symmetry

MDPI

Article

Almost-Minimal-Round BBB-Secure Tweakable Key-Alternating Feistel Block Cipher

Ming Jiang and Lei Wang *

Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; trojan1019@sjtu.edu.cn

* Correspondence: wanglei_hb@sjtu.edu.cn

Abstract: This paper focuses on designing a tweakable block cipher via by tweaking the Key-Alternating Feistel (KAF for short) construction. Very recently Yan et al. published a tweakable KAF construction. It provides a birthday-bound security with 4 rounds and Beyond-Birthday-Bound (BBB for short) security with 10 rounds. Following their work, we further reduce the number of rounds in order to improve the efficiency while preserving the same level of security bound. More specifically, we rigorously prove that 6-round tweakable KAF cipher is BBB- secure. The main technical contribution is presenting a more refined security proof framework, which makes significant efforts to deal with several subtle and complicated sub-events. Note that Yan et al. showed that 4-round KAF provides exactly Birthday-Bound security by a concrete attack. Thus, 6 rounds are (almost) minimal rounds to achieve BBB security for tweakable KAF construction.

Keywords: beyond-birthday-bound security; H-coefficient technique; key-alternating Feistel cipher; provable security; tweakable block cipher

Citation: Jiang, M.; Wang, F. Almost-Minimal-Round BBB-Secure Tweakable Key-Alternating Feistel Block Cipher. *Symmetry* **2021**, *13*, 649. https://doi.org/10.3390/sym13040649

Academic Editors: Alexei Kanel-Belov and Lorentz JÄNTSCHI

Received: 26 February 2021
Accepted: 7 April 2021
Published: 11 April 2021

1. Introduction

A *block cipher*, also known as a *pseudorandom permutation*, which is a pair of algorithms (E, D). A block cipher has two important parameters: block length and key length. If the block length is n bits and the key length is k bits, for a mathematical point of view, the block cipher can be seen as a mapping

$$\{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n.$$

E represents a mapping that from the key space and the message space to the message space, and D is the opposite direction of the mapping in E. In addition, we call E is *encryption*, and D is *decryption*. The schemes of block cipher are roughly separated into two main classes, which are named Feistel networks and substitution—permutation networks (SPNs).

The *tweakable block cipher* is formalized by Liskov et al. [1]. It introduces to the block cipher an extra public input parameter *tweak*. The tweak provides inherent variability for building higher higher-level cryptographic schemes, namely modes of operation. So far, the tweakable block cipher has got received wide applications. Examples include Message Encryption, Message Authentication Code [1,2], and Authenticated Encryption Mode [3–5], etc. Now designing secure tweakable block ciphers has become a very important research topic. Cryptographers build tweakable block ciphers either from the scratch [6–8], or based on existing cryptographic primitives such as block ciphers or permutations [2,9–11]. Among these approaches, one is introducing the tweak to general structures of classical block ciphers, namely the Feistel construction [12] and the Even—Mansour construction [13]. We refer the interested readers to [9,14–18] for tweaking the Even—Mansour construction.

This paper mainly focuses on tweaking the Feistel construction. Since invented by Horst Feistel in 1973 [12], the Feistel construction has been a mainstream class of block

ciphers. More specifically, there are several Feistel construction variants, such as Luby—Rackoff [19], Generalized Feistel [20], Key-Alternating Feistel [21], etc. They have been adopted in dedicated block ciphers including international and national standards. In 2007, Goldenberg et al. published the first paper of incorporating tweak to the Feistel constructions [22]. In particular, they paid attention to the Luby—Rackoff ciphers, and XOR tweaks to the dataflow branches. We write such tweak injection as *linear* tweak injection in this paper. Goldenberg et al. found that 6 rounds and more are secure (against polynomial adversaries). Moreover, they showed that 10 rounds are secure against 2^n adversarial queries, that is i.e., fully secure with n as the branch bit size of Luby—Rackoff structure. After that, Mitsuda and Iwata analyzed tweaking Generalized Feistel Structures with similar linear tweak injection [20]. They proved that $2d$ rounds are birthday-bound secure with d as the number of branches of Generalized Feistel Structure. Very recently, Yan et al. published a result of tweaking the Key-Alternating Feistel (KAF) Cipher [23]. They introduced the tweak by mixing round keys, and proved that 4 rounds have a birthday-bound security and 10 rounds enable a beyond-birthday-bound security of roughly $2^{2n/3}$ adversarial queries with n as the branch bit size. We will carry on the research of tweaking the KAF. (It is referred to as *Feistel-2* in IACR Tikz Library).

The Feistel network [12] is a popular structure of block ciphers. In the i-th round of the Feistel cipher, the intermediate state of input $x_i = L \parallel R$ is updated by the round function G_i, i.e., $L \parallel R \rightarrow R \parallel L \oplus G_i(k_i, R)$. After tweaking the generalized Feistel ciphers by Mitsuda and Iwata [20], there is are only a few works about tweaking the Feistel cipher. The most mainstream research is tweaking KAF ciphers as Yan et al did recently [23]. They introduced the tweak with several round keys by using a universal hash function $H(\cdot)$, that is, $tk_i \leftarrow H_{k_i}(t)$, where k_i is the secret key, t is the tweak. By tweaking KAF with the i-th round function, the input is updated through

$$m_L \parallel m_R \leftarrow m_R \parallel m_L \oplus F_i(H_{k_i}(t) \oplus m_R),$$

where $F(\cdot)$ is the ith-round function. Yan et al. presented a 4-round minimized structure with two round keys and a single random function, proved that it achieves Birthday-Bound security. Meanwhile, they presented a 10-round tweakable KAF (TKAF for short) construction (depict in Figure 1) that can achieve BBB security. In this work, we aim to optimize Yan et al's 10-round structure, and adopt other distinct construction of tweakable block ciphers. Then we give the proof that the new construction still meets the BBB security. We compared with Yan et al's work [23] which lists in Table 1.

Table 1. Comparison with related works.

Key Size	Rounds	Number of Round Functions	Bound	Reference
$10n$	10	10	$2n/3$	Yan et al. [23]
$6n$	6	6	$2n/3$	Guo et al. [24] (without tweak)
$6n$	6	6	$2n/3$	Section 3 (tweaked)

1.1. Our Contributions

In this paper, we present a 6-round TKAF cipher which meets the BBB security, with tweaking the additional outer four rounds based on based on Guo et al.'s 6-round KAF [24]. Unlike Yan et al.'s research, we adopt the approach of introducing tweak into the 6-round KAF directly. By utilizing Guo et al's proof methodology, we introduce the tweak via using a universal hash function. We prove when the adversary makes distinct queries with different tweaks, due to the uniformity of the mentioned hash function, it still meets BBB security.

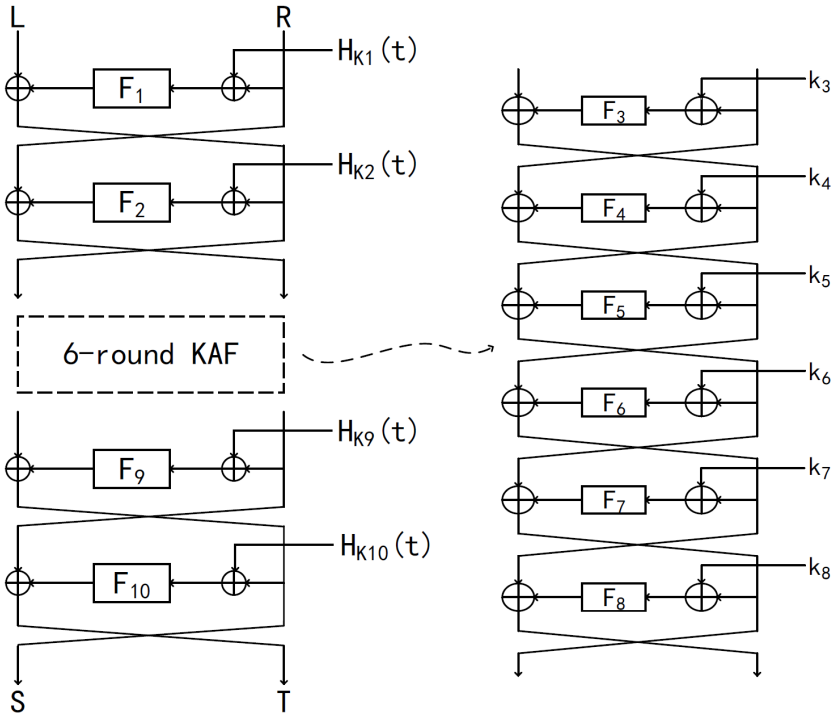

Figure 1. 10-round tweakable Key-Alternating Feistel cipher presented by Yan et al.

1.2. Structure of This Paper

Section 2 is the preliminaries of notations and definitions. Section 3 is the overview of proofs and core contribution. Section 4 is the proof of our conclusion. Section 5 is the future work.

2. Preliminaries

2.1. Notations and General Definitions

Let n denote a positive integer. Then $N = 2^n$ and $\mathcal{N} = \{0,1\}^n$. $\mathcal{F}(n)$ denotes the set of all functions mapping from $\mathcal{N}$ to $\mathcal{N}$. $\mathcal{P}(2n)$ denotes the set of all permutations in the range of $\{0,1\}^{2n}$. Let $\theta(s)$ be a random variable relying on one another random variable s. Then we denote by $\mathbb{E}_{s\in\mathcal{S}}[\theta(s)]$ the expectation of $\theta(s)$ taken over all $s \in \mathcal{S}$. For $X, Y \in \mathcal{N}$, denote $X \parallel Y$ or simply XY as their concatenation.

2.1.1. Block Cipher

A block cipher is a family of permutations indexed by the secret key. It is denoted as $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$, where $\mathcal{K}$ is the key space, $\mathcal{M}$ is the message space, and $\mathcal{C}$ is the ciphertext space. Hence for each $K \in \mathcal{K}$, $E(K,\cdot)$ or simply $E_K(\cdot)$ is a permutation from $\mathcal{M}$ to $\mathcal{C}$. In this paper, $\mathcal{M} = \mathcal{C} = \{0,1\}^{2n}$.

2.1.2. Tweakable Block Cipher

A tweakable block cipher is a family of permutations indexed by the secret key and the public tweak. It is denoted as $\widetilde{E} : \mathcal{K} \times \mathcal{T} \times \mathcal{M} \to \mathcal{C}$, where $\mathcal{K}$ is the key space, $\mathcal{T}$ is the tweak space, $\mathcal{M}$ is the message space, and $\mathcal{C}$ is the ciphertext space. Hence for each $K \in \mathcal{K}$ and each $T \in \mathcal{T}$, $E(K,T,\cdot)$ or simply $E_{K,T}(\cdot)$ is a permutation from $\mathcal{M}$ to $\mathcal{C}$. Similarly, $\mathcal{M} = \mathcal{C} = \{0,1\}^{2n}$. We denote $\widetilde{\Pi}(\mathcal{T}, 2n)$ as the set of all tweakable permutations with $\mathcal{M} = \mathcal{C} = \{0,1\}^{2n}$.

2.1.3. Key-Alternating Feistel (KAF) Cipher

A KAF is a block cipher with $\mathcal{M} = \mathcal{C} = \{0,1\}^{2n}$. It has an iterative structure. The i-th round function has the form $\Psi^{\boldsymbol{F}}_{k_i}(L \parallel R) = (R \parallel L \oplus \boldsymbol{F}_i(R \oplus k_i))$, where L and R are the left half and the right half of the inputs respectively, k_i is the i-th secret round key, and $\boldsymbol{F}_i$ is

the i-th public round function. We denote the r-round KAF with r public round functions $\boldsymbol{F} = (\boldsymbol{F}_1, \ldots, \boldsymbol{F}_r)$ in $\mathcal{F}(n)$ and a round-key vector $k = (k_1, \ldots, k_r)$ by

$$\mathsf{KAF}_k^{\boldsymbol{F}}(L \parallel R) = \Psi_{k_r}^{\boldsymbol{F}_r} \circ \ldots \circ \Psi_{k_1}^{\boldsymbol{F}_1}(L \parallel R).$$

2.1.4. Uniform AXU Hash Functions

A set of hash functions is denoted as $\mathcal{H} : \mathcal{K} \times \mathcal{T} \to \mathcal{N}$. For each key $k \in \mathcal{K}$, a keyed hash function $H(k, \cdot)$ or simply $H_k(\cdot)$ maps the tweak space $\mathcal{T}$ to $\mathcal{N}$. $\mathcal{H}$ is said to be *uniform* hash function if for any $t \in \mathcal{T}$ and $y \in \mathcal{N}$,

$$\Pr\left[k \xleftarrow{\$} \mathcal{K} : H(k, t) = y\right] = 2^{-n}.$$

Moreover, it is said to be ϵ-almost XOR-universal (ϵ-AXU) if for any $t, t' \in \mathcal{T}$ with $t \neq t'$ and any $y \in \mathcal{N}$,

$$\Pr\left[k \xleftarrow{\$} \mathcal{K} : H_k(t) \oplus H_k(t') = y\right] \leq \epsilon.$$

2.2. Security Definitions

A *distinguisher* $\mathcal{D}$ can be thought as a fundamental attacker, and it can make queries to one (or more) "oracle" which can be the block ciphers or the random permutations. The *advantage* of a distinguisher $\mathcal{D}$ in distinguishing two oracles $\mathcal{O}$ and $\mathcal{Q}$ can be defined as:

$$\boldsymbol{Adv}(\mathcal{D}) = \left|\Pr\left[\mathcal{D}^{\mathcal{O}} \to 1\right] - \Pr\left[\mathcal{D}^{\mathcal{Q}} \to 1\right]\right|.$$

We discuss this under the Random Permutation model. Firstly, we define two worlds– "the real world" and "the ideal world". When the distinguisher $\mathcal{D}$ interacts with the oracle $(\mathcal{O}, \boldsymbol{F})$, the real world means $\mathcal{O}$ is a tweakable block cipher $\widetilde{E}(k, \cdot)$, $\boldsymbol{F} = (F_1, \ldots, F_r)$ is a public random function or permutation of $\widetilde{E}$, where k is uniformly taken from $\mathcal{K}$. In addition, in the ideal world, $\mathcal{O}$ is a tweakable permutation $\widetilde{\Pi}$ and $\boldsymbol{F} = (F_1, \ldots, F_r)$ is a public random function or permutation of $\widetilde{\Pi}$. We call $\mathcal{O}$ *construction oracle* and $\boldsymbol{F}$ *inner component oracles*. The security of a tweakable block cipher is measured by the advantage of the distinguisher $\mathcal{D}$ that distinguishes the two worlds: $(\widetilde{E}(k, \cdot), \boldsymbol{F})$ and $(\widetilde{\Pi}, \boldsymbol{F})$(depict in Figure 2). We write

$$\boldsymbol{Adv}(\mathcal{D}) = \left|\Pr\left[\mathcal{D}^{\widetilde{E}(k,\cdot),\boldsymbol{F}} \to 1\right] - \Pr\left[\mathcal{D}^{\widetilde{\Pi},\boldsymbol{F}} \to 1\right]\right|.$$

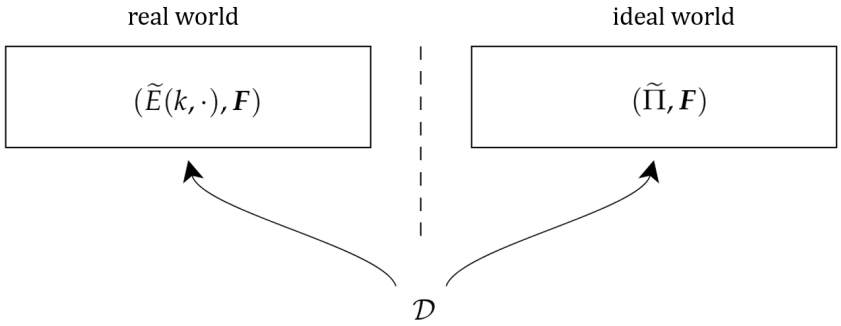

Figure 2. A distinguisher $\mathcal{D}$ distinguish the real world and the ideal world.

