BBB PRP Security of the Lai-Massey Mode

Ritam Bhaumik[1] and Mohammad Amin Raeisi[2]

[1] EPFL, Lausanne, Switzerland ritam.bhaumik@epfl.ch
[2] Sharif University of Technology, Tehran, Iran m.aminra81@gmail.com

Abstract. In spite of being a popular technique for designing block ciphers, Lai-Massey networks have received considerably less attention from a security analysis point-of-view than Feistel networks and Substitution-Permutation networks. In this paper we study the beyond-birthday-bound (BBB) security of Lai-Massey networks with independent random round functions against chosen-plaintext adversaries. Concretely, we show that five rounds are necessary and sufficient to achieve BBB security.

Keywords: Beyond-Birthday-Bound security · Block ciphers · Lai-Massey · Provable Security

1 Introduction

Background. Block ciphers, being a crucial cryptographic primitive, have been the subject of study for decades. The Feistel scheme, considered one of the earliest studied block ciphers, was demonstrated to be pseudorandom and a strong pseudorandom permutation in the breakthrough work of Luby and Rackoff [LR88] with three and four rounds, respectively. Since then, numerous studies have focused on analyzing the security of the many-rounds Feistel scheme. Patarin and related authors have established various birthday and beyond-birthday security bounds and designed generic attacks for different numbers of rounds in Feistel schemes in [Pat98, Pat01, Pat03, Pat04, PNB06, PNB07, TP09, VNP10], along with numerous other related works.

Lai-Massey Scheme. The Lai-Massey scheme, considered another significant block cipher, was initially used in [LM90] to propose a cipher known as PES (Proposed Encryption Standard) by Lai and Massey. In essence, the Lai-Massey scheme, operating on the algebraic group $(G, +)$, is a permutation on G^2 characterized by multiple rounds. One round of Lai-Massey consists of the transformation

$$(x, y) \mapsto (\sigma(x + F(x - y)), y + F(x - y)),$$

where F is a round function and σ is an orthomorphism. Subsequently, this work was adapted to develop the block cipher IDEA (International Data Encryption Algorithm) as discussed in [Lai92]. However, Vaudenay [Vau99] was the first one who provided this scheme to construct (strong) pseudorandom permutations at Asiacrypt'99. Later, other ciphers inspired by this scheme, including the WIDEA family [JM09], FOX [JV04], and the MESH family [JRPV03], were built.

Known Results. While there have been some studies on the (strong) pseudorandomness of the Lai-Massey scheme, it has received comparatively less attention regarding security analysis against general attacks when compared to the Feistel scheme and there are many open questions in this area. In [Vau99], it was proved that the three and four-round schemes are respectively pseudorandom and strong pseudorandom permutations up to

$O\left(2^{\frac{n}{2}}\right)$ queries. Later, it was shown in [LLG10] that three and four rounds are necessary for this matter. In [YPL11], Yun et al. introduced a new notion called the quasi-Feistel cipher and exploited that both the Feistel and Lai-Massey schemes belong to this family. This suggested that the Lai-Massey scheme has no advantage over the Feistel scheme. In a subsequent work [LLZ17], Luo et al. presented some generic attacks on the Lai–Massey scheme. They conjectured an advantage of the Lai–Massey scheme over the Feistel scheme within five rounds, based on the higher complexity required by these attacks.

Furthermore, the Beyond-Birthday-Bound security of this scheme has been an open question for many years. In Crypto'10, Hoang and Rogaway proved that most of the well-known types of generalized Feistel schemes achieve BBB security with enough number of rounds [HR10]. Luo et al. adapted this proof idea to analyze the Lai-Massey scheme [LLH15]. They used the coupling technology to find a general security bound for the Lai-Massey scheme. However, their result suggests that we need at least six rounds to achieve BBB security which is improved in our work.

1.1 Our Contributions

In this paper, we close a gap in known results for the BBB security analysis of Lai-Massey modes. It was known before that four rounds are necessary [LLZ17] and six rounds are sufficient [LLH15] for achieving beyond-birthday security. We improve both and show that five rounds are necessary and sufficient for BBB security (when we are using independent round functions).

Attack. Our first contribution is a chosen-plaintext distinguishing attack on four-round Lai-Massey (LM4) with independent round functions, which uses only $O(2^{n/2})$ queries when the round function has width n bits. Our attack is inspired by the birthday attack on four-round Feistel [Pat01] and relies on observing that the probability of certain collisions is doubled in LM4 when compared to a random permutation.

New Proofs. Our second contribution is a proof that a distinguisher making q chosen-plaintext queries to five-round Lai-Massey (LM5) with independent round functions cannot have a significant distinguishing advantage until $q \in \Omega(2^{2n/3})$. The security analysis presented some non-trivial challenges, owing to the somewhat complicated inter-dependencies of the internal variables. We overcome this using a novel technique where we first sample some *equality patterns*, use these to fix the internal equations, and then proceed to sample the internal values. We complete the proof using Jha and Nandi's formalisation [JN18] of Patarin's H-Coefficient Technique [Pat08].

1.2 Related Work

Until now, the security bounds and generic attacks outlined in [LLH15] and [LLZ17] represent the best-known results in the literature. Yet, there has been some growing attention to this scheme in non-classical settings [ZWSW23, CS22] in recent years. Additionally, There have been some works in the recent years to generalize the idea of the Lai-Massey scheme. Shamsabad and Dehanvi [SD20] generalized the ideas of this scheme into a new cipher, called the Generalized Lai-Massey scheme (GLM). Independently, Grassi [Gra22] proposed the generalized Amaryllises construction as a generalization of the Lai-Massey scheme, in which the linear combination in the Lai-Massey scheme is replaced by a non-linear one.

1.3 Outline

The subsequent sections of this paper are structured as follows. In Section 2, we provide essential preliminaries, including notation, security notions, and the proof technique we

are using in our main proof. Section 3 introduces our new attack on the four-round scheme. The proof of BBB security for the five-round scheme, employing the H-Coefficient Technique, is presented in Section 4. Finally, Section 5 provides a concise summary and conclusion of our research findings.

2 Preliminaries

Notation. In our notation, the set $\{1, 2, \ldots, m\}$ is denoted as $[m]$. Tuples of the form $(x_1, x_2, \ldots, x_q)$ are represented as x^q, with each element x_i in the tuple denoted by $x^q|_i$. An orthomorphism σ on an algebraic group $(G, +)$ is characterized as a permutation $x \to \sigma(x)$ such that the transformation $\tau(x) := \sigma(x) - x$ is also a permutation on G. Random selection from a finite set $\mathcal{S}$ is denoted by $S \leftarrow\!\!\$\ \mathcal{S}$. The execution of an algorithm $\mathcal{A}$ with an oracle accessing the function $\mathbf{F}$ is denoted as $\mathcal{A}^{\mathbf{F}(\cdot)}$. A pair of tuples (x^q, y^q) is considered function-compatible if, whenever $x_i = x_j$, it follows that $y_i = y_j$; if the reverse holds as well, we term the tuple permutation-compatible. Lastly, the expression $N(N-1)\ldots(N-r+1)$ is denoted as $(N)_r$.

2.1 Security Notions

We will employ the H-Coefficient Technique to analyze the security of the Lai-Massey scheme. This method relies on the outcomes of the interaction between the distinguisher and the oracle. Therefore we need some mathematical tools to formalize this interaction, and for this purpose, we use probabilistic functions.

Definition 1 (Probabilistic Function)**.** A probabilistic function with an input space $\mathcal{X}$ and an output space $\mathcal{Y}$ is a function $f : \mathcal{R} \times \mathcal{X} \to \mathcal{Y}$ for some finite set $\mathcal{R}$, called random coin space. We also simply write (abusing notation) $f : \mathcal{X} \xrightarrow{*} \mathcal{Y}$ suppressing the notation for random coin space.

Now we can use definition 1 to establish the definitions of joint response and joint query functions. Here, we model the query-asking part of the distinguisher as a joint query function, asking queries to the oracle interactively and the oracle is modeled as a joint response function, providing answers to the distinguisher's queries.

Definition 2 (Joint Response Function)**.** A q-joint $(\mathcal{X}, \mathcal{Y})$ response function is a probabilistic function $\mathbf{F} : \mathcal{X}^q \xrightarrow{*} \mathcal{Y}^q$ such that for all random coin r, the mapping $x^q \mapsto \mathbf{F}(r, x^q)|_i$ is functionally independent of $x_{i+1}, \ldots, x_q$.

