Hindawi
Security and Communication Networks
Volume 2021, Article ID 2751797, 26 pages
https://doi.org/10.1155/2021/2751797

Research Article

New Approach towards Generalizing Feistel Networks and Its Provable Security

Jiajie Liu, Bing Sun, and Chao Li

College of Liberal Arts and Sciences, National University of Defense Technology, Changsha 410073, China

Correspondence should be addressed to Jiajie Liu; l.jiajie@yahoo.com

Received 28 May 2021; Accepted 21 August 2021; Published 23 September 2021

Academic Editor: Majid Khan

This paper proposes a new approach to generalizing Feistel networks, which unifies the classical (balanced) Feistel network and the Lai–Massey structure. We call the new structure extended Feistel (E-Feistel) network. To justify its soundness, we investigate its indistinguishability using Patarin's H-coefficient technique. As a result, it is proved that the 4-round key-alternating E-Feistel (KAEF) cipher with adequately derived keys and identical round functions is secure up to $2^{n/2}$ queries, i.e., birthday-bound security. In addition, when adjacent round keys are independent and independent round functions are used, the 6-round KAEF is secure up to beyond-birthday-bound $2^{2n/3}$ queries. Our results indicate that the E-Feistel structure is secure and reliable and can be adopted in designing practical block ciphers.

1. Introduction

In recent years, as a result of more attention paid on privacy protection and information security, research on the design and cryptanalysis of block ciphers has become a research hotpot. The design of block ciphers strictly highlights efficiency and security, which deeply rely on iterative structures they choose. According to whether encryption is consistent with decryption, iterative structures can be divided into two categories. The structures that have consistent encryption and decryption are beneficial in hardware implementation because decryption does not take up extra storage. This kind of structures contains the Feistel structure, the SM4 structure, the Mars structure, and the Lai–Massey structure as specific instances. Another kind of structures mainly consists of the substitution-permutation networks (SPN).

The Feistel structure was proposed by Feistel and Tuchman of IBM when designing Lucifer in the late 1960s [1]. The Feistel structure became popular after the widespread use of the data encryption standard (DES) [2]. The input of the Feistel structure is divided into two blocks whose length is equal. The round function is applied to one half, using a subkey, and then, the output is XORed with the other half. Then, two halves are exchanged with each other. As a result, the diffusion of the Feistel structure is relatively slow. The Feistel structure has consistent encryption with decryption, which is efficiently beneficial in hardware implementation. Several well-known block ciphers adopt the Feistel structure, for example, SIMON [3] and SIMECK [4]. In addition, there are many extensions of the Feistel structure, such as the SM4 structure [5], the Mars structure [6], and the generalized Feistel structure [7, 8].

The Lai–Massey scheme was proposed by Lai and Massey in the International Data Encryption Algorithm (IDEA) [9]. Similar to the Feistel structure, the Lai–Massey scheme takes two equal-sized plaintexts as its input. Unlike the Feistel structure, the round function is applied to the sum of the two pieces, and the result is then added to both half blocks. Furthermore, an orthomorphism is always introduced to one of the halves, to compensate the existence of a differential covering any rounds with probability 1. Generally, the Lai–Massey structure also has consistent encryption and decryption. There are several instances, such as MESH [10] and FOX [11], that utilize this structure.

Different from the Feistel structure and the Lai–Massey scheme, one round function of the SPN structure is composed of an invertible nonlinear function S layer controlled by a subkey and an invertible linear transformation P layer.

Compared to the Feistel structure, the diffusion of the SPN structure might be faster. However, the decryption of an SPN structure is usually different from that of the encryption; thus, more resources might be required for the implementation. In recent years, there have been many ciphers using the SPN structure such as AES [12] and PHOTON [13]. Additionally, the SPN structure is often adopted as round functions in the design of Feistel ciphers such as SM4 [5] and Camellia [14].

Traditionally, the security of block ciphers is defined as the indistinguishability from random permutations when the keys are random and secret. This is known as pseudorandomness. However, in some block cipher-based schemes such as hash functions, the underlying block ciphers may not have secret keys. The security arguments for these schemes thereby cannot be based on the pseudorandomness of the underlying ciphers. To remedy this, theory community usually assumes their underlying ciphers as ideal ciphers, i.e., sets of independent random permutations indexed by the keys, and then argues the security of the whole schemes. By this, we expect practical block ciphers to be as secure as ideal ciphers, and a proof for such security is the final goal of the theory community. However, given the state of the art, the only choice is to replace some of the underlying components of the block ciphers by idealized oracles and then argue the closeness of the obtained idealized block ciphers and the ideal ciphers. The security proved in this way cannot be used as a guarantee for the actual security of practical block ciphers, but it is still helpful to deepen the understanding of that and promote theoretical research to the ultimate goal.

Luby and Rackoff [15] proposed Luby–Rackoff (LR) scheme in 1985. They also proved that the balanced Feistel structure covering 3 rounds is a pseudorandom permutation and that covering 4 rounds is a super-pseudorandom permutation. Later, the LR scheme has attracted a lot of attention and become the most popular model for Feistel ciphers in which the round functions are set as random functions. Following [15], a long series of work established better security (maybe using a larger number of rounds) [16–18]. Significantly, Gentry and Ramzan [19] proved that the Feistel cipher covering 4 rounds without keys is secure against $2^{n/2}$ queries. On the basis, Guo and Wang [20] showed that 4-round key-alternating Feistel (KAF) using the same round function is secure against $2^{n/2}$ queries, i.e., birthday-bound security. Guo also proved that the 6-round key-alternating Feistel (KAF) using the independent round functions and appropriate keys is secure up to $2^{2n/3}$ queries, i.e., beyond-birthday-bound security.

Iterative structures in block ciphers mainly adopt the three structures: the Feistel structure, the SPN structure, and the Lai–Massey scheme. The solidification of iterative structure may bring security risks; once backdoors of some iterative structures are founded, a series of block ciphers that adopt the structure would be effectively attacked and will not be secure any more. This promotes the diversity research of iterative structures. In this paper, we extend the original Feistel structure and propose an extended Feistel (E-Feistel) structure. This paper mainly focuses on studying the security of the E-Feistel structure from the perspective of theoretical security of the structure using H-coefficient technique. The main contributions are as follows:

(1) We propose a new iterative structure, the E-Feistel structure, which also has consistent encryption with decryption

(2) For birthday-bound security, we prove that the 4-round KAEF with the same round function is secure up to $2^{n/2}$ queries

(3) For beyond-birthday-bound security, we prove that the 6-round KAEF with independent round functions is secure up to $2^{2n/3}$ queries

This paper is organized as follows. We introduce the notations in Section 2. Section 3 presents the extended Feistel structure. Then, Sections 4 and 5, respectively, present our results on 4-round KAESF and 6-round KAEF and their security proofs. Section 6 concludes this paper.

2. Notations

Let $\mathbb{F}_2$ denote the binary field and $\mathbb{F}_2^n$ denote the n–dimensional vector space over $\mathbb{F}_2$. Throughout this paper, $(x_1, x_2, \ldots, x_n)$ always corresponds to a column vector. Let $c = (c_1, c_2, \ldots, c_n) \in \mathbb{F}_2^n$, $d = (d_1, d_2, \ldots, d_n) \in \mathbb{F}_2^n$. Then, the point product of two vectors c and d is calculated as

$$c \cdot d = c_1 d_1 \oplus c_2 d_2 \oplus \cdots \oplus c_n d_n, \tag{1}$$

denoted as $c^T d$.

2.1. The H-Coefficient Technique.

We use Patarin's H-coefficient technique [21, 22] to prove the birthday-bound security of 4-round KAEF and the beyond-birthday-bound security of 6-round KAEF. Therefore, we sum up the interaction between distinguisher D and its oracles in the queries of transcripts. Suppose D enquiring the ith oracle ($\mathrm{KAEF}_{k^{(i)}}^{F}$ or $P^{(i)}$) for q times; a record $\mathcal{Q}_{E_i} = \left\{ (L_1 R_1, S_1 T)_1, \ldots, (L_{q_i} R_{q_i}, S_{q_i} T_{q_i}) \right\}$ can be obtained, where $(L_j R_j, S_j T_j)$ represents the queries and answers of KAEF. Similarly, the queries made to f_i are recorded as $\mathcal{Q}_{f_i} = \left\{ (x_{i,1}, y_{i,1}), \ldots, (x_{i,p}, y_{i,p}) \right\}$, denoting the answer $y_{i,j}$ obtained by querying f_i with $x_{i,j}$. Let $\mathcal{Q}_E = (\mathcal{Q}_{E_1}, \ldots, \mathcal{Q}_{E_m})$ and $\mathcal{Q}_F = (\mathcal{Q}_{f_1}, \ldots, \mathcal{Q}_{f_t})$. The transcript of the distinguisher D is denoted as $\tau = (\mathcal{Q}_E, \mathcal{Q}_F)$.

Given a set $\mathcal{Q}_{f_i}$ of function queries and a function f_i, we say that f_i extends $\mathcal{Q}_{f_i}$, denoted $f_i \vdash \mathcal{Q}_{f_i}$ if $f_i(x) = y$, for all $(x, y) \in \mathcal{Q}_{f_i}$. Similarly, given a transcript of permutation queries $\mathcal{Q}_{E_i}$ and a permutation $P^{(i)}$, we say $P^{(i)}$ extends $\mathcal{Q}_{E_i}$, denoted $P^{(i)} \vdash \mathcal{Q}_{E_i}$, if $P^{(i)}(LR) = ST$, for all $(LR, ST) \in \mathcal{Q}_{E_i}$. The latter definition also extends to the t-round KAEF built on F and a key $k^{(i)}$; in that case, we write $\mathrm{KAEF}_{k^{(i)}}^{F} \vdash \mathcal{Q}_{E_i}$. Finally, for $\mathcal{Q}_F = (\mathcal{Q}_{f_1}, \ldots, \mathcal{Q}_{f_t})$ and $F = (f_1, \ldots, f_t)$, if $f_1 \vdash \mathcal{Q}_{f_1} \wedge \cdots \wedge f_t \vdash \mathcal{Q}_{f_t}$, then $F \vdash \mathcal{Q}_F$.

With regard to all achievable transcripts τ which represent an achievable record for queries of a series of oracles $(F, P^{(1)}, \ldots, P^{(m)})$ or $(F, \mathrm{KAEF}_{k^{(1)}}^{F}, \ldots, \mathrm{KAEF}_{k^{(m)}}^{F})$, the probability that D interacts with the real world and the ideal world is denoted as $\Pr_{\mathrm{re}}(\tau)$ and $\Pr_{\mathrm{id}}(\tau)$, respectively. $\Pr_{\mathrm{re}}(\tau)$ and $\Pr_{\mathrm{id}}(\tau)$ are defined formally as follows:

$$
\begin{aligned}
\mathrm{Pr}_{\mathrm{re}}(\tau) &= \mathrm{Pr}\left[\left(k^{(1)},\ldots,k^{(m)}\right)\overset{\$}{\leftarrow}(\mathscr{K})^{m}, F\overset{\$}{\leftarrow}(\mathscr{F}(n))^{t}\colon \mathrm{KAEF}^{F}_{k^{(1)}}\vdash\mathcal{Q}_{E_1}\wedge\cdots\wedge\mathrm{KAEF}^{F}_{k^{(m)}}\vdash\mathcal{Q}_{E_m}\wedge F\vdash Q_F\right],\\
\mathrm{Pr}_{\mathrm{id}}(\tau) &= \mathrm{Pr}\left[\left(P^{(1)},\ldots,P^{(m)}\right)\overset{\$}{\leftarrow}(\mathscr{P}(2n))^{m}, F\overset{\$}{\leftarrow}(\mathscr{F}(n))^{t}\colon P^{(1)}\vdash\mathcal{Q}_{E_1}\wedge\cdots\wedge P^{(m)}\vdash\mathcal{Q}_{E_m}\wedge F\vdash Q_F\right].
\end{aligned}
\tag{2}
$$

We can estimate the distinguishing advantage by calculating $\mathrm{Pr}_{\mathrm{re}}(\tau)$ and $\mathrm{Pr}_{\mathrm{id}}(\tau)$. The lemma is stated as follows:

Lemma 1 (see [23]). *Assume that there exists a function $\varepsilon(q_f, q_e) > 0$ such that, for every achievable transcript τ with q_e and q_f queries of two types, the following equation holds:*

$$\mathrm{Pr}_{\mathrm{id}}(\tau) - \mathrm{Pr}_{\mathrm{re}}(\tau) \le \mathrm{Pr}_{\mathrm{id}}(\tau)\cdot\varepsilon\left(q_f, q_e\right). \tag{3}$$

Then, the following equation holds:

$$\mathrm{Adv}_{\mathrm{KAEF}}\left(q_f, q_e\right) \le \varepsilon\left(q_f, q_e\right). \tag{4}$$

Lemma 2 (see [23]). *Fix a transcript with $\mathrm{Pr}_{id}(\tau) > 0$. Assume that (i) $\mathrm{Pr}[k \in \mathscr{K}_{bad}] \le \delta$ and (ii) there is a function $g\colon \mathscr{K} \longrightarrow [0, t\infty)$ such that, for all $k \in \mathscr{K}_{good}$, the following equation holds $\mathrm{Pr}_{re}(\tau, k)/\mathrm{Pr}_{id}(\tau, k) \ge 1 - g(k)$. Then, we have*

$$\mathrm{Pr}_{\mathrm{id}}(\tau) - \mathrm{Pr}_{\mathrm{re}}(\tau) \le \mathrm{Pr}_{\mathrm{id}}(\tau)\cdot\left(\delta + \mathbb{E}_{k\in\mathscr{K}}[g(k)]\right). \tag{5}$$

3. The Extended Feistel Structure

The extended Feistel structure (E-Feistel structure) denoted as $\mathscr{F}_{A,B}$ is defined in the following.

The round function and the branch have the same length, denoted as n. Let $A = [A_1, A_2]$ and $B = [B_1, B_2]$, where $A_i, B_j \in \mathbb{F}_2^{n\times n}$ and $A_1B_1 \oplus A_2B_2 = 0$. Let f be a map over $\mathbb{F}_2^n$. The input and output of $\mathscr{F}_{A,B,f}$ are denoted as $x = (x_1, x_2) \in \mathbb{F}_2^{n\times 2}$ and $y = (y_1, y_2) \in \mathbb{F}_2^{n\times 2}$, respectively. Then, the map $\mathscr{F}_{A,B,f}\colon \mathbb{F}_2^{n\times 2} \longrightarrow \mathbb{F}_2^{n\times 2}$ is defined as

$$\left(y_1, y_2\right) = \left(x_2, x_1\right) \oplus \left(B_2 f(g), B_1 f(g)\right), \tag{6}$$

where $g = A_1x_1 \oplus A_2x_2$, as illustrated in Figure 1.

Denote by $\mathscr{M}_n$ all the maps from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$. Then, the structure $\mathscr{F}_{A,B}$ is defined as $\mathscr{F}_{A,B} = \left\{\mathscr{F}_{A,B,f} \mid f \in \mathscr{M}_n\right\}$. $\mathscr{F}_{A,B}$ has similar procedures of decryption and encryption.

Let $A^* = [A_1^T, A_2^T]$ and $B^* = [B_1^T, B_2^T]$. We further associate A and B and branch swapping with the following two matrices:

$$
\begin{aligned}
\mathscr{A} &= \begin{bmatrix} A_1 & A_2 \\ A_2 & A_1 \end{bmatrix},\\
\mathscr{B} &= \begin{bmatrix} B_1^T & B_2^T \\ B_2^T & B_1^T \end{bmatrix}.
\end{aligned}
\tag{7}
$$

4. Four Rounds for Birthday-Bound Security

As seen in Figure 2, we assume that the 4-round idealized KAEF cipher uses the same random round function f and 4 independent random round keys $k = (k_1, k_2, k_3, k_4)$.

Firstly, we give the definition of round-key vectors in the 4-round key-alternating extended Feistel using the same round function (KAEFSF). The constraint on round keys does not need the round key to satisfy the condition of complete independence and randomness.

Definition 1 (suitable round-key vector for 4 rounds). A round-key vector $k = (k_1, k_2, k_3, k_4)$ is suitable if it satisfies the following conditions:

(i) k_1 and k_4 are uniform in $\{0,1\}^n$ (but they need not to be independent)

(ii) $k_1 \oplus k_4$ is also uniformly distributed in $\{0,1\}^n$

Proposition 1. *For $\mathscr{F}_{A,B}$, assume $\mathit{rank}(\mathscr{A}) = \mathit{rank}(\mathscr{B}) = 2n$, and $A_1B_2 \oplus A_2B_1$ is a nonsingular matrix.*

Proof. Since $\mathrm{rank}(\mathscr{A}) = \mathrm{rank}(\mathscr{B}) = 2n$, $\mathscr{A}$ and $\mathscr{B}$ are both nonsingular matrices. So,

$$\mathscr{A}\cdot\mathscr{B}^T = \begin{bmatrix} 0 & A_1B_2 \oplus A_2B_1 \\ A_2B_1 \oplus A_1B_2 & 0 \end{bmatrix} \tag{8}$$

is a nonsingular matrix. As a result, $A_1B_2 \oplus A_2B_1$ is a nonsingular matrix. For convenience, $A_1B_2 \oplus A_2B_1$ is denoted as O. □

Theorem 1. *Assuming $\mathit{rank}(\mathscr{A}) = \mathit{rank}(\mathscr{B}) = 2n$, for the 4-round idealized KAEFSF, the following equation holds:*

$$\mathrm{Adv}_{\mathrm{KAEFSF}}\left(q_f, q_e\right) \le \frac{9q_e^2 + 4q_eq_f}{N}. \tag{9}$$

Proof. According to Lemma 2, we can translate the proof of equation (9) into the proof of the following equation:

$$\mathrm{Pr}_{\mathrm{id}}(\tau) - \mathrm{Pr}_{\mathrm{re}}(\tau) \le \mathrm{Pr}_{\mathrm{id}}(\tau)\cdot\frac{9q_e^2 + 4q_eq_f}{N}. \tag{10}$$

For a fixed transcript $\tau = (Q_E, Q_F)$ with $|Q_E| = q_e$ and $|Q_F| = q_f$, key vectors can be distinguished whether good or bad concerning τ. Then, the probability $\mathrm{Pr}_{\mathrm{re}}(\tau, k)$ for good key k can be estimated. For convenience's sake, for any $x \in \{0,1\}^n$, $x \in \mathrm{Dom}\mathscr{F}$ ($x \notin \mathrm{Dom}\mathscr{F}$ otherwise) denotes whether there exists a corresponding record (x, y) in Q_F or not.