Theoretically, we only consider the information-theoretic distinguisher whose computation power is unlimited, i.e., it is determined, and only with limited information, that which means the number of access to the oracle is limited. We assume that the distinguishers do not make redundant queries. We also consider the distinguishers are under the

chosen-ciphertext-attack (CCA) model, meanwhile they can choose tweaks, where they have the ability to query all the oracles either forward or backward.

We denote q_e as the quantity of queries to the construction oracle and q_f as the number of queries to each inner component oracle, then the definition of insecurity of the tweakable block cipher $\widetilde{E}$ is

$$\boldsymbol{Adv}_{\widetilde{E}}(q_e, q_f) = \max_{\mathcal{D}}\{\boldsymbol{Adv}(\mathcal{D})\}.$$

H-Coefficient Technique

We utilizuse the H-coefficient technique [25,26] to evaluate the upper bound of the advantage of the adversary mentioned above.

Definition 1 (Transcript)**.** *A transcript $\tau = (\mathcal{Q}_E, \mathcal{Q}_{\boldsymbol{F}})$ is the response-tuple when the distinguisher $\mathcal{D}$ interacts with its oracle, where $\mathcal{Q}_E$ contains the tuples of the form $(t, LR, ST) \in \mathcal{T} \times \{0,1\}^{2n} \times \{0,1\}^{2n}$ which interacts with the construction oracle and $\mathcal{Q}_{\boldsymbol{F}}$ contains the tuples (x, y) which interacts with the inner component oracle.*

By definition, we can see that $\mathcal{D}$ either makes the direct query (t, LR) to the construction oracle with x to the inner component oracle, receiving answer ST and y, or makes the inverse query (t, ST) to the construction oracle with y to the inner component oracle, receiving answer LR and x. Suppose that $|\mathcal{Q}_E| = q_e$, and there are m distinct tweaks in the $\mathcal{Q}_E$. We assume there exist $q_i (1 \le i \le m)$ distinct queries for the i-th tweak, hence $\sum_{i=1}^{m} q_i = q_e$. That means $\mathcal{Q}_E = \bigcup \mathcal{Q}_{E_i}, 1 \le i \le m$, where $\mathcal{Q}_{E_i}$ are the corresponding queries of the i-th tweak. Similarly, we have $|\mathcal{Q}_{F_j}| = q_f$ and $\mathcal{Q}_{\boldsymbol{F}} = \bigcup \mathcal{Q}_{F_j}, 1 \le j \le r$.

We note that all the transcripts of queries are directionless and disordered form, but according to our hypothesis that the distinguisher $\mathcal{D}$ is deterministic. Thus, there is a one-to-one mapping between this statement and the primitive transcript of the interaction of $\mathcal{D}$ with its oracles. Meanwhile, the output of $\mathcal{D}$ is a deterministic function of τ.

In addition, for the function $\boldsymbol{F}_j$ and its set of queries $\mathcal{Q}_{F_j}$, if for each $(x, y) \in \mathcal{Q}_{F_j}$, $\boldsymbol{F}_j(x) = y$, we say that $\boldsymbol{F}_j$ *extends* $\mathcal{Q}_{F_j}$, denoted by $\boldsymbol{F}_j \vdash \mathcal{Q}_{F_j}$. Similarly, for the permutation $\boldsymbol{P}^{(i)}$ and its transcript sets $\mathcal{Q}_{E_i}$, if for each $(t, LR, ST) \in \mathcal{Q}_{E_i}$, $\boldsymbol{P}^{(i)}(t, LR) = ST$, we say that $\boldsymbol{P}^{(i)}$ extends $\mathcal{Q}_{E_i}$, denoted by $\boldsymbol{P}^{(i)} \vdash \mathcal{Q}_{E_i}$. With the above definition of "extend", we can define $\mathsf{KAF}^{\boldsymbol{F}}_{k^{(i)}} \vdash \mathcal{Q}_{E_i}$. Finally, for $\mathcal{Q}_{\boldsymbol{F}} = (\mathcal{Q}_{F_1}, \ldots, \mathcal{Q}_{F_t})$ and $\boldsymbol{F} = (\boldsymbol{F}_1, \ldots, \boldsymbol{F}_t)$, if $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1} \wedge \ldots \wedge \boldsymbol{F}_t \vdash \mathcal{Q}_{F_t}$, then we have $\boldsymbol{F} \vdash \mathcal{Q}_{\boldsymbol{F}}$.

We further define the probability that the interactions of the distinguisher $\mathcal{D}$ with the real world and the ideal world. In addition, we respectively denote them by $\Pr_{re}(\tau)$ and $\Pr_{id}(\tau)$, where τ is a transcript of these interactions.

With these definitions, we give the core lemma of the H-coefficient technique, and the distinguishing advantage could be inferred by the ratio of $\Pr_{re}(\tau)$ and $\Pr_{id}(\tau)$.

Lemma 1 (From [27])**.** *Assume that there is a function $\varphi(q_f, q_e) > 0$ such that for every possible transcript τ with q_e and q_f queries of the two types it holds*

$$|\Pr_{id}(\tau) - \Pr_{re}(\tau)| \le \Pr_{id}(\tau) \cdot \varphi(q_f, q_e),$$

then it holds

$$\boldsymbol{Adv}_{\mathsf{KAF}}(q_f, q_e) \le \varphi(q_f, q_e).$$

According to [27], the upper bound of $|\Pr_{id}(\tau) - \Pr_{re}(\tau)|$ is named "φ-point-wise proximity" of τ, which was raised by Hoang and Tessaro (HT) [27]. We let $\mathcal{K} = \mathcal{K}_{good} \cup \mathcal{K}_{bad}$, where $\mathcal{K}_{good}$ and $\mathcal{K}_{bad}$ are mutual exclusive subsets. Denote $\Pr_{re}(\tau, k)$ as the probability that $\mathcal{D}$ interacts with the real world, where $k \in \mathcal{K}$, and $\Pr_{id}(\tau, k)$ is that $\mathcal{D}$ interacts with the ideal world, where k is a "virtual" key uniformly selected from the key space $\mathcal{K}$. With the above definition, HT provided a lemma to establish point-wise proximity.

Lemma 2 (Lemma 1 of [27]). *Fix a transcript τ with $\Pr_{id}(\tau) > 0$. Assume that: (i) $\Pr[k \in \mathcal{K}_{bad}] \leq \delta$, and (ii) there is a function $g : \mathcal{K} \to [0, \infty)$ such that for all $k \in \mathcal{K}_{good}$, it holds $\frac{\Pr_{re}(\tau,k)}{\Pr_{id}(\tau,k)} \geq 1 - g(k)$. Then we have*

$$\Pr_{id}(\tau) - \Pr_{re}(\tau) \leq \Pr_{id}(\tau) \cdot (\delta + \mathbb{E}_{k \in \mathcal{K}}[g(k)]).$$

3. Overview

3.1. Beyond Birthday-Bound Security for Six Rounds

In the beginning, we need to guarantee that tweaking the KAF ciphers does not break its construction, and the influence on efficiency of the scheme execution can not cannot be enormous. For study of the execution efficiency and security, Liskov et al. [1] thought the cost of changing tweaks should be less than that of changing keys. However, the study by Jean et al. [14] showed that the adversary can hardly obtain the key, but has the ability to completely control the tweak.

In this paper, we use a nonlinear compound mode for tweaking the Feistel structure, instead of tweaking dependent or independent keys. As we known, the four rounds of KAF cipher do not meet BBB security [24], Yan's [23] work showed that tweaking 10 rounds KAF cipher can meet BBB security. Our work shows a method for tweaking the KAF cipher by the nonlinear pattern, and reduces the rounds of the scheme. For requirement of security, we consider to introduce the tweak with the round-key vectors by using a universal hash function.

Firstly, we use the suitable round-key vector which was defined by Guo [24]:

Definition 2 (Suitable Round-Key Vector for 6 Rounds [24]). *A round-key vector $k = (k_1, k_2, k_3, k_4, k_5, k_6)$ is suitable if it satisfies the following conditions:*

(i) *$k_1, k_2, k_3, k_4, k_5, k_6$ are uniformly distributed in $\{0,1\}^n$;*
(ii) *for $(i, j) \in \{(1,2), (2,3), (4,5), (5,6), (1,6)\}$, k_i and k_j are independent.*

Yan's [23] work used the minimized 6-round KAF as a "core", with additional four more rounds on the first and last sides of the "core", meanwhile introducing the tweak into these four rounds. They gave a 10-round TKAF construction with BBB security. In our work, we aim to"tweak" the first and last two rounds of the "core", and use a universal hash function to merge the tweak into round-key vectors.

Next, we denote this 6-round construction by

$$\mathsf{TKAF}_k^{\boldsymbol{F}}(t, x) = \Psi_{k_6,t}^{F_6} \circ \Psi_{k_5,t}^{F_5} \circ \Psi_{k_4}^{F_4} \circ \Psi_{k_3}^{F_3} \circ \Psi_{k_2,t}^{F_2} \circ \Psi_{k_1,t}^{F_1}(x),$$

where $\boldsymbol{F} = (F_1, F_2, F_3, F_4, F_5, F_6)$ are random functions, $k = (k_1, k_2, k_3, k_4, k_5, k_6)$ are the corresponding round keys, $t \in \mathcal{T}$ is a tweak and $x \in \{0,1\}^{2n}$ is a message (depict in Figure 3).

Finally, we upper- bound the advantage of an adversary to attack this scheme. By utilizing the H-coefficient technique which is in Lemma 2, we firstly upper upper-bound the bad key event δ, then upper- bound the expectation of the function $g(k)$, which holds $\frac{\Pr_{re}(\tau,k)}{\Pr_{id}(\tau,k)} \geq 1 - g(k)$. By Lemma 1, we could obtain the advantage. Thus, we have this theorem:

Theorem 1. *For the 6-round tweakable* KAF *cipher with a suitable round-key vector as specified in Definition 2, it holds*

$$\begin{aligned} \boldsymbol{Adv}_{\mathsf{TKAF}}(q_f, q_e) \leq\ & (7q_e^3 + 24q_e^2 q_f + 20q_e q_f^2)\frac{1}{N^2} \\ & + (4q_e^3 + 4q_e^2 q_f + 2q_e q_f^2 + 4q_e^2 + 6q_e q_f)\frac{\varepsilon}{N} \\ & + 4q_e^2 q_f \varepsilon^2 + 4q_e^2 \varepsilon^2. \end{aligned} \tag{1}$$

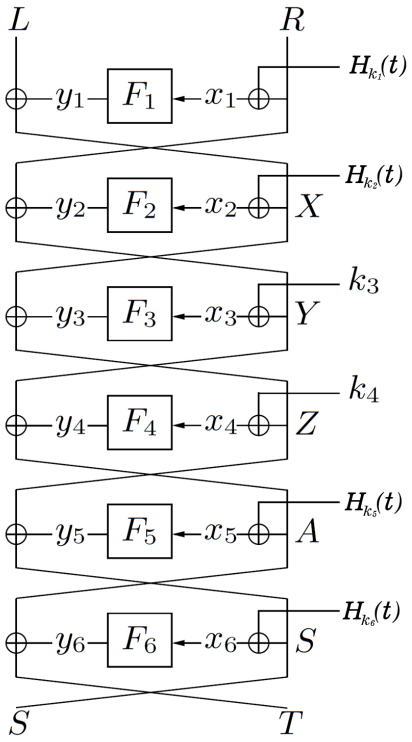

Figure 3. A tweakable Key-Alternating Feistel cipher with 6 rounds

3.2. Core Contribution

In our work, we analyze the influence of tweaking KAF ciphers on security. We tweak the outer four rounds of Guo et al's 6-round KAF and the proof of BBB security is the major research work we have done.

4. Security Proof of Theorem 1

In the following subsections, we present the methodology to prove Theorem 1. We fix a transcript $\tau = (\mathcal{Q}_E, \mathcal{Q}_F)$ with $\mathcal{Q}_F = (\mathcal{Q}_{F_1}, \mathcal{Q}_{F_2}, \mathcal{Q}_{F_3}, \mathcal{Q}_{F_4}, \mathcal{Q}_{F_5}, \mathcal{Q}_{F_6})$, where $|\mathcal{Q}_E| = q_e$ and $|\mathcal{Q}_{F_i}| = q_f$, $i = 1, \ldots, 6$. We divide the analysis of this claim into two parts: (i) define bad key vectors, then (ii) lower bound the probability $\Pr_{re}(\tau, k)$. We analyze these two parts respectively.

4.1. Bad Key Vectors and Probability

Definition 3 (Bad Key Vectors for 6 rounds)**.** *A suitable key vector $k \in \mathcal{K}$ is bad, for a transcript $\tau = (\mathcal{Q}_E, \mathcal{Q}_F)$, if one of the follow conditions is met:*

- **(A-1)** *there exists $(t, LR, ST) \in \mathcal{Q}_E$, $(x_1, y_1) \in \mathcal{Q}_{F_1}$, $(x_6, y_6) \in \mathcal{Q}_{F_6}$, such that $H_{k_1}(t) = R \oplus x_1$, $H_{k_6}(t) = S \oplus x_6$;*
- **(A-2)** *there exists $(t, LR, ST) \in \mathcal{Q}_E$, $(x_1, y_1) \in \mathcal{Q}_{F_1}$, $(x_2, y_2) \in \mathcal{Q}_{F_2}$, such that $H_{k_1}(t) = R \oplus x_1$, $H_{k_2}(t) = L \oplus y_1 \oplus x_2$;*
- **(A-3)** *there exists $(t, LR, ST) \in \mathcal{Q}_E$, $(x_5, y_5) \in \mathcal{Q}_{F_5}$, $(x_6, y_6) \in \mathcal{Q}_{F_6}$, such that $H_{k_6}(t) = S \oplus x_6$, $H_{k_5}(t) = T \oplus y_6 \oplus x_5$.*

otherwise, k is good. We denote $\mathcal{K}_{bad}$ for the set of bad key vectors, and $\mathcal{K}_{good}$ for the good key vectors.

