LFSR in $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle$, and the size of involved bits is too large to evaluate the correlation with brute force. Therefore, we decompose this Boolean function into six Boolean functions G_1, G_2, G_3, G_4, G_5, and G_6, i.e., $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle = G_1 \oplus G_2 \oplus G_3 \oplus G_4 \oplus G_5 \oplus G_6$.

$$
\begin{aligned}
G_1 &= n_{t+20} \oplus n_{t+31} \oplus n_{t+10}n_{t+20} \oplus n_{t+20}n_{t+31}, \\
G_2 &= n_{t+1} \oplus n_{t+1}n_{t+24} \oplus n_{t+1}n_{t+33}\ell_{t+42} \oplus \Lambda'[42] \cdot \ell_{t+42}, \\
G_3 &= n_{t+14} \oplus n_{t+14}n_{t+25} \oplus n_{t+14}n_{t+5} \oplus n_{t+14}n_{t+24} \oplus n_{t+5}n_{t+23}n_{t+31} \oplus n_{t+7}n_{t+25}n_{t+33}, \\
G_4 &= n_{t+16} \oplus n_{t+4}n_{t+27} \oplus n_{t+13}n_{t+4} \oplus n_{t+16}n_{t+27} \oplus n_{t+4}n_{t+36}\ell_{t+45} \oplus \Lambda'[45] \cdot \ell_{t+45}, \\
G_5 &= n_{t+6} \oplus n_{t+32} \oplus n_{t+38} \oplus n_{t+9} \oplus n_{t+19} \oplus n_{t+18}n_{t+9} \oplus n_{t+9}n_{t+19} \oplus n_{t+38}\ell_{t+30} \\
&\quad \oplus n_{t+8}n_{t+18} \oplus n_{t+8}n_{t+40}\ell_{t+49} \oplus n_{t+6}n_{t+24}n_{t+32} \oplus n_{t+8}n_{t+31} \\
&\quad \oplus n_{t+28}n_{t+30}n_{t+32}n_{t+34} \oplus n_{t+30}n_{t+32}n_{t+34}n_{t+36} \oplus n_{t+34}n_{t+36}n_{t+38}n_{t+40} \\
&\quad \oplus \Lambda'[30] \cdot \ell_{t+30} \oplus \Lambda'[49] \cdot \ell_{t+49}, \\
G_6 &= n_{t+3} \oplus n_{t+11} \oplus n_{t+12} \oplus n_{t+29} \oplus n_{t+37} \oplus \ell_{t+8}\ell_{t+17} \oplus \ell_{t+8}\ell_{t+29} \oplus n_{t+37}\ell_{t+29} \\
&\quad \oplus n_{t+35}\ell_{t+27} \oplus n_{t+12}n_{t+3} \oplus n_{t+3}n_{t+26} \oplus n_{t+15}n_{t+26} \oplus n_{t+3}n_{t+35}\ell_{t+44} \\
&\quad \oplus n_{t+11}n_{t+29}n_{t+37} \oplus n_{t+29}n_{t+31}n_{t+33}n_{t+35} \oplus \Lambda'[29] \cdot \ell_{t+29} \oplus \Lambda'[27] \cdot \ell_{t+27} \\
&\quad \oplus \Lambda'[8] \cdot \ell_{t+8} \oplus \Lambda'[17] \cdot \ell_{t+17} \oplus \Lambda'[44] \cdot \ell_{t+44}.
\end{aligned}
$$

Six Boolean functions G_1, G_2, G_3, G_4, G_5, and G_6 involve 3, 5, 8, 7, 18, and 20 bits, respectively. These involved bits are independent except for n_{t+24}, n_{t+31}, n_{t+33}, and n_{t+36}, where these four bits are colored by red. Therefore, we compute the conditional correlations of G_1, G_2, G_3, G_4, G_5, and G_6.