Bad keys are now defined as follows. □

Definition 2 (bad round-key vector for 4 rounds). In regard of $\tau = (Q_E, Q_F)$, a suitable round-key vector k is bad, as long as either of the two following conditions is satisfied:

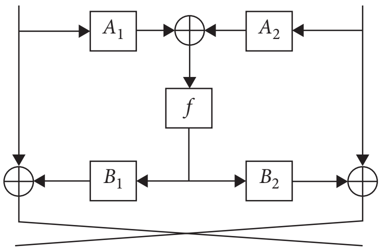

Figure 1: The extended feistel structure.

- (B-1): there exists $(LR, ST) \in Q_E$ that satisfies either $A_1L \oplus A_2R \oplus k_1 \in \mathrm{Dom}\mathscr{F}$ or $A_1T \oplus A_2S \oplus k_4 \in \mathrm{Dom}\mathscr{F}$
- (B-2): there exist two (not necessarily different) (LR, ST) and $(L'R', S'T')$ in Q_E that satisfies $A_1L \oplus A_2R \oplus k_1 = A_1T' \oplus A_2S' \oplus k_4$

The other transcripts are defined as good round-key vectors, denoted by $\mathscr{K}_{\mathrm{good}}$.

For each of the q_e records (LR, ST), since both k_1 and k_4 are uniformly distributed in $\{0,1\}^n$ and since $|\mathrm{Dom}\mathscr{F}| = |\mathrm{Q}_F| = q_f$, the probability that it satisfies (B-1) does not exceed $2q_f/N$. On the contrary, for each of the q_e^2 pairs of records (LR, ST) and $(L'R', S'T')$, since $k_1 \oplus k_4$ is uniform, the probability that the pair satisfies (B-2) does not exceed $1/N$. Therefore,

$$\Pr\left[k \xleftarrow{\$} \mathscr{K}\colon k \in \mathscr{K}_{\mathrm{bad}}\right] \leq \frac{2q_e q_f + q_e^2}{N}. \tag{11}$$

4.1. *Lowering Bounding the Probability for Good Keys.* We now lower bound the probability $\Pr_{\mathrm{re}}(\tau, k)$ for an arbitrary good round-key vector k. For this, we follow a clean "predicate" approach from [24]. We define a "bad" predicate $\mathrm{Bad}(F)$ on F. When conditions of $\mathrm{Bad}(F)$ are not fulfilled, the event $\mathrm{KAEFSF}_k^F \vdash \mathrm{Q}_E$ occurs if and only if $2q_e$ new and distinct equations on the random round function F are satisfied. For convenience's sake, we first define

$$\mathrm{Ext}\mathscr{F} \overset{\mathrm{def}}{=} \left\{ \begin{array}{c} x \in \{0,1\}^n\colon\ (LR, ST) \in \mathrm{Q}_E \text{ for } A_1L \oplus A_2R = k_1 \oplus x \text{ and some } S, T, \text{ or} \\ (LR, ST) \in \mathrm{Q}_E \text{ for } A_1T \oplus A_2S = k_4 \oplus x \text{ and some } L, R \end{array} \right\}. \tag{12}$$

Clearly, $|\mathrm{Ext}\mathscr{F}| \leq 2q_e$. Then, for any $n-$ to $-n$ bit function $f \vdash \mathrm{Q}_F$, the predicate $\mathrm{Bad}(F)$ holds if one of the following conditions is fulfilled:

- (C-1): $\exists (LR, ST) \in \mathrm{Q}_E$ such that $k_2 \oplus A_1R \oplus A_2L \oplus \mathrm{OF}(A_1L \oplus A_2R \oplus k_1) \in \mathrm{Dom}\mathscr{F} \cup \mathrm{Ext}\mathscr{F}$ or $k_3 \oplus A_1S \oplus A_2T \oplus \mathrm{OF}(A_1T \oplus A_2S \oplus k_4) \in \mathrm{Dom}\mathscr{F} \cup \mathrm{Ext}\mathscr{F}$
- (C-2): there exist two (not necessarily different) (LR, ST) and $(L'R', S'T')$ in Q_E such that $k_2 \oplus A_1R \oplus A_2L \oplus \mathrm{OF}(A_1L \oplus A_2R \oplus k_1) = k_3 \oplus A_1S' \oplus A_2T' \oplus \mathrm{OF}(A_1T' \oplus A_2S' \oplus k_4)$
- (C-3): there exist two different $(LR, ST) \in \mathrm{Q}_E$ and $(L'R', S'T') \in \mathrm{Q}_E$ such that $A_1R \oplus A_2L \oplus \mathrm{OF}(A_1L \oplus A_2R \oplus k_1)) = A_1R' \oplus A_2L' \oplus \mathrm{OF}(A_1L' \oplus A_2R' \oplus k_1)$ or $A_1S \oplus A_2T \oplus \mathrm{OF}(A_1T \oplus A_2S \oplus k_4) = A_1S' \oplus A_2T' \oplus \mathrm{OF}(A_1T' \oplus A_2S' \oplus k_4)$

To compute $\Pr[F \xleftarrow{\$} \mathscr{F}(n)\colon \mathrm{Bad}(F) | F \vdash \mathrm{Q}_F]$, we consider the conditions in turn. First, as k is good, for any $(LR, ST) \in \mathrm{Q}_E$, we have $k_1 \oplus A_1L \oplus A_2R \notin \mathrm{Dom}\mathscr{F}$ and $k_4 \oplus A_1T \oplus A_2S \notin \mathrm{Dom}\mathscr{F}$. Thus, based on $f \vdash \mathrm{Q}_F$, the values $F(k_1 \oplus A_1L \oplus A_2R)$ and $F(k_4 \oplus A_1T \oplus A_2S)$ remain uniformly distributed, and thus,

$$\begin{aligned} &-\Pr\left[k_2 \oplus A_1R \oplus A_2L \oplus \mathrm{OF}\left(A_1L \oplus A_2R \oplus k_1\right) \in \mathrm{Dom}\mathscr{F} \cup \mathrm{Ext}\mathscr{F}\right] \leq \frac{q_f + 2q_e}{N}, \\ &-\Pr\left[k_3 \oplus A_1S \oplus A_2T \oplus \mathrm{OF}\left(A_1T \oplus A_2S \oplus k_4\right) \in \mathrm{Dom}\mathscr{F} \cup \mathrm{Ext}\mathscr{F}\right] \leq \frac{q_f + 2q_e}{N}. \end{aligned} \tag{13}$$

So, $\Pr[(C-1)] \leq 2q_e(q_f + 2q_e)/N$. Second, for any two tuples (LR, ST) and $(L'R', S'T')$ from Q_E, the two function values $F(A_1L \oplus A_2R \oplus k_1)$ and $F(A_1T' \oplus A_2S' \oplus k_4)$ are independent by $(B-2)$. Then, as argued, we have

$$-\Pr\left[k_2 \oplus A_1R \oplus A_2L \oplus \mathrm{OF}\left(A_1L \oplus A_2R \oplus k_1\right) = k_3 \oplus A_1S' \oplus A_2T' \oplus \mathrm{OF}\left(A_1T' \oplus A_2S' \oplus k_4\right)\right] = \frac{1}{N}, \tag{14}$$

and thus, $\Pr[(C-2)] \leq q_e^2/N$.

Third, for any tuples (LR, ST) and $(L'R', S'T')$, if $A_1L \oplus A_2R = A_1L' \oplus A_2R'$, then $A_1R \oplus A_2L \oplus \mathrm{OF}(A_1L \oplus A_2R \oplus k_1) = A_1R' \oplus A_2L' \oplus \mathrm{OF}(A_1L' \oplus A_2R' \oplus k_1)$ is not possible. Otherwise, since $F(A_1L \oplus A_2R \oplus k_1)$ is uniform (as argued before), the probability to have $A_1R \oplus A_2L \oplus \mathrm{OF}$

$(A_1 L \oplus A_2 R \oplus k_1) = A_1 R' \oplus A_2 L' \oplus \mathrm{OF}\,(A_1 L' \oplus A_2 R' \oplus k_1)$ is $1/N$. It is similar for the other condition $A_1 S \oplus A_2 T \oplus \mathrm{OF}\,(A_1 T \oplus A_2 S \oplus k_4) = A_1 S' \oplus A_2 T' \oplus \mathrm{OF}\quad (A_1 T' \oplus A_2 S' \oplus k_4)$; thus, $\Pr[(C-3)] \leq q_e^2/N$ and

$$\begin{aligned}\Pr\left[F \overset{\$}{\leftarrow} \mathscr{F}(n)\colon \mathrm{Bad}(F) | F \vdash Q_F\right] &\leq \frac{2q_e\left(q_f + 2q_e\right)}{N} + \frac{q_e^2}{N} + \frac{2q_e^2}{N} \\ &\leq \frac{7q_e^2 + 2q_e q_f}{N}.\end{aligned} \tag{15}$$

Using an arbitrary order, write $\mathscr{Q}_E = \left\{(L_1 R_1, S_1 T)_1, \ldots, (L_{q_e} R_{q_e}, S_{q_e} T_{q_e})\right\}$, and for a given F, let

$$\begin{aligned} x_2^{(1)} &= k_2 \oplus A_1 R_1 \oplus A_2 L_1 \oplus OF\left(A_1 L_1 \oplus A_2 R_1 \oplus k_1\right) \\ &\vdots \\ x_2^{(q_e)} &= k_2 \oplus A_1 R_{q_e} \oplus A_2 L_{q_e} \oplus OF\left(A_1 L_{q_e} \oplus A_2 R_{q_e} \oplus k_1\right) \\ x_3^{(1)} &= k_3 \oplus A_1 S_1 \oplus A_2 T_1 \oplus OF\left(A_1 T_1 \oplus A_2 S_1 \oplus k_4\right) \\ &\vdots \\ x_3^{(q_e)} &= k_3 \oplus A_1 S_{q_e} \oplus A_2 T_{q_e} \oplus OF\left(A_1 T_{q_e} \oplus A_2 S_{q_e} \oplus k_4\right). \end{aligned} \tag{16}$$

For each $(L_i R_i, S_i T_i) \in \mathscr{Q}_E$, we assume that the outputs of F in the 2 and 3 rounds are denoted as X_i and Y_i, respectively:

$$\Pr\left[\mathrm{FAEFSF}_k^F\,(L_i R_i) = S_i T_i\right] = \Pr\left[F\left(x_2^{(i)}\right) = X_i \wedge F\left(x_3^{(i)}\right) = Y_i\right]. \tag{17}$$

Additionally, conditioned on $F \vdash \mathscr{Q}_F$ and $\mathrm{Bad}(F)$, (a) the induced values $x_2^{(1)}, \ldots, x_2^{(q_e)}, x_3^{(1)}, \ldots, x_3^{(q_e)}$ are $2q_e$ distinct ones (otherwise if $x_2^{(i)} = x_2^{(j)}$ or $x_3^{(i)} = x_3^{(j)}$, for some $i \neq j$, then (C-3) is fulfilled; if $x_2^{(i)} = x_3^{(j)}$, then (C-2) is fulfilled) and (b) the $2q_e$ images $F(x_2^{(1)}), \ldots, F(x_2^{(q_e)}), F(x_3^{(1)}), \ldots, F(x_3^{(q_e)})$ remain fully undetermined and thus uniform, otherwise (C-1) is fulfilled. Therefore, for each $i \in \{1, \ldots, q_e\}$, we have

$$\Pr\left[F\left(x_2^{(i)}\right) = X_i \wedge F\left(x_3^{(i)}\right) = Y_i\right] = \left(\frac{1}{N^2}\right), \tag{18}$$

and for any $k \in \mathscr{K}_{\mathrm{good}}$,

$$\begin{aligned}\frac{\Pr_{\mathrm{re}}(\tau, k)}{\Pr_{\mathrm{id}}(\tau, k)} &= \frac{\left(\left(1/|\mathscr{K}| \cdot N^{q_f}\right) \cdot \Pr\left[\mathrm{KAEFSF}_k^F \vdash Q_F | F \vdash Q_F\right]\right)}{\left(\left(1/ \cdot |\mathscr{K}| N^{q_f}\right) \cdot \prod_{i=0}^{q_e-1} 1/N^2 - i\right)} \\ &\geq \frac{\Pr\left[\mathrm{KAEFSF}_k^F \vdash Q_F | F \vdash Q_F \wedge \mathrm{Bad}(F)\right] \cdot \left(1 - \Pr\left[\mathrm{Bad}(F) | F \vdash Q_F\right]\right)}{\prod_{i=0}^{q_e-1} 1/N^2 - i} \\ &\geq 1 - \frac{\left(7q_e^2 + 2q_e q_f/N\right)\left(1/N^{2q_e}\right)}{\left(\prod_{i=0}^{q_e-1} 1/N^2 - i\right)} \\ &\geq \left(1 - \frac{7q_e^2 + 2q_e q_f}{N}\right)\left(1 - \frac{q_e^2}{N^2}\right) \geq 1 - \frac{7q_e^2 + 2q_e q_f}{N} - \frac{q_e^2}{N^2}.\end{aligned} \tag{19}$$

Gathering this and equation (11) and Lemma 2 yields the claim of equation (10).

5. Six Rounds for Beyond-Birthday-Bound Security

As seen in Figure 3, we assume that the 6-round idealized KAEF cipher uses the 6 independent random round functions $F = (F_1, F_2, F_3, F_4, F_5, F_6)$ and 6 random round keys $k = (k_1, k_2, k_3, k_4, k_5, k_6)$.

Similarly, we first define suitable round-key vectors.

Definition 3 (suitable round-key vector for 6 rounds). 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_3, and k_5 are uniformly distributed in $\{0,1\}^n$

(ii) k_2, k_4, and k_6 are uniformly distributed in 2^{n-r} possibilities

(iii) For $(i, j) \in \{(1,2), (2,3), (4,5), (5,6), (1,6)\}$, k_i and k_j are independent

Different from the proof of birthday-bound security, the upper bound of probability of various collisions should be small enough. As a result, all of key vectors need to be uniform. It is the same for the independence of key vectors. Beyond-birthday-bound security of KAEF can be proved under the instantiation with suitable round-key vectors.

Theorem 2. *For the 6-round idealized cipher KAEF with the suitable round-key vector, the following equation holds:*

$$\mathrm{Adv}_{\mathrm{KAEF}}\left(q_f, q_e\right) \leq \frac{7q_e^3 + 13q_eq_f^2 + 22q_e^2q_f}{N^2} + \frac{2^r\left(8q_eq_f^2 + 2q_e^2q_f\right)}{N^2}. \tag{20}$$

It is worth noting that when $r < n/2$, the security is beyond-birthday security. For example, when $r = 0$, the bound is of "typical" beyond-birthday form $O(q^3/N^2)$.

Similarly, we can translate the proof into the proof of the result; for any transcript τ, the following equation holds:

$$\mathrm{Pr}_{\mathrm{id}}(\tau) - \mathrm{Pr}_{\mathrm{re}}(\tau) \leq \mathrm{Pr}_{\mathrm{id}}(\tau) \cdot \frac{7q_e^3 + 13q_eq_f^2 + 22q_e^2q_f + 2^r\left(8q_eq_f^2 + 2q_e^2q_f\right)}{N^2}. \tag{21}$$

For a fixed a transcript $\tau = (Q_E, Q_F)$ with $Q_F = (Q_{F_1}, Q_{F_2}, Q_{F_3}, Q_{F_4}, Q_{F_5}, Q_{F_6})$, $|Q_E| = q_e$, and $|Q_{F_i}| = q_{f_i}$, for $i = 1, \ldots, 6$, we follow similar procedures in Section 4 to complete the proof.

5.1. Bad Round-Key Vectors and Probability. For convenience, $x_i \in \mathrm{Dom}\mathscr{F}_i$ (and $x_i \notin \mathrm{Dom}\mathscr{F}_i$ otherwise) denotes, for any $x_i \in \{0,1\}^n$, whether there exists a corresponding record (x_i, y_i) in Q_{F_i} or not. Additionally, $\mathrm{Img}\mathscr{F}_i(x_i)$ denotes the corresponding y_i. The bad round-key vector is defined as follows.

Definition 4 (bad round-key vector for 6 rounds). With respect to $\tau = (Q_E, Q_F)$, a suitable key vector k is bad if it satisfies one of the following conditions:

(B-1): there exist $(LR, ST) \in Q_E$, $(x_1, y_1) \in Q_{F_1}$, and $(x_6, y_6) \in Q_{F_6}$ such that $k_1 = A_1L \oplus A_2R \oplus x_1$ and $k_6 = A_1T \oplus A_2S \oplus x_6$

(B-2): there exist $(LR, ST) \in Q_E$, $(x_1, y_1) \in Q_{F_1}$, and $(x_2, y_2) \in Q_{F_2}$ such that $k_1 = A_1L \oplus A_2R \oplus x_1$ and $k_2 = A_1R \oplus A_2L \oplus Oy_1 \oplus x_2$

(B-3): there exist $(LR, ST) \in Q_E$, $(x_5, y_5) \in Q_{F_5}$, and $(x_6, y_6) \in Q_{F_6}$ such that $k_6 = A_1T \oplus A_2S \oplus x_6$ and $k_5 = A_1S \oplus A_2T \oplus Oy_6 \oplus x_5$

These bad key vectors are denoted as $\mathscr{K}_{\mathrm{bad}}$. The other key vectors are defined as good key vectors, denoted by $\mathscr{K}_{\mathrm{good}}$.

Because there are at most $q_eq_f^2$ choices for $(LR, ST) \in Q_E$, $(x_1, y_1) \in Q_{F_1}$, and $(x_6, y_6) \in Q_{F_6}$ and since k_1, resp. k_6, is uniform in 2^n, resp. 2^{n-r} possibilities, and furthermore, since k_1 and k_6 are independent, the following equation holds $\Pr[(B-1)] \leq q_eq_f^2/2^{2n-r} \leq 2^r q_eq_f^2/N^2$.

Similarly, we have $\Pr[(B-2)] \leq 2^r q_eq_f^2/N^2$. By symmetry, $\Pr[(B-3)] \leq 2^r q_eq_f^2/N^2$. As a result,

$$\Pr\left[k \xleftarrow{\$} \mathscr{K}\colon k \in \mathscr{K}_{\mathrm{bad}}\right] \leq \frac{3 \cdot 2^r \cdot q_eq_f^2}{N^2}. \tag{22}$$

5.2. Analysis for Good Keys. For a fixed good round-key vector k, we would get a lower bound for the probability $\Pr[F \xleftarrow{\$} (\mathscr{F}(n))^6\colon \mathrm{KAEF}_k^F \vdash Q_E | F \vdash Q_F]$. Inspired by Cogliati et al. [25, 26], it takes two steps. Firstly, through defining certain "bad" functions on (F_1, F_6), we would lower bound the probability under the condition. Secondly, under the assumption of "good" functions (F_1, F_6), the outer two rounds are removed; we only need to analyze the induced 4-round transcript to yield the final bounds.