Definition 3 (Joint Query Function)**.** A probabilistic function $\mathcal{A} : \mathcal{Y}^q \xrightarrow{*} \mathcal{X}^q$ is called q-joint $(\mathcal{X}, \mathcal{Y})$ query function if for all random coin r, the mapping $y^q \mapsto \mathcal{A}(r, y^q)|_i$ is functionally independent of $y_i, \ldots, y_q$.

One can observe that there exist functions $\mathcal{A}_i$ and $\mathbf{F}_i$, $i \in [q]$, such that for all y^q, $\mathcal{A}(r, y^q)|_i = \mathcal{A}_i(r, y^{i-1})$ and for all x^q, $\mathbf{F}(r', x^q)|_i = \mathbf{F}_i(r', x^i)$.

Next, we aim to define transcripts. A transcript is the outcome of the interaction between the distinguisher and the oracle, representing two tuples of queries and their corresponding responses.

Definition 4 (Transcript)**.** Let $\mathcal{A}$ and $\mathbf{F}$ be $(\mathcal{X}, \mathcal{Y})$ joint query function and joint response function respectively. Let $\mathcal{A}_i$ and $\mathbf{F}_i$ be defined as before. We define the transcript random variable as $\tau(\mathcal{A}^{\mathbf{F}}) = (X^q, Y^q)$ where X_i's and Y_i's are defined recursively as follows:

$$X_i = \mathcal{A}_i(R, Y^{i-1}), Y_i = \mathbf{F}_i(R', X^i), 1 \le i \le q$$

and R and R' are random coins of $\mathcal{A}$ and $\mathbf{F}$ respectively.

Now, let's consider a scenario where the response function provides additional information S when interacting with the adversary. This, in a sense, enhances the distinguisher's advantage in compromising the oracle's security. We will formalize this concept through extended transcripts, which we will later use in subsection 2.2 to define the H-Coefficient Technique.

Definition 5 (Extended Transcript)**.** An $\mathcal{S}$-extended $(\mathcal{X}, \mathcal{Y})$ joint response function is a probabilistic function $\bar{\mathbf{F}} = (\mathbf{F}, S) : \mathcal{X}^q \xrightarrow{*} \mathcal{Y}^q \times \mathcal{S}$. For any $(\mathcal{X}, \mathcal{Y})$ joint query function $\mathcal{A}$, we define the (extended) transcript of $\mathcal{A}^{\bar{\mathbf{F}}}$ as

$$\bar{\tau}(\mathcal{A}^{\mathbf{F}}) = \tau(\mathcal{A}^{\bar{\mathbf{F}}}) = (\tau(\mathcal{A}^{\mathbf{F}}, S(X^q)) := (\tau(\mathcal{A}^{\mathbf{F}(R,.)}), S(R, X^q))$$

where R denotes the random coin of $\bar{\mathbf{F}}$ and $\tau(\mathcal{A}^{\mathbf{F}}) = (X^q, Y^q)$. We call S adjoined random variable to $\mathbf{F}$.

We now possess the necessary tools for the formal definition of the distinguisher. At a high level, a distinguisher is a combination of a joint query function and a decision function. The joint query function interacts with the joint response function and generates some (extended) transcripts. Then, the decision function outputs some decision values based on the result of the interaction.

Definition 6 (Distinguisher)**.** Let $\mathbf{F}$ and $\mathbf{G}$ be two $(\mathcal{X}, \mathcal{Y})$ joint response functions and $\mathcal{A}$ be an $(\mathcal{X}, \mathcal{Y})$ joint query system with random coin space R. Let $b : R \times \mathcal{X}^q \times \mathcal{Y}^q \to \{0, 1\}$ be a binary function (also called decision function) We call the pair $(\mathcal{A}, b)$, denoted as $\mathcal{A}_b$, a distinguisher.

- the algorithm $\mathcal{A}$ interacts with a joint response function and obtains a transcript $\tau = (x^q, y^q)$.
- The function b finally makes a decision based on the transcript and the random cain initially sampled by $\mathcal{A}$.

Furthermore, the advantage of $\mathcal{A}_b$ over response functions $\mathbf{F}$ and $\mathbf{G}$ is defined as

$$\Delta_{\mathcal{A}_b}(\mathbf{F}; \mathbf{G}) := |Pr[\mathcal{A}_b^{\mathbf{F}} \to 1] - Pr[\mathcal{A}_b^{\mathbf{G}} \to 1]|. \tag{1}$$

Before introducing the H-Coefficient Technique, a few points should be noted:

- We consider adversaries with unbounded time, subject to limitations on the complexity associated with the oracle calls.
- As we want to provide some upper bounds for the advantage of any distinguisher over the two response functions, we assume that the decision-making function b is optimum and hence $\Delta_{\mathcal{A}_b}(\mathbf{F}; \mathbf{G}) = \Delta((R, \tau(\mathcal{A}^{\mathbf{F}})); (R, \tau(\mathcal{A}^{\mathbf{G}})))$. As a result, we can ignore the b notation.
- One can easily show that nondeterministic distinguishers have no advantage in comparison to the deterministic ones. Therefore we assume that the distinguishers are deterministic throughout the paper.
- We assume non-redundancy in distinguisher queries, meaning the response to one query cannot be derived from the responses to previous queries.
- Our ultimate goal is to analyze the PRP security of the Lai-Massey scheme. This involves determining upper bounds on the distinguishing advantage of any distinguisher over the Lai-Massey scheme and a random permutation when asking for encryption queries. We introduce the notation $\mathbf{Adv}_{\mathbf{F}}^{prp}(\theta_D) = \max_{\mathcal{A} \in \mathbb{A}(\theta_D)} \Delta_{\mathcal{A}}(\mathbf{F}; \pi)$, where $\mathbb{A}(\theta_D)$ represents the set of adversaries with data complexity at most θ_D, and π stands for the random permutation joint response function. In this context, the only data complexity we consider is the number of queries.

2.2 H-Coefficient Technique

Patarin first explained this technique in his Ph.D. thesis written in French. Later, he formally described it in SAC 2008 [Pat08]. Before that, Vaudenay mentioned this tool publicly in his decorrelation theory [Vau03], referring to Patarin's thesis. Later, Jha and Nandi published an in-depth survey [JN18] on this technique. They provided H-technique-based proofs for various popular symmetric-key designs across different paradigms, and we will use the notations from their survey.

Here, we introduce an extended version of this tool to analyze the security of the Lai-Massey scheme. The H-Coefficient technique examines attainable transcripts to find some upper bounds on the maximum advantage any distinguisher can achieve. At a high level, we categorize the set of all possible transcripts into two groups: good and bad transcripts, with the former demonstrating favorable characteristics.

Lemma 1 (Extended H-Coefficient Techinque)**.** *Suppose that* $\bar{\mathbf{F}} := (\mathbf{F}, S)$ *and* $\bar{\mathbf{G}} := (\mathbf{G}, S')$ *are two* $\mathcal{S}$*-extended* $(\mathcal{X}, \mathcal{Y})$ *response systems. Let* Ω *denote the set of all attainable transcripts, i.e., the support of* $Pr_{\bar{\mathbf{G}}}$*. Suppose there is a set* $\Omega_{bad} \subseteq \Omega$ *such that for all* $(x^q, y^q, s) \notin \Omega_{bad}$*,*

$$\frac{Pr[\mathbf{F}(x^q) = y^q, S = s]}{Pr[\mathbf{G}(x^q) = y^q, S' = s]} \geq 1 - \epsilon$$

for some $\epsilon \geq 0$*. Then, for any* $(\mathcal{X}, \mathcal{Y})$ *adversary* A*,*

$$\Delta_A(\mathbf{F}, \mathbf{G}) \leq Pr[(\tau(\mathcal{A}^{\mathbf{G}}, S') \in \Omega_{bad}] + \epsilon.$$

We will employ Lemma 1 to prove our main results. The key idea is to define $\mathbf{F}$ as the response function of the oracle accessing the Lai-Massey scheme and $\mathbf{G}$ as the response function corresponding to a random permutation. We will use this technique to derive bounds for $\mathbf{Adv}^{prp}_{\mathsf{LM5}}(q)$ based on the number of encryption queries.

3 Lai-Massey Mode and New Attack

Here, we formally describe the Lai-Massey scheme and present our attack on its four-round variant.