In the beginning, we upper- bound the probability of the bad key vectors. Firstly, we analyze the above three conditions respectively, consider **(A-1)** first. Since we have the key k_1 and k_6 picked from the key space $\mathcal{K}$ uniformly and randomly, for the properties of *suitable*, k_1 and k_6 are independent of each other (Definition 2). By the uniformity of H, H_{k_1} and

H_{k_6} are also independent. Thus Thus, there are N^2 possible choices. For $(t, LR, ST) \in \mathcal{Q}_E$, $(x_1, y_1) \in \mathcal{Q}_{F_1}$ and $(x_6, y_6) \in \mathcal{Q}_{F_6}$, we have at most $q_e q_f^2$ choices, as $|\mathcal{Q}_E| = q_e$, $|Dom\mathcal{F}_1| = |Dom\mathcal{F}_6| = q_f$, where $Dom\mathcal{F}$ is a set of x that there exists $(x, y) \in \mathcal{Q}_F$ such that $F(x) = y$, i.e., $Dom\mathcal{F} \stackrel{def}{=} \{x \in \{0,1\}^n : \exists (x, y) \in \mathcal{Q}_F,\ F(x) = y\}$. Therefore, the probability of condition **(A-1)** is at most $\frac{q_e q_f^2}{N^2}$.

Similarly, by definition of suitable key vector (Definition 2), it also holds that (k_1, k_2), (k_5, k_6) are independent, and for the uniformity of H, we have $\Pr[\textbf{(A-2)}] = \Pr[\textbf{(A-3)}] \leq \frac{q_e q_f^2}{N^2}$.

To sum up, we can upper- bound the probability of the bad key vectors with

$$\Pr\left[k \xleftarrow{\$} \mathcal{K} :\ k \in \mathcal{K}_{bad}\right] \leq \frac{3 q_e q_f^2}{N^2}. \tag{2}$$

4.2. Analysis for Good Keys

In the following, we fix the round- key vectors $k \in \mathcal{K}_{good}$, and aim to lower bound the probability $\Pr\left[\boldsymbol{F} \xleftarrow{\$} (\mathcal{F}(n))^6 : \mathsf{TKAF}_k^{\boldsymbol{F}} \vdash \mathcal{Q}_E | \boldsymbol{F} \vdash \mathcal{Q}_F\right]$. By the analytical method of Cogliati et al. [9,15], we divide this proof process into two steps: (i) upper bounding the probability that a pair of functions $(\boldsymbol{F}_1, \boldsymbol{F}_6)$ satisfies "bad" conditions. By these means, the "good" conditions of the function -pair can transfer the transcripts of the distinguisher on 6 rounds to a special transcripts on 4 rounds, it can be said that we "peel off" the outer two rounds [24]; then (ii) assuming that $(\boldsymbol{F}_1, \boldsymbol{F}_6)$ is good, by bounding the inner 4 rounds, we will prove the claim of Theorem 1.

Peeling Off the Outer Two Rounds

We pick a pair of round functions $(\boldsymbol{F}_1, \boldsymbol{F}_6)$ such that $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}$ and $\boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}$. For each transcript $(t, LR, ST) \in \mathcal{Q}_E$, denote $X \leftarrow L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ and $A \leftarrow T \oplus \boldsymbol{F}_6(H_{k_6}(t) \oplus S)$. From this, we obtain q_e transcripts with the form of (t, RX, AS). For convenience, we denote a new set including all these introduced transcript tuples by $\mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$. Furthermore, we define two subsets of $\mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, the transcripts that collide at the positions of X and A, respectively. Denote them by $\mathcal{ID}(X)$ and $\mathcal{ID}(A)$:

$$\mathcal{ID}(X) = \{(t, RX, AS) : (t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6),\ X \textit{ is identical}\}$$
$$\mathcal{ID}(A) = \{(t, RX, AS) : (t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6),\ A \textit{ is identical}\}$$

In order to characterize τ, we define four key-dependent quantities:

$$n_{(1)}(k) \stackrel{def}{=} |\{((t, LR, ST), (x_1, y_1)) \in \mathcal{Q}_E \times \mathcal{Q}_{F_1} : H_{k_1}(t) = R \oplus x_1\}|$$
$$n_{(6)}(k) \stackrel{def}{=} |\{((t, LR, ST), (x_6, y_6)) \in \mathcal{Q}_E \times \mathcal{Q}_{F_6} : H_{k_6}(t) = S \oplus x_6\}|$$
$$n_{(2,3)}(k) \stackrel{def}{=} |\{((t, LR, ST), (x_2, y_2), (x_3, y_3)) \in \mathcal{Q}_E \times \mathcal{Q}_{F_2} \times \mathcal{Q}_{F_3} : k_3 = R \oplus y_2 \oplus x_3\}|$$
$$n_{(4,5)}(k) \stackrel{def}{=} |\{((t, LR, ST), (x_4, y_4), (x_5, y_5)) \in \mathcal{Q}_E \times \mathcal{Q}_{F_4} \times \mathcal{Q}_{F_5} : k_4 = S \oplus y_5 \oplus x_4\}|$$

Now we define the "bad event" on the pair $(\boldsymbol{F}_1, \boldsymbol{F}_6)$. If the corresponding set $\mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$ of the pair $(\boldsymbol{F}_1, \boldsymbol{F}_6)$ fulfills one of the following "collision" conditions, we say that the predicate is bad, denoted by $\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$:

- **(B-1)** there exists $(t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_2, y_2) \in \mathcal{Q}_{F_2}$, $(x_5, y_5) \in \mathcal{Q}_{F_5}$, such that $H_{k_2}(t) = X \oplus x_2$, $H_{k_5}(t) = A \oplus x_5$;
- **(B-2)** there exists $(t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_2, y_2) \in \mathcal{Q}_{F_2}$, $(x_3, y_3) \in \mathcal{Q}_{F_3}$, such that $H_{k_2}(t) = X \oplus x_2$, $k_3 = R \oplus y_2 \oplus x_3$;
- **(B-3)** there exists $(t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_4, y_4) \in \mathcal{Q}_{F_4}$, $(x_5, y_5) \in \mathcal{Q}_{F_5}$, such that $H_{k_5}(t) = A \oplus x_5$, $k_4 = S \oplus y_5 \oplus x_4$;

- **(B-4)** there exist two distinct $(t, RX, AS), (t', R'X', A'S') \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_2, y_2) \in \mathcal{Q}_{F_2}$, such that $X = X'$ and $H_{k_2}(t) = X \oplus x_2$; or symmetrically two distinct (t, RX, AS), $(t', R'X', A'S') \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_5, y_5) \in \mathcal{Q}_{F_5}$, such that $A = A'$ and $H_{k_5}(t) = A \oplus x_5$;
- **(B-5)** there exist two distinct $(t, RX, AS), (t', R'X', A'S') \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_2, y_2) \in \mathcal{Q}_{F_2}$, such that $A = A'$ and $H_{k_2}(t) = X \oplus x_2$; or symmetrically two distinct (t, RX, AS), $(t', R'X', A'S') \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, $(x_5, y_5) \in \mathcal{Q}_{F_5}$, such that $X = X'$ and $H_{k_5}(t) = A \oplus x_5$;

If the predicate $\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$ does not hold, then we can deem that $(\boldsymbol{F}_1, \boldsymbol{F}_6)$ is *good*. Now we bound the probability of $\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$.

Lemma 3. *It holds*

$$\begin{aligned}&\Pr\left[\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6) | \boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1} \wedge \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}\right] \\ &\leq \frac{4q_e^2 q_f + q_e q_f^2}{N^2} + \frac{q_f(n_{(1)}(k) + n_{(6)}(k))}{N} \\ &+ \frac{n_{(2,3)}(k) + n_{(4,5)}(k)}{N} + 4q_e^2 q_f \varepsilon^2 + q_f \varepsilon (n_{(1)}(k) + n_{(6)}(k)).\end{aligned}$$

Proof. We prove the above 5 cases of $\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$ on the condition of $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1} \wedge \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}$:

(B-1) For arbitrary $(t, RX, AS) \in \mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$, if there exists $(x_2, y_2) \in \mathcal{Q}_{F_2}$ and $(x_5, y_5) \in \mathcal{Q}_{F_5}$, such that $H_{k_2}(t) = X \oplus x_2$ and $H_{k_5}(t) = A \oplus x_5$. Then for the corresponding $(t, LR, ST) \in \mathcal{Q}_E$, we have $L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) = H_{k_2}(t) \oplus x_2$ and $T \oplus \boldsymbol{F}_6(H_{k_6}(t) \oplus S) = H_{k_5}(t) \oplus x_5$. On account of the uniformity of H, it must hold $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$ (if $H_{k_1}(t) \oplus R \in Dom\mathcal{F}_1$ and $H_{k_2}(t) = X \oplus x_2$, then the condition **(A-2)** is fulfilled). Similarly, it must be $H_{k_6}(t) \oplus S \notin Dom\mathcal{F}_6$. Thus, on the condition of $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1} \wedge \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}$, $\boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ and $\boldsymbol{F}_6(H_{k_6}(t) \oplus S)$ keep uniform. SoSo, the probability of both $L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) = H_{k_2}(t) \oplus x_2$ and $T \oplus \boldsymbol{F}_6(H_{k_6}(t) \oplus S) = H_{k_5}(t) \oplus x_5$ holding is at most $\frac{1}{N^2}$. AndIn addition, the choices of all 3-tuples (t, LR, ST), (x_2, y_2), (x_5, y_5) do not exceed $q_e q_f^2$. Therefore, we have $\Pr[\textbf{(B-1)}] \leq \frac{q_e q_f^2}{N^2}$.

(B-2) and (B-3) We consider **(B-2)** firstly.

There exists a 3-tuple $((t, LRX, AST), (x_2, y_2), (x_3, y_3))$, such that the number of $k_3 = R \oplus y_2 \oplus x_3$ is $n_{(2,3)}(k)$, where (t, LRX, AST) is a joint notation of (t, LR, ST) and its corresponding induced X and A. Moreover, $H_{k_2}(t) = X \oplus x_2$ means $L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) = H_{k_2}(t) \oplus x_2$. When $H_{k_1}(t) \oplus R \in Dom\mathcal{F}_1$, then it can not cannot hold $L \oplus Img\boldsymbol{F}_1(H_{k_1}(t) \oplus R) = H_{k_2}(t) \oplus x_2$, otherwise **(A-2)** is fulfilled. Furthermore Furthermore, when $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$, on the condition of $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}$, then $\boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ keeps uniform. Meanwhile H also keeps uniform, thus we have the probability of $L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) = H_{k_2}(t) \oplus x_2$ is at most $\frac{1}{N}$. Therefore, $\Pr[\textbf{(B-2)}] \leq \frac{n_{(2,3)}(k)}{N}$. The condition **(B-3)** is symmetric with **(B-2)**, so with the similar analysis, we have $\Pr[\textbf{(B-3)}] \leq \frac{n_{(4,5)}(k)}{N}$.

(B-4) For the given pair of distinct merged transcripts (t, LRX, AST) and $(t', L'R'X', A'S'T')$ together with $(x_2, y_2) \in \mathcal{Q}_{F_2}$, we discuss the cases in three conditions:

- **Case 1:** when $t \neq t'$, if it holds $H_{k_1}(t) \oplus R = H_{k_1}(t') \oplus R'$, i.e., for the ε-AUX property of H function, the probability of $H_{k_1}(t) \oplus H_{k_1}(t') = R \oplus R'$ is at most ε. If $H_{k_1}(t) \oplus R \neq H_{k_1}(t') \oplus R'$, we note that $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$, $H_{k_1}(t') \oplus R' \notin Dom\mathcal{F}_1$, otherwise **(A-2)** is fulfilled. Thus, on the condition of $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}$, $\boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ and $\boldsymbol{F}_1(H_{k_1}(t') \oplus R')$ are independent with each other, also keep uniformly random. Then it holds $\Pr\left[\boldsymbol{F}_1(H_{k_1}(t) \oplus R) = L \oplus L' \oplus \boldsymbol{F}_1(H_{k_1}(t') \oplus R')\right] \leq \varepsilon + (1 - \varepsilon)\frac{1}{N} \leq \varepsilon + \frac{1}{N}$. Therefore, the probability of the collision at the position $H_{k_2}(t) \oplus X$ and $X = X'$ is at most $(\varepsilon + \frac{1}{N})\varepsilon \leq \varepsilon^2 + \frac{1}{N}\varepsilon$.
- **Case 2:** if $t = t'$ and $R \neq R'$, for $X = X'$, the probability of $\boldsymbol{F}_1(H_{k_1}(t) \oplus R) \oplus L = \boldsymbol{F}_1(H_{k_1}(t') \oplus R') \oplus L'$ is at most $\frac{1}{N}$. AndIn addition, for $H_{k_2}(t) = X \oplus x_2$, the probability of $H_{k_2}(t) = \boldsymbol{F}_1(H_{k_1}(t) \oplus R) \oplus L \oplus x_2$ is at most $\frac{1}{N}$. For the property of H, we have the probability of the collision at the position X is at most $\frac{1}{N^2}$.

- **Case 3:** if $t = t'$ and $R = R'$ but $L \neq L'$, it can not cannot be held that $X = X'$ and $H_{k_2}(t) = X \oplus x_2$.

To sum up, the probability of "former" part of **(B-4)** can not cannot exceed $\varepsilon^2 + \frac{1}{N^2}$, and the analysis of "latter" part is similar to the former part. We consider all possible pairs of transcripts, the quantity of these pairs can not cannot exceed $q_e^2 q_f$. Therefore, $\Pr[\textbf{(B-4)}] \leq 2q_e^2 q_f \varepsilon^2 + \frac{2q_e^2 q_f}{N^2}$.

(B-5) For the given transcripts $(t, LRX, AST) \neq (t^*, L^*R^*X^*, A^*S^*T^*)$ and $(x_2, y_2) \in \mathcal{Q}_{F_2}$, due to the conditions on good key vector, it holds $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$. The same as **(B-4)**, we consider the front part of this condition. According to the state of S, we respectively discuss in three cases:

- **Case 1:** it holds $H_{k_6}(t) \oplus S \notin Dom\mathcal{F}_6$, then for the distinct (t, LRX, AST) and $(t^*, L^*R^*X^*, A^*S^*T^*)$, they all have q_e choices.
  - If $t \neq t^*$, if it holds $H_{k_6}(t) \oplus S = H_{k_6}(t^*) \oplus S^*$, then the probability of $H_{k_6}(t) \oplus H_{k_6}(t^*) = S \oplus S^*$ is at most ε;
  - If $t \neq t^*$, if it holds $H_{k_6}(t) \oplus S \neq H_{k_6}(t^*) \oplus S^*$, then $\boldsymbol{F}_6(H_{k_6}(t) \oplus S)$ and $\boldsymbol{F}_6(H_{k_6}(t^*) \oplus S^*)$ are independent and uniformly random. Thus, on the condition of $\boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}$, we have

$$\Pr\left[T \oplus \boldsymbol{F}_6(H_{k_6}(t) \oplus S) = T^* \oplus \boldsymbol{F}_6(H_{k_6}(t^*) \oplus S^*)\right] \leq \varepsilon + (1-\varepsilon)\frac{1}{N} \leq \varepsilon + \frac{1}{N}.$$

    On the condition of $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}$, $\boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ is also uniform. Hence, similar with **(B-4)**, we have

$$\Pr\left[H_{k_2}(t) \oplus X = H_{k_2}(t^*) \oplus X^*\right] \leq \varepsilon^2 + \frac{1}{N}\varepsilon.$$

  - If $t = t^*$ but $S \neq S^*$, if $A = A^*$, then it holds

$$\Pr\left[\boldsymbol{F}_6(H_{k_6}(t) \oplus S) \oplus T = \boldsymbol{F}_6(H_{k_6}(t^*) \oplus S^*) \oplus T^*\right] \leq \frac{1}{N},$$

    and for $H_{k_2}(t) = X \oplus x_2$, the probability of $H_{k_2}(t) = \boldsymbol{F}_1(H_{k_1}(t) \oplus R) \oplus L \oplus x_2$ is at most $\frac{1}{N}$;
  - If $t = t^*$ and $S = S^*$ but $T \neq T^*$, it could not be held that $A = A^*$ or $H_{k_2}(t) = X \oplus x_2$.