For a fixed pair of (F_1, F_6) such that $F_1 \vdash Q_{F_1}$ and $F_6 \vdash Q_{F_6}$ and for each $(LR, ST) \in Q_E$, we set $X \leftarrow A_1R \oplus A_2L \oplus OF_1(A_1L \oplus A_2R \oplus k_1)$ and $U \leftarrow A_1S \oplus A_2T \oplus OF_6(A_1T \oplus A_2S \oplus k_6)$. Then, q_e tuples (LRX, UST) are obtained. The induced tuples are denoted by $Q_E^*(F_1, F_6)$. For convenience, $\varepsilon Q(x)$ is used to denote all tuples (LRX, UST) whose third coordinate equals X. Similarly, we define $\varepsilon Q(U)$:

$$\begin{aligned} -\varepsilon Q(X) &= \left\{(LRX, UST)\colon\ (LRX, UST) \in Q_E^*\left(F_1, F_6\right)\right\}, \\ -\varepsilon Q(U) &= \left\{(LRX, UST)\colon\ (LRX, UST) \in Q_E^*\left(F_1, F_6\right)\right\}. \end{aligned} \tag{23}$$

Additionally, we define several key-independent quantities characterizing τ:

$$\begin{aligned} \alpha_1(k) &\overset{\mathrm{def}}{=} \left|\left((LR, ST), (x_1, y_1)\right) \in Q_E \times Q_{F_1}\colon k_1 = A_1L \oplus A_2R \oplus x_1\right|, \\ \alpha_2(k) &\overset{\mathrm{def}}{=} \left|\left((LR, ST), (x_6, y_6)\right) \in Q_E \times Q_{F_6}\colon k_6 = A_1T \oplus A_2S \oplus x_6\right|, \\ \alpha_{2,3}(k) &\overset{\mathrm{def}}{=} \left|\left((LR, ST), (x_2, y_2), (x_3, y_3)\right) \in Q_E \times Q_{F_2} \times Q_{F_3}\colon k_3 = A_1L \oplus A_2R \oplus Oy_2 \oplus x_3\right|, \\ \alpha_{4,5}(k) &\overset{\mathrm{def}}{=} \left|\left((LR, ST), (x_4, y_4), (x_5, y_5)\right) \in Q_E \times Q_{F_4} \times Q_{F_5}\colon k_4 = A_1T \oplus A_2S \oplus Oy_5 \oplus x_4\right|. \end{aligned} \tag{24}$$

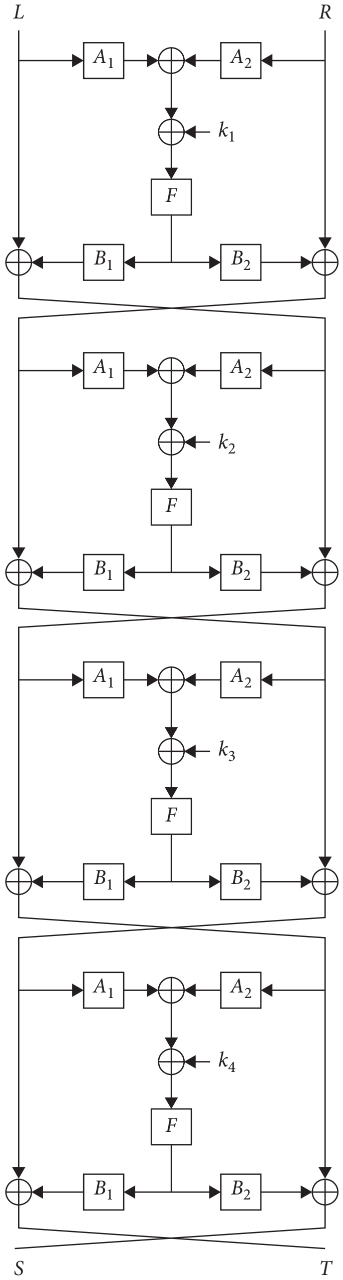

Figure 2: The 4-round KAEF with the same round function.

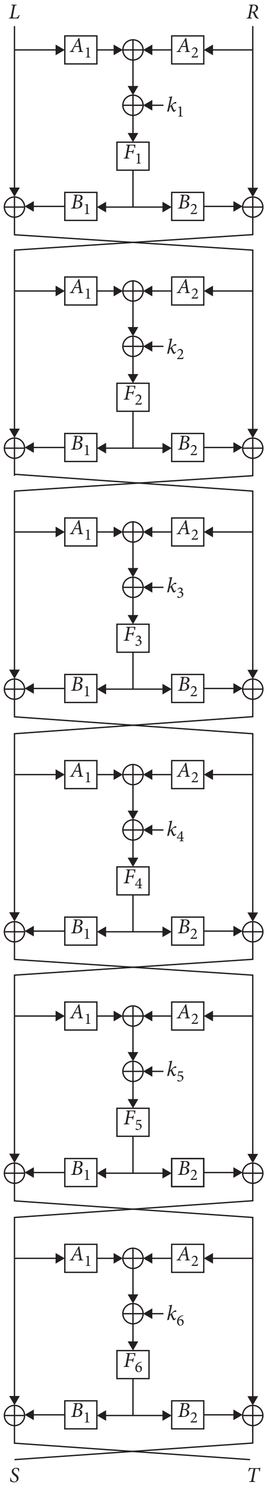

Figure 3: The 6-round KAEF with independent round functions.

The predicate $\mathrm{Bad}(F_1, F_6)$ holds if the induced set $Q_E^*(F_1, F_6)$ satisfies one of the following conditions:

(C-1): there exist three records $(LRX, UST) \in Q_E^*(F_1, F_6)$, $(x_2, y_2) \in Q_{F_2}$, and $(x_5, y_5) \in Q_{F_5}$ such that $k_2 = X \oplus x_2$ and $k_5 = U \oplus x_5$

(C-2): there exist three records $(LRX, UST) \in Q_E^*(F_1, F_6)$, $(x_2, y_2) \in Q_{F_2}$, and $(x_3, y_3) \in Q_{F_3}$ such that $k_2 = X \oplus x_2$ and $k_3 = A_1 L \oplus A_2 R \oplus O y_2 \oplus x_3$

(C-3): there exist three records $(LRX, UST) \in Q_E^*(F_1, F_6)$, $(x_4, y_4) \in Q_{F_4}$, and $(x_5, y_5) \in Q_{F_5}$ such that $k_5 = A \oplus x_5$ and $k_4 = A_1 T \oplus A_2 S \oplus O y_5 \oplus x_4$

(C-4): there exist two distinct (LRX, UST) and $(L'R'X', U'S'T')$ in $Q_E^*(F_1, F_6)$, and a pair (x_2, y_2) in Q_{F_2} such that $X = X'$ and $k_2 = X \oplus x_2$, or symmetrically, two distinct (LRX, UST) and $(L'R'X', A'S'T')$ in $Q_E^*(F_1, F_6)$ and a pair (x_5, y_5) in Q_{F_5} such that $A = A'$ and $k_5 = A \oplus x_5$

(C-5): there exist two distinct (LRX, UST) and $(L'R'X', U'S'T')$ in $Q_E^*(F_1, F_6)$ and a pair (x_2, y_2) in Q_{F_2} such that $U = U'$ and $k_2 = X \oplus x_2$, or symmetrically, two distinct (LRX, UST) and $(L'R'X', U'S'T')$ in $Q_E^*(F_1, F_6)$ and a pair (x_5, y_5) in Q_{F_5} such that $X = X'$ and $k_5 = U \oplus x_5$

(F_1, F_6) is defined as good if $\mathrm{Bad}(F_1, F_6)$ does not hold.

Lemma 3. *We have*

$$\Pr_{F_1,F_6}\left[\mathrm{Bad}(F_1,F_6) | F_1 \vdash Q_{F_1} \wedge F_6 \vdash Q_{F_6}\right] \leq \frac{q_e q_f^2}{N^2} + \frac{4q_e^2 q_f}{N^2} + \frac{\alpha_{2,3}(k) + \alpha_{4,5}(k)}{N} + \frac{q_f(\alpha_1(k) + \alpha_2(k))}{N}. \tag{25}$$

Proof. Due to page limits, see Appendix A. □

5.2.1. Analyzing the Inner Four Rounds. Let $F^* = (F_2, F_3, F_4, F_5)$. We denote

$$p(\tau, F_1, F_6) = \Pr\left[F^* \overset{\$}{\leftarrow} (\mathscr{F})^n\text{: } \mathrm{KAEF}_E^* \vdash Q_E^*(F_1, F_6) | F_i \vdash Q_{F_i}, \quad i = 1, 2, 3, 4, 5, 6\right]. \tag{26}$$

This captures the probability that the inner four rounds of KAEF "extend" the tuples in $Q_E^*(F_1, F_6)$. The probability $\Pr_{\mathrm{re}}(\tau, k)$ can be related to it.

Lemma 4 (see [17]). *Assume that there exists a function* ε: $(\mathscr{F}(n))^2 \times \mathscr{K} \longrightarrow [0, \infty]$ *such that, for any good* (F_1, F_6), *the following equation holds:*

$$\frac{p(\tau, F_1, F_6)}{\prod_{i=0}^{q_e-1} 1/N^2 - i} \geq 1 - \varepsilon(F_1, F_6, k). \tag{27}$$

Then, we have

$$\frac{\Pr_{\mathrm{re}}(\tau, k)}{\Pr_{\mathrm{id}}(\tau, k)} \geq 1 - \Pr\left[\mathrm{Bad}(F_1, F_6) | F_1 \vdash Q_{F_1}, F_6 \vdash Q_{F_6}\right] - \mathbb{E}_{F_1,F_6}\left[\varepsilon(F_1, F_6, k) | F_1 \vdash Q_{F_1}, F_6 \vdash Q_{F_6}\right]. \tag{28}$$

Now, we prove the assumption of Lemma 4.

Lemma 5. *For any fixed good tuple* (F_1, F_6), *there exists a function* $\varepsilon(F_1, F_6, k)$ *of the function pair and the round-key vector k such that inequality (20) mentioned in Lemma 4 holds. Moreover,*

$$\mathbb{E}_{F_1,F_6,k}\left[\varepsilon(F_1, F_6, k)\right] \leq \frac{7q_e^3 + 10q_e q_f^2 + 18q_e^2 q_f + 3 \cdot 2^r \cdot q_e q_f^2 + 2 \cdot 2^r \cdot q_e^2 q_f}{N^2}. \tag{29}$$

Proof. The general expression of $\varepsilon(F_1, F_6, k)$ is a function of several variables defined before, which suffers from a bad readability. Therefore, we directly establish (and present) the bound on its expectation. However, due to space constraints, for the full proof, refer to Appendix B.

Below, we present a sketch and the core results. According to the type of the involved collisions, we divide the tuples in $Q_E^*(F_1, F_6)$ into four groups:

$$\begin{aligned}
-\mathscr{G}_1 &= \left\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\text{: } |\varepsilon\mathcal{Q}(X)| = |\varepsilon\mathcal{Q}(U)| = 1, k_2 \oplus X \notin \mathrm{Dom}\mathscr{F}_2 \wedge k_5 \oplus U \notin \mathrm{Dom}\mathscr{F}_5\right\},\\
-\mathscr{G}_2 &= \left\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\text{: } k_2 \oplus X \in \mathrm{Dom}\mathscr{F}_2\right\},\\
-\mathscr{G}_3 &= \left\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\text{: } k_5 \oplus U \in \mathrm{Dom}\mathscr{F}_5\right\},\\
-\mathscr{G}_4 &= \left\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\text{: } |\varepsilon\mathcal{Q}(X)| \geq 2, \text{ or } |\varepsilon\mathcal{Q}(U)| \geq 2\right\}.
\end{aligned} \tag{30}$$

Let $\beta_1 = |\mathscr{G}_2|$, $\beta_2 = |\mathscr{G}_3|$, and $\beta_3 = |\mathscr{G}_4|$. Note that, by definition, these sets form a partition of $Q_E^*(F_1, F_6)$:

$-\mathscr{G}_1 \cap \mathscr{G}_2 = \mathscr{G}_1 \cap \mathscr{G}_3 = \mathscr{G}_1 \cap \mathscr{G}_4 = \varnothing$ by definition

$-\mathscr{G}_2 \cap \mathscr{G}_3 = \varnothing$, since otherwise $Q_E^*(F_1, F_6)$ would satisfy (C-1)$-\mathscr{G}_2 \cap \mathscr{G}_4 = \varnothing$, since, for any $(LRX, UST) \in \mathscr{G}_3$, $|\varepsilon\mathscr{Q}(X)| \geq 2$ would imply $Q_E^*(F_1, F_6)$ fulfilling (C-4), while $|\varepsilon\mathscr{Q}(U)| \geq 2$ would imply (C-5)

$-\mathscr{G}_3 \cap \mathscr{G}_4 = \varnothing$, since, for any $(LRX, UST) \in \mathscr{G}_3$, $|\varepsilon\mathscr{Q}(X)| \geq 2$ would imply (C-5), while $|\varepsilon\mathscr{Q}(U)| \geq 2$ would imply (C-4)

We denote $E_{\mathscr{G}_1}, E_{\mathscr{G}_2}, E_{\mathscr{G}_3}$, and $E_{\mathscr{G}_4}$ the event $\mathrm{KAEF}_k^{F^*} \vdash \mathscr{G}_1, \mathscr{G}_2, \mathscr{G}_3$ and $\mathscr{G}_4$. It can be seen that

$$p(\tau, F_1, F_6) = \Pr\left[E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}3} \wedge E_{\mathscr{G}_4} | F \vdash \mathscr{Q}_F\right] \qquad (31)$$

We next analyze the four groups in turn. The first one, i.e., $\Pr[E_{\mathscr{G}_1} | F \vdash \mathscr{Q}_F]$, involves the most complicated analysis. Briefly, for each tuple (LRX, UST) in $\mathscr{G}_1$, it consists of three cases.

In the first case, neither of the two corresponding intermediate values Y and Z derived from F_2 and F_5 collides with values that have been in the history. The probability that KAEF_k^F extends (LRX, UST) in this case is roughly at least

$$\left(1 - \frac{q_f + q_e + \beta_1}{N}\right)\left(1 - \frac{q_f + q_e + \beta_2}{N}\right)\frac{1}{N^2}. \qquad (32)$$

In the second case, the corresponding intermediate value Y collides with some "existing" values, yet the further derived Z is "free." The probability that KAEF_k^F extends (LRX, UST) in this case is roughly at least

$$\left(\frac{q_f + q_e}{N} - O\left(\frac{2^r \cdot q_f^2}{N^2} + \frac{\left(q_f + q_e\right)^2}{N^2}\right)\right)\frac{1}{N^2}. \qquad (33)$$

The third case is symmetrical to the second one: Z collides with "existing" values, yet Y is "free." The probability is roughly at least

$$\left(\frac{q_f + q_e}{N} - O\left(\frac{\left(q_f + q_e\right)^2}{N^2}\right)\right)\frac{1}{N^2}. \qquad (34)$$

Summing over the above, we obtain

$$\Pr\left[E_{\mathscr{G}_1} | F \vdash \mathscr{Q}_F\right] \geq \prod_{l=1}^{|\mathscr{G}_1|}\left(1 - \frac{\beta_1}{N} - \frac{\beta_2}{N} - O\left(\frac{2^r \cdot q_f^2}{N^2} + \frac{\left(q_f + q_e\right)^2}{N^2}\right)\right)\frac{1}{N^2}. \qquad (35)$$

The concrete bound is

$$\mathbb{E}_k\left[\Pr\left[E_{\mathscr{G}_1} | F \vdash \mathscr{Q}_F\right]\right] \geq \left(1 - \frac{2^r \cdot q_e q_f^2}{N^2} - \frac{2q_e\left(2q_f + q_e\right)\left(q_f + q_e\right)}{N^2} - \frac{\left(q_f + 2q_e\right)\left(\beta_1 + \beta_2\right)}{N}\right)\frac{1}{N^{2|\mathscr{G}_1|}}. \qquad (36)$$

To analyze $E_{\mathscr{G}_2}, E_{\mathscr{G}_3}$, and $E_{\mathscr{G}_4}$, we again apply the bad predicate approach. These groups involve collisions and have relatively small sizes: $|E_{\mathscr{G}_2}|, |E_{\mathscr{G}_3}|, |E_{\mathscr{G}_4}| = O(2^r \cdot q^2/N)$ (will be proved later). Therefore, any collisions between tuples in these groups and values related to Q_F or $E_{\mathscr{G}_1}$ can be included in the bad predicates; for each tuple, in these three groups, the probability would be $O(q/N)$ with $q = \max(q_e, q_f)$, yet it remains $O(q/N) \cdot O(2^r \cdot q^2/N) = O(2^r \cdot q^3/N)$ in total. In all, the results are

$$\Pr\left[E_{\mathscr{G}_2} \wedge E_{\mathscr{G}3} | E_{\mathscr{G}_1} \wedge F \vdash \mathscr{Q}_F\right] \geq \left(1 - \frac{\left(\beta_1 + \beta_2\right)\left(q_f + q_e\right)}{N}\right)\frac{1}{N^{2\left(|\mathscr{G}_2| + |\mathscr{G}_3|\right)}},$$

$$\Pr\left[E_{\mathscr{G}_4} | E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}3} \wedge F \vdash \mathscr{Q}_F\right] \geq \left(1 - \frac{2\beta_3\left(q_f + q_e\right)}{N}\right)\frac{1}{N^{2|\mathscr{G}_4|}}. \qquad (37)$$

□

Proof. It is analyzes for $\mathscr{G}_2, \mathscr{G}_3$, and $\mathscr{G}_4$.

