Two-Key Variant of the Four-Round Cascading LRW1

Shreya Dey[1,2], Avijit Dutta[1,3] and Kazuhiko Minematsu[4]

[1] Institute for Advancing Intelligence (IAI), TCG CREST, Salt Lake, Kolkata 700091, India
[2] Ramakrishna Mission Vivekananda Educational and Research Institute, Belur, Howrah 711202, India
[3] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India
[4] NEC Corporation, Kawasaki, Kanagawa, Japan
shreya.dey@tcgcrest.org, avijit.dutta@tcgcrest.org, k-minematsu@nec.com

Abstract. In EUROCRYPT'20, Bao et al. have proved that three rounds of cascaded LRW1 construction provide security up to $2^{2n/3}$ queries. However, in a recent work by Khairallah et al., it has been shown that the construction provides only birthday bound security via exhibiting a distinguishing attack on the construction, and thereby invalidating the claim of Bao et al. In an independent and contemporaneous work, Datta et al. have shown that four rounds of cascading of the LRW1 construction, dubbed as CLRW1^4—based on four independent keyed block ciphers—achieves $3n/4$-bit CCA security. In this paper, we have shown that a key reduced variant of the CLRW1^4 construction, dubbed as R-CLRW1^4 based on two independent keyed block ciphers, achieves $2n/3$-bit CCA security. The security proof of our construction relies on a heavy use of the H-Coefficient technique and non-trivial analysis in lower-bounding the real interpolation probability for good transcripts.

Keywords: Tweakable Block Cipher, Cascaded LRW, Beyond Birthday Bound, Mirror Theory, H-Coefficients Technique

1 Introduction

Block cipher, a ubiquitous object in cryptography, is a family of keyed permutations. Although block ciphers have received widespread adoption over the years as a fundamental cryptographic object, they inherently lack flexibility because they strictly limit the variability of the cipher in different modes of operation. To avoid this limitation, many applications of block ciphers are either implicitly or explicitly designed from a tweakable block cipher (TBC), another fundamental cryptographic primitive that introduces variability in the cipher. A TBC is defined as a family of permutations $\widetilde{\mathsf{E}} : \mathcal{K} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$, indexed by secret key $k \in \mathcal{K}$ and public tweak $t \in \mathcal{T}$. Liskov, Rivest, and Wagner [LRW02] first formalized the security notion of TBC and later it was found a broad range of its applications in the direction of designing authenticated encryption schemes OCB [RBB03], AEZ [HKR15], Deoxys [JNPS21], disk encryption scheme XTS [Mar10], and numerous MAC schemes such as PMAC [Rog04] and ZMAC [IMPS17], either in a direct way (assuming a dedicated TBC as its core) or an indirect way (the scheme is a block cipher mode, however, it internally builds a TBC based on a block cipher).
Recent standardization processes of CAESAR [CAE14] and NIST [NIS18] also accelerated this trend. The first construction of TBC was proposed by Liskov, Rivest, and Wanger in [LRW02], which was later renamed to LRW1 and LRW2 by Landecker et al. [LST12]. LRW1 transforms a block cipher into a TBC by masking the encryption output of the input message with the given tweak which is again re-encrypted to produce the ciphertext,

i.e., for a given block cipher E with key space $\{0,1\}^n$ and message space $\{0,1\}^n$, LRW1 construction is defined as follows:

$$\mathsf{LRW1}[\mathsf{E}]_K(T, M) \triangleq \mathsf{E}_K(\mathsf{E}_K(M) \oplus T),$$

where $T \in \{0,1\}^n$ is the tweak and $M \in \{0,1\}^n$ is the input message. It has been proved that LRW1 offers a tight $n/2$-bit security against chosen plaintext adversaries (CPA) [LRW02]. On the other hand, LRW2 transforms a block cipher into a TBC by masking the input and output of the given block cipher with hash of the given tweak, i.e., for a given block cipher E with key space $\{0,1\}^n$ and message space $\{0,1\}^n$. For a given almost-XOR-universal (AXU) keyed hash function[1] H, LRW2 construction is defined as follows:

$$\mathsf{LRW2}_{K,K'}[\mathsf{E}, H](T, M) \triangleq \mathsf{E}_K(M \oplus H_{K'}(T)) \oplus H_{K'}(T).$$

Liskov et al. [LRW02] have shown that LRW2 achieves a tight $n/2$-bit security against chosen-ciphertext adversaries. LRW2 has later been extended by Landecker et al. to provide beyond birthday-bound security. In particular, Landecker et al. [LST12] have shown that two-round cascading of the LRW2 construction, called CLRW2, achieves $2n/3$-bit CCA security. This bound was later improved to a tight $3n/4$-bit security in two independent works by Mennink [Men18], and Jha and Nandi [JN20]. Lampe and Seurin [LS13] have shown that r-round cascading of the LRW2 construction achieves $rn/r+2$-bit CCA security.