  Under the above cases, we have the probability of the collision at the position $H_{k_2}(t) \oplus X$ and $A = A^*$ is at most $\varepsilon^2 + \frac{1}{N^2}$. In addition, for $H_{k_6}(t) \oplus S \notin Dom\mathcal{F}_6$, the probability of **(B-5)**'s front part is at most $q_e^2 q_f \varepsilon^2 + \frac{q_e^2 q_f}{N^2}$.
- Case 2: For $H_{k_6}(t) \oplus S \in Dom\mathcal{F}_6$, the choices of (t, LRX, AST) are $n_{(6)}(k)$. Similar with **Case 1**, we have $\Pr\left[L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) \oplus H_{k_2}(t) \in \boldsymbol{F}_2\right] \leq \frac{q_f}{N} + q_f\varepsilon$. Therefore, the probability of holding at least one such transcript $(t, LRX,$ $AST)$ is at most $\frac{q_f \cdot n_{(6)}(k)}{N} + q_f n_{(1)}(k)\varepsilon$.

To sum up the above two cases, the probability that the former part of **(B-5)** holding is at most $q_e^2 q_f \varepsilon^2 + \frac{q_f \cdot n_{(6)}(k)}{N} + \frac{q_e^2 q_f}{N^2} + q_f n_{(1)}(k)\varepsilon$. Similarly, the latter part of **(B-5)** is symmetric with the former part. Therefore, we have

$$\Pr[\textbf{(B-5)}] \leq 2q_e^2 q_f \varepsilon^2 + \frac{q_f \cdot (n_{(1)}(k) + n_{(6)}(k))}{N} + \frac{2q_e^2 q_f}{N^2} + q_f\varepsilon(n_{(1)}(k) + n_{(6)}(k)).$$

We sum up all the five conditions, it holds

$$\begin{aligned}&\Pr\left[\mathsf{Bad}(\boldsymbol{F}_1,\boldsymbol{F}_6)|\boldsymbol{F}_1\vdash\mathcal{Q}_{F_1}\wedge\boldsymbol{F}_6\vdash\mathcal{Q}_{F_6}\right]\\&\leq\frac{4q_e^2q_f+q_eq_f^2}{N^2}+\frac{q_f\cdot(n_{(1)}(k)+n_{(6)}(k))}{N}\\&+\frac{n_{(2,3)}(k)+n_{(4,5)}(k)}{N}+4q_e^2q_f\varepsilon^2+q_f\varepsilon(n_{(1)}(k)+n_{(6)}(k)).\end{aligned}$$

Now we prove the Lemma 3. □

4.3. Analysis of the Inner Four Rounds

In the following section, we analyze the inner four rounds of TKAF which depicts in Figure 4. We denote $\mathcal{Q}_E^*(\boldsymbol{F}_1,\boldsymbol{F}_6)$ the set of tuples in the form (t,RX,AS), which is induced by peeling off outer two rounds. Similar with [24], we also write $\boldsymbol{F}^*=(\boldsymbol{F}_2,\boldsymbol{F}_3,\boldsymbol{F}_4,\boldsymbol{F}_5)$, further denote

$$\mathsf{p}(\tau,\boldsymbol{F}_1,\boldsymbol{F}_6)=\Pr\left[\boldsymbol{F}^*\xleftarrow{\$}(\mathcal{F}(n))^4:\mathsf{TKAF}_k^{\boldsymbol{F}^*}\vdash\mathcal{Q}_E^*(\boldsymbol{F}_1,\boldsymbol{F}_6)|\boldsymbol{F}_i\vdash\mathcal{Q}_{F_i},i=1,2,3,4,5,6\right].$$

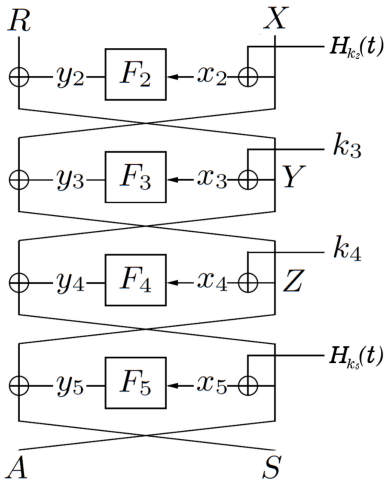

Figure 4. Inner 4 rounds of the tweakable Key-Alternating Feistel cipher.

Lemma 4 (From [24]). *Assume that there exists a function $\varphi:(\mathcal{F}(n))^2\times\mathcal{K}\to[0,\infty)$ such that for any good $(\boldsymbol{F}_1,\boldsymbol{F}_6)$, it holds*

$$\mathsf{p}(\tau,\boldsymbol{F}_1,\boldsymbol{F}_6)\Big/\prod_{i=0}^{q_e-1}\left(\frac{1}{N^2-i}\right)\geq 1-\varphi(\boldsymbol{F}_1,\boldsymbol{F}_6,k). \tag{3}$$

Then we have

$$\begin{aligned}\frac{\mathrm{Pr}_{re}(\tau,k)}{\mathrm{Pr}_{id}(\tau,k)}\geq 1-&\quad\Pr\left[\mathsf{Bad}(\boldsymbol{F}_1,\boldsymbol{F}_6)|\boldsymbol{F}_1\vdash\mathcal{Q}_{F_1},\boldsymbol{F}_6\vdash\mathcal{Q}_{F_6}\right]\\&-\quad\mathbb{E}_{\boldsymbol{F}_1,\boldsymbol{F}_6}\left[\varphi(\boldsymbol{F}_1,\boldsymbol{F}_6,k)|\boldsymbol{F}_1\vdash\mathcal{Q}_{F_1},\boldsymbol{F}_6\vdash\mathcal{Q}_{F_6}\right].\end{aligned}$$

Lemma 5. *For any fixed good tuple $(\boldsymbol{F}_1,\boldsymbol{F}_6)$, there exists a function $\varphi(\boldsymbol{F}_1,\boldsymbol{F}_6,k)$ of the function pair and the round- key vector k such that the inequality (3) mentioned in Lemma 4. Then,*

$$\begin{aligned}\mathbb{E}_{\boldsymbol{F}_1,\boldsymbol{F}_6,k}[\varphi(\boldsymbol{F}_1,\boldsymbol{F}_6,k)]\leq 4q_e^2\varepsilon^2+&\frac{7q_e^3+20q_e^2q_f+12q_eq_f^2}{N^2}\\&+\frac{4q_e^3\varepsilon+4q_e^2q_f\varepsilon+4q_e^2\varepsilon+6q_eq_f\varepsilon}{N}.\end{aligned} \tag{4}$$

Proof. Due to the space constraints, the full proof must be deferred to Appendix A. In the following, we only present a proof sketch and the core conclusions. At the beginning of the proof, we define some notations and values in order to present the proof process.

We divide the transcripts in $\mathcal{Q}_E^*(\boldsymbol{F}_1, \boldsymbol{F}_6)$ into four sets:

- $\mathcal{G}_1 = \{|\mathcal{ID}(X)| = |\mathcal{ID}(A)| = 1, and\ H_{k_2}(t) \oplus X \notin Dom\mathcal{F}_2 \wedge H_{k_5}(t) \oplus X \notin Dom\mathcal{F}_5\}$;
- $\mathcal{G}_2 = \{H_{k_2}(t) \oplus X \in Dom\mathcal{F}_2\}$;
- $\mathcal{G}_3 = \{H_{k_5}(t) \oplus A \in Dom\mathcal{F}_5\}$;
- $\mathcal{G}_4 = \{|\mathcal{ID}(X)| \geq 2\ or\ |\mathcal{ID}(A)| \geq 2\}$.

Then we denote $\mathsf{E}_{\mathcal{G}_1}$, $\mathsf{E}_{\mathcal{G}_2}$, $\mathsf{E}_{\mathcal{G}_3}$ and $\mathsf{E}_{\mathcal{G}_4}$ by the events that $\mathsf{TKAF}^{\boldsymbol{F}^*}_{H_k(t)} \vdash \mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$ and $\mathcal{G}_4$ respectively, and let $\beta_1 = |\mathcal{G}_2|$, $\beta_2 = |\mathcal{G}_3|$, $\beta_3 = |\mathcal{G}_4|$. We list $\mathcal{G}_i = \{(t, RX, AS), \ldots, (t, R_{|\mathcal{G}_i|}X_{|\mathcal{G}_i|}, A_{|\mathcal{G}_i|}S_{|\mathcal{G}_i|})\}$ with some arbitrary orders. Denote $\mathsf{E}_{|\mathcal{G}_i|}$ the event that $\mathsf{TKAF}^{\boldsymbol{F}^*}_{H_k(t)}$ extends the i-th tuple (t, R_iX_i, A_iS_i). We define four sets of "collision position":

$$
\begin{aligned}
Ext\mathcal{F}_3^{(l)} &\stackrel{def}{=} \{x_3 : \exists (t, R_iX_i, A_iS_i) \in \mathcal{G}_1,\ i \leq l, s.t.\ x_3 = k_3 \oplus R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i)\};\\
\mathcal{G}_2\mathcal{F}_3 &\stackrel{def}{=} \{x_3 : \exists (t, RX, AS) \in \mathcal{G}_2, s.t.\ x_3 = k_3 \oplus R \oplus ImgF_2(H_{k_2}(t) \oplus X)\};\\
Ext\mathcal{F}_4^{(l)} &\stackrel{def}{=} \{x_4 : \exists (t, R_iX_i, A_iS_i) \in \mathcal{G}_1,\ i \leq l, s.t.\ x_4 = k_4 \oplus S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)\};\\
\mathcal{G}_3\mathcal{F}_4 &\stackrel{def}{=} \{x_4 : \exists (t, RX, AS) \in \mathcal{G}_3, s.t.\ x_4 = k_4 \oplus S \oplus ImgF_5(H_{k_5}(t) \oplus A)\}.
\end{aligned}
$$

For convenience, we denote two values $e_3^{(l)} = |Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3|$, and $e_4^{(l)} = |Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4|$, which are the quantities of choices in the sets. Finally, the function $\mathsf{Num}_3^{(l)}(y_3)$ is the number of pre-images y_3, which belongs to the set $Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)}$. That is $\mathsf{Num}_3^{(l)}(y_3) \stackrel{def}{=} |\{x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)} : \boldsymbol{F}_3(x_3) = y_3\}|$.

Since we have these definitions mentioned above, we can lower bound

$$
\mathsf{p}(\tau, \boldsymbol{F}_1, \boldsymbol{F}_6) = \Pr\left[\mathsf{E}_{\mathcal{G}_1} \wedge \mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3} \wedge \mathsf{E}_{\mathcal{G}_4} | \boldsymbol{F} \vdash \mathcal{Q}_F\right].
$$

Analyzing these four sets in turn. First, we consider $\Pr\left[\mathsf{E}_{\mathcal{G}_1} | \boldsymbol{F} \vdash \mathcal{Q}_F\right]$. There are three cases for each transcript $(t, RX, AS) \in \mathcal{G}_1$:

(i) The two intermediate values Y and Z derived from $\boldsymbol{F}_2$ and $\boldsymbol{F}_5$ will not collide with the values that have been queried in the past time. So, the probability of this case is at least

$$
\left(1 - \frac{q_f + e_3^{(l+1)} + |\mathcal{G}_2\mathcal{F}_3|}{N}\right)\left(1 - \frac{q_f + e_4^{(l+1)} + |\mathcal{G}_3\mathcal{F}_4|}{N}\right)\frac{1}{N^2},
$$

(ii) The intermediate value Y collides with some values of the past queries, but Z is still "free". So, the probability of this case is at least

$$
\left(\frac{q_f + e_3^{(l)}}{N} - \frac{\sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)}{N} - \frac{q_f^2}{N^2} - \frac{(2q_f + q_e)(q_f + q_e)}{N^2}\right)\frac{1}{N^2}.
$$

(iii) This case is symmetrical to the second one, where Z collides with some past values, but Y is "free". The probability is at least

$$
\left(\frac{q_f + e_4^{(l)}}{N} - \frac{\sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3)}{N} - \frac{(2q_f + q_e)(q_f + q_e)}{N^2}\right)\frac{1}{N^2}.
$$

Summing over the above five cases, we have

$$\mathbb{E}_k\Big[\Pr\big[\mathsf{E}_{\mathcal{G}_1}|\boldsymbol{F}\vdash\mathcal{Q}_F\big]\Big]\geq\Big(1-\frac{q_eq_f^2}{N^2}-\frac{2q_e(2q_f+q_e)(q_f+q_e)}{N^2}-\frac{(q_f+2q_e)(\beta_1+\beta_2)}{N}\Big)\frac{1}{N^{2|\mathcal{G}_1|}}.$$

Then, we analyze $\mathsf{E}_{\mathcal{G}_2}$, $\mathsf{E}_{\mathcal{G}_3}$, and $\mathsf{E}_{\mathcal{G}_4}$. The events $\mathsf{E}_{\mathcal{G}_2}$ and $\mathsf{E}_{\mathcal{G}_3}$ can be considered simultaneously. For the rest events, we need to upper- bound the corresponding "bad" events, then consider the efficiency of introducing tweak. Through this method, we can lower bound these three events. See Appendix A for more details about the proof.