Lowering bounding $\Pr[E_{\mathcal{G}_2} \wedge E_{\mathcal{G}3} | E_{\mathcal{G}_1} \wedge F \vdash \mathcal{Q}_F]$. Consider $E_{\mathcal{G}_2}$ first; we lower bound that probability that is equivalent to F_4 and F_5, satisfying $2|\mathcal{G}_2|$ new and distinct equations. To this end, again, we define a predicate $\mathrm{Bad}_1(F_3)$, which holds if there exists $(LRX, UST) \in \mathcal{G}_2$ that fulfills one of the following conditions:

(i) The x_4 value derived using F_3 is in $\mathrm{Dom}\mathcal{F}_4$, i.e., $k_4 \oplus OF_3(x_3) \in \mathrm{Dom}\mathcal{F}_4$, where $x_3 = k_3 \oplus A_1 L \oplus A_2 R \oplus \mathrm{OIm}gF_2(k_2 \oplus X)$

(ii) The Z value derived using F_3 collides with Z' value of another tuple in $\mathcal{G}_2$, i.e., there exists $(L'R'X', U'S'T') \in \mathcal{G}_2$ such that $X \oplus OF_3(x_3) = X' \oplus OF_3(x_3')$, where $x_3' = k_3 \oplus A_1 L' \oplus A_2 R' \oplus \mathrm{OIm}gF_2(k_2 \oplus X')$

(iii) The Z value derived using F_3 collides with Z' value of another tuple in $\mathcal{G}_2$ or $\mathcal{G}_3$, i.e., there exists $(L^* R^* X^*, U^* S^* T^*) \in \mathcal{G}_1 \cup \mathcal{G}_3$ such that $X \oplus OF_3(x_3) = A_1 T^* \oplus A_2 S^* \oplus OF_5(k_5 \oplus U^*)$

We note that, for each $(LRX, UST) \in \mathcal{G}_2$, let $x_3 = k_3 \oplus A_1 L \oplus A_2 R \oplus \mathrm{OIm}gF_2(k_2 \oplus X)$; then, the following equation holds $x_3 \notin \mathrm{Dom}\mathcal{F}_3$(otherwise fulfilling (C-2)) and $x_3 \in \mathrm{Ext}\mathcal{F}_3^{(|\mathcal{G}_1|)}$(according to the analysis of $E_{\mathcal{G}_1}$). Thus, conditioned on $E_{\mathcal{G}_1} \wedge F_3 \vdash \mathcal{Q}_{F_3}$, the value $F_3(x_3)$ remains uniform. Therefore, for this (LRX, UST),

(i) The probability that condition (i) is fulfilled at most q_f/N

(ii) For each $(L'R'X', U'S'T') \in \mathcal{G}_2$, if the corresponding x_3' does not equal x_3, then the probability of $X \oplus OF_3(x_3) = X' \oplus OF_3(x_3')$ is at most $1/N$; otherwise, since the two tuples are distinct, it has to be $X \neq X'$, and thus, $X \oplus OF_3(x_3) \neq X' \oplus OF_3(x_3')$

(iii) For each $(L^* R^* X^*, U^* S^* T^*) \in \mathcal{G}_1 \cup \mathcal{G}_3$, the probability of $X \oplus OF_3(x_3) = A_1 T^* \oplus A_2 S^* \oplus OF_5(k_5 \oplus U^*)$ is at most $1/N$

Summing over the above yields

$$\Pr\left[\mathrm{Bad}_1(F_3) | E_{\mathcal{G}_1} \wedge F \vdash \mathcal{Q}_F\right] \leq \frac{|\mathcal{G}_2| \cdot \left(q_f + |\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|\right)}{N} \leq \frac{\beta_1 \cdot \left(q_f + q_e\right)}{N}. \tag{38}$$

It is not hard to see that conditioned on $\mathrm{Bad}_1(F_3)$, the $|\mathcal{G}_2|$ tuples in $\mathcal{G}_2$ indeed give rise to $|\mathcal{G}_2|$ distinct values $Z_1, \ldots, Z_{|\mathcal{G}_2|}$ (otherwise condition (ii) is fulfilled), for which $F_4(Z_1 \oplus k_4), \ldots, F_4(Z_{|\mathcal{G}_2|} \oplus k_4)$, all remain undetermined (otherwise condition (i) or (iii) is fulfilled)). Furthermore, at the "right side," they also give rise to $|\mathcal{G}_2|$ distinct values $U_1, \ldots, U_{|\mathcal{G}_2|}$ with $F_5(U_1 \oplus k_5), \ldots, F_5(U_{|\mathcal{G}_2|} \oplus k_5)$, all remain undetermined:

(i) $U_1, \ldots, U_{|\mathcal{G}_2|}$ are also distinct, otherwise fulfilling (C-5)

(ii) None of $U_1 \oplus k_5, \ldots, U_{|\mathcal{G}_2|} \oplus k_5$ is in $\mathrm{Dom}\mathcal{F}_5$, otherwise fulfilling (C-1)

(iii) Conditioned on $E_{\mathcal{G}_1}$, $F_5(U_1 \oplus k_5), \ldots, F_5(U_{|\mathcal{G}_2|} \oplus k_5)$, all remain undetermined, otherwise some U_i is shared between tuples in $\mathcal{G}_1$ and $\mathcal{G}_2$

Thus, in this case, the event $E_{\mathcal{G}_2}$ is equivalent to F_4 and F_5 satisfying $2|\mathcal{G}_2|$ new and distinct equations, and the probability does not exceed $1/N^{2|\mathcal{G}_2|}$.

We then consider $E_{\mathcal{G}_3}$. The analysis is similar to $E_{\mathcal{G}_2}$ by symmetry; we define a predicate $\mathrm{Bad}_1(F_4)$ on F_4, which holds if there exists $(LRX, UST) \in \mathcal{G}_3$ such that one of the following conditions is fulfilled:

(i) The induced value x_3 is in $\mathrm{Dom}\mathcal{F}_3$, i.e., $k_3 \oplus OF_4(x_4) \in \mathrm{Dom}\mathcal{F}_3$, where $x_4 = k_4 \oplus A_1 T \oplus A_2 S \oplus \mathrm{OIm}gF_5(k_5 \oplus U)$. The probability is at most q_f/N in total.

(ii) The induced value Y collides with the Y' value of another tuple in $\mathcal{G}_3$, i.e., there exists $(L'R'X', U'S'T') \in \mathcal{G}_3$ such that $U \oplus OF_4(x_4) = U' \oplus OF_4(x_4')$, where $x_4' = k_4 \oplus A_1 T' \oplus A_2 S' \oplus \mathrm{OIm}gF_5(k_5 \oplus U')$. The probability is at most $|\mathcal{G}_3|/N$.

(iii) The induced value Y collides with Y' value of another tuple in $\mathcal{G}_1$ or $\mathcal{G}_2$, i.e., there exists $(L^* R^* X^*, U^* S^* T^*) \in \mathcal{G}_1 \cup \mathcal{G}_2$ such that $U \oplus OF_4(x_4) = A_1 L^* \oplus A_2 R^* \oplus OF_2(k_2 \oplus X^*)$. The probability is at most $|\mathcal{G}_1| + |\mathcal{G}_2|/N$ in total.

Similar to $\mathrm{Bad}_1(F_3)$,

$$\Pr\left[\mathrm{Bad}_2(F_4) | E_{\mathcal{G}_1} \wedge F \vdash \mathcal{Q}_F\right] \leq \frac{|\mathcal{G}_3| \cdot \left(q_f + |\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|\right)}{N} \leq \frac{\beta_2 \cdot \left(q_f + q_e\right)}{N}. \tag{39}$$

And, conditioned on $\mathrm{Bad}_1(F_4)$, tuples in $\mathcal{G}_3$ give rise to $|\mathcal{G}_3|$ distinct values $Y_1, \ldots, Y_{|\mathcal{G}_3|}$, while the assumption $\mathrm{Bad}_1(F_1, F_6)$ ensures that they give rise to $|\mathcal{G}_3|$ distinct values $X_1, \ldots, X_{|\mathcal{G}_3|}$. Thus, the event $E_{\mathcal{G}_3}$ is equivalent to F_2 and F_3 satisfying $2|\mathcal{G}_3|$ new and distinct equations. Therefore, conditioned on $E_{\mathcal{G}_1} \wedge F \vdash \mathcal{Q}_F$, we have

$$\begin{aligned}\Pr\left[E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}|E_{\mathcal{G}_1}\wedge F\vdash\mathcal{Q}_F\right] &\geq (1-\Pr[\mathrm{Bad}_1(F_2)]-\Pr[\mathrm{Bad}_1(F_3)])\cdot\Pr\left[E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}|\mathrm{Bad}_1(F_2)\wedge\mathrm{Bad}_1(F_3)\right]\\ &\geq \left(1-\frac{(\beta_1+\beta_2)\cdot\left(q_f+q_e\right)}{N}\right)\frac{1}{N^{2(|\mathcal{G}_2|+|\mathcal{G}_3|)}}.\end{aligned} \tag{40}$$

Lower bounding $\Pr[E_{\mathcal{G}_4}|E_{\mathcal{G}_1}\wedge E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}\wedge F\vdash\mathcal{Q}_F]$: by definition, for any tuple $(LRX, UST)\in\mathcal{G}_4$, let $x_2 = k_2\oplus X$ and $x_5 = k_5\oplus U$; then, we have both $x_2\notin\mathrm{Dom}\mathcal{F}_2$ and $x_5\notin\mathrm{Dom}\mathcal{F}_5$. Moreover, conditioned on $E_{\mathcal{G}_1}\wedge E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}$, the two values $F_2(x_2)$ and $F_5(x_5)$ remain "undetermined" and uniform (otherwise, if $E_{\mathcal{G}_1}$, $E_{\mathcal{G}_2}$, and $E_{\mathcal{G}_3}$ imply $F_2(x_2)$ being fixed, then a tuple in $\mathcal{G}_1$, $\mathcal{G}_2$, or $\mathcal{G}_3$ would share the same X value with a tuple in $\mathcal{G}_4$, contradicting the definition of $\mathcal{G}_1$ or fulfilling (C-4) or (C-5), respectively).

For these tuples, we would lower bound the probability that they induce $2|\mathcal{G}_4|$ new and distinct equations on F_3 and F_4. To this end, we define a predicate $\mathrm{Bad}_3(F_2, F_5)$ on F_2 and F_5, which holds if there exists a tuple $(LRX, UST)\in\mathcal{G}_4$ such that if we let $x_2 = k_2\oplus X$ and $x_5 = k_5\oplus U$, then one of the following conditions is fulfilled:

At the "left side," concerning $F_2(x_2)$,

(i) The induced x_3 value falls in $\mathrm{Dom}\mathcal{F}_3$, i.e., $k_3\oplus A_1L\oplus A_2R\oplus OF_2(x_2)\in\mathrm{Dom}\mathcal{F}_3$. As discussed, $F_2(x_2)$ remains random; thus, the probability is clearly at most q_f/N.

(ii) The induced Y value collides with some "previously determined" Y', i.e., there exists another tuple $(L'R'X', U'S'T')\in\mathcal{G}_1\cup\mathcal{G}_2\cup\mathcal{G}_3$ such that $A_1L\oplus A_2R\oplus OF_2(x_2) = A_1L'\oplus A_2R'\oplus OF_2(x_2')$. It necessarily is $X\neq X'$; again, using the randomness of $F_2(x_2)$, we obtain the upper bound $|\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|/N\leq q_e/N$, for each $(LRX, UST)\in\mathcal{G}_4$.

At the "right side," concerning $F_5(x_5)$, similar to the above symmetry,

(i) $k_4\oplus A_1T\oplus A_2S\oplus HF_5(x_5)\in\mathrm{Dom}\mathcal{F}_4$: the probability is clearly at most q_f/N.

(ii) There exists another tuple $(L'R'X', U'S'T')\in\mathcal{G}_1\cup\mathcal{G}_2\cup\mathcal{G}_3$ such that $A_1T\oplus A_2S\oplus OF_5(x_5) = A_1T'\oplus A_2S'\oplus OF_5(x_5')$. The upper bound is $|\mathcal{G}_1| + |\mathcal{G}_2| + |\mathcal{G}_3|/N\leq q_e/N$, for each $(LRX, UST)\in\mathcal{G}_4$.

Thus, using $|\mathcal{G}_4| = \beta_3$, we obtain

$$\Pr\left[\mathrm{Bad}_3(F_2, F_5)|E_{\mathcal{G}_1}\wedge E_{\mathcal{G}2}\wedge E_{\mathcal{G}_3}\wedge F\vdash\mathrm{Q}_F\right]\leq 2\frac{|\mathcal{G}_4|\cdot\left(q_f+q_e\right)}{N}\leq\frac{2\beta_3\cdot\left(q_f+q_e\right)}{N}. \tag{41}$$

Similar to the analysis for $E_{\mathcal{G}_2}$ and $E_{\mathcal{G}_3}$, conditioned on $\mathrm{Bad}_3(F_2, F_5)$, the event $E_{\mathcal{G}_4}$ is equivalent to F_3 and F_4 satisfying $2|\mathcal{G}_4|$ new and distinct equations. Therefore,

$$\Pr\left[E_{\mathcal{G}_4}|E_{\mathcal{G}_1}\wedge E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}\wedge F\vdash\mathcal{Q}_F\right]\geq(1-\Pr[\mathrm{Bad}_3(F_2, F_5)])\cdot\frac{1}{N^{2|\mathcal{G}_4|}}\geq\left(1-\frac{2\beta_3\left(q_f+q_e\right)}{N}\right)\cdot\frac{1}{N^{2|\mathcal{G}_4|}}. \tag{42}$$

Summing up would yield a lower bound of the form

$$\begin{aligned}p(\tau, F_1, F_6) = \Pr\left[E_{\mathcal{G}_1}\wedge E_{\mathcal{G}_2}\wedge E_{\mathcal{G}_3}\wedge E_{\mathcal{G}_4}|F\vdash\mathcal{Q}_F\right] &\geq (1-\varepsilon_1)(1-\varepsilon_2)(1-\varepsilon_3)\frac{1}{N^{2(|\mathcal{G}_1|+|\mathcal{G}_2|+|\mathcal{G}_3|+|\mathcal{G}_4|)}}\\ &\geq (1-(\varepsilon_1+\varepsilon_2+\varepsilon_3))\frac{1}{N^{2q_e}}\left(\text{Since }\left|\mathcal{G}_1\right|+\left|\mathcal{G}_2\right|+\left|\mathcal{G}_3\right|+\left|\mathcal{G}_4\right| = q_e\right).\end{aligned} \tag{43}$$

We note $1/N^{2q_e} / \prod_{i=0}^{q_e-1} 1/N^2 - i \geq (1-(q_e/N^2))^{q_e} \geq 1-(q_e^2/N^2) \geq 1-(q_e^3/N^2)$. Thus, using $(1-A)(1-B) \geq 1-(A+B)$, we obtain

$$\frac{p(\tau, F_1, F_6)}{\prod_{i=0}^{q_e-1} 1/N^2 - i} \geq 1 - \varepsilon(F_1, F_6, k), \tag{44}$$

for which

$$\mathbb{E}_k[\varepsilon(F_1, F_6, k)] \leq \frac{(2q_f + 3q_e)(\beta_1 + \beta_2) + 2\beta_3(q_f + q_e)}{N} + \frac{2^r \cdot q_e q_f^2}{N^2} + \frac{2q_e(2q_f + q_e)(q_f + q_e) + q_e^3}{N^2}. \tag{45}$$

We now derive $\mathbb{E}_{F_1,F_6}[\mathbb{E}_k[\varepsilon(F_1,F_6,k)] | F_1 \vdash \mathcal{Q}_{F_1}, F_6 \vdash \mathcal{Q}_{F_6}]$. To this end, note that, by definition, β_1, β_2, and β_3 are quantities that depend on (F_1, F_6):

$$\begin{aligned}
\beta_1 &= \left|\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\colon k_2 \oplus X = k_2 \oplus A_1 R \oplus A_2 L \oplus OF_1(k_1 \oplus A_1 L \oplus A_2 R) \in \mathrm{Dom}\mathcal{F}_2\}\right|,\\
\beta_2 &= \left|\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\colon k_5 \oplus U = k_5 \oplus A_1 S \oplus A_2 T \oplus OF_6(k_6 \oplus A_1 T \oplus A_2 S) \in \mathrm{Dom}\mathcal{F}_5\}\right|,\\
\beta_3 &= \left|\{(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)\colon \exists(L'R'X', U'S'T') \text{ such that } X = X', \text{ or}: \exists(L'R'X', U'S'T') \in \mathcal{Q}_E^*(F_1, F_6) \text{ such that } U = U'\}\right|.
\end{aligned} \tag{46}$$

We consider β_1 first. For each $(LRX, UST) \in \mathcal{Q}_E^*(F_1, F_6)$, if $k_1 \oplus A_1 L \oplus A_2 R \in \mathrm{Dom}\mathcal{F}_1$, then $k_2 \oplus X \notin \mathrm{Dom}\mathcal{F}_2$ by $(B-2)$. Thus, conditioned on $F_1 \vdash \mathcal{Q}_{F_1}$, $F_1(k_1 \oplus R)$ remains uniform, and $\Pr[k_2 \oplus A_1 R \oplus A_2 L \oplus OF_1(k_1 \oplus A_1 L \oplus A_2 R) \in \mathrm{Dom}\mathcal{F}_2] \leq q_f/N$. Therefore,

$$\mathbb{E}_k[\beta_1] \leq \frac{q_e q_f}{N}. \tag{47}$$

Similarly, by symmetry, using the randomness supplied by F_6, $\mathbb{E}_k[\beta_2] \leq q_e q_f/N$.

Then, we consider β_3. We fix a record (LR, ST) such that $k_1 \oplus A_1 L \oplus A_2 R \notin \mathrm{Dom}\mathcal{F}_1$, and consider another $(L'R', S'T')$. If $A_1 L \oplus A_2 R = A_1 L' \oplus A_2 R'$, then it has to be 2^n, and thus, $X = X'$. Otherwise, as $k_1 \oplus A_1 L \oplus A_2 R \notin \mathrm{Dom}\mathcal{F}_1$, $F_1(k_1 \oplus A_1 L \oplus A_2 R)$ remains random conditioned on $F_1 \vdash \mathcal{Q}_{F_1}$, and

$$\Pr[X = X'] = \Pr[OF_1(k_1 \oplus A_1 L \oplus A_2 R) \oplus OF_1(k_1 \oplus A_1 L' \oplus A_2 R') = A_1 R \oplus A_2 L \oplus A_1 R' \oplus A_2 L'] = \frac{1}{N}. \tag{48}$$

The number of distinct pairs of such tuples is at most q_e^2. Thus, we know the expectation of the number of pairs:

$$\mathbb{E}_k[\{(LRX, UST)\colon k_1 \oplus A_1 L \oplus A_2 R \notin \mathrm{Dom}\mathcal{F}_1, \exists(L'R'X', U'S'T') \text{ s.t. } X = X'\}|] \leq \frac{q_e^2}{N}. \tag{49}$$

As the number of (LR, ST) such that $k_1 \oplus A_1 L \oplus A_2 R \in \mathrm{Dom}\mathcal{F}_1$ is $\alpha_1(k)$, we obtain

$$\mathbb{E}_k[\{(LRX, UST)\colon \exists(L'R'X', U'S'T') \text{ s.t. } X = X'\}|] \leq \frac{q_e^2}{N} + \alpha_1(k). \tag{50}$$

Symmetrically, $\mathbb{E}_k[\{(LRX, UST)\colon \exists(L'R'X', U'S'T') \text{ s.t. } U = U'\}|] \leq q_e^2/N + \alpha_2(k)$. Thus, $\mathbb{E}_k[\beta_3] \leq (2q_e^2/N) + \alpha_1(k) + \alpha_2(k)$. Finally, since k_1 and resp. k_6 are uniform in 2^n and resp. 2^{n-r} possibilities,

$$\mathbb{E}_k[\alpha_1(k)] = \sum_{(LR,ST)\in\mathcal{Q}_E}\ \sum_{(x_1,y_1)\in\mathcal{Q}_{F_1}} \Pr[k_1 = A_1L\oplus A_2R\oplus x_1] \le \frac{q_eq_f}{N}. \tag{51}$$

and $\mathbb{E}_k[\alpha_2(k)] \le 2^r \cdot q_eq_f/N$. Gathering all the above yields

$$\begin{aligned}\mathbb{E}_{F_1,F_6,k}[\varepsilon(F_1,F_6,k)] &\le \frac{4q_eq_f^2 + 6q_e^2q_f}{N^2} + \frac{2\left(q_f+q_e\right)\left(2^rq_eq_f + q_eq_f + 2q_e^2\right)}{N^2} + \frac{2^r\cdot q_eq_f^2}{N^2}\\ &+ \frac{2q_e\left(2q_f+q_e\right)\left(q_f+q_e\right)+q_e^3}{N^2} = \frac{7q_e^3 + 10q_eq_f^2 + 18q_e^2q_f + 3\cdot 2^r\cdot q_eq_f^2 + 2\cdot 2^r\cdot q_e^2q_f}{N^2},\end{aligned} \tag{52}$$

as claimed in (8). □

5.3. Concluding the Pointwise Proximity Proof. Gathering Lemma 2, Lemma 4, and equation (22), we obtain

$$\frac{\Pr_{\mathrm{re}}(\tau)}{\Pr_{\mathrm{id}}(\tau)} \ge 1 - \left(\frac{3\cdot 2^r\cdot q_eq_f^2}{N^2} + \mathbb{E}_k\left[\Pr\left[\mathrm{Bad}(F_1,F_6)|F_1\vdash\mathcal{Q}_{F_1},F_6\vdash\mathcal{Q}_{F_6}\right] + \mathbb{E}_k\left[\mathbb{E}_{F_1,F_6}\left[\varepsilon(F_1,F_6,k)|F_1\vdash\mathcal{Q}_{F_1},F_6\vdash\mathcal{Q}_{F_6}\right]\right]\right), \tag{53}$$

where $\varepsilon(F_1,F_6,k)$ is the function specified in equation (44). Note that its expectation has been bounded in Lemma 5.