Table 1: Comparison table for the tweakable block ciphers which are designed from block ciphers based on the pseudorandom permutation assumption. "$*$" denotes that the tweak length of the corresponding construction is of arbitrary length. "# of Hash" denotes the number of times the hash function evaluations take a tweak ("–" denotes there is no hash function). "†" denotes the corresponding security bound is tight. "$\star$" denotes the corresponding security bound of TNT proven in [BGGS20] is flawed, demonstrated by an attack of Khairallah [Kha23], and Jha et al. [JNS23, JKNS23]. Security bound of all these constructions are mentioned in CCA setting, except LRW1, which is proven secure only under CPA.

Constructions	(# of BC, Key)	Tweak	# of Hash	Security
LRW1 [LRW02]	$(2, k)$	n	–	$n/2$ (CPA) (†)
LRW2 [LRW02]	$(1, 2)$	$*$	1	$n/2$ (†)
CLRW2 [LST12]	$(2, 2)$	$*$	2	$3n/4$ [JN20] (†)
CLRW2^r [LS13]	(r, r): r odd	$*$	r	$(r-1)n/(r+1)$
	(r, r): r even	$*$	r	$rn/(r+2)$
TNT [BGGS20]	$(3, 3)$	n	–	$2n/3$ ($\star$)
CLRW1^r [ZQG22]	(r, r): r odd	n	–	$(r-1)n/(r+1)$
	(r, r): r even	n	–	$(r-2)n/r$
1k-TNT [JKNS23]	$(3, 1)$	n	–	$n/2$ †
LRW+ [JKNS23]	$(2, 2)$	$*$	2	$3n/4$
CLRW1^4 [DDDM23]	$(4, 4)$	n	–	$3n/4$
R-CLRW1^4 [This Paper]	$(4, 2)$	n	–	$2n/3$

1.1 Recent Progress on Cascading LRW1

In Eurocrypt 2020, Bao et al. [BGGS20] have shown that the 3-round cascading of the LRW1 construction, a.k.a TNT, achieves $2n/3$-bit CCA security. TNT is the extension of the basic LRW1 construction by masking its output with the given tweak and re-encrypting

[1] A keyed function $H : \mathcal{K} \times \mathcal{X} \to \{0,1\}^n$ is ϵ-AXU if $\max_{x \neq x, y \in \{0,1\}^n} \Pr[H_K(x) \oplus H_K(x') = y] \leq \epsilon$.

it with an independent keyed block cipher to produce the ciphertext, i.e., for a given block cipher family $\mathsf{E} : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$, indexed by κ-bit secret key, the construction TNT is defined as follows:

$$\mathsf{TNT}_{K_1,K_2,K_3}[\mathsf{E}](T,M) \stackrel{\Delta}{=} \mathsf{E}_{K_3}(T \oplus \underbrace{\mathsf{E}_{K_2}(T \oplus \mathsf{E}_{K_1}(M))}_{\mathsf{LRW1}_{K_1,K_2}}).$$

Later in [GGLS20], Guo et al. have shown that TNT achieves $3n/4$-bit security bound against all possible information theoretic CPA adversaries. In [ZQG22], Zhang et al. have studied the security analysis of the r-round cascading of the $\mathsf{LRW1}$ construction, called $\mathsf{CLRW1}^r$, defined as follows:

$$\mathsf{CLRW1}^r_{K_1,K_2,\ldots,K_r}[\mathsf{E}](T,M) \stackrel{\Delta}{=} \mathsf{E}_{K_r}(T \oplus \mathsf{E}_{K_{r-1}}(T \oplus \cdots (T \oplus \mathsf{E}_{K_2}(T \oplus \mathsf{E}_{K_1}(M))))).$$