For the proof, we have the results of the following three events:

$$\begin{aligned}\Pr\big[\mathsf{E}_{\mathcal{G}_2}\wedge\mathsf{E}_{\mathcal{G}_3}|\mathsf{E}_{\mathcal{G}_1}\wedge\boldsymbol{F}\vdash\mathcal{Q}_F\big]&\geq(1-\Pr[\mathsf{Bad}_1(\boldsymbol{F}_3)]-\Pr[\mathsf{Bad}_2(\boldsymbol{F}_4)])\\&\cdot\Pr\big[\mathsf{E}_{\mathcal{G}_2}\wedge\mathsf{E}_{\mathcal{G}_3}|\neg\mathsf{Bad}_1(\boldsymbol{F}_3)\wedge\neg\mathsf{Bad}_2(\boldsymbol{F}_4)\big]\\&\geq\Big(1-\frac{(\beta_1+\beta_2)(q_f+q_e)}{N}-(\beta_1+\beta_2)\varepsilon\Big)\\&\cdot\frac{1}{N^{2(|\mathcal{G}_2|+|\mathcal{G}_3|)}};\end{aligned}$$

$$\begin{aligned}\Pr\big[\mathsf{E}_{\mathcal{G}_4}|\mathsf{E}_{\mathcal{G}_1}\wedge\mathsf{E}_{\mathcal{G}_2}\wedge\mathsf{E}_{\mathcal{G}_3}\wedge\boldsymbol{F}\vdash\mathcal{G}_F\big]&\geq\Big(1-\Pr[\mathsf{Bad}_3(\boldsymbol{F}_2,\boldsymbol{F}_5)]\Big)\cdot\frac{1}{N^{2|\mathcal{G}_4|}}\\&\geq\Big(1-\frac{2\beta_3(q_f+q_e)}{N}-2\beta_3\varepsilon\Big)\cdot\frac{1}{N^{2|\mathcal{G}_4|}}.\end{aligned}$$

Finally, we sum up all four events, i.e.,

$$\begin{aligned}\mathsf{p}(\tau,\boldsymbol{F}_1,\boldsymbol{F}_6)&=\Pr\big[\mathsf{E}_{\mathcal{G}_1}\wedge\mathsf{E}_{\mathcal{G}_2}\wedge\mathsf{E}_{\mathcal{G}_3}\wedge\mathsf{E}_{\mathcal{G}_4}|\boldsymbol{F}\vdash\mathcal{G}_F\big]\\&\geq(1-\theta_1)(1-\theta_2)(1-\theta_3)\frac{1}{N^{2(|\mathcal{G}_1|+|\mathcal{G}_2|+|\mathcal{G}_3|+|\mathcal{G}_4|)}}\\&\geq(1-(\theta_1+\theta_2+\theta_3))\frac{1}{N^{2q_e}},\end{aligned}$$

where θ_1, θ_2, θ_3 are (A1), (A2) and (A3) respectively, furthermore $|\mathcal{G}_1|+|\mathcal{G}_2|+|\mathcal{G}_3|+|\mathcal{G}_4|=q_e$. We note that

$$\frac{1}{N^{2q_e}}\Big/\prod_{i=0}^{q_e-1}\frac{1}{N^2-i}\geq(1-\frac{q_e}{N^2})^{q_e}\geq1-\frac{q_e^2}{N^2}\geq1-\frac{q_e^3}{N^2},$$

then for (3), we have

$$\begin{aligned}\mathbb{E}_k[\varphi(\boldsymbol{F}_1,\boldsymbol{F}_6,k)]\leq&\frac{(3q_e+2q_f)(\beta_1+\beta_2)+2\beta_3(q_e+q_f)}{N}\\&+\frac{2q_e(q_e+2q_f)(q_e+q_f)+q_e^3}{N^2}+\frac{q_eq_f^2}{N^2}+(\beta_1+\beta_2+2\beta_3)\varepsilon.\end{aligned}$$

We know that β_1, β_2, and β_3 depend on $(\boldsymbol{F}_1,\boldsymbol{F}_6)$. We consider them respectively, focusing on β_1 firstly. For each $(t,RX,AS)\in\mathcal{Q}_E^*(\boldsymbol{F}_1,\boldsymbol{F}_6)$, if $H_{k_1}(t)\oplus R\in Dom\mathcal{F}_1$, then it must be $H_{k_2}(t)\oplus X\notin Dom\mathcal{F}_2$ because of ¬**(A-2)**. Thus, on the condition of $\boldsymbol{F}_1\vdash\mathcal{Q}_{F_1}$, $\boldsymbol{F}_1(H_{k_1}(t)\oplus R)$ keeps uniform, then we have

$$\Pr\big[H_{k_2}(t)\oplus L\oplus\boldsymbol{F}_1(H_{k_1}(t)\oplus R)\in Dom\mathcal{F}_2\big]\leq\frac{q_f}{N}.$$

Therefore, $\mathbb{E}_k[\beta_1]\leq\frac{q_eq_f}{N}$. The analysis method of β_2 is symmetric with β_1, by the uniformity of $\boldsymbol{F}_6$, we have $\mathbb{E}_k[\beta_2]\leq\frac{q_eq_f}{N}$.

To this end, we consider β_3. For the fixed transcript (t, LR, ST) such that $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$, give a distinct $(t', L'R', S'T')$. If $t \neq t'$ but $LR = L'R'$, for the uniformity of H, we have

$$\Pr\left[X = X'\right] = \Pr\left[L \oplus \boldsymbol{F}_1(H_{k_1}(t) \oplus R) = L' \oplus \boldsymbol{F}_1(H_{k_1}(t') \oplus R')\right] \leq \varepsilon;$$

if $t = t'$ and $R = R'$, then it must be $L \neq L'$, thus $X = X'$ is impossible; if $t = t'$ and $L = L'$ but $R \neq R'$, on account of $H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1$, then $\boldsymbol{F}_1(H_{k_1}(t) \oplus R)$ keeps uniformly random conditioned on $\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}$, therefore $\Pr[X = X'] = \frac{1}{N}$. In addition, the choices of distinct pairs (t, LR, ST) and $(t', L'R', S'T')$ are at most q_e^2. Thus Thus, we have

$$\begin{aligned}\mathbb{E}_k\left[\left|\{(t, LR, ST) : H_{k_1}(t) \oplus R \notin Dom\mathcal{F}_1, \text{ and } \exists (t', L'R', S'T') \text{ s.t. } X = X'\}\right|\right] \\ \leq q_e^2\varepsilon + q_e^2(1-\varepsilon)\frac{1}{N} \leq q_e^2\varepsilon + \frac{q_e^2}{N}.\end{aligned}$$

For $H_{k_1}(t) \oplus R \in Dom\mathcal{F}_1$, the number of the transcripts (t, LR, ST) which meet the above conditions is $n_{(1)}(k)$. We have

$$\mathbb{E}_k\left[\left|\{(t, LR, ST) : \exists (t', L'R', S'T') \text{ s.t. } X = X'\}\right|\right] \leq \frac{q_e^2}{N} + q_e^2\varepsilon + n_{(1)}(k).$$

Symmetrically,

$$\mathbb{E}_k\left[\left|\{(t, LR, ST) : \exists (t', L'R', S'T') \text{ s.t. } A = A'\}\right|\right] \leq \frac{q_e^2}{N} + q_e^2\varepsilon + n_{(6)}(k).$$

Thus, we have

$$\mathbb{E}_k[\beta_3] \leq \frac{2q_e^2}{N} + 2q_e^2\varepsilon + n_{(1)}(k) + n_{(6)}(k).$$

Finally, $H_{k_1}(t)$ and $H_{k_6}(t)$ are uniform in 2^n possible choices,

$$\mathbb{E}_k\left[n_{(1)}(k)\right] = \mathbb{E}_k\left[n_{(6)}(k)\right] = \sum_{(t,LR,ST)\in\mathcal{Q}_E} \sum_{(x_1,y_1)\in\mathcal{Q}_{F_1}} \Pr\left[H_{k_1}(t) = R \oplus x_1\right] \leq \frac{q_e q_f}{N}.$$

Gathering all the above yields, we have

$$\begin{aligned}\mathbb{E}_{\boldsymbol{F}_1,\boldsymbol{F}_6,k}[\varphi(\boldsymbol{F}_1, \boldsymbol{F}_6, k)] \leq & \frac{6q_e^2q_f + 4q_eq_f^2}{N^2} + \frac{4q_e(q_e+q_f)^2}{N^2} \\ & + \frac{2q_e(q_e+2q_f)(q_e+q_f) + q_e^3}{N^2} + \frac{q_eq_f^2}{N^2} \\ & + \frac{4q_e^2(q_e+q_f)\varepsilon}{N} + \frac{2q_eq_f\varepsilon}{N} \\ & + \frac{4q_e(q_e+q_f)\varepsilon}{N} + 4q_e^2\varepsilon^2 \\ & = 4q_e^2\varepsilon^2 + \frac{7q_e^3 + 20q_e^2q_f + 12q_eq_f^2}{N^2} \\ & + \frac{4q_e^3\varepsilon + 4q_e^2q_f\varepsilon + 4q_e^2\varepsilon + 6q_eq_f\varepsilon}{N},\end{aligned}$$

as claimed in (4). □

Now we have Lemma 2, Lemma 4, and (2), we obtain

$$\frac{\Pr_{re}(\tau)}{\Pr_{id}(\tau)} \geq 1 - \Big(\frac{3q_e q_f^2}{N^2} + \mathbb{E}_k\big[\Pr\big[\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)|\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}, \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}\big]\big] + \mathbb{E}_k\big[\mathbb{E}_{\boldsymbol{F}_1,\boldsymbol{F}_6}\big[\varphi(\boldsymbol{F}_1, \boldsymbol{F}_6, k)|\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}, \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}\big]\big]\Big).$$

For the expectation $\mathbb{E}_k\big[\Pr\big[\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)|\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}, \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}\big]\big]$, we note that k_3 and k_4 are uniformly picked from 2^n possibilities, then

$$\mathbb{E}_k\Big[n_{(2,3)}(k)\Big] = \mathbb{E}_k\Big[n_{(4,5)}(k)\Big] \leq \frac{q_e q_f^2}{N}.$$

It has been shown that $\mathbb{E}_k\Big[n_{(1)}(k)\Big] = \mathbb{E}_k\Big[n_{(6)}(k)\Big] \leq \frac{q_e q_f}{N}$. Then Lemma 3 yields

$$\mathbb{E}_k\big[\Pr\big[\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)|\boldsymbol{F}_1 \vdash \mathcal{Q}_{F_1}, \boldsymbol{F}_6 \vdash \mathcal{Q}_{F_6}\big]\big] \leq \frac{4q_e^2 q_f + 5q_e q_f^2}{N^2} + \frac{2q_e q_f^2 \varepsilon}{N} + 4q_e^2 q_f \varepsilon^2.$$

From all above, by Lemmas 1 and 2, we have proved the conclusion of Theorem 1.

5. Conclusions and Future Work

This paper presents a result of constructing a tweakable block cipher from the KAF construction. Our work is based on based on the study by Guo et al. [24], we introduce the tweak into their optimized 6-round scheme KAF in order to achieve the Beyond Birthday-Bound security. We utilize a universal hash function which is called ε-almost XOR-universal hash function, with tweak and round-key vector, we rebuild a new tweakable KAF scheme TKAF which meets the security of beyond birthday-bound. Finally Finally, by using the H-coefficient technique [25], we prove the security requirement and obtain a better conclusion with fewer rounds. Our approach is to introduce the tweak into the first and last two rounds of Guo's 6-round KAF structure, and utilize the universal hash function as the operation method. Can we introduce the tweak directly into the round function without using the universal hash function, and still meeting the beyond birthday-bound security? Or can we use another linear method to introduce a tweak? We leave these as future work.

Author Contributions: Conceptualization, M.J. and L.W.; methodology, M.J.; validation, M.J. and L.W.; formal analysis, M.J.; investigation, M.J.; writing—original draft preparation, M.J.; writing—review and editing, M.J. and L.W.; supervision, L.W.; project administration, L.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Key Research and Development Program of China No. 2018YFB0803400.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to thank Yaobin Shen for some valuable guidance and advice. At the same time, thanks to other students in our laboratory for their great help in writing the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Lemma 5

Appendix A.1. $\Pr\left[\mathsf{E}_{\mathcal{G}_1}|\boldsymbol{F}\vdash\mathcal{Q}_F\right]$

Firstly, we consider the event $\mathsf{E}_{\mathcal{G}_1}$, i.e., lower bounding $\Pr\left[\mathsf{E}_{\mathcal{G}_1}|\boldsymbol{F}\vdash\mathcal{Q}_F\right]$. By the definition, there must be $\mathsf{E}_{\mathcal{G}_1}=\mathsf{E}_{|\mathcal{G}_1|}\wedge\ldots\wedge\mathsf{E}_1$. So, Therefore, we consider to lowering bound the probability of E_{l+1} on the condition of $\mathsf{E}_l\wedge\ldots\wedge\mathsf{E}_1\wedge\boldsymbol{F}\vdash\mathcal{Q}_F$.

We note that on the condition of $\mathsf{E}_l\wedge\ldots\wedge\mathsf{E}_1$, for arbitrary $x_3\in ExtF_3^{(l)}$ and $x_4\in ExtF_4^{(l)}$, $F_3(x_3)$ and $F_4(x_4)$ will be considered to be "fixed". For convenience, we denote $x_2^{(l+1)}=H_{k_2}(t)\oplus X_{l+1}$ and $x_5^{(l+1)}=H_{k_5}(t)\oplus A_{l+1}$, furthermore denote $Y_{l+1}=R_{l+1}\oplus\boldsymbol{F}_2(x_2^{(l+1)})$ and $Z_{l+1}=S_{l+1}\oplus\boldsymbol{F}_5(x_5^{(l+1)})$. Depending on the states of two intermediate values Y_{l+1} and Z_{l+1}, we consider the event E_{l+1} in three cases:

- CASE 1-no collision: Y_{l+1} and Z_{l+1} satisfy

$$k_3\oplus Y_{l+1}\notin Dom\mathcal{F}_3\cup Ext\mathcal{F}_3^{(l)}\cup\mathcal{G}_2\mathcal{F}_3$$
$$k_4\oplus Z_{l+1}\notin Dom\mathcal{F}_4\cup Ext\mathcal{F}_4^{(l)}\cup\mathcal{G}_3\mathcal{F}_4.$$

  It holds $\boldsymbol{F}_3(k_3\oplus Y_{l+1})=X_{l+1}\oplus Z_{l+1}$ and $\boldsymbol{F}_4(k_4\oplus Z_{l+1})=Y_{l+1}\oplus A_{l+1}$;
- CASE 2-left collision: Y_{l+1} satisfies $k_3\oplus Y_{l+1}\in Dom\mathcal{F}_3\cup Ext\mathcal{F}_3^{(l)}$, but Z_{l+1} satisfies $k_4\oplus Z_{l+1}\notin Dom\mathcal{F}_4\cup Ext\mathcal{F}_4^{(l)}\cup\mathcal{G}_3\mathcal{F}_4$. It holds $\boldsymbol{F}_4(k_4\oplus Z_{l+1})=Y_{l+1}\oplus A_{l+1}$ and $\boldsymbol{F}_5(x_5^{(l+1)})=Z_{l+1}\oplus S_{l+1}$;
- CASE 3-right collision: Z_{l+1} satisfies $k_4\oplus Z_{l+1}\in Dom\mathcal{F}_4\cup Ext\mathcal{F}_4^{(l)}$, but Y_{l+1} satisfies $k_3\oplus Y_{l+1}\notin Dom\mathcal{F}_3\cup Ext\mathcal{F}_3^{(l)}\cup\mathcal{G}_2\mathcal{F}_3$. It holds $\boldsymbol{F}_2(x_2^{(l+1)})=Y_{l+1}\oplus R_{l+1}$ and $\boldsymbol{F}_3(k_3\oplus Y_{l+1})=Z_{l+1}\oplus X_{l+1}$.

By these, accumulating all probabilities of above three cases, we have

$$\Pr[\mathsf{E}_{l+1}|\mathsf{E}_l\wedge\ldots\wedge\mathsf{E}_1\wedge\boldsymbol{F}\vdash\mathcal{Q}_F].$$

Now we consider these three cases respectively.