3.1 Formal Description

As previously mentioned, the Lai-Massey scheme is a modification of the block cipher IDEA. Suppose that $(G, +)$ is a group. The scheme is a permutation on G^2 and consists of r rounds. Using independent random functions $F_1, F_2, \ldots, F_r$ and an orthomorphism σ on G, the input $(x_0, y_0) \in G^2$ of the scheme undergoes sequential processing through the r rounds. In each i-th round, where $1 \leq i < r$, the scheme generates a new tuple $(x_i, y_i) \in G^2$ from (x_{i-1}, y_{i-1}) in the following way:

$$x_i := \sigma(x_{i-1} + F_i(x_{i-1} - y_{i-1})), \tag{2}$$
$$y_i := y_{i-1} + F_i(x_{i-1} - y_{i-1}). \tag{3}$$

The r-th and final round is similar to the previous rounds, with the modification that the σ is omitted when determining the value of x_r, leading to the modified expression

$$x_r := x_{r-1} + F_r(x_{r-1} - y_{r-1}). \tag{4}$$

The tuple (x_r, y_r) obtained after the r-th round is the output of the scheme.

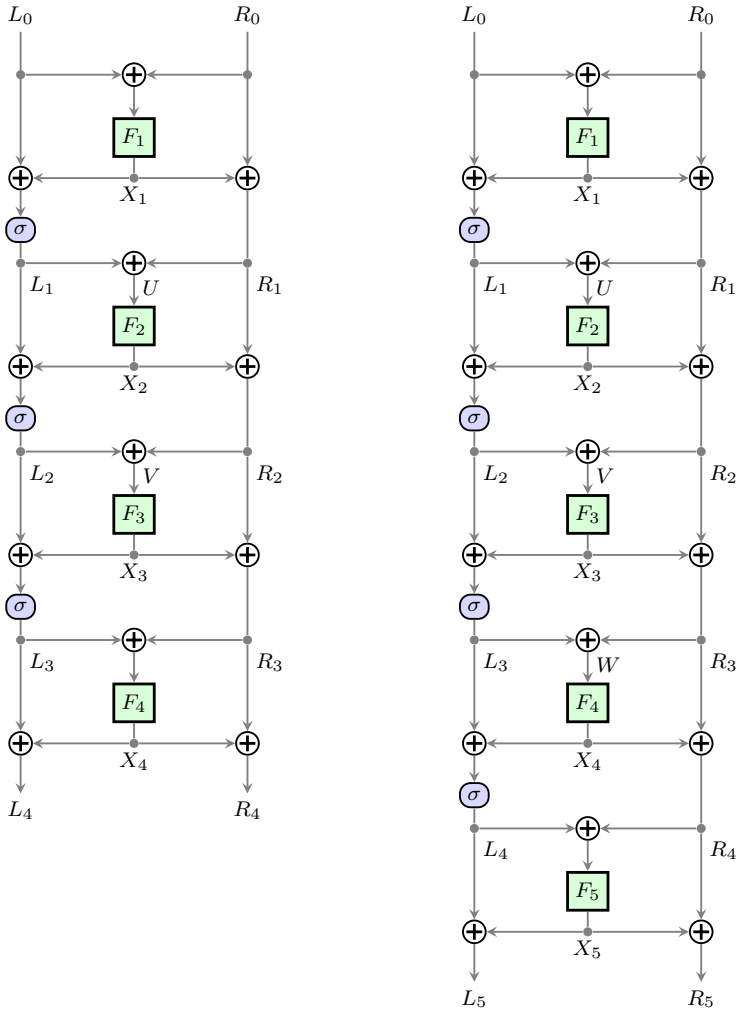

Figure 1: Diagram of the Lai-Massey scheme with internal variables labelled as used in the text. **Left:** 4-round Lai-Massey. **Right:** 5-round Lai-Massey.

In this paper, we consider G as the set of n-bit binary strings, denoted as $\{0,1\}^n$, with the group operation $\oplus$. Thus, the $+$ and $-$ operations in (2), (3), and (4) are both replaced by $\oplus$ operations. Additionally, we take the orthomorphism

$$\sigma(x) = \sigma(x^L, x^R) := (x^R, x^L \oplus x^R),$$

where x^L and x^R respectively denote the left and right halves of x. This definition of σ is commonly employed in the literature and is a standard choice for an orthomorphism on n-bit binary strings. One can easily deduce a few convenient properties of this σ as outlined below in Lemma 2.

Lemma 2. *For the σ defined above and any x, x', we have*

$$\sigma(\sigma(x)) = \sigma^{-1}(x) = \sigma(x) \oplus x,$$
$$\sigma(x \oplus x') = \sigma(x) \oplus \sigma(x').$$

Throughout the paper, we represent the tuple of x_t and y_t values for various scheme inputs as L_t^q and R_t^q, respectively. The tuple of outputs of F_t is denoted as X_t^q. Additionally,

we use L_{ti} and R_{ti} to represent the i-th values for x_t and y_t, respectively. Similarly, X_{ti} denotes the i-th output of F_t.

3.2 Our Attack

In this subsection, we explain our attack on the four-round scheme in detail. One can notice that this attack can be easily changed to compromise the security of the three-round scheme too.

Our attacks use $O\left(2^{\frac{n}{2}}\right)$ chosen plaintexts. The main idea is to identify a random variable for which the expected value within the Lai-Massey scheme significantly differs from its expected value in a random permutation. Patarin originally introduced this idea in [Pat91] and applied it in [Pat01] to design generic attacks on multi-round Feistel schemes. This idea was also independently rediscovered in [AV96]. These results indicate that the four-round scheme does not provide any advantage over the three-round scheme when it comes to CPA attacks.

Before explaining the attack, we present the following lemma to describe the dependencies between the values in the four-round scheme.

Lemma 3. *Let U_i and V_i be the i-th inputs of F_2 and F_3 respectively. We can write the internal values within the scheme in terms of $L_0^q, R_0^q, L_4^q, R_4^q, U^q$ and V^q in the following way:*

$$L_{1i} = \sigma^{-1}(L_{0i} \oplus R_{0i} \oplus U_i), \tag{5}$$
$$R_{1i} = \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus \sigma(U_i), \tag{6}$$
$$L_{2i} = \sigma^{-1}(U_i \oplus V_i), \tag{7}$$
$$R_{2i} = \sigma^{-1}(U_i) \oplus \sigma(V_i), \tag{8}$$
$$X_{2i} = \sigma(V_i) \oplus U_i \oplus \sigma^{-1}(L_{0i} \oplus R_{0i}), \tag{9}$$
$$X_{3i} = \sigma^{-1}(U_i) \oplus V_i \oplus \sigma(L_{4i} \oplus R_{4i}). \tag{10}$$

Lemma 3 follows from the proof of Lemma 4 in Section 4.3, so we skip the proof here. Now, we can describe how our attack operates. We introduce a distinguisher $\mathcal{A}$ (an algorithm) with access to an oracle $\mathcal{O}$, which can be either the four-round scheme or a random permutation, and the goal is to determine which oracle it is interacting with. This algorithm works as described in Algorithm 1.

Algorithm 1 Algorithm to distinguish the four-round scheme from a random permutation

1: Sample m different values $(x_1, x_2, \ldots, x_m)$ from $\{0,1\}^n$ uniformly at random
2: **for** i $\leftarrow 1$ to m **do**
3: $\quad (y_i^1, y_i^2) \leftarrow \mathcal{O}(x_i, x_i)$
4: **end for**
5: $N \leftarrow |\{(i,j) \mid 1 \le i < j \le m,\ y_i^1 \oplus y_i^2 \oplus x_i = y_j^1 \oplus y_j^2 \oplus x_j\}|$
6: Based on the value N, decide whether $\mathcal{O}$ is a random permutation or the four-round scheme

Theorem 1. *The attack described in Algorithm 1 can distinguish the four-round scheme from a random permutation using $m \in O\left(2^{\frac{n}{2}}\right)$ encryption queries.*

Proof. We aim to show that the expected value of N during $\mathcal{A}$'s interaction with the four-round Lai-Massey scheme is approximately twice as much as when it interacts with a random permutation and both of these values can get significant if $\mathcal{A}$ asks up to $m \in O\left(2^{\frac{n}{2}}\right)$ queries.