For $\mathbb{E}_k[\Pr[\mathrm{Bad}(F_1,F_6)|F_1\vdash\mathcal{Q}_{F_1},F_6\vdash\mathcal{Q}_{F_6}]$, since k_3 and resp. k_4 are both uniformly distributed in 2^n and resp. 2^{n-r} possibilities, we have

$$\begin{aligned}\mathbb{E}_k\left[\alpha_{2,3}(k)\right] &\le \frac{q_eq_f^2}{N},\\ \mathbb{E}_k\left[\alpha_{4,5}(k)\right] &\le \frac{2^rq_eq_f^2}{N}.\end{aligned} \tag{54}$$

At the end of previous section, we have shown $\mathbb{E}_k[\alpha_1(k)] \le q_eq_f/N$ and $\mathbb{E}_k[\alpha_1(k)] \le 2^rq_eq_f/N$. Injecting them into the bound of Lemma 3 yields

$$\mathbb{E}_k\left[\Pr\left[\mathrm{Bad}(F_1,F_6)|F_1\vdash\mathcal{Q}_{F_1},F_6\vdash\mathcal{Q}_{F_6}\right] \le \frac{3q_eq_f^2}{N^2} + \frac{2^rq_eq_f^2}{N^2} + \frac{4q_e^2q_f}{N^2}. \tag{55}$$

Gathering all the above eventually establishes equation (21).

6. Conclusion

For diversity of iterative structures, we propose an extended Feistel structure. The new iterative structure also has similar encryption with decryption. This paper mainly investigates the security of the new structure from studying the distinguishability between the ideal cipher adopting this structure and a random permutation. Results show that, for birthday-bound security, the 4-round KAEF is secure against $2^{n/2}$ queries and for beyond-birthday-bound security, the 6-round KAEF is secure against $2^{2n/3}$ queries. As a result, the new iterative is a reliable structure and can provide more choices for cipher designing.

Appendix

A. Proof of Lemma 3

We upper bound the probabilities of the bad conditions in turn.

A.1. Condition (C-1). For any $(LRX, AST) \in Q_E^*(F_1,F_6)$, if there exist $(x_2,y_2)\in Q_{F_2}$ and $(x_5,y_5)\in Q_{F_5}$ such that $k_2 = X\oplus x_2$ and $k_5 = A\oplus x_5$, then we would have $A_1R\oplus A_2L\oplus OF_1(A_1L\oplus A_2R\oplus k_1) = k_2\oplus x_2$ and $A_1S\oplus A_2T \oplus OF_6(A_1T\oplus A_2S\oplus k_6) = k_5\oplus x_5$ for the corresponding $(LR,ST)\in Q_E$. It cannot be $A_1L\oplus A_2R\oplus k_1 \in \mathrm{Dom}\mathscr{F}_1$, as otherwise $A_1L\oplus A_2R\oplus k_1 \in \mathrm{Dom}\mathscr{F}_1$ along with fulfilling (B-2) in Definition 3; similarly, it cannot be $A_1T\oplus A_2S\oplus k_6 \in \mathrm{Dom}\mathscr{F}_6$. Thus, conditioned on $F_1\vdash Q_{F_1}$ and $F_6\vdash Q_{F_6}$, the two values $F_1(A_1L\oplus A_2R\oplus k_1)$ and $F_6(A_1T\oplus A_2S\oplus k_6)$ remain uniform. Thus, for each 3-tuple $((LR,ST),(x_2,y_2),(x_5,y_5))$, the probability that the case of both $A_1R\oplus A_2L\oplus OF_1(A_1L\oplus A_2R\oplus k_1) = k_2\oplus x_2$ and $A_1S\oplus A_2T\oplus OF_6(A_1T\oplus A_2S\oplus k_6) = k_5\oplus x_5$ hold is at most $1/N^2$. Since we have at most $q_eq_f^2$ such 3-tuples, the total probability does not exceed $q_eq_f^2/N^2$.

A.2. Conditions (C-2) and (C-3). Consider (C-2) first. By definitions, the number of triplets $((LRX,AST),(x_2,y_2),(x_3,y_3))$ such that $k_3 = A_1L\oplus A_2R\oplus Oy_2\oplus x_3$ is $\alpha_{2,3}(k)$, where (LRX,AST) is a "merged" notation for (LR,ST) and the corresponding induced X and A. On the contrary, $k_2 = X\oplus x_2$ would imply $A_1R\oplus A_2L\oplus OF_1(A_1L\oplus A_2R\oplus k_1) = k_2\oplus x_2$. Now, if $A_1L\oplus A_2R\oplus k_1\in\mathrm{Dom}\mathscr{F}_1$, then it cannot be $A_1R\oplus A_2L\oplus O\mathrm{Img}F_1(A_1L\oplus A_2R\oplus k_1) = k_2\oplus x_2$, otherwise (B-2) is fulfilled. Whereas when $A_1L\oplus A_2R\oplus k_1\notin\mathrm{Dom}\mathscr{F}_1$, then conditioned on $F_1 \vdash Q_{F_1}$, the value $F_1(A_1L \oplus A_2R \oplus k_1)$ is uniform, and thus, $\Pr[A_1R \oplus A_2L \oplus OF_1(A_1L \oplus A_2R \oplus k_1) = k_2 \oplus x_2] = 1/N$. As a result, we have $\Pr[(C-2)] \le \alpha_{2,3}(k)/N$. For condition (C-3), it is similar by symmetry, resulting in $\Pr[(C-3)] \le \alpha_{4,5}(k)/N$.

A.3. Condition (C-4). Consider the first half of (C-4) first, and consider such two tuples (LRX, AST) and $(L'R'X', A'S'T')$. We note that neither $A_1L \oplus A_2R \oplus k_1$ nor $A_1L' \oplus A_2R' \oplus k_1$ can be in $\mathrm{Dom}\mathcal{F}_1$, as otherwise it satisfies (B-2). Thus, conditioned on $F_1 \vdash Q_{F_1}$, both $F_1(A_1L \oplus A_2R \oplus k_1)$ and $F_1(A_1L' \oplus A_2R' \oplus k_1)$ remain uniform. Thus,

$$\begin{aligned}\Pr[X = X' \wedge k_2 \oplus X = x_2] &\le \Pr[OF_1(A_1L \oplus A_2R \oplus k_1) = OF_1(A_1L' \oplus A_2R' \oplus k_1) \oplus A_1(R \oplus R') \oplus A_2(L \oplus L')] \\ &\cdot \Pr[OF_1(A_1L \oplus A_2R \oplus k_1) = A_1R \oplus A_2L \oplus k_2 \oplus x_2] \le \frac{1}{N^2}.\end{aligned} \tag{A.1}$$

The number of choices of (LRX, AST), $(L'R'X', A'S'T')$, and (x_2, y_2) is at most $q_e^2 q_f$; thus, the probability of the first half is at most $q_e^2 q_f / N^2$ in total. For the second half, it is similar by symmetry, leading to the same bound $q_e^2 q_f / N^2$. Thus, $\Pr[(C-4)] \le 2q_e^2 q_f / N^2$.

A.4. Condition (C-5). Consider the first half of the condition, and consider such three tuples (LRX, AST), $(L'R'X', A'S'T')$, and $(x_2, y_2) \in Q_{F_2}$, respectively. By $(B-2)$, $A_1L \oplus A_2R \oplus k_1 \notin \mathrm{Dom}\mathcal{F}_1$, depending on the state of S and T, we distinguish two cases.

(i) Case 1: $A_1T \oplus A_2S \oplus k_6 \notin \mathrm{Dom}\mathcal{F}_6$. Then, we have at most q_e choices for (LRX, AST) and at most q_e choices for $(L'R'X', A'S'T')$. Conditioned on $F_6 \vdash Q_{F_6}$, $F_6(A_1T \oplus A_2S \oplus k_6)$ remains random; thus,

$$\Pr[A = A'] = \Pr[A_1S \oplus A_2T \oplus OF_6(A_1T \oplus A_2S \oplus k_6) = A_1S' \oplus A_2T' \oplus OF_6(A_1T' \oplus A_2S' \oplus k_6)] \le \frac{1}{N}. \tag{A.2}$$

Conditioned on $F_1 \vdash Q_{F_1}$, the value $F_1(A_1L \oplus A_2R \oplus k_1)$ is also random; thus, we similarly have $\Pr[A_1R \oplus A_2L \oplus OF_1(A_1L \oplus A_2R \oplus k_1) \oplus k_2 \in \mathrm{Dom}\mathcal{F}_2] \le q_f/N$. Thus, the probability is at most $q_e^2 q_f / N^2$ in total;

(ii) Case 2: $A_1T \oplus A_2S \oplus k_6 \in \mathrm{Dom}\mathcal{F}_6$. Then, we have $\alpha_2(k)$ choices for (LRX, AST). Similar to Case 1, $\Pr[A_1R \ \oplus A_2L \oplus OF_1(A_1L \oplus A_2R \oplus k_1) \oplus k_2 \in \mathrm{Dom}\,\mathcal{F}_2] \le q_f/N$. Thus, the probability that there exists at least one such tuple (LRX, AST) is at most $q_f \alpha_2(k)/N$.

Summing over the two cases results in $q_e^2 q_f / N^2 + q_f \alpha_2(k)/N$, the analysis for the second half is similar by symmetry, giving $q_e^2 q_f / N^2 + q_f \alpha_1(k)/N$. Thus, $\Pr[(C-5)] \le 2q_e^2 q_f / N^2 + q_f(\alpha_2(k) + \alpha_1(k))/N$.

B. Proof of Lemma 5

Lower bounding the probability is $\Pr[E_{\mathcal{G}_1} | F \vdash Q_F]$. We write $\mathcal{G}_1 = \{(L_1R_1X_1, A_1S_1T_1), \ldots, (L_{|\mathcal{G}_1|}R_{|\mathcal{G}_1|}X_{|\mathcal{G}_1|}, A_{|\mathcal{G}_1|}S_{|\mathcal{G}_1|}T_{|\mathcal{G}_1|})\}$ using some arbitrary order. Let E_l be the event that KAEF_k^F extends the lth tuple $(L_lR_lX_l, A_lS_lT_l)$. Then, $E_{\mathcal{G}_1} = E_{|\mathcal{G}_1|} \wedge \cdots \wedge E_1$.

We next focus on lower bounding $\Pr[E_{l+1} | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F]$ for the $(l+1)$th tuple $(L_{l+1}R_{l+1}X_{l+1}, A_{l+1}S_{l+1}T_{l+1})$. The approach is to lower bound the probability that E_{l+1} is equivalent to 2 new and distinct equations on F_2, F_3, F_4, and F_5. For this, we define four sets for positions "occupied by previous tuples:"

$$\begin{aligned}
\mathrm{Ext}\mathcal{F}_3^{(l)} &\overset{\mathrm{def}}{=} \left\{ \begin{matrix} x_3\colon \exists (L_iR_iX_i, A_iS_iT_i) \in \mathcal{G}_1 & i \le l \\ \text{s.t.} & x_3 = k_3 \oplus A_1L_i \oplus A_2R_i \oplus OF_2(k_2 \oplus X_i) \end{matrix} \right\}, \\
\mathcal{G}_2\mathcal{F}_3 &\overset{\mathrm{def}}{=} \{x_3\colon \exists (LRX, AST) \in \mathcal{G}_2 \text{ s.t. } x_3 = k_3 \oplus A_1L \oplus A_2R \oplus \mathrm{OIm}gF_2(k_2 \oplus X)\}, \\
\mathrm{Ext}\mathcal{F}_4^{(l)} &\overset{\mathrm{def}}{=} \left\{ \begin{matrix} x_4\colon \exists (L_iR_iX_i, A_iS_iT_i) \in \mathcal{G}_1 & i \le l \\ \text{s.t.} & x_4 = k_4 \oplus A_1T_i \oplus A_2S_i \oplus OF_5(k_5 \oplus A_i) \end{matrix} \right\}, \\
\mathcal{G}_3\mathcal{F}_4 &\overset{\mathrm{def}}{=} \{x_4\colon \exists (LRX, AST) \in \mathcal{G}_3 \text{ s.t. } x_4 = k_4 \oplus A_1T \oplus A_2S \oplus \mathrm{OIm}gF_5(k_5 \oplus A)\}.
\end{aligned} \tag{B.1}$$

We note that, for any $x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)}$, conditioned on $E_l \wedge \cdots \wedge E_1$, the value $F_3(x_3)$ has been "fixed" according to a corresponding tuple and cannot be deemed random, for $F_4(x_4)$ with $x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)}$.

Let $x_2^{(l+1)} = k_2 \oplus X_{l+1}$ and $x_5^{(l+1)} = k_5 \oplus A_{l+1}$. Then, given the round functions F_2, F_3, F_4, and F_5, the two intermediate values Y_{l+1} and Z_{l+1} would be determined. Depending on their state, the event E_{l+1} consists of at least three cases:

(i) Case 1 ("no collision"): the two induced values $Y_{l+1} = A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus OF_2(x_2^{(l+1)})$ and $Z_{l+1} = A_1 T_{l+1} \oplus A_2 S_{l+1} \oplus OF_5(x_5^{(l+1)})$ satisfy

$$\begin{aligned} k_3 \oplus Y_{l+1} &\notin \mathrm{Dom}\mathscr{F}_3 \cup \mathrm{Ext}\mathscr{F}_3^{(l)} \cup \mathscr{G}_2\mathscr{F}_3, \\ k_4 \oplus Z_{l+1} &\notin \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4. \end{aligned} \tag{B.2}$$

Then, the following equation holds $F_3(k_3 \oplus Y_{l+1}) = O^{-1}(X_{l+1} \oplus Z_{l+1})$ and $F_4(k_4 \oplus Z_{l+1}) = O^{-1}(Y_{l+1} \oplus A_{l+1})$.

(ii) Case 2 ("left collision"): the induced $Y_{l+1} = A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus OF_2(x_2^{(l+1)})$ satisfies

$$k_3 \oplus Y_{l+1} \in \mathrm{Dom}\mathscr{F}_3 \cup \mathrm{Ext}\mathscr{F}_3^{(l)}, \tag{B.3}$$

but the further induced value $Z_{l+1} = X_{l+1} \oplus OF_3(k_3 \oplus Y_{l+1})$ satisfies

$$k_4 \oplus Z_{l+1} \notin \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4. \tag{B.4}$$

Then, the following equation holds $F_4(k_4 \oplus Z_{l+1}) = O^{-1}(Y_{l+1} \oplus A_{l+1})$ and $F_5(x_5^{(l+1)}) = O^{-1}(A_1 T_{l+1} \oplus A_2 S_{l+1} \oplus Z_{l+1})$.

(iii) Case 3 ("right collision"): similar to Case 2 by symmetry, the induced value $Z_{l+1} = A_1 T_{l+1} \oplus A_2 S_{l+1} \oplus OF_5(x_5^{(l+1)})$ satisfies $k_4 \oplus Z_{l+1} \in \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}\mathscr{F}_4^{(l)}$, but the further induced $Y_{l+1} = A_{l+1} \oplus OF_4(k_4 \oplus Z_{l+1})$ satisfies $k_3 \oplus Y_{l+1} \notin \mathrm{Dom}\mathscr{F}_3 \cup \mathrm{Ext}\mathscr{F}_3^{(l)} \cup \mathscr{G}_2\mathscr{F}_3$. Then, the following equation holds $F_2(x_2^{(l+1)}) = O^{-1}(A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus Y_{l+1})$ and $F_3(k_3 \oplus Y_{l+1}) = O^{-1}(X_{l+1} \oplus Z_{l+1})$.

By these, we have

$$\Pr\left[E_{l+1} | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] = \sum_{i=1,2,3} \Pr\left[E_{l+1} \wedge \mathrm{CASE}i | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right]. \tag{B.5}$$

Let $e_3^{(l)} = |\mathrm{Ext}\mathscr{F}_3^{(l)}/\mathrm{Dom}\mathscr{F}|$ and $e_4^{(l)} = |\mathrm{Ext}\mathscr{F}_4^{(l)}/\mathrm{Dom}\mathscr{F}_4|$. We use three sections to bound each probability in turn.