Appendix A.1.1. Case 1

For $(t,R_{l+1}X_{l+1},A_{l+1}S_{l+1})\in\mathcal{G}_1$, by definition, we have $x_2^{(l+1)}=H_{k_2}(t)\oplus X_{l+1}\notin Dom\mathcal{F}_2$. With tuples in $\mathcal{Q}_E^*(\boldsymbol{F}_1,\boldsymbol{F}_6)$, X_{l+1} does not collide with other corresponding positions since $|\mathcal{ID}(X_{l+1})|=1$. Thus Thus, $\boldsymbol{F}_2(x_2^{(l+1)})$ remains uniformly random on the condition of $\mathsf{E}_l\wedge\ldots\wedge\mathsf{E}_1\wedge\boldsymbol{F}\vdash\mathcal{Q}_F$. Moreover, $\Pr\left[k_3\oplus Y_{l+1}\in Dom\mathcal{F}_3\cup Ext\mathcal{F}_3^{(l)}\cup\mathcal{G}_2\mathcal{F}_3\right]\leq\frac{q_f+e_3^{(l+1)}+|\mathcal{G}_2\mathcal{F}_3|}{N}$. Symmetrically, we have $\Pr\left[k_4\oplus Z_{l+1}\in Dom\mathcal{F}_4\cup Ext\mathcal{F}_4^{(l)}\cup\mathcal{G}_3\mathcal{F}_4\right]\leq\frac{q_f+e_4^{(l+1)}+|\mathcal{G}_3\mathcal{F}_4|}{N}$. Then, the probability that these two equations $\boldsymbol{F}_3(k_3\oplus Y_{l+1})=X_{l+1}\oplus Z_{l+1}$ and $\boldsymbol{F}_4(k_4\oplus Z_{l+1})=Y_{l+1}\oplus A_{l+1}$ are simultaneously fulfilled is $\frac{1}{N^2}$.

From above,

$$\Pr[\mathsf{E}_{l+1}\wedge\text{CASE 1}|\mathsf{E}_l\wedge\ldots\wedge\mathsf{E}_1\wedge\boldsymbol{F}\vdash\mathcal{Q}_F]$$
$$\geq\left(1-\frac{q_f+e_3^{(l+1)}+|\mathcal{G}_2\mathcal{F}_3|}{N}\right)\left(1-\frac{q_f+e_4^{(l+1)}+|\mathcal{G}_3\mathcal{F}_4|}{N}\right)\frac{1}{N^2}.$$

Appendix A.1.2. Case 2

We consider the opposite case of CASE 2, and upper- bound the probability on this condition. Let pcoll be the probability of the contrary case. We have

$$\mathsf{pcoll} = \Pr\Big[\exists x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)},\ \exists x_4 \in Dom\mathcal{F}_4 \cup Ext\mathcal{F}_4^{(l)} \cup \mathcal{G}_3\mathcal{F}_4 : \mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F],$$

where $\mathsf{Coll}(x_3, x_4)$ stands for the collision event

$$X_{l+1} \oplus y_3 = (k_4 \oplus x_4) \wedge R_{l+1} \oplus \boldsymbol{F}_2(x_2^{(l+1)}) = k_3 \oplus x_3,$$

Then, we consider five subcases of the opposite CASE 2 respectively, and upper- bound for each in turn.

- **Subcase 2.1:** $x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)}$, *and* $x_4 \in \mathcal{G}_3\mathcal{F}_4$.
  For each $x_4 \in \mathcal{G}_3\mathcal{F}_4$, by definition, we have the number of $x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)}$ which satisfies the collision $X_{l+1} \oplus y_3 = k_4 \oplus x_4$ is $\sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)$. In addition, similar with CASE 1, we can still deem $\boldsymbol{F}_2(x_2^{(l+1)})$ as uniformly random. Thus, it holds $\Pr\Big[\boldsymbol{F}_2(x_2^{(l+1)}) = R_{l+1} \oplus k_3 \oplus x_3\Big] \leq \frac{1}{N}$. Therefore, the upper bound of **Subcase 2.1** is

$$\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)} \\ x_4 \in \mathcal{G}_3\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \leq \frac{\sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)}{N}.$$

- **Subcase 2.2:** $x_3 \in Dom\mathcal{F}_3$, *and* $x_4 \in Dom\mathcal{F}_4$. Define a key-dependent value:

$$\mathsf{Num}_{3,4}^{+}(k, X) \stackrel{def}{=} |\{((x_3, y_3), (x_4, y_4)) \in \mathcal{Q}_{F_3} \times \mathcal{Q}_{F_4} : k_4 = X \oplus y_3 \oplus x_4\}|.$$

  Then we have the quantity of (x_3, x_4) which satisfies the collision condition $X_{l+1} \oplus y_3 = k_4 \oplus x_4$ is $\mathsf{Num}_{3,4}^{+}(k, X_{l+1})$. Same as **Subcase 2.1**,

$$\Pr\Big[\boldsymbol{F}_2(x_2^{(l+1)}) = R_{l+1} \oplus k_3 \oplus x_3\Big] \leq \frac{1}{N}.$$

  Thus, we have

$$\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ x_4 \in Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \leq \frac{\mathsf{Num}_{3,4}^{+}(k, X_{l+1})}{N}.$$

  It can be seen, k_4 is uniform in N values. So, the expectation of $\mathsf{Num}_{3,4}^{+}(k, X_{l+1})$ is at most $\frac{q_f^2}{N}$. Thus Thus, the upper bound of the probability on the condition of **Subcase 2.2** is at most $\frac{q_f^2}{N^2}$.
- **Subcase 2.3:** $x_3 \in Dom\mathcal{F}_3$, *and* $x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4$. By definition, we write

$$\begin{aligned} &\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \\ = &\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ i = 1, \ldots, l}} \mathsf{sgn}(i) \cdot \Pr\Big[\mathsf{Coll}(x_3, x_4^{(i)}) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big], \end{aligned}$$

where $x_4^{(i)} = k_4 \oplus Z_i$, and for i-th tuple $(t, R_i X_i, A_i S_i) \in \mathcal{G}_4$, we have $Z_i = S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)$. In addition, $\mathsf{sgn}(i) = 1$ if and only if i is the smallest index that satisfies $x_4^{(i)} \in Ext\mathcal{F}_4^{(i)} \backslash Dom\mathcal{F}_4$, since $x_4^{(i)} \notin Ext\mathcal{F}_4^{(i-1)}$.

First, we focus on $\Pr\left[\mathsf{Coll}(x_3, x_4^{(i)}) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\right]$. Considering the probability on the condition that E_i fits into CASE 1,2 and 3. It can be seen that if E_i fits into CASE 3, then we have $x_4^{(i)} \in Dom\mathcal{F}_4$, it contradicts the **Subcase 2.3**. Let $y_3 = ImgF_3(x_3)$, write $Y_i = R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i)$.

(i) E_i fits into CASE 1

We derive Z_i from $Z_i = S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)$, and $\boldsymbol{F}_5(x_5^{(i)})$ keeps uniform. Then we have

$$\begin{aligned}\Pr\left[X_{l+1} \oplus y_3 = (k_4 \oplus x_4^{(i)})\right] &= \Pr[X_{l+1} \oplus y_3 = Z_i] \\ &= \Pr\left[\boldsymbol{F}_5(x_5^{(i)}) = X_{l+1} \oplus y_3 \oplus S_i\right] \leq \frac{1}{N}.\end{aligned}$$

Furthermore, we have $\Pr\left[\boldsymbol{F}_2(x_2^{(i)}) = R_{l+1} \oplus k_3 \oplus x_3\right] \leq \frac{1}{N}$. Thus

$$\Pr\left[\mathsf{Coll}(x_3, x_4^{(i)}) | \mathsf{E}_i \; fits \; into \; \text{CASE } 1 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F)\right] \leq \frac{1}{N^2}.$$

(ii) E_i fits into CASE 2

Let $x_3^{(i)} = k_3 \oplus Y_i$, $y_3^{(i)} = \boldsymbol{F}_3(x_3^{(i)})$. We have $X_{l+1} \oplus y_3 = Z_i = X_i \oplus y_3^{(i)}$. By definition, the number of choices for such $y_3^{(i)}$ is $\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3 \oplus X_i)$. Furthermore, for these choices of $y_3^{(i)}$, the probability of the following two collisions is at most $\frac{1}{N}$, i.e.,

$$R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i) = k_3 \oplus x_3^{(i)}, \; R_{l+1} \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_{l+1}) = k_3 \oplus x_3.$$

Thus

$$\begin{aligned}&\Pr\left[\mathsf{Coll}(x_3, x_4^{(i)}) | \mathsf{E}_i \; fits \; into \; \text{CASE } 2 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F)\right] \\ &\qquad \leq \frac{\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3 \oplus X_i)}{N^2}.\end{aligned}$$

From the above,

$$\begin{aligned}&\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \\ &\leq \sum_{x_3 \in Dom\mathcal{F}_3} \sum_{i=,\ldots,l} \mathsf{sgn}(i) \cdot \frac{\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3 \oplus X_i)}{N^2} \\ &\leq \sum_{x_3 \in Dom\mathcal{F}_3} \frac{q_f + e_3^{(l)}}{N^2} \leq \frac{q_f(q_f + q_e)}{N^2}.\end{aligned}$$

- **Subcase 2.4:** $x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3$, *and* $x_4 \in Dom\mathcal{F}_4$. By definition, we write

$$\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3 \\ x_4 \in Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]$$
$$= \sum_{\substack{i=1,\ldots,l \\ x_4 \in Dom\mathcal{F}_4}} \mathsf{sgn}'(i) \cdot \Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big],$$

where $x_3^{(i)} = k_3 \oplus Y_i$, and for i-th tuple $(t, R_i X_i, A_i S_i) \in \mathcal{G}_4$, we have $Y_i = R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i)$. In addition, $\mathsf{sgn}'(i) = 1$ if and only if i is the smallest index that satisfies $x_3^{(i)} \in Ext\mathcal{F}_3^{(i)} \backslash Dom\mathcal{F}_3$, since $x_3^{(i)} \notin Ext\mathcal{F}_3^{(i-1)}$.
First, we focus on $\Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big]$. Let $y_4 = ImgF_4(x_4)$, write $Z_i = S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)$. That is $y_3^{(i)} = X_i \oplus Z_i$. Thus, the collision $X_{l+1} \oplus y_3^{(i)} = (k_4 \oplus x_4)$ can be seen as $X_{l+1} \oplus X_i = Z_i \oplus (k_4 \oplus x_4)$. Same as **Subcase 2.3**, we only need to consider two cases on E_i.

(i) E_i fits into CASE 1
We know that $Z_i = S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)$ and $\boldsymbol{F}_5(H_{k_5}(t) \oplus A_i)$ keep uniform. Then it holds

$$\Pr[X_{l+1} \oplus X_i = Z_i \oplus (k_4 \oplus x_4)]$$
$$= \Pr\big[\boldsymbol{F}_5(H_{k_5}(t) \oplus A_i) = S_i \oplus X_{l+1} \oplus X_i \oplus k_4 \oplus x_4\big] \leq \frac{1}{N}.$$

Then, we have $\Pr\Big[\boldsymbol{F}_2(H_{k_2}(t) \oplus X_{l+1}) = R_{l+1} \oplus k_3 \oplus x_3^{(i)}\Big] \leq \frac{1}{N}$. Thus

$$\Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4) | \mathsf{E}_i \textit{ fits into } \textsc{Case } 1 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F)\Big] \leq \frac{1}{N^2}.$$

(ii) E_i fits into CASE 3
Let $x_4^{(i)} = k_4 \oplus Z_i$. We have $X_{l+1} \oplus X_i = x_4^{(i)} \oplus x_4$ because of $X_{l+1} \oplus X_i = Z_i \oplus k_4 \oplus x_4$. We note that if X_{l+1}, x_i and x_4 are "fixed", then the possibility of choices of $x_4^{(i)}$ is at most 1. Therefore, if Y_{l+1} collides with $x_3^{(i)}$, the following two collisions have to happen:

$$S_i \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_i) = k_4 \oplus x_4^{(i)}, R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_{l+1}) = k_3 \oplus x_3^{(i)}.$$

Thus

$$\Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4) | \mathsf{E}_i \textit{ fits into } \textsc{Case } 3 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big] \leq \frac{1}{N^2}.$$

According to **Subcase 2.3**, we have

$$\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3 \\ x_4 \in Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]$$
$$\leq \sum_{i=,\ldots,l;\ x_4 \in Dom\mathcal{F}_4} \mathsf{sgn}'(i) \cdot \frac{1}{N^2} \leq \frac{q_f e_3^{(l)}}{N^2} \leq \frac{q_f q_e}{N^2}.$$

- **Subcase 2.5:** $x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3$, *and* $x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4$. By definition, we write

$$\begin{aligned} &\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \\ &= \sum_{i,j=1,\ldots,l} \mathsf{sgn}(i) \cdot \mathsf{sgn}'(j) \cdot \Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4^{(j)}) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big], \end{aligned}$$

where $x_3^{(i)} = k_3 \oplus Y_i$, and for i-th tuple $(t, R_i X_i, A_i S_i) \in \mathcal{G}_4$, we have $Y_i = R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i)$. In addition, $\mathsf{sgn}(i) = 1$ if and only if i is the smallest index that satisfies $x_3^{(i)} \in Ext\mathcal{F}_3^{(i)} \backslash Dom\mathcal{F}_3$, since $x_3^{(i)} \notin Ext\mathcal{F}_3^{(i-1)}$. In addition, $x_4^{(j)} = k_4 \oplus Z_j$, and for j-th tuple $(t, R_j X_j, A_j S_j) \in \mathcal{G}_4$, we have $Z_j = S_j \oplus \boldsymbol{F}_5(H_{k_5}(t) \oplus A_j)$. In addition, $\mathsf{sgn}'(j) = 1$ if and only if j is the smallest index that satisfies $x_4^{(j)} \in Ext\mathcal{F}_4^{(j)} \backslash Dom\mathcal{F}_4$, since $x_4^{(j)} \notin Ext\mathcal{F}_4^{(j-1)}$.

  (i) When $j > i$, due to $X_{l+1} \oplus y_3^{(i)} = X_j \oplus y_3^{(j)}$, according to **Subcase 2.3**, the number of choices for such $y_3^{(i)}$ is $\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3^{(i)} \oplus X_j)$. Furthermore, for each $(x_3^{(i)}, x_4^{(j)})$, the upper bound of the probability is $\frac{\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3^{(i)} \oplus X_j)}{N^2}$.

  (ii) When $i > j$, due to $X_{l+1} \oplus X_i = x_4^{(i)} \oplus x_4^{(j)}$, according to **Subcase 2.4**, the upper bound of the probability for each $(x_3^{(i)}, y_3^{(j)})$ is $\frac{1}{N^2}$.