To demonstrate this, consider the interaction of the distinguisher with the four-round scheme oracle, i.e., the execution of $\mathcal{A}^{\mathbf{F}(.)}$. Here, the i-th scheme input is $L_{0i} = R_{0i} = x_i$, and the i-th output is $L_{4i} = y_i^1$ and $R_{4i} = y_i^2$. Now, first, recalling the identity $\sigma^{-1}(x) = \sigma(x) \oplus x$ from Lemma 2, observe that

$$\begin{aligned} L_{4i} \oplus R_{4i} \oplus x_i &= L_{3i} \oplus R_{3i} \oplus x_i \\ &= \sigma(L_{2i}) \oplus R_{2i} \oplus \sigma^{-1}(X_{3i}) \oplus x_i \\ &= \sigma(\sigma^{-1}(U_i \oplus V_i)) \oplus \sigma^{-1}(U_i) \oplus \sigma(V_i) \oplus \sigma^{-1}(X_{3i}) \oplus x_i \\ &= \sigma(U_i) \oplus \sigma^{-1}(V_i) \oplus \sigma^{-1}(X_{3i}) \oplus x_i. \end{aligned} \tag{11}$$

So we aim to determine the values U_i, V_i, and X_{3i} in (11) and identify an equivalent condition for the equality $y_i^1 \oplus y_i^2 \oplus x_i = y_j^1 \oplus y_j^2 \oplus x_j$ to hold. Based on (5), we have

$$L_{1i} = \sigma^{-1}(x_i \oplus x_i \oplus U_i) = \sigma^{-1}(U_i). \tag{12}$$

Furthermore, by using the definition of L_{1i}, we can write

$$L_{1i} = \sigma(L_{0i} \oplus F_1(L_{0i} \oplus R_{0i})) = \sigma(x_i \oplus F_1(0)). \tag{13}$$

Combining (12) and (13), we obtain

$$U_i = \sigma(\sigma(x_i \oplus F_1(0))) = \sigma^{-1}(x_i \oplus F_1(0)). \tag{14}$$

From (11) and (14) we get

$$\begin{aligned} L_{4i} \oplus R_{4i} \oplus x_i &= \sigma(U_i) \oplus \sigma^{-1}(V_i) \oplus \sigma^{-1}(X_{3i}) \oplus x_i \\ &= x_i \oplus F_1(0) \oplus \sigma^{-1}(V_i) \oplus \sigma^{-1}(X_{3i}) \oplus x_i \\ &= F_1(0) \oplus \sigma^{-1}(V_i \oplus X_{3i}). \end{aligned} \tag{15}$$

From (15) we see that $L_{4i} \oplus R_{4i} \oplus x_i = L_{4j} \oplus R_{4j} \oplus x_j$ if and only if $V_i \oplus X_{3i} = V_j \oplus X_{3j}$. This demonstrates that, within the execution of $\mathcal{A}^{\mathbf{F}(.)}$, the equality $y_i^1 \oplus y_i^2 \oplus x_i = y_j^1 \oplus y_j^2 \oplus x_j$ holds in two cases:

1. when $V_i = V_j$, the outputs of the F_3 function are also equal due to identical inputs, leading to $X_{3i} = X_{3j}$, and the equality holds;
2. In the second case, when $V_i \neq V_j$, but the function outputs can still accidentally satisfy the given equality.

Therefore, the expected value of N during the execution of $\mathcal{A}^{\mathbf{F}(.)}$ is roughly double that during the execution of $\mathcal{A}^{\mathbf{\Pi}(.)}$. Furthermore, since each equality holds with a probability of approximately 2^{-n}, we can choose $m \in O\left(2^{\frac{n}{2}}\right)$ to ensure that both expected values get significant. As a result, the distinguisher can count the number of equalities to distinguish the four-round scheme from a random permutation. □

4 Main Result and Proof

In this section, we present some proofs demonstrating that the five-round scheme is secure up to complexity $2^{\frac{2n}{3}}$ when the underlying functions are pseudorandom.

The main idea is to use the H-Coefficient Technique and release the inputs of F_2, F_3 and F_4 for extended transcripts. Before explaining the proof in detail, we express all information in terms of the transcript values and the inputs of the aforementioned internal functions.

4.1 Internal Values and Compatibility Conditions

Let the i-th-query inputs of F_2, F_3, F_4 be denoted as U_i, V_i, W_i respectively. These values represent the extended parameters within the extended transcripts, serving as the foundational elements upon which we will establish all other values. Let X_{ti} denote the i-th-query output of F_t. We first recall the following defining equations:

$$\begin{aligned} L_{ti} &:= \sigma(L_{(t-1)i} \oplus X_{ti}), \\ R_{ti} &:= R_{(t-1)i} \oplus X_{ti}, \\ U_i &:= L_{1i} \oplus R_{1i}, \\ V_i &:= L_{2i} \oplus R_{2i}, \\ W_i &:= L_{3i} \oplus R_{3i}. \end{aligned}$$

By rearranging the above equations through repeated applications of Lemma 2, we obtain the following lemma.

Lemma 4. *If for some i, in addition to the input blocks L_{0i} and R_{0i}, and the output blocks L_{5i} and R_{5i}, we know the inner function inputs U_i, V_i, and W_i, we can determine all the outputs of F for the i-th query by the following equations:*

$$\begin{aligned} X_{1i} &= \sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i} \oplus U_i), \\ X_{2i} &= \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus U_i \oplus \sigma(V_i), \\ X_{3i} &= \sigma^{-1}(U_i) \oplus V_i \oplus \sigma(W_i), \\ X_{4i} &= \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(V_i) \oplus W_i, \\ X_{5i} &= \sigma(L_{5i}) \oplus \sigma^{-1}(R_{5i} \oplus W_i). \end{aligned}$$

We can also determine the intermediate round outputs by the following equations:

$$\begin{aligned} L_{1i} &= \sigma^{-1}(L_{0i} \oplus R_{0i} \oplus U_i), & R_{1i} &= \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus \sigma(U_i), \\ L_{2i} &= \sigma^{-1}(U_i \oplus V_i), & R_{2i} &= \sigma^{-1}(U_i) \oplus \sigma(V_i), \\ L_{3i} &= \sigma^{-1}(V_i \oplus W_i), & R_{3i} &= \sigma^{-1}(V_i) \oplus \sigma(W_i), \\ L_{4i} &= \sigma^{-1}(L_{5i} \oplus R_{5i} \oplus W_i), & R_{4i} &= \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(W_i). \end{aligned}$$

Function-compatibility demands that we have $X_{tj} = X_{tj}$ whenever the i-th query and the j-th query have the same input to F_t. Then the following corollary follows directly from Lemma 4.

Corollary 1. *The following conditions are necessary and sufficient for the internal functions' inputs and outputs to be function-compatible:*

1. *if $L_{0i} \oplus R_{0i} = L_{0j} \oplus R_{0j}$, then $\sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i} \oplus U_i) = \sigma^{-1}(L_{0j}) \oplus \sigma(R_{0j} \oplus U_j)$;*
2. *if $U_i = U_j$, then $\sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus \sigma(V_i) = \sigma^{-1}(L_{0j} \oplus R_{0j}) \oplus \sigma(V_j)$;*
3. *if $V_i = V_j$, then $\sigma^{-1}(U_i) \oplus \sigma(W_i) = \sigma^{-1}(U_j) \oplus \sigma(W_j)$;*
4. *if $W_i = W_j$, then $\sigma^{-1}(V_i) \oplus \sigma(L_{5i} \oplus R_{5i}) = \sigma^{-1}(V_j) \oplus \sigma(L_{5j} \oplus R_{5j})$;*
5. *if $L_{5i} \oplus R_{5i} = L_{5j} \oplus R_{5j}$, then $\sigma(L_{5i}) \oplus \sigma^{-1}(R_{5i} \oplus W_i) = \sigma(L_{5j}) \oplus \sigma^{-1}(R_{5j} \oplus W_j)$.*

A proof of Lemma 4 is given in Section 4.3.

4.2 BBB Security Proof

We can now prove the BBB security of the five-round scheme. As previously described, We are going to apply the H-Coefficient Technique and release the inputs associated with functions F_2, F_3, and F_4. The main challenge is to find a suitable algorithm for releasing these values in the ideal scheme because they are not well-defined in this context. Our approach involves initially releasing some of these values randomly, followed by the construction of graphs based on the newly established conditions. When we already know that a solution is unattainable within the system, as indicated by the graphs, we will release specific *bad* extended transcripts. Conversely, in cases where a solution is feasible, we will choose a solution within the system, resulting in a *good* transcript.

Theorem 2.

$$\mathbf{Adv}^{prp}_{\mathsf{LM5}}(q) \in O\left(\binom{q}{2} \cdot 2^{-2n} + \binom{q}{3} \cdot (2^{-2n} + 2^{-3n}) + \binom{q}{4} \cdot (2^{-3n} + 2^{-4n})\right).$$

Proof. We break down the proof into three main parts. First, we describe extended transcripts and how they are released in the ideal and real schemes. Next, we analyze the *bad* transcripts and the probability of them occurring in the ideal scheme. Finally, we examine the *good* transcripts and determine lower bounds for the ratio discussed in the H-Coefficient Technique for them.