Definition 2 (Conditional correlation). *Let G be a Boolean function from n bits to 1 bit, and let x be the input of G. We add a condition for bits $x_i \in \mathbb{I}$, and these bits are fixed to v_i. Then, the conditional correlation of G is defined as*

$$\sum_{x \in \{\{0,1\}^n, x_i = v_i \text{ for all } x_i \in \mathbb{I}\}} (-1)^{G(x)}.$$

We add conditions for four bits n_{t+24}, n_{t+31}, n_{t+33}, and n_{t+36}. Then, we compute the conditional correlations of the six Boolean functions, and then, compute the conditional correlation of G by using the piling-up lemma. Finally, the correlation of G is computed by summing conditional correlations of G over all conditions.

Table 5. Case that $\Lambda'[8, 17, 27, 29, 30, 42, 44, 45, 49] = 000100000$.

n_{t+24}	n_{t+31}	n_{t+33}	n_{t+36}	G_1	G_2	G_3	G_4	G_5	G_6	correlation
0	0	0	0	2^{-1}	0	2^{-1}	2^{-2}	0	-2^{-6}	0
0	0	0	1	2^{-1}	0	2^{-1}	2^{-2}	0	-2^{-6}	0
0	0	1	0	2^{-1}	2^{-1}	2^{-2}	2^{-2}	0	-2^{-6}	0
0	0	1	1	2^{-1}	2^{-1}	2^{-2}	2^{-2}	0	-2^{-6}	0
0	1	0	0	-2^{-1}	0	2^{-2}	2^{-2}	0	-2^{-6}	0
0	1	0	1	-2^{-1}	0	2^{-2}	2^{-2}	0	-2^{-6}	0
0	1	1	0	-2^{-1}	2^{-1}	0	2^{-2}	0	-2^{-6}	0
0	1	1	1	-2^{-1}	2^{-1}	0	2^{-2}	0	-2^{-6}	0
1	0	0	0	2^{-1}	1	2^{-1}	2^{-2}	$2^{-5.415}$	-2^{-6}	$-2^{-15.415} \times 2^{-4}$
1	0	0	1	2^{-1}	1	2^{-1}	2^{-2}	$2^{-5.415}$	-2^{-6}	$-2^{-15.415} \times 2^{-4}$
1	0	1	0	2^{-1}	2^{-1}	2^{-2}	2^{-2}	$2^{-5.415}$	-2^{-6}	$-2^{-17.415} \times 2^{-4}$
1	0	1	1	2^{-1}	2^{-1}	2^{-2}	2^{-2}	$2^{-5.415}$	-2^{-6}	$-2^{-17.415} \times 2^{-4}$
1	1	0	0	-2^{-1}	1	2^{-2}	2^{-2}	$-2^{-5.415}$	-2^{-6}	$-2^{-16.415} \times 2^{-4}$
1	1	0	1	-2^{-1}	1	2^{-2}	2^{-2}	$-2^{-5.415}$	-2^{-6}	$-2^{-16.415} \times 2^{-4}$
1	1	1	0	-2^{-1}	2^{-1}	2^{-2}	2^{-2}	$-2^{-5.415}$	-2^{-6}	$-2^{-17.415} \times 2^{-4}$
1	1	1	1	-2^{-1}	2^{-1}	2^{-2}	2^{-2}	$-2^{-5.415}$	-2^{-6}	$-2^{-17.415} \times 2^{-4}$
sum										$-2^{-17.415}$

Table 5 shows the correlation of G when $\Lambda'[8, 17, 27, 29, 30, 42, 44, 45, 49] = 000100000$. Here, note that each conditional correlation must be divided by 2^4 because we add 4-bit condition. Finally, Table 6 summarizes each correlation,

where we picked the case whose absolute values of correlation are greater than 2^{-18}.

Table 6. Correlations of $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle$.