B.1. Case 1. As $(L_{l+1} R_{l+1} X_{l+1}, A_{l+1} S_{l+1} T_{l+1})$ is in $\mathscr{G}_1$, we have $x_2^{(l+1)} = k_2 \oplus X_{l+1} \notin \mathrm{Dom}\mathscr{F}_2$. Furthermore, X_{l+1} does not collide with any other tuples in $Q_E^*(F_1, F_6)$ since $\varepsilon Q(X_{l+1}) = 1$. So, conditioned on $E_l \wedge \cdots \wedge E_1 \wedge E_1 \wedge F \vdash Q_F$, $F_2(x_2^{(l+1)})$ remains random and

$$\Pr\left[k_3 \oplus Y_{l+1} \in \mathrm{Dom}\mathscr{F}_3 \cup \mathrm{Ext}F_3^{(l)} \cup \mathscr{G}_2\mathscr{F}_3\right] \le \frac{\left(q_f + e_3^{(l)} + \left|\mathscr{G}_2\mathscr{F}_3\right|\right)}{N}. \tag{B.6}$$

By symmetry,

$$\Pr\left[k_4 \oplus Z_{l+1} \in \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}F_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4\right] \le \frac{\left(q_f + e_4^{(l)} + \left|\mathscr{G}_3\mathscr{F}_4\right|\right)}{N}. \tag{B.7}$$

Then, it can be seen the two equations $F_3(k_3 \oplus Y_{l+1}) = O^{-1}(X_{l+1} \oplus Z_{l+1})$ and $F_4(k_4 \oplus Z_{l+1}) = O^{-1}(Y_{l+1} \oplus A_{l+1})$ are fulfilled with probability $1/N^2$, and thus,

$$\Pr\left[E_{l+1} \wedge \mathrm{CASE1} | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] \ge \left(1 - \frac{\left(q_f + e_3^{(l)} + \left|\mathscr{G}_2\mathscr{F}_3\right|\right)}{N}\right)\left(1 - \frac{\left(q_f + e_4^{(l)} + \left|\mathscr{G}_2\mathscr{F}_3\right|\right)}{N}\right)\frac{1}{N^2}. \tag{B.8}$$

One may notice that if we only consider this Case 1, then we would end up with an undesired birthday-type bound since $|\mathscr{G}_4| = O(q_e)$.

However, this gap is filled in by the other two cases analyzed below.

B.2. Case 2. Recall that, in this case,

$$k_3 \oplus Y_{l+1} = x_3 \in \mathrm{Dom}\mathscr{F}_3 \cup \mathrm{Ext}\mathscr{F}_3^{(l)}, \text{ while } k_4 \oplus Z_{l+1} = x_4 \notin \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4. \tag{B.9}$$

Instead of lower bounding $\Pr[E_{l+1} \wedge \text{CASE2} | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F]$, we upper bound the probability of the opposite case. It can be seen that Y_{l+1} colliding with the involved (x_3, x_4) implies that $X_{l+1} \oplus O y_3 = (k_4 \oplus x_4)$, where $y_3 = F_3(x_3)$. Therefore, we proceed to upper bound:

$$p\text{coll} = \Pr\left[\exists x_3 \in \text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}, \exists x_4 \in \text{Dom}\mathscr{F}_4 \cup \text{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4\colon \text{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right], \tag{B.10}$$

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

$$X_{l+1} \oplus O y_3 = (k_4 \oplus x_4) \wedge A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus O F_2\left(x_2^{(l+1)}\right) = (k_3 \oplus x_3), \tag{B.11}$$

where $x_2^{(l+1)} = k_2 \oplus X_{l+1}$.

In detail, the to-be-bounded probability could be written as

$$p\text{coll} = \sum_{\substack{x_3 \in \text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}, \\ x_4 \in \text{Dom}\mathscr{F}_4 \cup \text{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3\mathscr{F}_4}} \Pr[\text{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F]. \tag{B.12}$$

Let $x_2^{(l+1)} = k_2 \oplus X_{l+1}$. In the following, we distinguish five subcases and derive bound for each in turn.

B.2.1. Subcase 2.1. $x_3 \in \text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}$, and $x_4 \in \mathscr{G}_3\mathscr{F}_4$. Define $\text{Num}_3^l(y_3)$ as the number of preimages of y_3 under the map defined by $\text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}$, i.e.,

$$\text{Num}_3^l(y_3) = \left|\left\{x_3 \in \text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}\colon F_3(x_3) = y_3\right\}\right|. \tag{B.13}$$

By this and by the constraint that $X_{l+1} \oplus O y_3 = k_4 \oplus x_4$, for each $x_4 \in \mathscr{G}_3\mathscr{F}_4$, the number of $x_3 \in \text{Dom}\mathscr{F}_3 \cup \text{Ext}\mathscr{F}_3^{(l)}$ such that $X_{l+1} \oplus O y_3 = k_4 \oplus x_4$ is $\text{Num}_3^{(l)}(X_{L+1} \oplus k_4 \oplus x_4)$. Therefore, the number of such "bad" pair is $\sum_{x_4 \in \mathscr{G}_3\mathscr{F}_4} \text{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)$ in total. On the contrary, similar to Case 1, $F_2(x_2^{(l+1)})$ can still be deemed random; thus,

$$\Pr\left[F_2\left(x_2^{(l+1)}\right) = O^{-1}(A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus k_3 \oplus x_3)\right] \leq \frac{1}{N} \tag{B.14}$$

and

$$\sum_{\substack{x_3 \in Dom\mathscr{F}_3 \cup Ext\mathscr{F}_3^{(l)}, \\ x_4 \in \mathscr{G}_3\mathscr{F}_4}} \Pr[\text{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq \frac{\sum_{x_4 \in \mathscr{G}_3\mathscr{F}_4} \text{Num}_3^{(l)}(X_{l+1} \oplus k_4 \oplus x_4)}{N}. \tag{B.15}$$

B.2.2. Subcase 2.2. $x_3 \in \text{Dom}\mathscr{F}_3$, and $x_4 \in \text{Dom}\mathscr{F}_4$. For this, we introduce a new \$k-\$dependent quantity:

$$\alpha_{3,4}^{+}(k, X) \overset{\text{def}}{=} \left|\left\{((x_3, y_3), (x_4, y_4)) \in Q_{F_3} \times Q_{F_4}\colon k_4 = X \oplus O y_3 \oplus x_4\right|. \tag{B.16}$$

Thus, the number of such pairs (x_3, x_4) with $X_{l+1} \oplus O y_3 = k_4 \oplus x_4$ is $\alpha_{3,4}^{+}(k, X_{l+1})$. We also have $\Pr[F_2(x_2^{(l+1)}) = O^{-1}(A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus k_3 \oplus x_3)] \leq 1/N$. Therefore,

$$\sum_{x_3 \in \text{Dom}\mathscr{F}_3, x_4 \in \text{Dom}\mathscr{F}_4} \Pr[\text{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq \frac{\alpha_{3,4}^{+}(k, X_{l+1})}{N}. \tag{B.17}$$

Since k_4 is uniform in 2^{n-r} values, it can be seen $\mathbb{E}[\alpha_{3,4}^{+}(k, X_{l+1})] \leq 2^r \cdot q_f^2/N$. Therefore, the expectation of the probability is at most $2^r \cdot q_f^2/N^2$.

B.2.3 Subcase 2.3. $x_3 \in \mathrm{Dom}\mathscr{F}_3$ and $x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4$. By definition, we have

$$\sum_{x_3 \in \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}} \mathrm{Pr}\left[\mathrm{Coll}\left(x_3, x_4\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] = \sum_{x_3 \in \mathrm{Dom}\mathscr{F}_3, i=1,\ldots,l} \mathrm{sgn}\,(i) \cdot \mathrm{Pr}\left[\mathrm{Coll}\left(x_3, x_4\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right], \tag{B.18}$$

where $x_4^{(i)} = k_4 \oplus Z_i, Z_i = A_1 T_i \oplus A_2 S_i \oplus O F_5\,(k_5 \oplus A_i)$ are derived from the ith tuple $(L_i R_i X_i, A_i S_i T_i) \in \mathscr{G}_4$, and $\mathrm{sgn}\,(i) = 1$ if and only if i is the smallest index satisfying these conditions (i.e., $x_4^{(i)} \in \mathrm{Ext}\mathscr{F}_4^{(i)}$, while $x_4^{(i)} \notin \mathrm{Ext}\mathscr{F}_4^{(i-1)}$) and $x_4^{(i)} \notin \mathrm{Dom}\mathscr{F}_4$.

We focus on $\mathrm{Pr}[\mathrm{Coll}\,(x_3, x_4^{(i)}) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F]$. For convenience, we let $y_3 = \mathrm{Img}F_3\,(x_3)$ and write $Y_i = A_1 L_i \oplus A_2 R_i \oplus O F_2\,(k_2 \oplus X_i)$. We consider the conditional probabilities $\mathrm{Pr}[\mathrm{Coll}\,(x_3, x_4^{(i)}) | E_i \text{ fits into CASE } j \wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F]$, for $j = 1, 2, 3$. It can be seen if E_i fits into Case 3, then $x_4^{(i)} \in \mathrm{Dom}\mathscr{F}_4$, and this $x_4^{(i)}$ is the discussion of subcase 2.3. So, we consider $j = 1$ or 2:

(i) When E_i fits into Case 1, according to the corresponding analysis, Z_i was derived via $Z_i = A_1 T_i \oplus A_2 S_i \oplus O F_5\,(k_5 \oplus A_i)$, and $F_5\,(k_5 \oplus A_i)$ was uniform. Thus,

$$\begin{aligned} \mathrm{Pr}\left[X_{l+1} \oplus O y_3 = \left(k_4 \oplus x_4^{(i)}\right)\right] &= \mathrm{Pr}\left[X_{l+1} \oplus O y_3 = Z_i\right] \\ &= \mathrm{Pr}\left[F_5\,(k_5 \oplus A_i) = O^{-1}\left(X_{l+1} \oplus O y_3 \oplus A_1 T_i \oplus A_2 S_i\right)\right] \leq \frac{1}{N}. \end{aligned} \tag{B.19}$$

Since we further have $\mathrm{Pr}[F_2\,(k_2 \oplus X_{l+1}) = O - 1\,(A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus k_3 \oplus x_3)] \leq 1/N$, the following equation holds

$$\begin{aligned} &\mathrm{Pr}\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right) | E_i \text{ fits into CASE } 1 \wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] \\ &\leq \frac{1}{N^2} \end{aligned} \tag{B.20}$$

(ii) When E_i fits into Case 2, let $x_3^{(i)} = k_3 \oplus Y_i$ and $y_3^{(i)} = F_3\,(x_3^{(i)})$. Then, $X_{l+1} \oplus O y_3 = Z_i$ implies

$$X_{l+1} \oplus O y_3 = X_i \oplus O y_3^{(i)}, \tag{B.21}$$

meaning that, for each triple (X_{l+1}, X_i, y_3), the number of choices for such $y_3^{(i)}$ is $\mathrm{Num}_3^{(l)}\,(O^{-1} X_{l+1} \oplus O^{-1} X_i \oplus y_3)$. For each such $y_3^{(i)}$, the event $\mathrm{Coll}\,(x_3, x_4^{(i)})$ essentially implies two collisions, i.e.,

$$\begin{aligned} A_1 L_i \oplus A_2 R_i \oplus O F_2\left(k_2 \oplus X_i\right) &= k_3 \oplus x_3^{(i)}, \\ A_1 L_{l+1} \oplus A_2 R_{l+1} \oplus O F_2\left(k_2 \oplus X_{l+1}\right) &= k_3 \oplus x_3. \end{aligned} \tag{B.22}$$

Therefore,

$$\mathrm{Pr}\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right) | E_i \text{ fits into CASE } 2 \wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] \leq \frac{\mathrm{Num}_3^{(l)}\left(O^{-1} X_{l+1} \oplus y_3 \oplus O^{-1} X_i\right)}{N^2}. \tag{B.23}$$

By the above, for any j, we have

$$\mathrm{Pr}\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right) | E_i \text{ fits into CASE } j \wedge E_l \wedge \ldots \wedge E_1 \wedge F \vdash Q_F\right] \leq \frac{\mathrm{Num}\left(O^{-1} X_{l+1} \oplus y_3 \oplus O^{-1} X_i\right)}{N^2}, \tag{B.24}$$

where $\mathrm{Num}\,(O^{-1}X_{l+1}\oplus y_3\oplus O^{-1}X_i)/N^2$ is denoted as B.

Thus,

$$\begin{aligned}&\Pr\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right)|E_l\wedge\ldots\wedge E_1\wedge F\vdash Q_F\right]\\&=\frac{\sum_{j=1}^{3}\Pr\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right)\wedge E_i\wedge E_i \text{ fits into CASE } j|E_l\wedge\cdots\wedge E_{i-1}\wedge E_{i+1}\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]}{\Pr[E_i]}\\&\leq\sum_{j=1}^{3}B\cdot\frac{\Pr\left[E_i\wedge E_i \text{ fits into CASE } j|E_l\wedge\cdots\wedge E_{i-1}\wedge E_{i+1}\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]}{\Pr[E_i]}=B.\end{aligned}\tag{B.25}$$

This means

$$\begin{aligned}&\sum_{x_3\in\mathrm{Dom}\mathcal{F}_3, x_4\in Ext\mathcal{F}_4^{(l)}/\mathrm{Dom}\mathcal{F}_4}\Pr\left[\mathrm{Coll}\,(x_3,x_4)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\\&\leq\sum_{x_3\in\mathrm{Dom}\mathcal{F}_3}\sum_{i=1,\cdots,l}\mathrm{sgn}\,(i)\cdot\frac{\mathrm{Num}_3^{(l)}\left(O^{-1}X_{l+1}\oplus y_3\oplus O^{-1}X_i\right)}{N^2}\\&\leq\sum_{x_3\in\mathrm{Dom}\mathcal{F}_3}\frac{q_f+e_3^{(l)}}{N^2}\leq\frac{q_f\left(q_f+q_e\right)}{N^2}.\end{aligned}\tag{B.26}$$

B.2.4. Subcase 2.4. $x_3\in\mathrm{Dom}\mathcal{F}_3\backslash\mathrm{Dom}\mathcal{F}_3$, and $x_4\in\mathrm{Dom}\,\mathcal{F}_4$. By definition, we have

$$\begin{aligned}&\sum_{x_3\in\mathrm{Ext}\mathcal{F}_3^{(l)}/\mathrm{Dom}\mathcal{F}_3, x_4\in\mathrm{Dom}\mathcal{F}_4}\Pr\left[\mathrm{Coll}\,(x_3,x_4)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\\&=\sum_{i=1,\cdots,l, x_4\in\mathrm{Dom}\mathcal{F}_4}\mathrm{sgn}'\,(i)\cdot\Pr\left[\mathrm{Coll}\left(x_3^{(i)}, x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right],\end{aligned}\tag{B.27}$$

where $x_3^{(i)}=k_3\oplus Y_i$ and $Y_i=A_1L_i\oplus A_2R_i\oplus OF_2\,(k_2\oplus X_i)$ are derived from the ith tuple $(L_iR_iX_i, A_iS_iT_i)\in\mathcal{G}_4$, and $\mathrm{sgn}'\,(i)=1$ if and only if i is the smallest index satisfying these conditions $(x_3^{(i)}\notin\mathrm{Ext}\mathcal{F}_3$. We let $y_4=\mathrm{Im}gF_4\,(x_4)$ and write $Z_i=A_1T_i\oplus A_2S_i\oplus OF_5\,(k_5\oplus A_i)$. Now, the following equation holds $y_3^{(i)}=O^{-1}\,(X_i\oplus Z_i)$; thus, the collision relation $X_{l+1}\oplus Oy_3^{(i)}=(k_4\oplus x_4)$ translates into $X_{l+1}\oplus X_i=Z_i\oplus(k_4\oplus x_4)$. Similar to subcase 2.3, we distinguish two cases.

(i) When E_i fits into CASE 1, we have $Z_i=A_1T_i\oplus A_2S_i\oplus OF_5\,(k_5\oplus A_i)$ and $F_5\,(k_5\oplus A_i)$ was uniform. Thus,

$$\Pr\left[X_{l+1}\oplus X_i=Z_i\oplus(k_4\oplus x_4)\right]=\Pr\left[F_5\,(k_5\oplus A_i)=O^{-1}\left(X_{l+1}\oplus X_i\oplus k_4\oplus x_4\oplus A_1T_i\oplus A_2S_i\right)\right]\leq\frac{1}{N}.\tag{B.28}$$

This along with $\Pr[F_2\,(k_2\oplus X_{l+1})=O^{-1}\,(A_1L_{l+1}\oplus A_2R_{l+1}\oplus k_3\oplus x_3)]\leq 1/N$ yields

$$\Pr\left[\mathrm{Coll}\left(x_3^{(i)}, x_4\right)|E_i \text{ fits into CASE } 1|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\leq\frac{1}{N^2}.\tag{B.29}$$

(ii) When E_i fits into Case 3, let $x_4^{(i)}=k_4\oplus Z_i$. Then, $X_{l+1}\oplus X_i=Z_i\oplus k_4\oplus x_4$ implies $X_{l+1}\oplus X_i=x_4^{(i)}\oplus x_4$. Note that, for the fixed X_{l+1}, X_i, and x_4, the number of choices for $x_4^{(i)}$ is at most 1. And, for Y_{l+1} to collide with $x_3^{(i)}$, the two collisions $A_1T_i\oplus A_2S_i\oplus OF_5\,(k_5\oplus A_i)=k_4\oplus x_4^{(i)}$ and $A_1L_{l+1}\oplus A_2R_{l+1}\oplus OF_2\,(k_2\oplus X_{l+1})=k_3\oplus x_3^{(i)}$ are required to happen:

$$\Pr\left[\mathrm{Coll}\left(x_3^{(i)}, x_4\right)|E_i \text{ fits into CASE } 1|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\leq\frac{1}{N^2},\tag{B.30}$$

which follows.

Using a counting similar to subcase 2.3, we obtain

$$\sum_{x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)}/\mathrm{Dom}\mathscr{F}_3, x_4 \in Do\ m\mathscr{F}_4} \Pr\left[\mathrm{Coll}\left(x_3, x_4\right) | E_l \wedge \ldots \wedge E_1 \wedge F \vdash Q_F\right] \leq \sum_{i=1,\cdots,l, x_4 \in \mathrm{Dom}\mathscr{F}_4} \mathrm{sgn}'(i) \cdot \frac{1}{N^2} \leq \frac{q_f e_3^{(l)}}{N^2} \leq \frac{q_f q_e}{N^2}. \tag{B.31}$$

B.2.5. Subcase 2.5. $x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)} \backslash \mathrm{Dom}\mathscr{F}_3$, and $x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4$. By definition, we have

$$\sum_{x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)}/\mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)}/\mathrm{Dom}\mathscr{F}_4} \Pr\left[\mathrm{Coll}\left(x_3, x_4\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] = \sum_{i=1,\ldots,l} \mathrm{sgn}(i) \cdot \mathrm{sgn}(j) \cdot \Pr\left[\mathrm{Coll}\left(x_3^{(i)}, x_4^{(j)}\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right], \tag{B.32}$$

where

(1) $x_3^{(i)} = k_3 \oplus Y_i$ and $Y_i = A_1 L_i \oplus A_2 R_i \oplus OF_2\left(k_2 \oplus X_i\right)$ are derived from the ith tuple $\left(L_i R_i X_i, A_i S_i T_i\right) \in \mathscr{G}_4$ and $\mathrm{sgn}(i) = 1$ if and only if i is the smallest index satisfying these conditions and $x_3^{(i)} \notin \mathrm{Dom}\mathscr{F}_3$.