Extended Systems. Suppose that we have a $(L_0^q, R_0^q, L_5^q, R_5^q)$ tuple as transcript. Let $\mathbf{F}$ be the response system corresponding to LM5 and $\mathbf{\Pi}$ be the system corresponding to a random permutation. We define a $(\{0,1\}^n, \{0,1\}^n, \{0,1\}^n)^q$-extended system by adjoining the internal values U^q, V^q and W^q. In the case of $\mathbf{F}$, this is well-defined from the definition of LM5. In the ideal system $\mathbf{\Pi}$, we sample these values according to Algorithm 2.

Algorithm 2 Algorithm for releasing the internal values in LM5 BBB security proof

for i $\leftarrow$ 1 to q **do**
  if there is some $j < i$ s.t. $L_{0i} \oplus R_{0i} = L_{0j} \oplus R_{0j}$ **then**
    $U_i \leftarrow \sigma(L_{0i} \oplus L_{0j}) \oplus R_{0i} \oplus R_{0j} \oplus U_j$ ▷ According to the corollary 1
  else
    $U_i \leftarrow\!\!\$\ \{0,1\}^n$
  end if
end for
for i $\leftarrow$ 1 to q **do**
  $V_i \leftarrow\!\!\$\ \{0,1\}^n$
  $W_i \leftarrow\!\!\$\ \{0,1\}^n$
end for ▷ These are temporary values
Form condition graphs G_V and G_W
if $(L_0^q, R_0^q, L_5^q, R_5^q, U^q, G_V, G_W)$ is not a *good* tuple **then**
  Release some junk values for V^q and W^q
else
  Release one solution from the system of equations and non-equalities that we have
end if

Now, we proceed to explain the algorithm in detail. This algorithm operates through several stages:

- First, it selects U^q values. It iterates through all the q values, randomly sampling the variables when they are not constrained by previous variables due to the conditions outlined in Corollary 1. If prior values impose constraints on a variable, the algorithm

arbitrarily selects one of these values and chooses the predetermined value to assign to this variable.

- Subsequently, with the values of U^q fixed, we are left with only two sets of unknown variables, V^q and W^q. We proceed to generate temporary values with uniform sampling for V^q and W^q and identify the equality patterns among these values.
- Now that we have established the equality patterns for V^q and W^q and selected the U^q values, it becomes easy to see that all the conditions can be expressed in the form of $V_i \oplus V_j = C_{ij}$ or $W_i \oplus W_j = C'_{ij}$. This allows us to construct condition graphs for the V^q and W^q values, denoted as G_V and G_W respectively. The vertices within these graphs correspond to equality classes present in the equality patterns, and the weighted edges correspond to the conditions. These two graphs are then formed. If the tuple $(L_0^q, R_0^q, L_5^q, R_5^q, U^q, G_V, G_W)$ is deemed as not *good* (we will define this term later) we release some junk values for V^q and W^q such that the input/output pairs of functions do not get compatible. On the other hand, if this tuple is *good*, we proceed to sample a set of values for V^q and W^q that conform to the established equality patterns, along with the system of equations and non-equalities that we have formulated for V^q and W^q values. These values, in combination with U^q values, are then released.

Analysis of Bad Transcripts. Let us consider a fixed extended transcript, denoted as $(L_0^q, R_0^q, L_5^q, R_5^q, U^q, V^q, W^q)$. Our objective is to define *bad* transcripts to prevent function incompatibility within F_1, F_2, F_3, F_4, and F_5. The algorithm we use to sample internal values dictates that function incompatibility arises if and only if the tuple $(L_0^q, R_0^q, L_5^q, R_5^q, U^q, G_V, G_W)$ is not considered *good*. Consequently, we must establish the criteria for a *good* tuple of values and graphs, as well as calculate the probability that such a tuple is not *good*. We assert that this tuple is deemed not *good* if and only if at least one of the following conditions is met:

1. There exist three distinct indices, denoted as i, j, and k, such that $L_{5i} \oplus R_{5i} = L_{5j} \oplus R_{5j} = L_{5k} \oplus R_{5k}$, the probability of this event is approximately $\binom{q}{3} \cdot 2^{-2n}$ due to the randomness in the output values.
2. There are three indices i, j and k such that $U_i = U_j = U_k$. In this case, one can observe that no two values among $L_{0i} \oplus R_{0i}, L_{0j} \oplus R_{0j}$ and $L_{0k} \oplus R_{0k}$ are equal. This is because if any of them were equal, it would imply that their corresponding U values cannot be equal. Consequently, these U values have independent probability distributions, and the probability of this particular event occurring is given by $\binom{q}{3} \cdot 2^{-2n}$.
3. There is any loop in the G_V graph. In this case, there are some i and j such that they belong to the same equality group and either $U_i = U_j$ or i and j are in the same equality group in W^q. So this case holds with probability around $\binom{q}{2} \cdot 2^{-2n}$.
4. A loop is present in the G_W graph. This case is similar to the previous one. As a result, the probability of such an event occurring is approximately $\binom{q}{2} \cdot 2^{-2n}$.
5. There is a multi-edge in the G_V graph. Such an event may occur in the following situations:
   - When $U_i = U_j$ and both indices i and j belong to the same equality group within W^q, the probability of this scenario occurring is approximately $\binom{q}{2} \cdot 2^{-2n}$.
   - When indices j and k are members of the same equality group within V^q and two conditions hold: one involving indices i and j (where either $U_i = U_j$ or i

and j share the same equality group within W^q), and another one relating to indices i and k, the probability of this event is approximately $\binom{q}{3} \cdot 2^{-3n}$.

- This case involves the following conditions: indices i and j belong to the same equality group in V^q, as do indices k and l. Additionally, there should be one condition relating to indices i and k, and another condition for indices j and l. The probability of this scenario is approximately $\binom{q}{4} \cdot 2^{-4n}$.

So the overall probability of this event is at most around $\binom{q}{2} \cdot 2^{-2n} + \binom{q}{3} \cdot 2^{-3n} + \binom{q}{4} \cdot 2^{-4n}$.

6. There is a multi-edge in the W graph. The analysis of this case is similar to the previous case.

7. There is a P_3 subgraph in the G_V graph. Such subgraph can be found in the G_V graph in the following cases:

   - When two conditions are satisfied, one involving indices i and j (where either $U_i = U_j$ or both indices belong to the same equality group within W^q), and the other concerning indices j and k, the probability of this case is around $\binom{q}{3} \cdot 2^{-2n}$.
   - When two conditions are met: one involving indices i and j, and the other involving indices k and l, with the added requirement that indices j and k belong to the same equality group within V^q. the probability for this case is around $\binom{q}{4} \cdot 2^{-3n}$.

   As a result, the overall probability of this case is at most around $\binom{q}{3} \cdot 2^{-2n} + \binom{q}{4} \cdot 2^{-3n}$.

8. There is a P_3 subgraph in the G_W graph. The analysis of this case is similar to the previous case.

9. There are three indices i, j, and k such that they are all members of the same equality group within either V^q or W^q. The probability of this case is around $\binom{q}{3} \cdot 2^{-2n}$.