The authors have shown that $\mathsf{CLRW1}^r$ achieves CCA security up to $2^{(r-2)n/r}$ queries, for $r \geq 2$. In addition, the construction achieves an improved CCA security bound up to $2^{(r-1)n/(r+1)}$ queries, when r is odd, using the folklore result of the coupling technique [LPS12]. In a recent work, Khairallah et al. [JNS23] have demonstrated a birthday bound CCA distinguishing attack and shown a provable birthday bound guarantee on the TNT construction. Their result, in effect, has invalidated the security claim of Bao et al. [BGGS20]. Moreover, authors have also proposed a TBC construction, dubbed as $\mathsf{4\text{-}LRW1}$, that provably achieves $3n/4$-bit CCA security. Around the same time, Datta et al. [DDDM23] independently showed that a four-round cascading of the $\mathsf{LRW1}$ construction, dubbed $\mathsf{CLRW1}^4$, also achieves $3n/4$-bit CCA security. [2] For an n-bit keyed block cipher E_k, the construction $\mathsf{CLRW1}^4$ is defined as follows:

$$\mathsf{CLRW1}^4_{K_1,K_2,K_3,K_4}[\mathsf{E}](T,M) \stackrel{\Delta}{=} \mathsf{E}_{K_4}(T \oplus \underbrace{\mathsf{E}_{K_3}(T \oplus \mathsf{E}_{K_2}(T \oplus \mathsf{E}_{K_1}(M)))}_{\mathsf{CLRW1}^3_{K_1,K_2,K_3}[\mathsf{E}]}).$$

Note that, the output of the $\mathsf{CLRW1}^4$ construction can be realized by masking the output of the TNT construction with tweak T and then re-encrypt it with an independent block cipher E_{K_4}. We compare different constructions of tweakable block ciphers in Table 1.

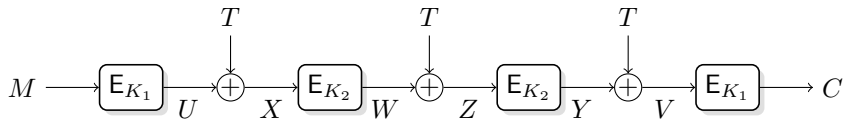


Figure 1: The R-$\mathsf{CLRW1}^4$ construction based on two block cipher keys.

1.2 Ideal Block Cipher-Based TBC Construction

Designing TBCs out of an ideal block cipher was first proposed by Mennink in [Men15], where two candidates were proposed $\widetilde{F}[1]$ and $\widetilde{F}[2]$ that are built out of an ideal block cipher with n-bit key and n-bit data. It has been shown that the former achieves $2n/3$ bit security whereas the latter achieves n-bit security. Later in [WGZ+16], Wang et al. identified a flaw in Mennink's design and mounted a key recovery attack on the construction in birthday query complexity. In addition to this, Wang et al. proposed 32 efficient constructions of TBCs with two invocations of an ideal block cipher and showed that each of them achieves n-bit security. Both the proposals by Mennink and Wang et al. admit tweaks of

[2] We would like to note that $\mathsf{4\text{-}LRW1}$ and $\mathsf{CLRW1}^4$ are the same constructions but with different names.

length n-bits. To admit a longer tweak, Jha et al. [JLM+17] proposed XHX and proved its security up to $(n+\kappa)/2$ bits when using an n-bit block cipher with κ-bit key. Later Lee and Lee [LL18] extended XHX by one round, dubbed it as XHX2, and proved its security up to $\min\{2(n+\kappa)/3, n+(\kappa/2)-\log_2 n\}$ bits. However, both XHX and XHX2 use an almost-XOR-universal hash to incorporate a large tweak. In [SS23], Shen and Standaert studied designing n-bit secure TBC with $2n$-bit tweaks from a block cipher. They have shown that three block cipher calls are necessary and sufficient to ensure beyond the birthday bound security for designing a TBC with $2n$-bit tweak out of an ideal block cipher.

1.3 Our Contribution

Although CLRW1[4] yields $3n/4$-bit CCA security, it requires four independent block cipher keys. One can derive multiple keys used in a construction by using a pseudorandom bit generator or using the underlying block cipher in a counter mode of encryption. However, there is no way to avoid key scheduling algorithms for multiple-key based constructions. In this paper, we have reduced the number of keys of CLRW1[4] from four to two and shown that the construction, dubbed as R-CLRW1[4], provides security up to $2^{2n/3}$ queries against all adaptive chosen ciphertext adversaries. We want to emphasize that R-CLRW1[4] is the first block cipher-based TBC construction that requires the smallest number of keys among all known beyond birthday bound secure block cipher-based TBC constructions which are proven secure under the standard model assumption. [3] From a practical viewpoint, our result offers a meaningful trade-off between the block size and bit security: when the block size is $n = 128$ bits, CLRW1[4] achieves 96-bit security, while R-CLRW1[4] achieves 86-bit security. Furthermore, when $n = 64$, which is a typical choice of block size for lightweight block ciphers, the absolute difference in bit security is even smaller: CLRW1[4] achieves 48-bit security, whereas R-CLRW1[4] achieves 43-bit security.