  To sum up,

$$\begin{aligned} &\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \\ &\leq \sum_{j=1,\ldots,l} \sum_{i=1,\ldots,l} \frac{\mathsf{Num}_3^{(l)}(X_{l+1} \oplus y_3^{(i)} \oplus X_j)}{N^2} \\ &\leq \sum_{j=1,\ldots,l} \frac{q_f + e_3^{(l)}}{N^2} \leq \frac{q_e(q_f + q_e)}{N^2}. \end{aligned}$$

- **Summing over all five subcases:** We have

$$\mathbb{E}_k[\mathsf{pcoll}] \leq \sum_{x_4 \in \mathcal{G}_3 \mathcal{F}_4} \frac{\mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)}{N} + \frac{q_f^2}{N^2} + \frac{(2q_f + q_e)(q_f + q_e)}{N^2}.$$

  The five cases above are opposite conditions to CASE 2. Moreover, if it holds $(t, R_{l+1} X_{l+1}, A_{l+1} S_{l+1}) \in \mathcal{G}_1$, then we have (i) $x_5^{(l+1)} \notin Dom\mathcal{F}_5$, (ii) $|\mathcal{ID}(A_{(l+1)})| = 1$, that implies the position of $x_5^{(l+1)}$ can be deemed as "new".
  For these arguments above, we have

$$\begin{aligned} &\mathbb{E}_k\Big[\Pr[\mathsf{E}_{l+1} \wedge \textsc{Case 2} | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]\Big] \\ &\geq \Big(\frac{q_f + e_3^{(l)}}{N} - \frac{\sum_{x_4 \in \mathcal{G}_3 \mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)}{N} \\ &- \frac{q_f^2}{N^2} - \frac{(2q_f + q_e)(q_f + q_e)}{N^2}\Big) \frac{1}{N^2}. \end{aligned}$$

Appendix A.1.3. Case 3

In this case, if it holds $x_2^{(l+1)} \notin Dom\mathcal{F}_2$ and $x_3 = k_3 \oplus Y_{l+1} \notin Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)} \cup \mathcal{G}_2\mathcal{F}_3$, then we have

$$\begin{aligned}&\Pr[\mathsf{TKAF}\ extends\ (t, R_{l+1}X_{l+1}, A_{l+1}S_{l+1})]\\ &= \Pr\Big[\boldsymbol{F}_2(x_2^{(l+1)}) = R_{l+1} \oplus Y_{l+1} \wedge \boldsymbol{F}_3(x_3) = X_{l+1} \oplus Z_{l+1}\Big] = \frac{1}{N_2}\end{aligned}$$

With the similar analysis of CASE 2, we denote

$$\mathsf{pcoll} = \sum_{\substack{x_3 \in Dom\mathcal{F}_3 \cup Ext\mathcal{F}_3^{(l)} \cup \mathcal{G}_2\mathcal{F}_3\\ x_4 \in Dom\mathcal{F}_4 \cup Ext\mathcal{F}_4^{(l)}}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F].$$

Also, we consider five subcases in turn.

- **Subcase 3.1:** $x_3 \in \mathcal{G}_2\mathcal{F}_3$, *and* $x_4 \in Dom\mathcal{F}_4 \cup Ext\mathcal{F}_4^{(l)}$. On this condition, as the constraint $A_{l+1} \oplus y_4 = k_3 \oplus x_3$, we have

$$\begin{aligned}&\sum_{\substack{x_3 \in \mathcal{G}_2\mathcal{F}_3\\ x_4 \in Dom\mathcal{F}_4 \cup Ext\mathcal{F}_4^{(l)}}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]\\ &\qquad \le \frac{\sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3)}{N},\end{aligned}$$

  where $\mathsf{Num}_4^{(l)}(y_4) = |\{x_4 \in Dom\mathcal{F}_4 \cup Ext\mathcal{F}_4^{(i)} : \boldsymbol{F}_4(x_4) = y_4\}|$.
- **Subcase 3.2:** $x_3 \in Dom\mathcal{F}_3$, *and* $x_4 \in Dom\mathcal{F}_4$. Define a key-dependent value:

$$\mathsf{Num}_{3,4}^{-}(k, A) \stackrel{def}{=} \Big|\{((x_3, y_3), (x_4, y_4)) \in \mathcal{Q}_{F_3} \times \mathcal{Q}_{F_4} : k_3 = A \oplus y_4 \oplus x_3\}\Big|.$$

  On account of the uniformity of k_3 in N choices, we have

$$\begin{aligned}&\mathbb{E}_k\Big[\sum_{x_3 \in Dom\mathcal{F}_3, x_4 \in Dom\mathcal{F}_4} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]\Big]\\ &\qquad \le \frac{\mathbb{E}_k\Big[\mathsf{Num}_{3,4}^{-}(k, A_{l+1})\Big]}{N} \le \frac{q_f^2}{N^2}.\end{aligned}$$

- **Subcase 3.3:** $x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3$, *and* $x_4 \in Dom\mathcal{F}_4$. By definition, we write

$$\begin{aligned}&\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3\\ x_4 \in Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]\\ &= \sum_{\substack{i = 1, \ldots, l\\ x_4 \in Dom\mathcal{F}_4}} \mathsf{sgn}'(i) \cdot \Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big],\end{aligned}$$

  where $x_3^{(i)} = k_3 \oplus Y_i$, and for i-th tuple $(t, R_iX_i, A_iS_i) \in \mathcal{G}_4$, we have $Y_i = R_i \oplus \boldsymbol{F}_2(H_{k_2}(t) \oplus X_i)$. In addition, $\mathsf{sgn}'(i) = 1$ if and only if i is the smallest index that satisfies $x_3^{(i)} \in Ext\mathcal{F}_3^{(i)} \backslash Dom\mathcal{F}_3$, since $x_3^{(i)} \notin Ext\mathcal{F}_3^{(i-1)}$. Similar with **Subcase 2.3**,

(i) E_i fits into CASE 1
We have

$$\Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4)|\mathsf{E}_i \textit{ fits into } \textsc{Case } 1 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F)\Big] \leq \frac{1}{N^2}.$$

(ii) E_i fits into CASE 3
We have $A_{l+1} \oplus y_4 = A_i \oplus y_4^{(i)}$. Therefore,

$$\Pr\Big[\mathsf{Coll}(x_3^{(i)}, x_4)|\mathsf{E}_i \textit{ fits into } \textsc{Case } 3 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F)\Big] \leq \frac{\mathsf{Num}_4^{(l)}(A_{l+1} \oplus y_4 \oplus A_i)}{N^2}.$$

From above with the similar calculation, we have

$$\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \setminus Dom\mathcal{F}_3 \\ x_4 \in Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]$$

$$\leq \sum_{x_4 \in Dom\mathcal{F}_4} \sum_{i=,\ldots,l} \mathsf{sgn}'(i) \cdot \frac{\mathsf{Num}_4^{(l)}(A_{l+1} \oplus y_4 \oplus A_i)}{N^2} \leq \frac{q_f(q_f + q_e)}{N^2}.$$

- **Subcase 3.4:** $x_3 \in Dom\mathcal{F}_3$, *and* $x_4 \in Ext\mathcal{F}_4^{(l)} \setminus Dom\mathcal{F}_4$. By definition, we write

$$\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \setminus Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]$$

$$= \sum_{\substack{x_4 \in Dom\mathcal{F}_4 \\ i=1,\ldots,l}} \mathsf{sgn}(i) \cdot \Pr\Big[\mathsf{Coll}(x_3, x_4^{(i)})|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big],$$

It also holds $\Pr\Big[\mathsf{Coll}(x_3, x_4^{(i)})|\mathsf{E}_i \textit{ fits into } \textsc{Case } 1 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big] \leq \frac{1}{N^2}$. When E_i fits into CASE 2, due to $A_{l+1} \oplus A_i = x_3^{(i)} \oplus x_3$, we have
$\Pr\Big[\mathsf{Coll}(x_3, x_4^{(i)})|\mathsf{E}_i \textit{ fits into } \textsc{Case } 2 \wedge \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\Big] = \frac{1}{N^2}$.
Therefore,

$$\sum_{\substack{x_3 \in Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \setminus Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4)|\mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \leq \frac{q_f e_4^{(l)}}{N^2} \leq \frac{q_f q_e}{N^2}.$$

- **Subcase 3.5:** $x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3$, *and* $x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4$. Similar to **Subcase 2.5**, we have

$$\begin{aligned}
&\sum_{\substack{x_3 \in Ext\mathcal{F}_3^{(l)} \backslash Dom\mathcal{F}_3 \\ x_4 \in Ext\mathcal{F}_4^{(l)} \backslash Dom\mathcal{F}_4}} \Pr[\mathsf{Coll}(x_3, x_4) | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F] \\
&\leq \sum_{j=1,\ldots,l} \sum_{i=1,\ldots,l} \frac{\mathsf{Num}_4^{(l)}(A_{l+1} \oplus y_4 \oplus A_j)}{N^2} \\
&\leq \sum_{j=1,\ldots,l} \frac{q_f + e_4^{(l)}}{N^2} \leq \frac{q_e(q_f + q_e)}{N^2}.
\end{aligned}$$

- **Summing over all five subcases:** We have

$$\begin{aligned}
&\mathbb{E}_k[\Pr[\mathsf{E}_{l+1} \wedge \textsc{Case } 3 | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]] \\
&\geq \Big(\frac{q_f + e_4^{(l)}}{N} - \frac{\sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3)}{N} - \frac{(2q_f + q_e)(q_f + q_e)}{N^2} \Big) \frac{1}{N^2}.
\end{aligned}$$

Appendix A.1.4. Conclusions of $\mathsf{E}_{\mathcal{G}_1}$

Summing over all the three cases:

$$\begin{aligned}
&\mathbb{E}_k\Big[\Pr[\mathsf{E}_{l+1} | \mathsf{E}_l \wedge \ldots \wedge \mathsf{E}_1 \wedge \boldsymbol{F} \vdash \mathcal{Q}_F]\Big] \\
&\geq \Big(\Big(1 - \frac{q_f + e_3^{(l)} + |\mathcal{G}_2\mathcal{F}_3|}{N}\Big) \Big(1 - \frac{q_f + e_4^{(l)} + |\mathcal{G}_3\mathcal{F}_4|}{N}\Big) \\
&+ \frac{2q_f + e_3^{(l)} + e_4^{(l)}}{N} - \frac{q_f^2}{N^2} - \frac{2(2q_f + q_e)(q_f + q_e)}{N^2} \\
&- \frac{1}{N} \sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4) - \frac{1}{N} \sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3) \Big) \frac{1}{N^2}.
\end{aligned}$$

We denote

$$B_l = \frac{1}{N} \Big(\sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4) + \sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3) \Big),$$

$|\mathcal{G}_2\mathcal{F}_3| \leq |\mathcal{G}_2| = \beta_1$, $|\mathcal{G}_3\mathcal{F}_4| \leq |\mathcal{G}_3| = \beta_2$, and $|\mathcal{G}_1| \leq q_e$. Then it holds

$$\begin{aligned}
&\mathbb{E}_k\Big[\Pr\big[\mathsf{E}_{\mathcal{G}_1} | \boldsymbol{F} \vdash \mathcal{Q}_F\big]\Big] \\
&\geq \prod_{l=0}^{|\mathcal{G}_1|-1} \Big(1 - \frac{q_f^2}{N^2} - \frac{2q_e(2q_f + q_e)(q_f + q_e)}{N^2} - \frac{|\mathcal{G}_2\mathcal{F}_3| + |\mathcal{G}_3\mathcal{F}_4|}{N} - B_l\Big) \cdot \frac{1}{N^{2|\mathcal{G}_1|}} \\
&\geq \Big(1 - \frac{q_e q_f^2}{N^2} - \frac{2q_e(2q_f + q_e)(q_f + q_e)}{N^2} - \frac{q_e(\beta_1 + \beta_2)}{N} - \sum_{l=0}^{q_e-1} B_l\Big) \cdot \frac{1}{N^{2|\mathcal{G}_1|}}.
\end{aligned}$$

Secondly, we consider $\sum_{l=0}^{q_e-1} B_l$. By definition, we have

$$\begin{aligned}
\sum_{y_3 \in \{0,1\}^n} \mathsf{Num}_3^{(l)}(y_3) = q_f + e_3^{(l)} \leq q_f + q_e \\
\sum_{y_4 \in \{0,1\}^n} \mathsf{Num}_4^{(l)}(y_4) = q_f + e_4^{(l)} \leq q_f + q_e.
\end{aligned}$$

Thus

$$\sum_{l=0}^{q_e-1} \sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} \mathsf{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4) \leq \sum_{x_4 \in \mathcal{G}_3\mathcal{F}_4} (q_f + q_e) \leq (q_f + q_e)\beta_2.$$

Similarly,

$$\sum_{l=0}^{q_e-1} \sum_{x_3 \in \mathcal{G}_2\mathcal{F}_3} \mathsf{Num}_4^{(l)}(A_{l+1} \oplus k_3 \oplus x_3) \leq (q_f + q_e)|\mathcal{G}_2\mathcal{F}_3| \leq (q_f + q_e)\beta_1.$$

Finally, we have the upper bound

$$\begin{aligned} \mathbb{E}_k\Big[\Pr\big[\mathsf{E}_{\mathcal{G}_1}|\boldsymbol{F} \vdash \mathcal{Q}_F\big]\Big] \geq \Big(1 - \frac{q_e q_f^2}{N^2} - \frac{2q_e(2q_f + q_e)(q_f + q_e)}{N^2} \\ - \frac{(q_f + 2q_e)(\beta_1 + \beta_2)}{N}\Big)\frac{1}{N^{2|\mathcal{G}_1|}}. \end{aligned} \tag{A1}$$

Appendix A.2. $\Pr\big[\mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3}|\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F} \vdash \mathcal{G}_F\big]$

Next, we analyze the event $\mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3}$, we firstly focus on $\mathsf{E}_{\mathcal{G}_2}$. Define the "bad" event on this condition, we denote by $\mathsf{Bad}_1(\boldsymbol{F}_3)$: there exists $(t, RX, AS) \in \mathcal{G}_2$, one of the following conditions is fulfilled:

(i) $x_4 = k_4 \oplus X \oplus \boldsymbol{F}_3(x_3) \in Dom\mathcal{F}_4$, where $x_3 = k_3 \oplus R \oplus ImgF_2(H_{k_2}(t) \oplus X)$;
(ii) there exists $(t', R'X', A'S') \in \mathcal{G}_2$, such that $X \oplus \boldsymbol{F}_3(x_3) = X' \oplus \boldsymbol{F}_3(x_3')$, where $x_3' = k_3 \oplus R' \oplus ImgF_2(H_{k_2}(t') \oplus X')$;
(iii) there exists $(t^*, R^*X^*, A^*S^*) \in \mathcal{G}_1 \cup \mathcal{G}_3$, such that $X \oplus \boldsymbol{F}_3(x_3) = S^* \oplus \boldsymbol{F}_5(H_{k_5}(t^*) \oplus A^*)$.