We can summarize the previous case studies into the following lemma:

Lemma 5. *If we release the extended transcripts as described, we obtain the following bound for the bad transcripts:*

$$\Pr[(\tau(\mathcal{A}^{\mathbf{\Pi}}, U^q, V^q, W^q) \in \Omega_{bad}] \in O\left(\binom{q}{2} \cdot 2^{-2n} + \binom{q}{3} \cdot (2^{-2n} + 2^{-3n}) + \binom{q}{4} \cdot (2^{-3n} + 2^{-4n})\right).$$

Analysis of Good Transcripts. Consider a fixed *good* transcript $(L_0^q, R_0^q, L_5^q, R_5^q, U^q, V^q, W^q)$, and let r_1, r_2, r_3, r_4, and r_5 represent the number of distinct inputs for the functions F_1, F_2, F_3, F_4, and F_5, respectively. Let r denote $r_1 + r_2 + r_3 + r_4 + r_5$. In the real scheme, the probability of observing this transcript is given by

$$\Pr[\mathbf{F}(L_0^q, R_0^q) = (L_5^q, R_5^q), \mathsf{S}' = (U^q, V^q, W^q)] = (2^{-n})^r. \tag{16}$$

In the ideal scheme, this probability is

$$\Pr[\mathbf{\Pi}(L_0^q, R_0^q) = (L_5^q, R_5^q), \mathsf{S} = (U^q, V^q, W^q)] = \frac{1}{(2^{2n})_q} \cdot \Pr[\mathsf{S} = (U^q, V^q, W^q)]. \tag{17}$$

To apply the H-Coefficient Technique, it is necessary to find some lower bounds for the ratio $\frac{(16)}{(17)}$. One can observe that

$$\frac{(16)}{(17)} = \frac{(2^{-n})^r \cdot (2^{2n})_q}{\Pr[\mathsf{S} = (U^q, V^q, W^q)]}$$

$$= \frac{(2^{-n})^{r} \cdot (2^{2n})_q}{(2^{-n})^{r_1} \cdot \Pr[(S_2, S_3) = (V^q, W^q) \mid \mathsf{S}_1 = U^q]}$$
$$= \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q}{\Pr[(\mathsf{S}_2, \mathsf{S}_3) = (V^q, W^q) \mid \mathsf{S}_1 = U^q]}. \tag{18}$$

To compute the probability $\Pr[(\mathsf{S}_2, \mathsf{S}_3) = (V^q, W^q) \mid \mathsf{S}_1 = U^q]$, observe that we first select the equality pattern, and then we choose one solution for the system of equations and inequalities. Therefore, it is essential to establish bounds for both the probability associated with the equality pattern of V^q and W^q and the number of solutions to the system of equations and inequalities.

The probability of obtaining a given equality pattern for V^q is $(2^{-n})^{q-r_3} \cdot \frac{2^n-1}{2^n} \cdot \frac{2^n-2}{2^n} \cdot \ldots \cdot \frac{2^n-r_3+1}{2^n}$. Similarly, the probability of obtaining a given equality pattern for W^q is $(2^{-n})^{q-r_4} \cdot \frac{2^n-1}{2^n} \cdot \frac{2^n-2}{2^n} \cdot \ldots \cdot \frac{2^n-r_4+1}{2^n}$. Therefore, we have

$$(18) = \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot P^*}{(2^{-n})^{q-r_3} \cdot \frac{2^n-1}{2^n} \cdot \frac{2^n-2}{2^n} \cdot \ldots \cdot \frac{2^n-r_3+1}{2^n} \cdot (2^{-n})^{q-r_4} \cdot \frac{2^n-1}{2^n} \cdot \frac{2^n-2}{2^n} \cdot \ldots \cdot \frac{2^n-r_4+1}{2^n}}.$$

Where P^* represents the number of solutions of the system. Thus we have

$$(18) = \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot P^*}{(2^{-n})^{q-r_3} \cdot (2^{-n})^{q-r_4} \cdot \frac{(2^n)_{r_3}}{(2^n)^{r_3}} \cdot \frac{(2^n)_{r_4}}{(2^n)^{r_4}}}$$
$$= \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot P^* \cdot (2^n)^{r_3} \cdot (2^n)^{r_4}}{(2^{-n})^{q-r_3} \cdot (2^{-n})^{q-r_4} \cdot (2^n)_{r_3} \cdot (2^n)_{r_4}}$$
$$= \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot P^*}{(2^{-2n})^{q} \cdot (2^n)_{r_3} \cdot (2^n)_{r_4}}. \tag{19}$$

We now aim to compute P^* and examine (19) to determine a lower bound for $\frac{(16)}{(17)}$.

In the graph G_V, there are r_3 vertices and exactly $m = 2q - (r_2 + r_4)$ edges. Initially, we assign values to single edges. For the first edge, we have 2^n choices. For the second edge, we have a minimum of $2^n - 4$ choices because none of the vertices of this edge can be equal to the previous one. Subsequently, for the third edge, we will have at least $2^n - 8$ choices, and so on. Following this, we choose values for the individual vertices. For the first single vertex, we have at least $2^n - 2m$ choices, and for the second single vertex, we have at least $2^n - 2m - 1$ choices, and so forth. Therefore, the total number of solutions for V^q, denoted as P_V^*, is at least:

$$P_V^* \geq 2^n(2^n - 4)(2^n - 8)\ldots(2^n - 4(m-1)) \cdot (2^n - 2m)_{r_3-2m}$$
$$= 4 \cdot (2^{n-2}) \cdot 4 \cdot (2^{n-2} - 1) \cdot \ldots \cdot 4 \cdot (2^{n-2} - (m-1)) \cdot (2^n - 2m)_{r_3-2m}$$
$$= 4^m \cdot (2^{n-2})_m \cdot (2^n - 2m)_{r_3-2m}. \tag{20}$$

Applying the same reasoning, we can determine a lower bound for the number of solutions for W^q, which is denoted as P_W^*

$$P_W^* \geq 4^{m'} \cdot (2^{n-2})_{m'} \cdot (2^n - 2m')_{r_4-2m'}. \tag{21}$$

By combining equations (20) and (21) to establish a lower bound for P^*, we obtain

$$(19) = \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot P_V^* \cdot P_W^*}{(2^{-2n})^{q} \cdot (2^n)_{r_3} \cdot (2^n)_{r_4}}$$
$$= \frac{(2^{-n})^{r-r_1} \cdot (2^{2n})_q \cdot 4^{m+m'} \cdot (2^{n-2})_m \cdot (2^{n-2})_{m'}}{(2^{-2n})^{q} \cdot (2^n)_{2m} \cdot (2^n)_{2m'}}$$

$$= \frac{(2^{-n})^{4q-(m+m')} \cdot (2^{2n})_q \cdot 4^{m+m'} \cdot (2^{n-2})_m \cdot (2^{n-2})_{m'}}{(2^{-2n})^q \cdot (2^n)_{2m} \cdot (2^n)_{2m'}}$$
$$\geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \frac{(2^{-n})^{4q-(m+m')} \cdot (2^{2n})^q \cdot 4^{m+m'} \cdot (2^{n-2})_m \cdot (2^{n-2})_{m'}}{(2^{-2n})^q \cdot (2^n)_{2m} \cdot (2^n)_{2m'}}$$
$$= \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \frac{(2^n)^{m+m'} \cdot 4^{m+m'} \cdot (2^{n-2})_m \cdot (2^{n-2})_{m'}}{(2^n)_{2m} \cdot (2^n)_{2m'}}. \tag{22}$$

Now, consider that increasing m by one multiplies the fraction by $\frac{2^n \cdot (2^n - 4m)}{(2^n - 2m)(2^n - 2m - 1)}$, which can be simplified to $\frac{2^{2n} - 4m \cdot 2^n}{2^{2n} - 4m \cdot 2^n + (4m^2 + 2m - 2^n)}$. As m is increased from 0 to $\frac{q}{2}$, this fraction first increases and then decreases. Therefore, we only need to consider two cases: either $m = m' = 0$ or $m = m' = \frac{q}{2}$ because increasing either m or m' from 0 to $\frac{q}{2}$ results in the fraction either increasing or decreasing.

If we set $m = m' = 0$, the probability is greater than or equal to $1 - \frac{\binom{q}{2}}{2^{2n}}$, which leads to the desired result. However, in the case where $m = m' = \frac{q}{2}$, we have

$$(22) \geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \frac{(2^n)^q \cdot 4^q \cdot (2^{n-2})_{\frac{q}{2}} \cdot (2^{n-2})_{\frac{q}{2}}}{(2^n)_q \cdot (2^n)_q}$$
$$\geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \frac{(2^n)^q \cdot (2^n)^2 (2^n - 4)^2 \dots (2^n - 4(\frac{q}{2} - 1))^2}{(2^n)_q \cdot (2^n)_q}$$
$$\geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \frac{(2^n)^q \cdot (2^n)(2^n - 2)(2^n - 4)(2^n - 6) \dots (2^n - 2(q-1))}{(2^n)_q \cdot (2^n)_q}$$
$$= \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \prod_{i=0}^{q-1} \frac{2^n \cdot (2^n - 2i)}{(2^n - i)^2}$$
$$= \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \prod_{i=0}^{q-1} (1 - \frac{i^2}{(2^n - i)^2})$$
$$\geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \left(1 - \sum_{i=0}^{q-1} \frac{i^2}{(2^n - i)^2}\right). \tag{23}$$