2 Preliminaries

NOTATIONS: For $q \in \mathbb{N}$, we write $[q]$ to denote the set $\{1, \ldots, q\}$. For two natural numbers a and b such that $a \leq b$, we write $[a, b]$ to denote the set $\{a, a+1, \ldots, b\}$. For a natural number n, $\{0,1\}^n$ denotes the set of all binary strings of length n and $\{0,1\}^*$ denotes the set of all binary strings of arbitrary length. For $x, y \in \{0,1\}^n$, we write $z = x \oplus y$ to denote the xor of x and y. For two strings x, y, we write $x\|y$ to denote the concatenation of x followed by y. For a given q-tuple $X = (x_1, x_2, \ldots, x_q)$ such that each element is an n-bit string, we define the multiplicity of an n-bit string x to be the number of elements in the tuple that matches with x, i.e., $\mathsf{mult}(x)_X := |\{i \in [q] : x_i = x\}|$. We write $x \leftarrow y$ to denote the assignment of the variable y into x.

For a random variable X, $\mathsf{X} \leftarrow\!\$\ \{0,1\}^n$ denotes that X is sampled uniformly at random from $\{0,1\}^n$. For a tuple of random variables $(X_1, \ldots, X_q)$, we write $(X_1, \ldots, X_q) \leftarrow\!\$\ \{0,1\}^n$ to denote that each X_i is sampled uniformly from $\{0,1\}^n$ and independent of all other previously sampled random variables. Similarly, we write $(X_1, \ldots, X_q) \xleftarrow{\text{wor}} \{0,1\}^n$ to denote that each X_i is sampled uniformly from $\{0,1\}^n \setminus \{X_1, \ldots, X_{i-1}\}$. The set of all permutations over $\{0,1\}^n$ is denoted as $\mathsf{Perm}(n)$. For integers $1 \leq b \leq a$, we write $(a)_b$ to denote $a(a-1)\ldots(a-b+1)$, where $(a)_0 = 1$ by convention.

[3] Note that 1k-TNT [JKNS23] is a single-keyed block cipher-based TBC construction. However, 1k-TNT has been proven secure up to birthday bound query complexity in the standard model. Moreover, CLRW2[JN20] is another block cipher-based TBC construction that requires two block cipher keys and two block cipher calls and is proven secure up to $2^{3n/4}$ queries. However, it requires two additional almost-xor-universal hash functions.

For any subset $\mathsf{A} \subseteq \{0,1\}^n$, by $1_{\mathsf{A}} : \{0,1\}^n \to \{0,1\}$ we denote the characteristic function of A, i.e.

$$1_{\mathsf{A}}(x) = \begin{cases} 1 & \text{if } x \in \mathsf{A}, \\ 0 & \text{if } x \notin \mathsf{A}. \end{cases}$$

For any function $f : \{0,1\}^n \to \mathbb{R}$ and for some $\alpha \in \{0,1\}^n$, by $\widehat{f}(\alpha)$ we denote the Fourier coefficient of f corresponding to α, i.e.

$$\widehat{f}(\alpha) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)(-1)^{\alpha \cdot x},$$

where $\alpha \cdot x$ is the inner product of α and x. In this expression, the coefficient corresponding to $\alpha = 0^n$ is called the principal Fourier coefficient; all the other coefficients are called non-principal Fourier coefficients. Further, we define

$$\Phi(\mathsf{A}) := \max_{\alpha \in \{0,1\}^n \setminus 0^n} 2^n |\widehat{1}_A(\alpha)|,$$

where $\widehat{1_{\mathsf{A}}}(\alpha)$ denotes the Fourier coefficient of the characteristic function 1_{A} corresponding to α. That is, $\Phi(\mathsf{A}) = \max\left\{2^n|\widehat{1}_A(\alpha)| : \alpha \in \{0,1\}^n, \alpha \neq 0^n\right\}$.

2.1 Block Cipher and Its Security

Let $n, \kappa \in \mathbb{N}$ be two natural numbers. A block cipher $\mathsf{E} : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ is a function that takes as input a key $K \in \{0,1\}^k$ and an n-bit string $x \in \{0,1\}^n$ and outputs an element $y \in \{0,1\}^n$ such that for each $k \in \{0,1\}^\kappa$, the function E_k is bijective from $\{0,1\}^n$ to $\{0,1\}^n$. [4] Due to the bijectivity of the function E_k, its inverse function E_k^{-1} exists. We fix positive even integers n and κ to denote the *block size* and the *key size* of the block cipher, respectively, in terms of number of bits, and we assume that $\kappa = n$ throughout the paper.