$\Lambda'[8]$	$\Lambda'[17]$	$\Lambda'[27]$	$\Lambda'[29]$	$\Lambda'[30]$	$\Lambda'[42]$	$\Lambda'[44]$	$\Lambda'[45]$	$\Lambda'[49]$	correlation
0	0	0	1	0	0	0	0	0	$-2^{-17.4150}$
0	0	0	1	0	0	0	0	1	$-2^{-17.4150}$
0	0	0	1	1	0	0	0	0	$2^{-17.4150}$
0	0	1	1	0	0	1	0	0	$-2^{-17.8301}$
0	0	1	1	0	0	1	0	1	$-2^{-17.8301}$
0	0	1	1	1	0	1	0	0	$2^{-17.8301}$
0	1	0	0	0	0	0	0	0	$-2^{-17.4150}$
0	1	0	0	0	0	0	0	1	$-2^{-17.4150}$
0	1	0	0	1	0	0	0	0	$2^{-17.4150}$
0	1	1	0	0	0	1	0	0	$-2^{-17.8301}$
0	1	1	0	0	0	1	0	1	$-2^{-17.8301}$
0	1	1	0	1	0	1	0	0	$2^{-17.8301}$
1	0	0	1	0	0	0	0	0	$-2^{-17.4150}$
1	0	0	1	0	0	0	0	1	$-2^{-17.4150}$
1	0	0	1	1	0	0	0	0	$2^{-17.4150}$
1	0	1	1	0	0	1	0	0	$-2^{-17.8301}$
1	0	1	1	0	0	1	0	1	$-2^{-17.8301}$
1	0	1	1	1	0	1	0	0	$2^{-17.8301}$
1	1	0	0	0	0	0	0	0	$2^{-17.4150}$
1	1	0	0	0	0	0	0	1	$2^{-17.4150}$
1	1	0	0	1	0	0	0	0	$-2^{-17.4150}$
1	1	1	0	0	0	1	0	0	$2^{-17.8301}$
1	1	1	0	0	0	1	0	1	$2^{-17.8301}$
1	1	1	0	1	0	1	0	0	$-2^{-17.8301}$

B Correlation of $g_t'' + \langle L^{(t)}, \Lambda' \rangle$ of Plantlet

Similarly to the case of Fruit-80, we compute the correlation of $g_t'' + \langle L^{(t)}, \Lambda' \rangle$ of Plantlet. After extracting independent terms from $g_t' \oplus \langle L^{(t)}, \Lambda \rangle$, $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle$ is described as

$$\begin{aligned} g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle = {} & n_{t+4}\ell_{t+6} \oplus \ell_{t+32}\ell_{t+17} \oplus n_{t+4}\ell_{t+32}n_{t+38} \\ & \oplus n_{t+16}\ell_{t+18} \oplus \ell_{t+44}\ell_{t+29} \oplus n_{t+16}\ell_{t+44}n_{t+50} \\ & \oplus \bigoplus_{j \in \mathbb{A}} n_{t+j} \oplus \bigoplus_{j \in \mathbb{A}} n_{t+12+j} \\ & \oplus \bigoplus_{i \in \{1,3,5,7,8,9,10\}} \left(n_{t+40+i} \oplus g_{t+i} \right) \oplus \langle L^{(t)}, \Lambda' \rangle, \end{aligned}$$

where

$$\langle L^{(t)}, \Lambda' \rangle = (\Lambda'[6]\ell_{t+6} \oplus \Lambda'[17]\ell_{t+17} \oplus \Lambda'[18]\ell_{t+18} \oplus \Lambda'[29]\ell_{t+29} \oplus \Lambda'[32]\ell_{t+32} \oplus \Lambda'[44]\ell_{t+44}).$$

Now, let us expand all terms in $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$ as