We aim to establish the security of the scheme for up to $2^{\frac{2n}{3}}$ queries. Consequently, we can assume that $2^n - i \geq 2^n - 2^{\frac{2n}{3}} \geq \frac{2^n}{2}$, and thus, we have

$$(23) \geq \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \left(1 - \sum_{i=0}^{q-1} \frac{4i^2}{2^{2n}}\right)$$
$$= \left(1 - \frac{\binom{q}{2}}{2^{2n}}\right) \left(1 - \frac{q^3}{2^{2n}}\right) \geq 1 - \left(\frac{\binom{q}{2}}{2^{2n}} + \frac{q^3}{2^{2n}}\right). \tag{24}$$

Inequality (24) demonstrates that we achieve the desired lower bound in this case as well. This leads us to the following lemma for the *good* transcripts:

Lemma 6. *We can establish the following lower bound for the previously discussed ratio*

$$\frac{\Pr[\mathbf{F}(L_0^q, R_0^q) = (L_5^q, R_5^q), \mathsf{S}' = (U^q, V^q, W^q)]}{\Pr[\mathbf{\Pi}(L_0^q, R_0^q) = (L_5^q, R_5^q), \mathsf{S} = (U^q, V^q, W^q)]} \geq 1 - \left(\frac{\binom{q}{2}}{2^{2n}} + \frac{q^3}{2^{2n}}\right).$$

By combining Lemma 5 and Lemma 6, we can derive the result using the H-Coefficient Technique. □

4.3 Proof of Lemma 4

We first express the value X_{1i} in terms of the provided values:

$$\begin{aligned} & U_i = \sigma(L_{0i}) \oplus R_{0i} \oplus \sigma^{-1}(X_{1i}) \\ \implies & \sigma^{-1}(X_{1i}) = \sigma(L_{0i}) \oplus R_{0i} \oplus U_i \\ \implies & X_{1i} = \sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i}) \oplus \sigma(U_i). \end{aligned} \tag{25}$$

Next, we see that

$$\begin{aligned} V_i &= \sigma(L_{1i}) \oplus R_{1i} \oplus \sigma^{-1}(X_{2i}) \\ &= \sigma(\sigma(L_{0i} \oplus X_{1i})) \oplus R_{0i} \oplus X_{1i} \oplus \sigma^{-1}(X_{2i}) \\ &= \sigma^{-1}(L_{0i}) \oplus \sigma^{-1}(X_{1i}) \oplus R_{0i} \oplus X_{1i} \oplus \sigma^{-1}(X_{2i}) \\ &= \sigma^{-1}(L_{0i}) \oplus \sigma(X_{1i}) \oplus R_{0i} \oplus \sigma^{-1}(X_{2i}) \\ &\overset{(25)}{=} \sigma^{-1}(L_{0i}) \oplus \sigma(\sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i}) \oplus \sigma(U_i)) \oplus R_{0i} \oplus \sigma^{-1}(X_{2i}) \\ &= \sigma^{-1}(L_{0i}) \oplus L_{0i} \oplus \sigma^{-1}(R_{0i}) \oplus \sigma^{-1}(U_i) \oplus R_{0i} \oplus \sigma^{-1}(X_{2i}) \\ &= \sigma(L_{0i} \oplus R_{0i}) \oplus \sigma^{-1}(U_i) \oplus \sigma^{-1}(X_{2i}). \end{aligned}$$

Then we can find the value of X_{2i} as

$$\begin{aligned} & \sigma^{-1}(X_{2i}) = \sigma(L_{0i} \oplus R_{0i}) \oplus \sigma^{-1}(U_i) \oplus V_i \\ \implies & X_{2i} = \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus U_i \oplus \sigma(V_i). \end{aligned} \tag{26}$$

Based on (25) and (26), we can deduce the values of L_{1i}, R_{1i}, L_{2i}, and R_{2i} too as

$$\begin{aligned} L_{1i} &= \sigma(L_{0i} \oplus X_{1i}) \\ &\overset{(25)}{=} \sigma(L_{0i} \oplus \sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i}) \oplus \sigma(U_i)) \\ &= \sigma(\sigma(L_{0i} \oplus R_{0i} \oplus U_i)) = \sigma^{-1}(L_{0i} \oplus R_{0i} \oplus U_i); \qquad (27) \\ R_{1i} &= R_{0i} \oplus X_{1i} \\ &\overset{(25)}{=} R_{0i} \oplus \sigma^{-1}(L_{0i}) \oplus \sigma(R_{0i}) \oplus \sigma(U_i) = \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus \sigma(U_i); \qquad (28) \\ L_{2i} &= \sigma(L_{1i} \oplus X_{2i}) \\ &\overset{(26),(27)}{=} \sigma(\sigma^{-1}(L_{0i} \oplus R_{0i} \oplus U_i) \oplus \sigma^{-1}(L_{0i} \oplus R_{0i}) \oplus U_i \oplus \sigma(V_i)) \\ &= \sigma(\sigma(U_i \oplus V_i))) = \sigma^{-1}(U_i \oplus V_i); \qquad (29) \\ R_{2i} &= \sigma^{-1}(U_i) \oplus \sigma(V_i). \qquad (30) \end{aligned}$$

Next we write W_i in terms of the known values:

$$\begin{aligned} & W_i = L_{3i} \oplus R_{3i} \\ & \quad = \sigma(L_{2i}) \oplus R_{2i} \oplus \sigma^{-1}(X_{3i}) \\ & \quad \overset{(29),(30)}{=} U_i \oplus V_i \oplus \sigma^{-1}(U_i) \oplus \sigma(V_i) \oplus \sigma^{-1}(X_{3i}) \\ & \quad = \sigma(U_i) \oplus \sigma(V_i) \oplus \sigma(X_{3i}) \\ \implies & \sigma^{-1}(X_{3i}) = \sigma(U_i) \oplus \sigma^{-1}(V_i) \oplus W_i \\ \implies & X_{3i} = \sigma^{-1}(U_i) \oplus V_i \oplus \sigma(W_i). \end{aligned} \tag{31}$$

Then we can find the values of L_{3i} and R_{3i}:

$$L_{3i} = \sigma(L_{2i} \oplus X_{3i})$$

$$\stackrel{(29),(31)}{=} \sigma(\sigma^{-1}(U_i \oplus V_i) \oplus \sigma^{-1}(U_i) \oplus V_i \oplus \sigma(W_i))$$
$$= \sigma(\sigma(V_i \oplus W_i)) = \sigma^{-1}(V_i \oplus W_i); \tag{32}$$
$$R_{3i} = \sigma^{-1}(V_i) \oplus \sigma(W_i). \tag{33}$$

One last observation is

$$\begin{aligned} L_{5i} \oplus R_{5i} &= L_{4i} \oplus R_{4i} \\ &= \sigma(L_{3i}) \oplus R_{3i} \oplus \sigma^{-1}(X_{4i}) \\ &\stackrel{(35),(33)}{=} V_i \oplus W_i \oplus \sigma^{-1}(V_i) \oplus \sigma(W_i) \oplus \sigma^{-1}(X_{3i}) \\ &= \sigma(V_i) \oplus \sigma^{-1}(W_i) \oplus \sigma^{-1}(X_{4i}) \\ \implies \quad \sigma^{-1}(X_{4i}) &= L_{5i} \oplus R_{5i} \oplus \sigma(V_i) \oplus \sigma^{-1}(W_i) \\ \implies \quad X_{4i} &= \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(V_i) \oplus W_i. \end{aligned} \tag{34}$$

This allows us to calculate

$$\begin{aligned} L_{4i} &= \sigma(L_{3i} \oplus X_{4i}) \\ &\stackrel{(35),(34)}{=} \sigma(\sigma^{-1}(V_i \oplus W_i) \oplus \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(V_i) \oplus W_i) \\ &= \sigma(\sigma(L_{5i} \oplus R_{5i} \oplus W_i)) = \sigma^{-1}(L_{5i} \oplus R_{5i} \oplus W_i); \end{aligned} \tag{35}$$
$$R_{4i} = \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(W_i). \tag{36}$$

Finally, we can determine the value of X_{5i} as

$$\begin{aligned} X_{5i} &= R_{5i} \oplus R_{4i} \\ &= R_{5i} \oplus R_{3i} \oplus X_{4i} \\ &\stackrel{(33),(34)}{=} R_{5i} \oplus \sigma^{-1}(V_i) \oplus \sigma(W_i) \oplus \sigma(L_{5i} \oplus R_{5i}) \oplus \sigma^{-1}(V_i) \oplus W_i \\ &= \sigma(L_{5i}) \oplus \sigma^{-1}(R_{5i}) \oplus \sigma^{-1}(W_i). \end{aligned} \tag{37}$$

This completes the proof of Lemma 4. □

5 Conclusion

In this work, we analyzed the Lai-Massey scheme as initially introduced in [LM90, Vau99] and improved the existing security bounds for this scheme. Specifically, we demonstrated that five rounds are necessary and sufficient for achieving BBB security. Our findings highlight new open problems in the literature, offering interesting directions for future research:

- Our results indicate that five rounds suffice for achieving BBB security against chosen-plaintext adversaries. However, the security against chosen-ciphertext adversaries remains unexplored. An essential open problem is determining the optimal number of rounds for achieving security in this context.
- Our four-round scheme attack suggests that $O\left(2^{\frac{n}{2}}\right)$ represents the optimal security bound for PRP security in the $LM4$ scheme. However, no tight security bounds are available for a greater number of rounds or different security notions, posing an open problem in the literature. Addressing this problem not only resolves the minimum necessary number of rounds for optimal security but also explores whether the Lai-Massey scheme holds any advantage over the Feistel scheme in terms of security bounds.