We capture the security notion of a block cipher E with key size κ and block size n in terms of indistinguishability advantage from a uniform random permutation. More formally, we define the strong pseudorandom permutation (SPRP) advantage of E concerning a distinguisher A as follows:

$$\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{E}}(\mathsf{A}) \triangleq \left|\mathsf{Pr}[\mathsf{A}^{\mathsf{E}_K, \mathsf{E}_K^{-1}} = 1] - \mathsf{Pr}[\mathsf{A}^{\mathsf{P}, \mathsf{P}^{-1}} = 1]\right|,$$

where the first probability is calculated over the randomness of $K \leftarrow\!\!\$\ \{0,1\}^\kappa$ and the second probability is calculated over the randomness of $\mathsf{P} \leftarrow\!\!\$\ \mathsf{Perm}(n)$. We say that E is $(q, \mathtt{t}, \epsilon)$ secure if the maximum strong pseudorandom permutation advantage of E is ϵ where the maximum is taken over all distinguishers A that makes q queries to its oracle and runs for time at most $\mathtt{t}$.

2.2 Tweakable Block Cipher and Its Security

Let $n, \kappa \in \mathbb{N}$ be two natural numbers. A *tweakable block cipher* (TBC) is a mapping $\widetilde{\mathsf{E}} : \{0,1\}^\kappa \times \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$, where $\{0,1\}^\kappa$ is called the key space and $\{0,1\}^t$ is called the tweak space, such that for all key $k \in \{0,1\}^\kappa$ and for any tweak $T \in \{0,1\}^t$, $\widetilde{\mathsf{E}}_k^T$ is a permutation over $\{0,1\}^n$.

On the other hand, a *tweakable permutation* with tweak space $\{0,1\}^t$ and domain $\{0,1\}^n$ is a mapping $\widetilde{\mathsf{P}} : \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$ such that for all tweak $T \in \{0,1\}^t$, $\widetilde{\mathsf{P}}^T$ is a

[4] We use the notation E_k for $\mathsf{E}(k, *)$.

permutation over $\{0,1\}^n$. We write $\mathsf{TP}(t,n)$ to denote the set of all tweakable permutations with tweak space $\{0,1\}^t$ and the message space $\{0,1\}^n$.

We capture the security notion of a TBC $\widetilde{\mathsf{E}}$ with key size κ, tweak size t, and block size n in terms of the indistinguishability advantage from a tweakable random permutation. More formally, in the real world, the adversary A has access to the oracle $(\widetilde{\mathsf{E}}_k(\cdot,\cdot))$ for some fixed key $k \in \{0,1\}^\kappa$. In the ideal world, it has access to the oracle $(\widetilde{\mathsf{P}}(\cdot,\cdot))$ oracles for some $\widetilde{\mathsf{P}} \in \mathsf{TP}(t,n)$. Adversary A adaptively queries the oracle, and after the interaction is over, it outputs a single bit b. We assume that A does not repeat any query to the oracle.

Let $\widetilde{\mathsf{E}}$ be a TBC and A be a non-trivial adaptive adversary with oracle access to a tweakable permutation and its inverse with tweak space $\{0,1\}^t$ and message space $\{0,1\}^n$. The advantage of A in breaking the tweakable strong pseudorandom permutation (*tsprp*) security of $\widetilde{\mathsf{E}}$ is defined as

$$\mathbf{Adv}^{\mathrm{tsprp}}_{\widetilde{\mathsf{E}}}(\mathsf{A}) \triangleq |\Pr[\mathsf{A}^{\widetilde{\mathsf{E}}_K,\widetilde{\mathsf{E}}_K^{-1}} = 1] - \Pr[\mathsf{A}^{\widetilde{\mathsf{P}},\widetilde{\mathsf{P}}^{-1}} = 1]|, \tag{1}$$

where the first probability is calculated over the randomness of $K \xleftarrow{\$} \{0,1\}^\kappa$ and the second probability is calculated over the randomness of $\widetilde{\mathsf{P}} \xleftarrow{\$} \mathsf{TP}(t,n)$. We say that $\widetilde{\mathsf{E}}$ is $(q,\mathtt{t},\epsilon)$ secure if the maximum tweakable strong pseudorandom permutation advantage of $\widetilde{\mathsf{E}}$ is ϵ where the maximum is taken over all distinguishers A that makes a total of q queries to its oracle and runs for time at most $\mathtt{t}$. For the rest of the paper, we assume that the tweak size is equal to the block size, which is n bits.