$$\begin{aligned}
&g''_t \oplus \langle L^{(t)}, \Lambda' \rangle \\
&= n_{t+4}\ell_{t+6} \oplus \ell_{t+32}\ell_{t+17} \oplus n_{t+4}\ell_{t+32}n_{t+38} \\
&\quad \oplus n_{t+16}\ell_{t+18} \oplus \ell_{t+44}\ell_{t+29} \oplus n_{t+16}\ell_{t+44}n_{t+50} \\
&\quad \oplus \cancel{n_{t+1}} \oplus n_{t+6} \oplus n_{t+15} \oplus n_{t+17} \oplus \cancel{n_{t+23}} \oplus \cancel{n_{t+28}} \oplus n_{t+34} \\
&\quad \oplus n_{t+13} \oplus \cancel{n_{t+18}} \oplus \cancel{n_{t+27}} \oplus \cancel{n_{t+29}} \oplus n_{t+35} \oplus \cancel{n_{t+40}} \oplus \cancel{n_{t+46}} \\
&\quad \oplus n_{t+41} \oplus (\cancel{n_{t+1}} \oplus n_{t+14} \oplus \cancel{n_{t+20}} \oplus n_{t+36} \oplus \cancel{n_{t+40}} \oplus n_{t+3}n_{t+26} \oplus n_{t+4}n_{t+6} \\
&\qquad \oplus n_{t+8}n_{t+9} \oplus n_{t+15}n_{t+22} \oplus n_{t+17}n_{t+19} \oplus \cancel{n_{t+23}n_{t+25}} \oplus n_{t+27}n_{t+33} \\
&\qquad \oplus n_{t+34}n_{t+37}n_{t+38}n_{t+39} \oplus n_{t+11}n_{t+12}n_{t+13} \oplus n_{t+28}n_{t+31}n_{t+32}) \\
&\quad \oplus \cancel{n_{t+43}} \oplus (n_{t+3} \oplus n_{t+16} \oplus \cancel{n_{t+22}} \oplus n_{t+38} \oplus \cancel{n_{t+42}} \oplus n_{t+5}n_{t+28} \oplus n_{t+6}n_{t+8} \\
&\qquad \oplus n_{t+10}n_{t+11} \oplus n_{t+17}n_{t+24} \oplus n_{t+19}n_{t+21} \oplus \cancel{n_{t+25}n_{t+27}} \oplus n_{t+29}n_{t+35} \\
&\qquad \oplus n_{t+36}n_{t+39}n_{t+40}n_{t+41} \oplus n_{t+13}n_{t+14}n_{t+15} \oplus n_{t+30}n_{t+33}n_{t+34}) \\
&\quad \oplus \cancel{n_{t+45}} \oplus (n_{t+5} \oplus \cancel{n_{t+18}} \oplus n_{t+24} \oplus n_{t+40} \oplus \cancel{n_{t+44}} \oplus n_{t+7}n_{t+30} \oplus n_{t+8}n_{t+10} \\
&\qquad \oplus n_{t+12}n_{t+13} \oplus n_{t+19}n_{t+26} \oplus n_{t+21}n_{t+23} \oplus n_{t+27}n_{t+29} \oplus n_{t+31}n_{t+37} \\
&\qquad \oplus n_{t+38}n_{t+41}n_{t+42}n_{t+43} \oplus n_{t+15}n_{t+16}n_{t+17} \oplus n_{t+32}n_{t+35}n_{t+36}) \\