We note that for each $(t, RX, AS) \in \mathcal{G}_2$, let $x_3 = k_3 \oplus R \oplus ImgF_2(H_{k_2}(t) \oplus X)$, we have $x_3 \notin Dom\mathcal{F}_3$ (for the condition of $\neg\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$) and $x_3 \notin Ext\mathcal{F}_3^{|\mathcal{G}_1|}$ (for the analysis of $\mathsf{E}_{\mathcal{G}_1}$). Then, on the condition of $\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F}_3 \vdash \mathcal{G}_{\boldsymbol{F}_3}$, the values of function $\boldsymbol{F}_3(x_3)$ keep uniform. Thus, for (t, RX, AS):

(i) the probability of condition (i) fulfilled is at most $\frac{q_f}{N}$;
(ii) for each $(t', R'X', A'S') \in \mathcal{G}_2$, if the corresponding $x_3 \neq x_3'$, we have

$$\Pr\big[X \oplus \boldsymbol{F}_3(x_3) = X' \oplus \boldsymbol{F}_3(x_3')\big] \leq \frac{1}{N};$$

If the two tuples are distinct, i.e., $(t, RX, AS) \neq (t', R'X', A'S')$: (a) $t \neq t'$, $X = X'$, and $x_3 = x_3'$, then $\Pr[X \oplus \boldsymbol{F}_3(x_3) = X' \oplus \boldsymbol{F}_3(x_3')] \leq \varepsilon$; (b) if $t = t'$, $X \neq X'$, and $x_3 = x_3'$, then it must be $X \oplus \boldsymbol{F}_3(x_3) \neq X' \oplus \boldsymbol{F}_3(x_3')$.
(iii) for each $(t^*, R^*X^*, A^*S^*) \in \mathcal{G}_1 \cup \mathcal{G}_3$, we have

$$\Pr\big[X \oplus \boldsymbol{F}_3(x_3) = S^* \oplus \boldsymbol{F}_5(H_{k_5}(t^*) \oplus A^*)\big] \leq \frac{1}{N}.$$

Summing up the above, we have the probability of $\mathsf{Bad}_1(\boldsymbol{F}_3)$:

$$\begin{aligned} \Pr\big[\mathsf{Bad}_1(\boldsymbol{F}_3)|\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\big] &\leq \frac{|\mathcal{G}_2| \cdot (q_f + |\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|)}{N} + |\mathcal{G}_2| \cdot \varepsilon \\ &\leq \frac{\beta_1(q_f + q_e)}{N} + \beta_1 \cdot \varepsilon. \end{aligned}$$

We can see that if $\mathsf{Bad}_1(\boldsymbol{F}_3)$ does not happen, there are $|\mathcal{G}_2|$ values $Z_1, \ldots, Z_{|\mathcal{G}_2|}$ in $\mathcal{G}_2$ which are distinct (otherwise (ii) is fulfilled). In addition, $\boldsymbol{F}_4(k_4 \oplus Z_1), \ldots, \boldsymbol{F}_4(k_4 \oplus Z_{|\mathcal{G}_2|})$ are all undetermined (otherwise (i) and (iii) are fulfilled).

Moreover, at the "right" part, there are also $|\mathcal{G}_2|$ values $A_1, \ldots, A_{|\mathcal{G}_2|}$, such that $\boldsymbol{F}_5(H_{k_5}(t) \oplus A_1), \ldots, \boldsymbol{F}_5(H_{k_5}(t) \oplus A_{|}\mathcal{G}_2|)$ are also undetermined.

Therefore, the event $\mathsf{E}_{\mathcal{G}_2}$ is equivalent to $\boldsymbol{F}_4$ and $\boldsymbol{F}_5$ satisfying $2|\mathcal{G}_2|$ new equations, so the probability does not exceed $\frac{1}{N^{2|\mathcal{G}_2|}}$.

Similar to the analysis of $\mathsf{E}_{\mathcal{G}_2}$, we consider the event $\mathsf{E}_{\mathcal{G}_3}$. Likewise, we define the bad event $\mathsf{Bad}_2(\boldsymbol{F}_4)$ that there exists $(t, RX, AS) \in \mathcal{G}_3$, one of the following conditions is fulfilled:

(i) $x_3 = k_3 \oplus A \oplus \boldsymbol{F}_4(x_4) \in Dom\mathcal{F}_3$, where $x_4 = k_4 \oplus S \oplus Img F_5(H_{k_5}(t) \oplus A)$, the probability is at most $\frac{q_f}{N}$;

(ii) there exists $(t', R'X', A'S') \in \mathcal{G}_3$, such that $A \oplus \boldsymbol{F}_4(x_4) = A' \oplus \boldsymbol{F}_4(x'_4)$, where $x'_4 = k_4 \oplus S' \oplus Img F_5(H_{k_5}(t') \oplus A')$, and the probability is at most $\frac{|\mathcal{G}_3|}{N} + \varepsilon$;

(iii) there exists $(t^*, R^*X^*, A^*S^*) \in \mathcal{G}_1 \cup \mathcal{G}_2$, such that $A \oplus \boldsymbol{F}_4(x_4) = R^* \oplus \boldsymbol{F}_2(H_{k_2}(t^*) \oplus X^*)$, and the probability is at most $\frac{|\mathcal{G}_1|+|\mathcal{G}_2|}{N}$.

Thus, we have the probability of $\mathsf{Bad}_2(\boldsymbol{F}_4)$:

$$\begin{aligned}\Pr\left[\mathsf{Bad}_2(\boldsymbol{F}_4)|\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\right] &\leq \frac{|\mathcal{G}_3| \cdot (q_f + |\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|)}{N} + |\mathcal{G}_3| \cdot \varepsilon \\ &\leq \frac{\beta_2(q_f + q_e)}{N} + \beta_2 \cdot \varepsilon.\end{aligned}$$

Same as $\mathsf{E}_{\mathcal{G}_2}$, the event $\mathsf{E}_{\mathcal{G}_3}$ is equivalent to $\boldsymbol{F}_2$ and $\boldsymbol{F}_3$ satisfying $2|\mathcal{G}_3|$ new equations.

Therefore, on the condition of $\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F} \vdash \mathcal{Q}_F$, we have

$$\begin{aligned}\Pr\left[\mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3}|\mathsf{E}_{\mathcal{G}_1} \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\right] &\geq (1 - \Pr[\mathsf{Bad}_1(\boldsymbol{F}_3)] - \Pr[\mathsf{Bad}_2(\boldsymbol{F}_4)]) \\ &\quad \cdot \Pr\left[\mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3}|\neg\mathsf{Bad}_1(\boldsymbol{F}_3) \wedge \neg\mathsf{Bad}_1(\boldsymbol{F}_4)\right] \\ &\geq \left(1 - \frac{(\beta_1 + \beta_2)(q_f + q_e)}{N} - (\beta_1 + \beta_2)\varepsilon\right) \\ &\quad \cdot \frac{1}{N^{2(|\mathcal{G}_2|+|\mathcal{G}_3|)}}.\end{aligned} \tag{A2}$$

Appendix A.3. $\Pr\left[\mathsf{E}_{\mathcal{G}_4}|\mathsf{E}_{\mathcal{G}_1} \wedge \mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3} \wedge \boldsymbol{F} \vdash \mathcal{Q}_F\right]$

Thirdly, we analyze the event $\mathsf{E}_{\mathcal{G}_4}$. By definition, for arbitrary $(t, RX, AS) \in \mathcal{G}_4$, we denote $x_2 = H_{k_2}(t) \oplus X$ and $x_5 = H_{k_5}(t) \oplus A$ such that $x_2 \notin Dom\mathcal{F}_2$ and $x_5 \notin Dom\mathcal{F}_5$. Furthermore, on the condition of $\mathsf{E}_{\mathcal{G}_1} \wedge \mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3}$, and the conditions of bad event $\mathsf{Bad}(\boldsymbol{F}_1, \boldsymbol{F}_6)$, the two values of functions $\boldsymbol{F}_2(x_2)$ and $\boldsymbol{F}_5(x_5)$ must be uniform and undetermined.

We also define the bad event $\mathsf{Bad}_3(\boldsymbol{F}_2, \boldsymbol{F}_5)$ that there exists $(t, RX, AS) \in \mathcal{G}_4$, such that x_2 and x_5 fulfill one of following conditions:

- left part: consider $\boldsymbol{F}_2(x_2)$:
  - (i) $x_3 = k_3 \oplus R \oplus \boldsymbol{F}_2(x_2) \in Dom\mathcal{F}_3$, on account of the randomness of $\boldsymbol{F}_2(x_2)$, for each $(t, RX, AS) \in \mathcal{G}_4$, the probability of which is at most $\frac{q_f}{N}$;
  - (ii) there exists $(t', R'X', A'S') \in \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, such that $R \oplus \boldsymbol{F}_2(x_2) = R' \oplus \boldsymbol{F}_2(H_{k_2}(t') \oplus X')$. For distinct two tuples in $\mathcal{G}_4$, (a) it might be $t \neq t'$, such that Y collides with some "previously-ly determined" Y', the probability of which is ε; (b) if $t = t'$ but $X \neq X'$ (it can not cannot be $R = R'$), by the randomness of $\boldsymbol{F}_2(x_2)$, for each $(t, RX, AS) \in \mathcal{G}_4$, the upper bound of the probability is $\frac{|\mathcal{G}_1|+|\mathcal{G}_2|+|\mathcal{G}_3|}{N} + \varepsilon \leq \frac{q_e}{N} + \varepsilon$.
- right part: consider $\boldsymbol{F}_5(x_5)$, similar to the above:
  - (i) $k_4 \oplus S \oplus \boldsymbol{F}_5(x_5) \in Dom\mathcal{F}_4$, for each $(t, RX, AS) \in \mathcal{G}_4$, the probability of which is at most $\frac{q_f}{N}$;
  - (ii) there exists another distinct $(t', R'X', A'S') \in \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, such that $S \oplus \boldsymbol{F}_5(x_5) = S' \oplus \boldsymbol{F}_5(H_{k_5}(t') \oplus A')$. For each $(t, RX, AS) \in \mathcal{G}_4$, the upper bound of the probability is $\frac{|\mathcal{G}_1|+|\mathcal{G}_2|+|\mathcal{G}_3|}{N} + \varepsilon \leq \frac{q_e}{N} + \varepsilon$.

Thus, denote $|\mathcal{G}_4| = \beta_3$, we have

$$\begin{aligned}\Pr\left[\mathsf{E}_{\mathcal{G}_4} | \mathsf{E}_{\mathcal{G}_1} \wedge \mathsf{E}_{\mathcal{G}_2} \wedge \mathsf{E}_{\mathcal{G}_3} \wedge \boldsymbol{F} \vdash \mathcal{G}_F\right] &\geq \left(1 - \Pr[\mathsf{Bad}_3(\boldsymbol{F}_2, \boldsymbol{F}_5)]\right) \cdot \frac{1}{N^{2|\mathcal{G}_4|}} \\ &\geq \left(1 - \frac{2\beta_3(q_f + q_e)}{N} - 2\beta_3\varepsilon\right) \cdot \frac{1}{N^{2|\mathcal{G}_4|}}.\end{aligned} \tag{A3}$$

References

1. Liskov, M.; Rivest, R.L.; Wagner, D. Tweakable block ciphers. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 18–22 August 2002; Springer: Berlin/Heidelberg, Germany, 2002; pp. 31–46.
2. Landecker, W.; Shrimpton, T.; Terashima, R.S. Tweakable blockciphers with beyond birthday-bound security. In Proceedings of the Annual Cryptology Conference, Santa Barbara, CA, USA, 19–23 August 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 14–30.
3. Andreeva, E.; Bogdanov, A.; Luykx, A.; Mennink, B.; Tischhauser, E.; Yasuda, K. Parallelizable and authenticated online ciphers. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Bangalore, India, 1–5 December 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 424–443.
4. Rogaway, P. Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Jeju Island, Korea, 5–9 December 2004; Springer: Berlin/Heidelberg, Germany, 2004; pp. 16–31.
5. Rogaway, P.; Bellare, M.; Black, J. OCB: A block-cipher mode of operation for efficient authenticated encryption. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **2003**, *6*, 365–403. [CrossRef]
6. Crowley, P. Mercy: A fast large block cipher for disk sector encryption. In Proceedings of the International Workshop on Fast Software Encryption, New York, NY, USA, 10–12 April 2000; Springer: Berlin/Heidelberg, Germany, 2000; pp. 49–63.
7. Ferguson, N.; Lucks, S.; Schneier, B.; Whiting, D.; Bellare, M.; Kohno, T.; Callas, J.; Walker, J. The Skein hash function family. *NIST (Round 3)* **2010**, *7*, 3, submitted.
8. Schroeppel, R. Hasty pudding cipher specification. In Proceedings of the First AES Candidate Workshop, Ventura, CA, USA, 20–22 August 1998.
9. Cogliati, B.; Seurin, Y. Beyond-birthday-bound security for tweakable Even-Mansour ciphers with linear tweak and key mixing. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, 29 November–3 December 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 134–158.
10. Mennink, B. XPX: Generalized tweakable even-mansour with improved security guarantees. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 14–18 August 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 64–94.
11. Naito, Y. Tweakable blockciphers for efficient authenticated encryptions with beyond the birthday-bound security. *IACR Trans. Symmetric Cryptol.* **2017**, 1–26. [CrossRef]
12. Feistel H. Cryptography and computer privacy. *Sci. Am.* **1973**, *228*, 15–23. [CrossRef]
13. Even, S.; Mansour, Y. A construction of a cipher from a single pseudorandom permutation. *J. Cryptol.* **1997**, *10*, 151–161. [CrossRef]
14. Jean, J.; Nikolić, I.; Peyrin, T. Tweaks and keys for block ciphers: The TWEAKEY framework. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, 7–11 December 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 274–288.
15. Cogliati, B.; Lampe, R.; Seurin, Y. Tweaking even-mansour ciphers. In Proceedings of the Annual Cryptology Conference, Santa Barbara, CA, USA, 16–20 August 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 189–208.
16. Cogliati, B.; Seurin, Y. On the provable security of the iterated Even-Mansour cipher against related-key and chosen-key attacks. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, 26–30 April 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 584–613.
17. Farshim, P.; Procter, G. The related-key security of iterated Even—Mansour ciphers. In Proceedings of the International Workshop on Fast Software Encryption, Istanbul, Turkey, 8–11 March 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 342–363.
18. Granger, R.; Jovanovic, P.; Mennink, B.; Neves, S. Improved masking for tweakable blockciphers with applications to authenticated encryption. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, 8–12 May 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 263–293.
19. Luby, M.; Rackoff, C. How to construct pseudorandom permutations from pseudorandom functions. *SIAM J. Comput.* **1988**, *17*, 373–386. [CrossRef]
20. Mitsuda, A.; Iwata, T. Tweakable pseudorandom permutation from generalized feistel structure. In Proceedings of the International Conference on Provable Security, Shanghai, China, 30 October–1 November 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 22–37.
21. Lampe, R.; Seurin, Y. Security analysis of key-alternating Feistel ciphers. In Proceedings of the International Workshop on Fast Software Encryption, London, UK, 3–5 March 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 243–264.

22. Goldenberg, D.; Hohenberger, S.; Liskov, M.; Schwartz, E.C.; Seyalioglu, H. On tweaking luby-rackoff blockciphers. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Kuching, Malaysia, 2–6 December 2007; Springer: Berlin/Heidelberg, Germany, 2007; pp. 342–356.
23. Yan, H.; Wang, L.; Shen, Y.; Lai, X. Tweaking Key-Alternating Feistel Block Ciphers. In Proceedings of the International Conference on Applied Cryptography and Network Security, Rome, Italy, 9–22 October 2020; Springer: Cham, Switzerland, 2020; pp. 69–88.
24. Guo, C.; Wang, L. Revisiting key-alternating Feistel ciphers for shorter keys and multi-user security. In Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, 2–6 December 2018; Springer: Cham, Switzerland, 2018; pp. 213–243.
25. Patarin J. The "coefficients H" technique. In Proceedings of the International Workshop on Selected Areas in Cryptography, Sackville, NB, Canada, 14–15 August 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 328–345.
26. Chen, S.; Steinberger, J. Tight security bounds for key-alternating ciphers. In Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, 11–15 May 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 327–350.
27. Hoang, V.T.; Tessaro, S. Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In Proceedings of the Annual International Cryptology Conference, Santa Barbara, CA, USA, 14–18 August 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3–32.