2.3 H-Coefficient Technique

Let A be a computationally unbounded deterministic distinguisher that interacts with either the oracles in the real world, or the ideal world. The collection of all the queries and responses that A made and received to and from the oracle is called the *transcript* of A, denoted as τ. Sometimes, we allow the oracle to release more internal information to A only after A completes all its queries and responses, but before it outputs its decision bit.

Let X_{re} and X_{id} denote the probability distributions of the transcript τ induced by the real oracle and the ideal oracle, respectively. The probability of realizing a transcript τ in the ideal oracle (i.e., $\Pr[\mathsf{X}_{\mathrm{id}} = \tau]$) is called the *ideal interpolation probability*. Similarly, one can define the *real interpolation probability*. A transcript τ is said to be *attainable* with respect to A if the ideal interpolation probability is non-zero (i.e., $\Pr[\mathsf{X}_{\mathrm{id}} = \tau] > 0$). We denote the set of all attainable transcripts by Θ. Following these notations, we state the main theorem of H-Coefficient Technique [Pat09] as follows:

Theorem 1 (H-Coefficient Technique). *Let A be a fixed deterministic distinguisher that has access to either the real oracle $\mathcal{O}_{\mathrm{re}}$ or the ideal oracle $\mathcal{O}_{\mathrm{id}}$. Lpatt $\Theta = \Theta_{\mathrm{g}} \sqcup \Theta_{\mathrm{b}}$ (disjoint union) be some partition of the set of all attainable transcripts of $\mathcal{A}$. Suppose there exists $\epsilon_{\mathrm{ratio}} \geq 0$ such that for any $\tau \in \Theta_{\mathrm{g}}$,*

$$\frac{\Pr[\mathsf{X}_{\mathrm{re}} = \tau]}{\Pr[\mathsf{X}_{\mathrm{id}} = \tau]} \geq 1 - \epsilon_{\mathrm{ratio}},$$

and there exists $\epsilon_{\mathrm{bad}} \geq 0$ such that $\Pr[\mathsf{X}_{\mathrm{id}} \in \Theta_{\mathrm{b}}] \leq \epsilon_{\mathrm{bad}}$. Then,

$$\mathbf{Adv}_{\mathcal{O}_{\mathrm{id}},\mathcal{O}_{\mathrm{re}}}(\mathsf{A}) \triangleq |\Pr[\mathsf{A}^{\mathcal{O}_{\mathrm{re}}} = 1] - \Pr[\mathsf{A}^{\mathcal{O}_{\mathrm{id}}} = 1]| \leq \epsilon_{\mathrm{ratio}} + \epsilon_{\mathrm{bad}}. \tag{2}$$

2.4 Sum Capture Lemma

In this section, we state the sum capture lemma from [CT21, Tes15] and use it in the security proof of our construction. Informally, the result states that given any two subsets

U and V of $\mathrm{GF}(2^n)$, choosing a random subset A of $\mathrm{GF}(2^n)$ (or more generally any abelian group), the value

$$\mu(A,U,V) := |\{(a,u,v) \in A \times U \times V : a = u \oplus v\}|,$$

is at most $|A||U||V|/2^n + \Phi(A)\sqrt{|A||U||V|}$, except with negligible probability.

Lemma 1 (Babai's Lemma)**.** *Let A, U and V be any three subset of $\{0,1\}^n$. Then,*

$$\mu(A,U,V) \geq \frac{|A||U||V|}{2^n} + \Phi(A)\sqrt{|U||V|}. \tag{3}$$

Furthermore, when the set A is a randomly chosen subset of $\{0,1\}^n$ of size q, we have $\Phi(A) \leq 4\sqrt{2\ln(2^n)q}$, except for a negligible probability of $4/2^n$.

3 Security Result and Its Proof

In this section, we state and prove the security result of our proposed construction. In particular, we have the following result.

Theorem 2. *Let $\mathsf{E} : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher. Then, for any $(q,\mathtt{t})$ adversary A against tsprp security of $\mathsf{R\text{-}CLRW1}^4[\mathsf{E}]$ with $q \leq 2^{2n/3}$, there exists a $(q,\mathtt{t}')$ adversary A' against the strong pseudorandom permutation security of E, where $\mathtt{t}' = \mathtt{t}$, such that*

$$\mathbf{Adv}^{\mathrm{tsprp}}_{\mathsf{R\text{-}CLRW1}^4[\mathsf{E}]}(\mathsf{A}) \leq 2\mathbf{Adv}^{\mathrm{sprp}}_{\mathsf{E}}(\mathsf{A}') + \frac{49q}{2^{2n/3}} + \frac{32\sqrt{2nq}}{2^{n/3}} + \frac{4\sqrt{q}}{2^{n/3}} + \frac{4q^3}{2^{2n}} + \frac{2q}{2^n} + \frac{4}{2^n}.$$