&\quad \oplus \cancel{n_{t+47}} \oplus (n_{t+7} \oplus \cancel{n_{t+20}} \oplus n_{t+26} \oplus \cancel{n_{t+42}} \oplus \cancel{n_{t+46}} \oplus n_{t+9}n_{t+32} \oplus n_{t+10}n_{t+12} \\
&\qquad \oplus n_{t+14}n_{t+15} \oplus n_{t+21}n_{t+28} \oplus \cancel{n_{t+23}n_{t+25}} \oplus n_{t+29}n_{t+31} \oplus n_{t+33}n_{t+39} \\
&\qquad \oplus n_{t+40}n_{t+43}n_{t+44}n_{t+45} \oplus n_{t+17}n_{t+18}n_{t+19} \oplus n_{t+34}n_{t+37}n_{t+38}) \\
&\quad \oplus \cancel{n_{t+48}} \oplus (n_{t+8} \oplus n_{t+21} \oplus \cancel{n_{t+27}} \oplus \cancel{n_{t+43}} \oplus \cancel{n_{t+47}} \oplus n_{t+10}n_{t+33} \oplus n_{t+11}n_{t+13} \\
&\qquad \oplus n_{t+15}n_{t+16} \oplus n_{t+22}n_{t+29} \oplus n_{t+24}n_{t+26} \oplus n_{t+30}n_{t+32} \oplus n_{t+34}n_{t+40} \\
&\qquad \oplus n_{t+41}n_{t+44}n_{t+45}n_{t+46} \oplus n_{t+18}n_{t+19}n_{t+20} \oplus n_{t+35}n_{t+38}n_{t+39}) \\
&\quad \oplus \cancel{n_{t+49}} \oplus (n_{t+9} \oplus \cancel{n_{t+22}} \oplus \cancel{n_{t+28}} \oplus \cancel{n_{t+44}} \oplus \cancel{n_{t+48}} \oplus n_{t+11}n_{t+34} \oplus n_{t+12}n_{t+14} \\
&\qquad \oplus n_{t+16}n_{t+17} \oplus n_{t+23}n_{t+30} \oplus \cancel{n_{t+25}n_{t+27}} \oplus n_{t+31}n_{t+33} \oplus n_{t+35}n_{t+41} \\
&\qquad \oplus n_{t+42}n_{t+45}n_{t+46}n_{t+47} \oplus n_{t+19}n_{t+20}n_{t+21} \oplus n_{t+36}n_{t+39}n_{t+40}) \\
&\quad \oplus n_{t+50} \oplus (n_{t+10} \oplus \cancel{n_{t+23}} \oplus \cancel{n_{t+29}} \oplus \cancel{n_{t+45}} \oplus \cancel{n_{t+49}} \oplus n_{t+12}n_{t+35} \oplus n_{t+13}n_{t+15} \\
&\qquad \oplus n_{t+17}n_{t+18} \oplus n_{t+24}n_{t+31} \oplus n_{t+26}n_{t+28} \oplus n_{t+32}n_{t+34} \oplus n_{t+36}n_{t+42} \\
&\qquad \oplus n_{t+43}n_{t+46}n_{t+47}n_{t+48} \oplus n_{t+20}n_{t+21}n_{t+22} \oplus n_{t+37}n_{t+40}n_{t+41}) \\
&\quad \oplus (\Lambda'[6]\ell_{t+6} \oplus \Lambda'[32]\ell_{t+32} \oplus \Lambda'[17]\ell_{t+17} \\
&\quad \oplus \Lambda'[18]\ell_{t+18} \oplus \Lambda'[44]\ell_{t+44} \oplus \Lambda'[29]\ell_{t+29}).
\end{aligned}$$