- We designed an attack for a commonly used orthomorphism permutation. A key question remains: can the ideas we have discussed here be applied to all linear or even all orthomorphism permutations?

References

[AV96] William Aiello and Ramarathnam Venkatesan. Foiling birthday attacks in length-doubling transformations - benes: A non-reversible alternative to feistel. In Ueli M. Maurer, editor, *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, volume 1070 of *Lecture Notes in Computer Science*, pages 307–320. Springer, 1996.

[CS22] Amit Kumar Chauhan and Somitra Sanadhya. Quantum security of FOX construction based on lai-massey scheme. *IACR Cryptol. ePrint Arch.*, page 1001, 2022.

[Gra22] Lorenzo Grassi. On generalizations of the lai-massey scheme: the birth of amaryllises. *IACR Cryptol. ePrint Arch.*, page 1245, 2022.

[HR10] Viet Tung Hoang and Phillip Rogaway. On generalized feistel networks. In Tal Rabin, editor, *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, volume 6223 of *Lecture Notes in Computer Science*, pages 613–630. Springer, 2010.

[JM09] Pascal Junod and Marco Macchetti. Revisiting the IDEA philosophy. In Orr Dunkelman, editor, *Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, February 22-25, 2009, Revised Selected Papers*, volume 5665 of *Lecture Notes in Computer Science*, pages 277–295. Springer, 2009.

[JN18] Ashwin Jha and Mridul Nandi. Applications of h-technique: Revisiting symmetric key security analysis. *IACR Cryptol. ePrint Arch.*, page 1130, 2018.

[JRPV03] Jorge Nakahara Jr., Vincent Rijmen, Bart Preneel, and Joos Vandewalle. The MESH block ciphers. In Kijoon Chae and Moti Yung, editors, *Information Security Applications, 4th International Workshop, WISA 2003, Jeju Island, Korea, August 25-27, 2003, Revised Papers*, volume 2908 of *Lecture Notes in Computer Science*, pages 458–473. Springer, 2003.

[JV04] Pascal Junod and Serge Vaudenay. FOX : A new family of block ciphers. In Helena Handschuh and M. Anwar Hasan, editors, *Selected Areas in Cryptography, 11th International Workshop, SAC 2004, Waterloo, Canada, August 9-10, 2004, Revised Selected Papers*, volume 3357 of *Lecture Notes in Computer Science*, pages 114–129. Springer, 2004.

[Lai92] Xuejia Lai. *On the design and security of block ciphers.* PhD thesis, ETH Zurich, Zürich, Switzerland, 1992.

[LLG10] Yiyuan Luo, Xuejia Lai, and Zheng Gong. Pseudorandomness analysis of the (extended) lai-massey scheme. *Inf. Process. Lett.*, 111(2):90–96, 2010.

[LLH15] Yiyuan Luo, Xuejia Lai, and Jing Hu. The pseudorandomness of many-round lai-massey scheme. *J. Inf. Sci. Eng.*, 31(3):1085–1096, 2015.

[LLZ17] Yiyuan Luo, Xuejia Lai, and Yujie Zhou. Generic attacks on the lai-massey scheme. *Des. Codes Cryptogr.*, 83(2):407–423, 2017.

[LM90] Xuejia Lai and James L. Massey. A proposal for a new block encryption standard. In Ivan Damgård, editor, *Advances in Cryptology - EUROCRYPT '90, Workshop on the Theory and Application of of Cryptographic Techniques, Aarhus, Denmark, May 21-24, 1990, Proceedings*, volume 473 of *Lecture Notes in Computer Science*, pages 389–404. Springer, 1990.

[LR88] Michael Luby and Charles Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM J. Comput.*, 17(2):373–386, 1988.

[Pat91] Jacques Patarin. New results on pseudorandom permutation generators based on the DES scheme. In Joan Feigenbaum, editor, *Advances in Cryptology - CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1991, Proceedings*, volume 576 of *Lecture Notes in Computer Science*, pages 301–312. Springer, 1991.

[Pat98] Jacques Patarin. About feistel schemes with six (or more) rounds. In Serge Vaudenay, editor, *Fast Software Encryption, 5th International Workshop, FSE '98, Paris, France, March 23-25, 1998, Proceedings*, volume 1372 of *Lecture Notes in Computer Science*, pages 103–121. Springer, 1998.

[Pat01] Jacques Patarin. Generic attacks on Feistel schemes. In Colin Boyd, editor, *ASIACRYPT 2001*, volume 2248 of *LNCS*, pages 222–238. Springer, Heidelberg, December 2001.

[Pat03] Jacques Patarin. Luby-rackoff: 7 rounds are enough for $2^{n(1\text{-epsilon})}$security. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, volume 2729 of *Lecture Notes in Computer Science*, pages 513–529. Springer, 2003.

[Pat04] Jacques Patarin. Security of random feistel schemes with 5 or more rounds. In Matthew K. Franklin, editor, *Advances in Cryptology - CRYPTO 2004, 24th Annual International CryptologyConference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings*, volume 3152 of *Lecture Notes in Computer Science*, pages 106–122. Springer, 2004.

[Pat08] Jacques Patarin. The "coefficients h" technique. In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *Selected Areas in Cryptography, 15th International Workshop, SAC 2008, Sackville, New Brunswick, Canada, August 14-15, Revised Selected Papers*, volume 5381 of *Lecture Notes in Computer Science*, pages 328–345. Springer, 2008.

[PNB06] Jacques Patarin, Valérie Nachef, and Côme Berbain. Generic attacks on unbalanced Feistel schemes with contracting functions. In Xuejia Lai and Kefei Chen, editors, *ASIACRYPT 2006*, volume 4284 of *LNCS*, pages 396–411. Springer, Heidelberg, December 2006.

[PNB07] Jacques Patarin, Valérie Nachef, and Côme Berbain. Generic attacks on unbalanced Feistel schemes with expanding functions. In Kaoru Kurosawa, editor, *ASIACRYPT 2007*, volume 4833 of *LNCS*, pages 325–341. Springer, Heidelberg, December 2007.

[SD20] M. R. Mirzaee Shamsabad and Seyed Mojtaba Dehnavi. Lai-massey scheme revisited. *IACR Cryptol. ePrint Arch.*, page 5, 2020.

[TP09] Joana Treger and Jacques Patarin. Generic attacks on Feistel networks with internal permutations. In Bart Preneel, editor, *AFRICACRYPT 09*, volume 5580 of *LNCS*, pages 41–59. Springer, Heidelberg, June 2009.

[Vau99] Serge Vaudenay. On the Lai-Massey scheme. In Kwok-Yan Lam, Eiji Okamoto, and Chaoping Xing, editors, *ASIACRYPT'99*, volume 1716 of *LNCS*, pages 8–19. Springer, Heidelberg, November 1999.

[Vau03] Serge Vaudenay. Decorrelation: A theory for block cipher security. *J. Cryptol.*, 16(4):249–286, 2003.

[VNP10] Emmanuel Volte, Valérie Nachef, and Jacques Patarin. Improved generic attacks on unbalanced Feistel schemes with expanding functions. In Masayuki Abe, editor, *ASIACRYPT 2010*, volume 6477 of *LNCS*, pages 94–111. Springer, Heidelberg, December 2010.

[YPL11] Aaram Yun, Je Hong Park, and Jooyoung Lee. On lai-massey and quasi-feistel ciphers. *Des. Codes Cryptogr.*, 58(1):45–72, 2011.

[ZWSW23] Zhongya Zhang, Wenling Wu, Han Sui, and Bolin Wang. Post-quantum security on the lai-massey scheme. *Des. Codes Cryptogr.*, 91(8):2687–2704, 2023.