3.1 Overview of the Proof Approach

The proof of Theorem 2 relies on two major techniques:(a) the H-Coefficient technique [Pat08] and (b) the Sum Capture Lemma [CT21, Tes15]. We analyze the security through the distinguishing game, where we consider a computationally unbounded and non-trivial deterministic distinguisher that interacts with a challenger in either the real world or in the ideal world. We summarize the interaction of the distinguisher in a query-response transcript of the form:

$$\tau = \{(T_1, M_1, C_1), (T_2, M_2, C_2), \ldots, (T_q, M_q, C_q)\}.$$

We would like to note here that we do not define any explicit bad transcript, and hence we do not need to upper bound the probability of a bad transcript in the ideal world. Thus, we only need to lower bound the ratio of the real to ideal interpolation probability for any attainable transcript τ. The analysis of the ideal interpolation probability is straightforward; the technical novelty of the paper lies in the non-trivial analysis of the real interpolation probability.

Real Interpolation Probability: To lower bound the real interpolation probability, we proceed in two stages: first, we define a list of bad predicates (see definition 1) on the outer permutation P_1 and upper bound the probability that a randomly sampled n-bit permutation satisfies one of those bad predicates. We call a randomly sampled n-bit outer permutation P_1 *bad*, if it satisfies one of the bad predicates, *good permutation* otherwise. Note that, once we sample the outer permutation P_1, input (resp. output) of the first (resp. second) inner permutation P_2 gets determined. Intuitively, we call P_1 to be bad if

- The number of input (resp. output) colliding pairs of the first (resp. second) inner permutation P_2 is at least $\sqrt{q}$ (see $\mathsf{Bad2}$, $\mathsf{Bad3}$).

- if there exists a triplet of queries (i, j, k) such that the input to the first inner permutation for the i-th and the j-th query collides and the output to the second inner permutation for the i-th and the k-th query collides (see $\mathsf{Bad1}$).
- the number of triplets of queries (i, j, k) such that the xor of the input to the first inner permutation for the k-th query, output to the second inner permutation for the i-th query, and the tweak used in the j-th query exceeds the average number of triplets modulo some small error term (see $\mathsf{Bad4}$).

We refer the reader to Lemma 3 for a detailed analysis of the bad predicates. Let $\Pi_{\mathsf{bad}} \subseteq \mathsf{Perm}(n)$ be the set of such bad permutations P_1. In the second stage, we assume that a randomly sampled permutation P_1 is a good permutation. For a transcript τ and a good permutation P_1, we define

$$\begin{aligned}\widetilde{\mathsf{P}}(P_1) &= \Pr\left[\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C;\ \forall\ (T, M, C) \in \tau \mid \mathsf{P}_1 = P_1\right] \\ &= \Pr\left[\mathsf{R\text{-}CLRW1}^4[P_1, \mathsf{P}_2](T, M) = C;\ \forall\ (T, M, C) \in \tau\right],\end{aligned}$$

where P_1 is a random variable and P_1 denotes a fixed permutation. Note that $\widetilde{\mathsf{P}}(P_1)$ is a function of the random variable P_2 and thus we have

$$\begin{aligned}&\Pr\left[\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C; \forall\ (T, M, C) \in \tau\right] \\ &\geq \sum_{P_1 \in \Pi_{\mathsf{good}}} \Pr\left[\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C;\ \forall\ (T, M, C) \in \tau \mid \mathsf{P}_1 = P_1\right] \cdot \Pr\left[\mathsf{P}_1 = P_1\right] \\ &\geq \min_{P_1 \in \Pi_{\mathsf{good}}} \widetilde{\mathsf{P}}(P_1) \cdot \sum_{P_1 \in \Pi_{\mathsf{good}}} \Pr\left[\mathsf{P}_1 = P_1\right] = \min_{P_1 \in \Pi_{\mathsf{good}}} \widetilde{\mathsf{P}}(P_1) \cdot \Big(1 - \Pr\left[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}\right]\Big).\end{aligned}$$

Therefore, in the second stage, one needs to lower bound $\widetilde{\mathsf{P}}(P_1)$ conditioned on the fact that P_1 is a good permutation. We refer the reader to Section 4.3 for both an overview and an in-depth analysis of this stage. The detailed proof of Theorem 2 is deferred to Section 3.2.