There are 46 bits in the NFSR and 6 bits in the LFSR in $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle$, and the size of involved bits is too large to evaluate the correlation with brute force. We decompose this Boolean function into four Boolean functions G_1, G_2, G_3, and

G_4, i.e., $g''_t \oplus \langle L^{(t)}, \Lambda' \rangle = G_1 \oplus G_2 \oplus G_3 \oplus G_4$.

$$
\begin{aligned}
G_1 = {} & n_{t+6} \oplus n_{t+8} \oplus n_{t+9} \oplus n_{t+10} \oplus n_{t+21} \oplus n_{t+38} \oplus n_{t+4}\ell_{t+6} \oplus \ell_{t+32}\ell_{t+17} \\
& \oplus n_{t+4}\ell_{t+32}n_{t+38} \oplus n_{t+4}n_{t+6} \oplus n_{t+8}n_{t+9} \oplus n_{t+6}n_{t+8} \\
& \oplus n_{t+8}n_{t+10} \oplus n_{t+9}n_{t+32} \oplus \Lambda'[6]\ell_{t+6} \oplus \Lambda'[17]\ell_{t+17} \oplus \Lambda'[32]\ell_{t+32} \\
G_2 = {} & n_{t+7} \oplus n_{t+34} \oplus n_{t+27}n_{t+33} \oplus n_{t+7}n_{t+30} \oplus n_{t+21}n_{t+23} \\
& \oplus n_{t+27}n_{t+29} \oplus n_{t+29}n_{t+31} \oplus n_{t+33}n_{t+39} \\
& \oplus n_{t+10}n_{t+33} \oplus n_{t+30}n_{t+32} \oplus n_{t+23}n_{t+30} \oplus n_{t+31}n_{t+33} \\
& \oplus n_{t+30}n_{t+33}n_{t+34} \\
G_3 = {} & n_{t+13} \oplus n_{t+14} \oplus n_{t+35} \oplus n_{t+36} \oplus n_{t+40} \oplus n_{t+41} \\
& \oplus n_{t+10}n_{t+11} \oplus n_{t+29}n_{t+35} \oplus n_{t+12}n_{t+13} \oplus n_{t+31}n_{t+37} \oplus n_{t+10}n_{t+12} \\
& \oplus n_{t+14}n_{t+15} \oplus n_{t+11}n_{t+34} \oplus n_{t+12}n_{t+14} \oplus n_{t+35}n_{t+41} \\
& \oplus n_{t+12}n_{t+35} \oplus n_{t+13}n_{t+15} \oplus n_{t+36}n_{t+42} \oplus n_{t+11}n_{t+13} \\
& \oplus n_{t+34}n_{t+40} \oplus n_{t+11}n_{t+12}n_{t+13} \oplus n_{t+13}n_{t+14}n_{t+15} \\
& \oplus n_{t+32}n_{t+35}n_{t+36} \oplus n_{t+35}n_{t+38}n_{t+39} \oplus n_{t+36}n_{t+39}n_{t+40} \\
& \oplus n_{t+37}n_{t+40}n_{t+41} \oplus n_{t+34}n_{t+37}n_{t+38} \\
& \oplus n_{t+41}n_{t+44}n_{t+45}n_{t+46} \oplus n_{t+34}n_{t+37}n_{t+38}n_{t+39} \\
& \oplus n_{t+36}n_{t+39}n_{t+40}n_{t+41} \oplus n_{t+40}n_{t+43}n_{t+44}n_{t+45} \\
& \oplus n_{t+42}n_{t+45}n_{t+46}n_{t+47} \oplus n_{t+43}n_{t+46}n_{t+47}n_{t+48} \\
& \oplus n_{t+38}n_{t+41}n_{t+42}n_{t+43} \\
G_4 = {} & n_{t+3} \oplus n_{t+5} \oplus n_{t+15} \oplus n_{t+16} \oplus n_{t+17} \oplus n_{t+24} \oplus n_{t+26} \oplus n_{t+50} \\
& \oplus n_{t+16}\ell_{t+18} \oplus \ell_{t+44}\ell_{t+29} \oplus n_{t+16}\ell_{t+44}n_{t+50} \oplus n_{t+3}n_{t+26} \\
& \oplus n_{t+15}n_{t+22} \oplus n_{t+17}n_{t+19} \oplus n_{t+5}n_{t+28} \oplus n_{t+17}n_{t+24} \\
& \oplus n_{t+19}n_{t+21} \oplus n_{t+19}n_{t+26} \oplus n_{t+21}n_{t+28} \oplus n_{t+15}n_{t+16} \\
& \oplus n_{t+22}n_{t+29} \oplus n_{t+24}n_{t+26} \oplus n_{t+16}n_{t+17} \oplus n_{t+17}n_{t+18} \\
& \oplus n_{t+24}n_{t+31} \oplus n_{t+26}n_{t+28} \oplus n_{t+32}n_{t+34} \\
& \oplus n_{t+17}n_{t+18}n_{t+19} \oplus n_{t+18}n_{t+19}n_{t+20} \oplus n_{t+19}n_{t+20}n_{t+21} \\
& \oplus n_{t+28}n_{t+31}n_{t+32} \oplus n_{t+15}n_{t+16}n_{t+17} \oplus n_{t+20}n_{t+21}n_{t+22} \\
& \oplus \Lambda'[18]\ell_{t+18} \oplus \Lambda'[29]\ell_{t+29} \oplus \Lambda'[44]\ell_{t+44}
\end{aligned}
$$

Four Boolean functions G_1, G_2, G_3, and G_4 involve 14, 12, 24, and 24 bits, respectively. These involved bits are independent except for n_{t+39}, n_{t+38}, n_{t+34}, n_{t+32}, n_{t+31}, n_{t+29}, n_{t+21}, n_{t+15}, and n_{t+10}, where these nine bits are colored by red. Therefore, we compute the conditional correlations of G_1, G_2, G_3, and G_4.