(2) $x_4^{(j)} = k_4 \oplus Z_j$ and $Z_j = A_1 T_j \oplus A_2 S_j \oplus OF_5\left(k_5 \oplus A_j\right)$ are derived from the jth tuple $\left(L_j R_j X_j, A_j S_j T_j\right) \in \mathscr{G}_4$ and $\mathrm{sgn}'(i) = 1$ if and only if j is the smallest index satisfying these conditions and $x_4^{(j)} \notin \mathrm{Dom}\mathscr{F}_4$.

(i) If $i < j$, then we utilize the constraint $X_{l+1} \oplus O y_3^{(i)} = X_j \oplus O y_3^{(j)}$ and follow the same line as the analysis of subcase 2.3. This shows the number of choices for $y_3^{(j)}$ is $\mathrm{Num}_3\left(O^{-1} X_{l+1} \oplus y_3^{(i)} \oplus X_j\right)$; thus, the upper bound $\mathrm{Num}_3\left(O^{-1} X_{l+1} \oplus y_3^{(i)} \oplus X_j\right)/N^2$ for each $\left(x_3^{(i)}, x_4^{(j)}\right)$.

(ii) If $i > j$, then we utilize the constraint $X_{l+1} \oplus O y_3^{(i)} = X_j \oplus O y_3^{(j)}$, and follow the same line as the analysis of subcase 2.4. Then, establish the bound $1/N^2$ for each $\left(x_3^{(i)}, x_4^{(j)}\right)$.

In all

$$\begin{aligned} &\sum_{x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)} \backslash \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4} \Pr\left[\mathrm{Coll}\left(x_3, x_4\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] \\ &\leq \sum_{i=1,\ldots,l} \sum_{i=1,\ldots,l} \frac{\mathrm{Num}_3\left(O^{-1} X_{l+1} \oplus y_3^{(i)} \oplus X_j\right)}{N^2} \leq \sum_{i=1,\ldots,l} \frac{q_f + e_3^{(l)}}{N^2} \leq \frac{q_f\left(q_f + q_e\right)}{N^2}, \end{aligned} \tag{B.33}$$

summing over the five cases yields

$$\begin{aligned} \mathbb{E}_k\left[p\mathrm{coll}\right] &\leq \frac{\sum_{x_4 \in \mathscr{G}_3 \mathscr{F}_4} \mathrm{Num}_3^{(l)}\left(O^{-1} X_{l+1} \oplus k_4 \oplus O^{-1} X_4\right)}{N} + \frac{2^r \cdot q_f^2}{N^2} + \frac{q_f\left(q_f + q_e\right)}{N^2} + \frac{q_f q_e}{N^2} + \frac{q_f\left(q_f + q_e\right)}{N^2} \\ &\leq \frac{\sum_{x_4 \in \mathscr{G}_3 \mathscr{F}_4} \mathrm{Num}_3^{(l)}\left(O^{-1} X_{l+1} \oplus k_4 \oplus O^{-1} X_4\right)}{N} + \frac{2^r \cdot q_f^2}{N^2} + \frac{\left(2q_f + q_e\right)\left(q_f + q_e\right)}{N^2}. \end{aligned} \tag{B.34}$$

Clearly, once such collisions do not happen, the mentioned requirements are met, and we have $x_4 \notin \mathrm{Dom}\mathscr{F}_4 \cup \mathrm{Ext}\mathscr{F}_4^{(l)} \cup \mathscr{G}_3 \mathscr{F}_4$. Moreover, as $\left(L_{l+1} R_{l+1} X_{l+1}, A_{l+1} S_{l+1} T_{l+1}\right) \in \mathscr{G}_1$, we have (a) $x_5^{(l+1)} \notin \mathrm{Dom}\mathscr{F}_5$ and (b) $\left|\varepsilon Q\left(A_{l+1}\right)\right| = 1$, i.e., the position of $x_5^{(l+1)}$ cannot be "taken" by previous tuples. By these,

$$\mathbb{E}_k\left[\Pr\left[E_{l+1}\wedge \text{CASE } 2|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\right] \geq \left(\frac{q_f + e_3^{(l)}}{N} - \frac{\sum_{x_4\in\mathcal{G}_3\mathcal{F}_4}\text{Num}_3^{(l)}\left(X_{l+1}+k_4+X_4\right)}{N} - \frac{2^r\cdot q_f^2}{N^2} - \frac{\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}\right)\frac{1}{N^2}. \tag{B.35}$$

B.3. Case 3. In this case, since $x_2^{(l+1)} \notin \text{Dom}\mathcal{F}_2$, $|\varepsilon Q(X_{l+1})| = 1$, and $x_3 = k_3 \oplus Y_{l+1} \notin \text{Dom}\mathcal{F}_3 \cup \text{Ext}F_3^{(l)} \cup \mathcal{G}_2\mathcal{F}_3$,

$$\begin{aligned}\Pr\left[\text{KAF extends }\left(L_{l+1}R_{L+1}X_{l+1}, A_{l+1}S_{l+1}T_{l+1}\right)\right] &= \Pr\left[F_2\left(x_2^{(l+1)}\right) = O^{-1}\left(A_1L_{l+1}\oplus A_2R_{l+1}\oplus Y_{l+1}\right)\wedge F_3\left(x_3\right) = O^{-1}\left(X_{l+1}\oplus Z_{l+1}\right)\right]\\ &= \frac{1}{N^2}.\end{aligned} \tag{B.36}$$

Thus, by lowering bounding the probability of colliding with such "bad" (x_3, x_4), we would derive the result for this case. Similar to Case 2 by symmetry, we write

$$p\text{coll} = \sum_{\substack{x_3\in Do\ m\mathcal{F}_3\cup ExtF_3^{(l)}\cup\mathcal{G}_2\mathcal{F}_3\\ x_4\in Do\ m\mathcal{F}_4\cup ExtF_4^{(l)}}} x_3 \in \text{Dom}\mathcal{F}_3\cup\text{Ext}F_3^{(l)}\Pr\left[\text{Coll}\left(x_3,x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]. \tag{B.37}$$

Let $x_5^{(l+1)} = k_5 \oplus A_{l+1}$. Similar to subsection B.2, we also distinguish five subcases, and the arguments are similar by symmetry:

(1) $x_3 \in \mathcal{G}_2\mathcal{F}_3$, and $x_4 \in \text{Dom}\mathcal{F}_4 \cup \text{Ext}\mathcal{F}_4^{(l)}$. In this case, utilizing the constraint $A_{l+1} \oplus Oy_4 = k_3 \oplus y_3$, we have

$$\sum_{x\in\mathcal{G}_2\mathcal{F}_3, x_4\in\text{Dom}\mathcal{F}_4\cup\text{Ext}\mathcal{F}_4^{(l)}}\Pr\left[\text{Coll}\left(x_3,x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right] \leq \frac{\sum_{x_3\in\mathcal{G}_3\mathcal{F}_3}\text{Num}_4^{(l)}\left(O^{-1}\left(A_{l+1}\oplus k_3\oplus x_3\right)\right)}{N}, \tag{B.38}$$

where $\text{Num}_4^{(l)}(y_4) = |\{x_4 \in \text{Dom}\mathcal{F}_4 \cup \text{Ext}F_4^{(l)}: F_4(x_4) = y_4\}|$.

(2) $x_3 \in \text{Dom}\mathcal{F}_3$, and $x_4 \in \text{Dom}\mathcal{F}_4$. Define $\alpha_{3,4}^{-}(k, A) \overset{\text{def}}{=} |\{((x_3, y_3), (x_4, y_4)) \in Q_{F_3}\times Q_{F_4}: k_3 = x_3 \oplus Oy_4 \oplus A\}|$. Since k_3 is uniform in 2^n values, we have

$$\mathbb{E}_k\left[\sum_{x_3\in\text{Dom}\mathcal{F}_3, x_4\in\text{Dom}\mathcal{F}_4}\Pr\left[\text{Coll}\left(x_3,x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\right] \leq \frac{\mathbb{E}_k\left[\alpha_{3,4}^{-}\left(k, A_{l+1}\right)\right]}{N} \leq \frac{q_f^2}{N^2}. \tag{B.39}$$

(3) $x_3 \in \text{Ext}\mathcal{F}_3^{(l)}\backslash\text{Dom}\mathcal{F}_3$, and $x_4 \in \text{Dom}\mathcal{F}_4$. By definition, we have

$$\sum_{x_3\in\text{Ext}\mathcal{F}_3\backslash\text{Dom}\mathcal{F}_3, x_4\in\text{Dom}\mathcal{F}_4}\Pr\left[\text{Coll}\left(x_3,x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right] = \sum_{i=1,\ldots,l, x_4\in Dom\mathcal{F}_4}\text{sgn}'(i)\cdot\Pr\left[\text{Coll}\left(x_3^{(i)},x_4\right)|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right], \tag{B.40}$$

where $x_3^{(i)} = k_3 \oplus Y_i$ and $Y_i = A_1 L_i \oplus A_2 R_i \oplus F_2 (k_2 \oplus X_i)$ are derived from the $i$th tuple $(L_i R_i X_i, A_i S_i T_i) \in \mathscr{G}_4$ and $\mathrm{sgn}'(i) = 1$ if and only if \$i\$ is the smallest index satisfying these conditions and $x_3^{(i)} \notin \mathrm{Dom}\mathscr{F}_3$.

Then, similarly to the analysis for the subcase 2.3 in subsection B.2,

$$\Pr\left[\mathrm{Coll}\left(x_3^{(i)}, x_4\right) | E_i \text{ fits into CASE } 2 \wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right] = \frac{\mathrm{Num}_4^{(l)}\left(O^{-1} A_{l+1} \oplus y_4 \oplus O^{-1} A_i\right)}{N^2}. \tag{B.41}$$

(i) When E_i fits into Case 1, it can be shown Pr $[\mathrm{Coll}(x_3^{(i)}, x4) | E_i$ fits into Case $1 \wedge E_l \wedge \cdots \wedge E_1 \ \wedge F \vdash Q_F] \leq 1/N^2$.

(ii) When E_i fits into Case 3, it can be shown $A_{l+1} \oplus O y_4 = A_i \oplus O y_4^{(i)}$. By this,

By the above and a similar calculation, we have

$$\sum_{x_3 \in \mathrm{Ext}\mathscr{F}_3 \backslash \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Dom}\mathscr{F}_4} \Pr[\mathrm{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq \sum_{x_4 \in \mathrm{Dom}\mathscr{F}_4} \sum_{i=1,\ldots,l} \frac{\mathrm{Num}_4^{(l)}\left(O^{-1} A_{l+1} \oplus y_4 \oplus O^{-1} A_i\right)}{N^2} \leq \frac{q_f\left(q_f + q_e\right)}{N^2}. \tag{B.42}$$

(4) $x_3 \in \mathrm{Dom}\mathscr{F}_3$, and $x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4$. By definition, we have

$$\sum_{x_3 \in \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4 \backslash \mathrm{Dom}\mathscr{F}_4} \Pr[\mathrm{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] = \sum_{i=1,\ldots,l, x_4 \in \mathrm{Dom}\mathscr{F}_3} \mathrm{sgn}(i) \cdot \Pr\left[\mathrm{Coll}\left(x_3, x_4^{(i)}\right) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F\right]. \tag{B.43}$$

It also holds $\Pr[\mathrm{Coll}(x_3, x4^{(i)}) | E_i$ fits into Case 1 $\wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq 1/N^2$. On the contrary, when E_i fits into CASE 2, it can be shown as $A_{l+1} \oplus A_i = x_3^{(i)} \oplus x_3$, which helps cinch $\Pr[\mathrm{Coll}(x_3, x4 \quad (i)) | E_i$ fits into Case $1 \wedge E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq 1/N^2$. Therefore,

$$\sum_{x_3 \in \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4} \Pr[\mathrm{Coll}(x3, x4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] \leq \frac{q_f e_4^{(l)}}{N^2} \leq \frac{q_f q_e}{N^2}. \tag{B.44}$$

(5) $x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)} \backslash \mathrm{Dom}\mathscr{F}_3$, and $x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4$. Similar to the last subcase in subsection B.2, it can be shown as

$$\begin{aligned} \sum_{x_3 \in \mathrm{Ext}\mathscr{F}_3^{(l)} \backslash \mathrm{Dom}\mathscr{F}_3, x_4 \in \mathrm{Ext}\mathscr{F}_4^{(l)} \backslash \mathrm{Dom}\mathscr{F}_4} \Pr[\mathrm{Coll}(x_3, x_4) | E_l \wedge \cdots \wedge E_1 \wedge F \vdash Q_F] &\leq \sum_{j=1,\ldots,l} \sum_{i=1,\ldots,l} \frac{\mathrm{Num}_4^{(l)}\left(O^{-1} A_{l+1} \oplus y_4 \oplus O^{-1} A_i\right)}{N^2} \\ &\leq \sum_{j=1,\ldots,l} \frac{q_f + e_4^{(l)}}{N^2} \leq \frac{q_f\left(q_f + q_e\right)}{N^2}. \end{aligned} \tag{B.45}$$

The above gives rise to the following bound:

$$\mathbb{E}_k\left[\Pr\left[E_{l+1}\wedge\text{CASE 3}|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\right]\geq\left(\frac{q_f+e_4^{(l)}}{N}-\frac{\sum_{x_3\in\mathcal{G}_2\mathcal{F}_3}\text{Num}_4^{(l)}\left(A_{l+1}\oplus k_3\oplus x_3\right)}{N}-\frac{\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}\right)\frac{1}{N^2}. \tag{B.46}$$

B.4. Summary for $E_{\mathcal{G}_1}$. Summing over the three cases results in

$$\begin{aligned}\mathbb{E}_k\left[\Pr\left[E_{l+1}|E_l\wedge\cdots\wedge E_1\wedge F\vdash Q_F\right]\right]\geq&\left(\left(1-\frac{q_f+e_3^{(l)}+\left|\mathcal{G}_2\mathcal{F}_3\right|}{N}\right)\right)\left(1-\frac{q_f+e_4^{(l)}+\left|\mathcal{G}_3\mathcal{F}_4\right|}{N}\right)\\&+\frac{2q_f+e_3^{(l)}+e_4^{(l)}}{N}-\frac{2^r\cdot q_f^2}{N^2}-\frac{2\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}\\&-\frac{\sum_{x_4\in\mathcal{G}_3\mathcal{F}_4}\text{Num}_3^{(l)}\left(O^{-1}\left(X_{l+1}\oplus k_4\oplus x_4\right)\right)+\sum_{x_3\in\mathcal{G}_2\mathcal{F}_3}\text{Num}_4^{(l)}\left(O^{-1}\left(A_{l+1}\oplus k_3\oplus x_3\right)\right)}{N}\frac{1}{N^2}\\\geq&\left(1-\frac{2^r\cdot q_f^2}{N^2}-\frac{2\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}-\frac{\left|\mathcal{G}_2\mathcal{F}\right|+\left|\mathcal{G}_3\mathcal{F}_4\right|}{N}-B_l\right)\frac{1}{N^2}.\end{aligned} \tag{B.47}$$

Note that $|\mathcal{G}_2\mathcal{F}_3|\leq|\mathcal{G}_2|=\beta_1$, $|\mathcal{G}_3\mathcal{F}_4|\leq|\mathcal{G}_3|=\beta_2$, and (b) $|G_1|\leq q_e$. Therefore,

$$\begin{aligned}\mathbb{E}_k\left[\Pr\left[E_{\mathcal{G}_1}|F\vdash Q_F\right]\right]\geq&\prod_{l=0}^{|\mathcal{G}_1|-1}\left(1-\frac{2^r\cdot q_f^2}{N^2}-\frac{2\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}-\frac{\left|\mathcal{G}_2\mathcal{F}_3\right|+\left|\mathcal{G}_3\mathcal{F}_4\right|}{N}-B_l\right)\cdot\frac{1}{N^{2|\mathcal{G}_1|}}\\\geq&\left(1-\frac{2^r\cdot q_eq_f^2}{N^2}-\frac{2q_e\left(2q_f+q_e\right)\left(q_f+q_e\right)}{N^2}-\frac{q_e\left(\beta_1+\beta_2\right)}{N}-\sum_{l=0}^{q_e-1}B_l\right)\cdot\frac{1}{N^{2|\mathcal{G}_1|}}.\end{aligned} \tag{B.48}$$

We finally consider $\sum_{l=0}^{q_e-1}B_l$. To this end, we note that, by definition, we have

$$\begin{aligned}&\sum_{y_3\in\{0,1\}^n}\text{Num}_3^{(l)}\left(y_3\right)=q_f+e_3^{(l)}\leq q_f+q_e,\\-&\sum_{y_4\in\{0,1\}^n}\text{Num}_4^{(l)}\left(y_4\right)=q_f+e_4^{(l)}\leq q_f+q_e.\end{aligned} \tag{B.49}$$

Therefore,

$$\sum_{l=0}^{q_e-1}\sum_{x_4\in\mathscr{G}_3\mathscr{F}_4}\mathrm{Num}_3^{(l)}\left(O^{-1}\left(X_{l+1}\oplus k_4\oplus x_4\right)\right)\leq\sum_{x_3\in\mathscr{G}_3\mathscr{F}_4}\left(q_f+q_e\right)\leq\left(q_f+q_e\right)\beta_2, \tag{B.50}$$

and similarly,

$$\sum_{l=0}^{q_e-1}\sum_{x_3\in\mathscr{G}_2\mathscr{F}_3}\mathrm{Num}_4^{(l)}\left(O^{-1}\left(A_{l+1}\oplus k_3\oplus x_3\right)\right)\leq\left(q_f+q_e\right)\left|\mathscr{G}_2\mathscr{F}\right|\leq\left(q_f+q_e\right)\beta_1. \tag{B.51}$$

B.5. Analysis for $\mathscr{G}_2$, $\mathscr{G}_3$, and $\mathscr{G}_4$. Lower bounding $\Pr[E_{\mathscr{G}_2}\wedge E_{\mathscr{G}_3}E_{\mathscr{G}_1}\wedge F\vdash Q_F]$: consider $E_{\mathscr{G}_2}$ first; we lower bound the probability that it is equivalent to F_4 and F_5 satisfying $2|\mathscr{G}_2|$ new and distinct equations. To this end, again we define a predicate $\mathrm{Bad}_1(F_3)$, which holds if there exists $(LRX, AST)\in G_2$ that fulfills one of the following conditions:

(i) The x_4 value derived using F_3 is in $\mathrm{Dom}\mathscr{F}_4$, i.e., $k_4\oplus X\oplus OF_3(x_3)\in\mathrm{Dom}\mathscr{F}_4$, where $x_3=k_3\oplus A_1L\oplus A_2R\oplus\mathrm{Img}F_2(k_2\oplus X)$

(ii) The Z value derived using F_3 collides with the Z' value of another tuple in $\mathscr{G}_2$, i.e., there exists $(L^*R^*X^*, A^*S^*T^*)\in\mathscr{G}_2$ such that $X\oplus OF_3(x_3)=X'\oplus F_3(x_{3'})$, where $x_{3'}=k_3\oplus A_1L'\oplus A_2R'\oplus\mathrm{Img}F_2(k_2\oplus X)$

(iii) The Z value derived using F_3 collides with the Z value of a tuple in $\mathscr{G}_1$ or $\mathscr{G}_3$, i.e., there exists $(L^*R^*X^*, A^*S^*T^*)\in\mathscr{G}_1\cup\mathscr{G}_3$ such that $X\oplus OF_3(x_3)=S*\oplus F_5(k_5\oplus A*)$

We note that, for each $(LRX, AST)\in\mathscr{G}_2$, let $x_3=k_3\oplus A_1L\oplus A_2R\oplus O\mathrm{Img}F_2(k2\oplus X)$; then, the following equation holds $x_3\notin\mathrm{Dom}\mathscr{F}_3$ (otherwise fulfilling (C-2)) and $x_3\in\mathrm{Ext}\mathscr{F}_3^{(|\mathscr{G}_1|)}$ (according to the analysis of $E_{\mathscr{G}_1}$). Thus, conditioned on $E_{\mathscr{G}_1}\wedge F_3\vdash Q_{F_3}$, the value $F_3(x_3)$ remains uniform. Therefore, for this (LRX, AST),

(i) The probability that condition (i) is fulfilled is at most q_f/N.