Table 7 shows the correlation of G when $\Lambda'[6, 17, 18, 29, 32, 44] = 001100$. Here, note that each conditional correlation must be divided by 2^9 because we add 9-bit condition. Table 8 summarizes each correlation, where we picked the case whose correlation is non-zero.

Table 7. Case that $\Lambda'[8, 17, 27, 29, 30, 42, 44, 45, 49] = 000100000$.

n_{t+10}	n_{t+15}	n_{t+29}	n_{t+31}	n_{t+34}	n_{t+38}	n_{t+39}	G_1	G_2	G_3	G_4	correlation
0	0	0	0	0	0	0	2^{-3}	2^{-2}	$2^{-6.4150}$	2^{-6}	$2^{-17.4150-9}$
0	0	0	0	0	0	1	2^{-3}	2^{-2}	2^{-8}	2^{-6}	2^{-19-9}
0	0	0	0	0	1	0	-2^{-3}	2^{-2}	$2^{-6.4150}$	2^{-6}	$-2^{-17.4150-9}$
0	0	0	0	0	1	1	-2^{-3}	2^{-2}	-2^{-8}	2^{-6}	2^{-19-9}
0	0	0	1	0	1	1	-2^{-3}	2^{-2}	2^{-7}	-2^{-7}	2^{-19-9}
0	0	1	0	0	0	0	2^{-3}	2^{-2}	$-2^{-4.5406}$	2^{-7}	$-2^{-16.5406-9}$
0	0	1	0	0	0	1	2^{-3}	-2^{-2}	$-2^{-4.5406}$	2^{-7}	$2^{-16.5406-9}$
0	0	1	0	0	1	0	-2^{-3}	2^{-2}	-2^{-8}	2^{-7}	2^{-20-9}
0	0	1	0	0	1	1	-2^{-3}	-2^{-2}	2^{-8}	2^{-7}	2^{-20-9}
0	1	0	0	0	0	0	2^{-3}	2^{-2}	$-2^{-5.1926}$	2^{-7}	$-2^{-17.1926-9}$
0	1	0	0	0	0	1	2^{-3}	2^{-2}	$-2^{-5.4150}$	2^{-7}	$-2^{-17.4150-9}$
0	1	0	0	0	1	0	-2^{-3}	2^{-2}	-2^{-7}	2^{-7}	2^{-19-9}
0	1	0	0	0	1	1	-2^{-3}	2^{-2}	2^{-8}	2^{-7}	-2^{-20-9}
0	1	1	0	0	0	0	2^{-3}	2^{-2}	$2^{-5.1926}$	2^{-6}	$2^{-16.1926-9}$
0	1	1	0	0	0	1	2^{-3}	-2^{-2}	$2^{-5.4150}$	2^{-6}	$-2^{-16.4150-9}$
0	1	1	0	0	1	0	-2^{-3}	2^{-2}	2^{-7}	2^{-6}	-2^{-18-9}
0	1	1	0	0	1	1	-2^{-3}	-2^{-2}	-2^{-8}	2^{-6}	-2^{-19-9}
0	1	1	1	0	0	0	2^{-3}	2^{-2}	$-2^{-5.6781}$	2^{-7}	$-2^{-17.6781-9}$
0	1	1	1	0	0	1	2^{-3}	-2^{-2}	$-2^{-5.6781}$	2^{-7}	$2^{-17.6781-9}$
0	1	1	1	0	1	0	-2^{-3}	2^{-2}	-2^{-8}	2^{-7}	2^{-20-9}
0	1	1	1	0	1	1	-2^{-3}	-2^{-2}	2^{-8}	2^{-7}	2^{-20-9}
1	0	0	0	1	0	0	2^{-3}	-2^{-2}	$-2^{-3.6077}$	2^{-6}	$2^{-14.6077-9}$
1	0	0	0	1	0	1	2^{-3}	-2^{-2}	$-2^{-3.4764}$	2^{-6}	$2^{-14.4764-9}$
1	0	0	0	1	1	1	-2^{-3}	-2^{-2}	2^{-8}	2^{-6}	2^{-19-9}
1	0	0	1	1	1	0	-2^{-3}	-2^{-2}	$-2^{-3.6077}$	-2^{-7}	$2^{-15.6077-9}$
1	0	0	1	1	1	1	-2^{-3}	-2^{-2}	2^{-7}	-2^{-7}	-2^{-19-9}
1	0	1	0	1	0	0	2^{-3}	-2^{-2}	$2^{-4.5406}$	2^{-7}	$-2^{-16.5406-9}$
1	0	1	0	1	0	1	2^{-3}	2^{-2}	$2^{-4.5406}$	2^{-7}	$2^{-16.5406-9}$
1	0	1	0	1	1	0	-2^{-3}	-2^{-2}	2^{-7}	2^{-7}	2^{-19-9}
1	0	1	0	1	1	1	-2^{-3}	2^{-2}	$-2^{-3.4764}$	2^{-7}	$2^{-15.4764-9}$
1	1	0	0	1	0	0	2^{-3}	-2^{-2}	$2^{-5.6781}$	2^{-7}	$-2^{-17.6781-9}$
1	1	0	0	1	0	1	2^{-3}	-2^{-2}	$2^{-5.4150}$	2^{-7}	$-2^{-17.4150-9}$
1	1	0	0	1	1	0	-2^{-3}	-2^{-2}	-2^{-8}	2^{-7}	-2^{-20-9}
1	1	0	0	1	1	1	-2^{-3}	-2^{-2}	$2^{-4.5406}$	2^{-7}	$2^{-16.5406-9}$
1	1	1	0	1	0	0	2^{-3}	-2^{-2}	$2^{-5.6781}$	2^{-6}	$-2^{-16.6781-9}$
1	1	1	0	1	0	1	2^{-3}	2^{-2}	$2^{-5.4150}$	2^{-6}	$2^{-16.4150-9}$
1	1	1	0	1	1	0	-2^{-3}	-2^{-2}	-2^{-8}	2^{-6}	-2^{-19-9}
1	1	1	0	1	1	1	-2^{-3}	2^{-2}	$2^{-4.5406}$	2^{-6}	$-2^{-15.5406-9}$
1	1	1	1	1	0	0	2^{-3}	-2^{-2}	$-2^{-5.6781}$	2^{-7}	$2^{-17.6781-9}$
1	1	1	1	1	0	1	2^{-3}	2^{-2}	$-2^{-5.6781}$	2^{-7}	$-2^{-17.6781-9}$
1	1	1	1	1	1	0	-2^{-3}	-2^{-2}	$2^{-4.6781}$	2^{-7}	$2^{-16.6781-9}$
1	1	1	1	1	1	1	-2^{-3}	2^{-2}	-2^{-8}	2^{-7}	2^{-20-9}
sum											$2^{-22.1420}$

When $n_{t+21} = 0$ or $n_{t+32} = 1$, the correlation is 0. Therefore, n_{t+21} must be 1, and n_{t+32} must be 0, and columns for n_{t+21} and n_{t+32} are omitted.

Table 8. Correlations of $g_t'' \oplus \langle L^{(t)}, \Lambda' \rangle$.

$\Lambda'[6]$	$\Lambda'[17]$	$\Lambda'[18]$	$\Lambda'[29]$	$\Lambda'[32]$	$\Lambda'[44]$	correlation
0	0	1	1	0	0	$2^{-22.142}$
0	0	1	1	0	1	$-2^{-22.142}$
0	0	1	1	1	0	$2^{-22.142}$
0	0	1	1	1	1	$-2^{-22.142}$
0	1	1	1	0	0	$2^{-22.142}$
0	1	1	1	0	1	$-2^{-22.142}$
0	1	1	1	1	0	$-2^{-22.142}$
0	1	1	1	1	1	$2^{-22.142}$
1	0	1	1	0	0	$2^{-23.678}$
1	0	1	1	0	1	$-2^{-23.678}$
1	0	1	1	1	0	$2^{-23.678}$
1	0	1	1	1	1	$-2^{-23.678}$
1	1	1	1	0	0	$2^{-22.142}$
1	1	1	1	0	1	$-2^{-22.142}$
1	1	1	1	1	0	$-2^{-22.142}$
1	1	1	1	1	1	$2^{-22.142}$