(ii) For each $(LRX, AST)\in\mathscr{G}_2$, if the corresponding $x_{3'}$ does not equal x_3, then the probability of $X\oplus OF_3(x_3)=X'\oplus OF_3(x_{3'})$ is at most $1/N$; otherwise, since the two tuples are distinct, it has to be $X\neq X$, and thus, $X\oplus OF_3(x_3)\neq X'\oplus OF_3(x_{3'})$.

(iii) For each $(L*R*X*, A*S*T*)\in\mathscr{G}_1\cup\mathscr{G}_3$, the probability of

$$X\oplus OF_3(x_3)=A_1T*\oplus A_2S*\oplus OF_5(k_5\oplus A*) \tag{B.52}$$

is at most $1/N$. Summing over the above yields

$$\Pr\left[\mathrm{Bad}_1(F_3)|E_{\mathscr{G}_1}\wedge F\vdash Q_F\right]\leq\frac{\left|\mathscr{G}_2\right|\cdot\left(q_f+\left|\mathscr{G}_1\right|+\left|\mathscr{G}_2\right|+\left|\mathscr{G}_3\right|\right)}{N}\leq\frac{\beta_1\left(q_f+q_e\right)}{N}. \tag{B.53}$$

It is not hard to see that conditioned on $\mathrm{Bad}_1(F_3)$, the $|\mathscr{G}_2|$ tuples in $\mathscr{G}_2$ indeed give rise to $|\mathscr{G}_2|$ distinct values $Z_1,\ldots,Z_{|\mathscr{G}_2|}$ (otherwise condition (ii) is fulfilled), for which $F_4(k_4\oplus Z_1),\ldots,F_4(k_4\oplus Z_{|\mathscr{G}_2|}|)$ all remain undetermined (otherwise condition (i) or (iii) fulfilled). Furthermore, at the "right side," they also give rise to $|\mathscr{G}_2|$ distinct values $A_1,\cdots,A_{|\mathscr{G}_2|}$ with $F_5(k_5\oplus A_1),\ldots,F_5(k_5\oplus A_{|\mathscr{G}_2|})$ all undetermined:

(i) $A_1,\ldots,A_{|\mathscr{G}_2|}$ are also distinct, otherwise fulfilling (C-5)

(ii) None of $k_5\oplus A_1,\ldots,k_5\oplus A_{|\mathscr{G}_2|}$ is in $\mathrm{Dom}\mathscr{F}_5$, otherwise fulfilling (C-1)

(iii) Conditioned on $E_{\mathscr{G}_1}, F_5(k_5\oplus A_1),\ldots,F_5(k_5\oplus A_{|\mathscr{G}_2|})$ remain undetermined, otherwise some A_i is shared between tuples in $\mathscr{G}_1$ and $\mathscr{G}_2$ and (C-5) is fulfilled

Thus, in this case, the event $E_{\mathscr{G}_2}$ is equivalent to F_4 and F_5 satisfying $2|\mathscr{G}_2|$ new equations, the probability of which does not exceed $1/N^{2|\mathscr{G}_2|}$.

We then consider $E_{\mathscr{G}_3}$. The analysis is similar to $E_{\mathscr{G}_2}$ by symmetry; we define a predicate $\mathrm{Bad}_1(F_4)$ on F_4, which holds if there exists $(LRX, AST)\in\mathscr{G}_3$ such that one of the following conditions is fulfilled:

(i) The induced value x_3 is in $\mathrm{Dom}\mathscr{F}_3$, i.e., $k_3\oplus A\oplus OF_4(x_4)\in\mathrm{Dom}\mathscr{F}_3$, where $x_4=k_4\oplus A_1T\oplus A_2S\oplus\mathrm{Img}F_5(k_5\oplus A)$. The probability is at most q_f/N in total.

(ii) The induced Y collides with the Y' value of another tuple in $\mathscr{G}_3$, i.e., there exists a tuple $(L'R'X', A'S'T') \in \mathscr{G}_3$ such that $A \oplus OF_4(x_4) = A' \oplus OF_4(x_{4'})$, where $x_{4'} = k_4 \oplus A_1 T' \oplus A_2 S' \oplus \mathrm{Im}gF_5(k_5 \oplus A')$. The probability is at most $|\mathscr{G}_3|/N$ in total.

(iii) The induced Y collides with the $Y*$ value of a tuple in $\mathscr{G}_1$ or $\mathscr{G}_2$; i.e., there exists $(L*R*X*, A*S*T*) \in \mathscr{G}_1 \cup \mathscr{G}_2$ such that $A \oplus OF_4(x_4) = A_1 L* \oplus A_2 R* \oplus F_2(k_2 \oplus X)$. The probability is at most $|\mathscr{G}_1| + |\mathscr{G}_2|/N$ in total.

Similar to $\mathrm{Bad}_1(F_3)$,

$$\Pr\left[\mathrm{Bad}_2(F_4) | E_{\mathscr{G}_1} \wedge F \vdash \mathrm{Q}_F\right] \le \frac{|\mathscr{G}_3| \cdot \left(q_f + |\mathscr{G}_1| + |\mathscr{G}_2| + |\mathscr{G}_3|\right)}{N} \le \frac{\beta_2\left(q_f + q_e\right)}{N}; \tag{B.54}$$

and conditioned on $\mathrm{Bad}_1(F_4)$, tuples in $\mathscr{G}_3$ give rise to $|\mathscr{G}_3|$ distinct values $Y_1, \ldots, Y_{|\mathscr{G}_3|}$, while the assumption $\mathrm{Bad}(F_1, F_6)$ ensures that they give rise to $|\mathscr{G}_3|$ distinct values $X_1, \ldots, X_{|\mathscr{G}_3|}$. Thus, the event $E_{\mathscr{G}_3}$ is equivalent to F_2 and F_3 satisfying $2|\mathscr{G}_3|$ new equations. Therefore, conditioned on $E_{\mathscr{G}_1} \wedge F \vdash \mathrm{Q}_F$, we have

$$\begin{aligned}
&\Pr\left[E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3} | E_{\mathscr{G}_1} \wedge F \vdash \mathrm{Q}_F\right] \\
&\quad \ge \left(1 - \Pr\left[\mathrm{Bad}_1(F_2)\right] - \Pr\left[\mathrm{Bad}_1(F_3)\right]\right) \cdot \Pr\left[E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3} | \mathrm{Bad}_1(F_2) \wedge \mathrm{Bad}_1(F_3)\right] \\
&\quad \ge \left(1 - \frac{(\beta_1 + \beta_2)\left(q_f + q_e\right)}{N}\right) \frac{1}{N^{2\left(|\mathscr{G}_2| + |\mathscr{G}_3|\right)}}.
\end{aligned} \tag{B.55}$$

Lower bounding $\Pr[E_{\mathscr{G}_4} | E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3} \wedge F \vdash \mathrm{Q}_F]$: by definition, for any tuple $(LRX, AST) \in \mathscr{G}_4$, let $x_2 = k_2 \oplus X$ and $x_5 = k_5 \oplus A$; then, we have both $x_2 \notin \mathrm{Dom}\mathscr{F}_2$ and $x_5 \notin \mathrm{Dom}\mathscr{F}_5$. Moreover, conditioned on $E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3}$, the two values $F_2(x_2)$ and $F_5(x_5)$ remain "undetermined" and uniform (otherwise, if $E_{\mathscr{G}_1}$, $E_{\mathscr{G}_2}$, or $E_{\mathscr{G}_3}$ implies $F_2(x_2)$ being fixed, then a tuple in $\mathscr{G}_1$, $\mathscr{G}_2$, or $\mathscr{G}_3$ would share the same X value with a tuple in $\mathscr{G}_4$, contradicting the definition of $\mathscr{G}_1$, or fulfilling (C-4) or (C-5), respectively).

For these tuples, we would lower bound the probability that they induce $2|\mathscr{G}_4|$ new and distinct equations on F_3 and F_4. To this end, we define a predicate $\mathrm{Bad}_3(F_2, F_5)$ on F_2 and F_5, which holds if there exists a tuple $(LRX, AST) \in \mathscr{G}_4$ such that if we let $x_2 = k_2 \oplus X$ and $x_5 = k_5 \oplus A$, then one of the following conditions is fulfilled.

(1) At the "left side," concerning $F_2(x_2)$,

(i) The induced x_3 value falls in $\mathrm{Dom}\mathscr{F}_3$, i.e., $k_3 \oplus A_1 L \oplus A_2 R \oplus F_2(x_2) \in \mathrm{Dom}\mathscr{F}_3$. As discussed, $F_2(x_2)$ remains random; thus, the probability is clearly at most q_f/N, for each $(LRX, AST) \in \mathrm{Dom}\mathscr{G}_4$.

(ii) The induced Y value collides with some "previously determined" Y'; i.e., there exists another tuple $(L'R'X', A'S'T') \in \mathscr{G}_1 \cup \mathscr{G}_2 \cup \mathscr{G}_3$ that $A_1 L \oplus A_2 R \oplus F_2(x_2) = A_1 L' \oplus A_2 R' \oplus F_2(x_{2'})$. It needs to be $X \ne X'$; again, using the randomness of $F_2(x_2)$, we obtain the upper bound $|\mathscr{G}_1| + |\mathscr{G}_2| + |\mathscr{G}_3|/N \le q_e/N$, for each $(LRX, AST) \in \mathscr{G}_4$;

(2) At the "right side," concerning $F_5(x_5)$, similar to the above by symmetry,

$$k_4 \oplus A_1 T \oplus A_2 S \oplus OF_5(x_5) \in \mathrm{Dom}\mathscr{F}_4 \tag{B.56}$$

For each $(LRX, AST) \in \mathscr{G}_4$, the probability is at most q_f/N:

(i) There exists another tuple $(L'R'X', A'S'T') \in \mathscr{G}_1 \cup \mathscr{G}_2 \cup \mathscr{G}_3$ such that $A_1 T \oplus A_2 S \oplus OF_5(x_5) = A_1 T' \oplus A_2 S' \oplus OF_5(x_{5'})$. The upper bound is $|\mathscr{G}_1| + |\mathscr{G}_2| + |\mathscr{G}_3|/N \le q_e/N$, for each (LRX, AST) in $\mathscr{G}_4$.

Thus, using $|\mathscr{G}_4| = \beta_3$, we obtain

$$\Pr\left[\mathrm{Bad}_3(F_2, F_5) | E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3} \wedge F \vdash \mathrm{Q}_F\right] \le 2\left(\frac{|\mathscr{G}_4| \cdot \left(q_f + q_e\right)}{N}\right) \le \frac{2\beta_3\left(q_f + q_e\right)}{N}. \tag{B.57}$$

Similar to the analysis for $E_{\mathscr{G}_2}$ and $E_{\mathscr{G}_3}$, conditioned on $\mathrm{Bad}_3(F_2, F_5)$, the event $E_{\mathscr{G}_4}$ is equivalent to F_3 and F_4 satisfying $2|\mathscr{G}_4|$ new and distinct equations. Therefore,

$$\begin{aligned}&\Pr\left[E_{\mathscr{G}_4} | E_{\mathscr{G}_1} \wedge E_{\mathscr{G}_2} \wedge E_{\mathscr{G}_3} \wedge F \vdash Q_F\right]\\ &\geq \left(1-\Pr\left[\mathrm{Bad}_1\left(F_2, F_5\right)\right]\right)\cdot \mathrm{frac1}N^{2|\mathscr{G}_4|} \geq \left(1-\frac{2\beta_3\left(q_f+q_e\right)}{N}\right)\cdot\frac{1}{N^{2|\mathscr{G}_4|}}.\end{aligned} \tag{B.58}$$

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Sorkin, "Lucifer, a cryptographic algorithm," *Cryptologia*, vol. 8, no. 1, pp. 22–42, 1984.

[2] D. Coppersmith, C. Holloway, S. M. Matyas, and N. Zunic, "The data encryption standard," *Information Security Technical Report*, vol. 2, no. 2, pp. 22–24, 1997.

[3] R. Beaulieu, D. Shors, J. Smith, S. Clark, and L. Wingers, "The Simon and Speck lightweight block ciphers." in *Proceedings of the 52nd Annual Design Automation Conference*, pp. 1–6, San Francisco, CA, USA, June 2015.

[4] G. Yang, B. Zhu, V. Suder, M. D. Aagaard, and G. Gong, "The simeck family of lightweight block ciphers," in *Proceedings of Cryptographic Hardware and Embedded Systems - CHES 2015 -17th International Workshop*, pp. 307–329, Saint-Malo, France, September 2015.

[5] W. Diffie and G. Ledin, "SMS4 encryption algorithm for wireless networks," *IACR Cryptol. ePrint Arch.*vol. 329, 2008.

[6] G. Bose, *The 128-bit block cipher MARS*, Master's thesis, Florida International University, Miami, FL, USA, 2003.

[7] K. Nyberg, "Generalized feistel networks," in *Proceedings of Advances in Cryptology -ASIACRYPT'96, International Conference on the Theory and Applications of Cryptology and Information Security*, pp. 91–104, Kyongju, South Korea, April 1996.

[8] T. P. Berger, J. Francq, M. Minier, and G. Thomas, "Extended generalized feistel networks using matrix representation to propose a new lightweight block cipher: lilliput," *IEEE Transactions on Computers*, vol. 65, no. 7, pp. 2074–2089, 2016.

[9] S. Vaudenay, "On the Lai-Massey scheme," in *Proceedings of Advances in Cryptology -ASIACRYPT99, International Conference on the Theory and Applications of Cryptology and Information Security*, pp. 8–19, Singapore, November 1999.

[10] J. N. Jr, V. Rijmen, B. Preneel, and J. Wandelle, "The mesh block ciphers," in *Proceedings of Information Security Applications, 4th International Workshop*, pp. 458–473, WISA, Jeju Island, Korea, August 2003.

[11] P. Junod and S. Vaudenay, "Fox : a new family of block ciphers," in *Proceedings of Selected Areas in Cryptography, 11th International Workshop, SAC*, pp. 114–129, Waterloo, Canada, August 2004.

[12] J. Daemen and V. Rijmen, *The Design of Rijndael - the Advanced Encryption Standard(AES)*, Springer, Berlin, Germany, 2nd edition, 2020.

[13] J. Guo, T. Peyrin, and A. Poschmann, "The Photon family of lightweight hash functions," in *Proceedings of Advances in Cryptology—CRYPTO 2011 - 31st Annual Cryptology Conference*, pp. 222–239, Santa Barbara, CA, USA, August 2011.

[14] K. Aoki, T. Ichikawa, M. Kanda et al., "Camellia: a 128-bit block cipher suitable for multiple platforms- design and analysis," *In*, in *Proceedings of Selected Areas in Cryptography, 7th Annual International Workshop, SAC*, pp. 39–56, Ontario, Canada, August 2000.

[15] M. Luby and C. Rackoff, "How to construct pseudo-random permutations from pseudo-random functions," in *Proceedings of the Advances in Cryptology—CRYPTO '85*, p. 447, Santa Barbara, CA, USA, August 1985.

[16] J. Patarin, "Security of random feistel schemes with 5 or more rounds," in *Proceedings of the Advances in Cryptology-CRYPTO 2004*, pp. 106–122, Santa Barbara, CA, USA, August 2004.

[17] V. T. Hoang and P. Rogaway, "On generalized feistel networks," in *Proceedings of Advances in Cryptology-CRYPTO 2010*, pp. 613–630, Springer, Santa Barbara, CA, USA, August 2010.

[18] M. Barbosa and P. Farshim, "The related-key analysis of feistel constructions," in *Proceedings of the International Workshop on Fast Software Encryption FSE 2014*, pp. 265–284, London, UK, March 2014.

[19] C. Gentry and Z. Ramzan, "Eliminating random permutation oracles in the even-mansour cipher," in *Proceedings of Advances in Cryptology - ASIACRYPT 2004, 10th International Conference on the Theory and Application of Cryptology and Information Security*, pp. 32–47, Jeju Island, Korea, December 2004.

[20] C. Guo and L. Wang, "Revisiting key-alternating feistel ciphers for shorter keys and multi-user security," in *Proceedings of Advances in Cryptology—ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security*, pp. 213–243, Brisbane, Australia, December 2018.

[21] J. Patarin, "How to construct pseudorandom and super pseudorandom permutations from one single pseudorandom function," in *Proceedings of Advances in Cryptology - EUROCRYPT '92, Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 256–266, Balatonfured, Hungary, May 1992.

[22] S. Chen and J. Steinberger, "Tight security bounds for key-alternating ciphers," in *Proceedings of the Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 327–350, Copenhagen, Denmark, May 2014.

[23] V. T. Hoang and S. Tessaro, "Key-alternating ciphers and key-length extension: exact bounds and multi-user security," in *Proceedings of the CRYPTO 2016*, pp. 3–32, Berlin, Germany, August 2016.

[24] Y. Dodis, J. Katz, J. Steinberger, A. Thiruvengadam, and Z. Zhang, "Provable security of substitution-permutation networks" cryptology eprint archive," 2017, http://eprint.iacr.org/2017/016.pdf Report 2017/016.

[25] B. Cogliati and Y. Seurin, "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 SecurityASIACRYPT 2015*, pp. 134–158, Auckland, New Zealand, December 2015.

[26] B. Cogliati, R. Lampe, and Y. Seurin, "Tweaking even-mansour ciphers," in *Proceedings of the Annual Cryptology Conference CRYPTO 2015*, pp. 189–208, Santa Barbara, CA, USA, August 2015.