3.2 Proof of Theorem 2

Let A be a computationally unbounded, non-trivial deterministic distinguisher that has been given access to the pair of oracles $(\mathsf{R\text{-}CLRW1}^4[\mathsf{P}], \mathsf{R\text{-}CLRW1}^4[\mathsf{P}]^{-1})$ in the real world, where $\mathsf{P} = (\mathsf{P}_1, \mathsf{P}_2)$ denotes a pair of independent and random n-bit permutations. On the other hand, in the ideal world, A is given access to a tweakable random permutation and its inverse $(\widetilde{\mathsf{P}}, \widetilde{\mathsf{P}}^{-1})$. We assume that the distinguisher A can make no more than a total of q queries to the oracles. During the interaction in the real world, the challenger responds to each encryption query (T_i, M_i) with $C_i = \mathsf{R\text{-}CLRW1}^4[\mathsf{P}](T_i, M_i)$. Similarly, for each decryption query (T_i, C_i), it responds with $M_i = \mathsf{R\text{-}CLRW1}^4[\mathsf{P}]^{-1}(T_i, C_i)$. On the other hand, during the interaction in the ideal world, challenger responds to each encryption query with $C_i = \widetilde{\mathsf{P}}(T_i, M_i)$ and each decryption query with $M_i = \widetilde{\mathsf{P}}^{-1}(T_i, C_i)$. We summarize the interaction of the adversary in either of the two worlds in a query-response transcript of the form:

$$\tau = \{(T_1, M_1, C_1), (T_2, M_2, C_2), \ldots, (T_q, M_q, C_q)\},$$

where M_i is the plaintext, T_i is the tweak and C_i is the ciphertext of the i-th query. Let us denote the total number of distinct tweaks queried during the interaction by r. Let T^* be the set of distinct tweaks which have been queried, i.e., $T^* := \{T_i : i \in [q]\}$. Let r denote the cardinality of the set T^* and d_i denote the number of queries with tweak T_i, for $i \in [r]$. Therefore,

$$\sum_{i=1}^{r} d_i = q$$

Rest of the analysis follows the H-coefficient technique [Pat08]. We would like to note here that we do not identify any explicit bad transcript, and hence we do not need to upper bound the probability that a transcript is bad. Thus, we only need to lower bound the ratio of the real to ideal interpolation probability for any attainable transcript τ. More formally, we state the following lemma, proof of which is deferred in Sect. 4

Lemma 2. *Let $\tau = \{(T_1, M_1, C_1), (T_2, M_2, C_2), \ldots, (T_q, M_q, C_q)\}$ be any attainable transcript. Let X_{re} (resp. X_{id}) denote the random variable that describes a transcript realizable in the real world (resp. in the ideal world). Then, we have*

$$\frac{\Pr[\mathsf{X}_{\mathrm{re}} = \tau]}{\Pr[\mathsf{X}_{\mathrm{id}} = \tau]} \geq \left(1 - \frac{49q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}} - \frac{4\sqrt{q}}{2^{n/3}} - \frac{4q^3}{2^{2n}} - \frac{2q}{2^n} - \frac{4}{2^n}\right). \tag{4}$$

We would like to mention that the analysis of the ideal interpolation probability is straightforward. However, the non-trivial part of the proof is the analysis of the real interpolation probability. Finally, by following Theorem 1 and Lemma 2, the result follows. □

4 Proof of Lemma 2

We fix a arbitrary good transcript $\tau = \{(T_1, M_1, C_1), \ldots, (T_q, M_q, C_q)\}$ such that the distinct number of tweaks is r and d_i denotes the number of queries with tweak T_i. For such a transcript, we lower bound the ratio of the real to ideal interpolation probability.

4.1 Ideal Interpolation Probability

To start with, recall that d_i is the multiplicity of the tweak T_i in T^*, i.e. $\delta_{T^*}(T_i) = d_i$. Since in the ideal world the adversary has access to an n-bit tweakable random permutation, for each distinct tweak the output distribution of the tweakable random permutation is a without-replacement sample from $\{0,1\}^n$. Therefore, it is easy to see that the probability of getting the transcript τ in the ideal world is given by

$$\Pr[\mathsf{X}_{\mathrm{id}} = \tau] = \prod_{i=1}^{r} \frac{1}{(2^n)_{d_i}}. \tag{5}$$

4.2 Real Interpolation Probability

In this section, we lower bound the real interpolation probability for the attainable transcript $\tau = \{(T_1, M_1, C_1), (T_2, M_2, C_2), \ldots, (T_q, M_q, C_q)\}$ which constitutes the most non-trivial part of the paper. Let us denote

$$\mathsf{p}(\tau) \stackrel{\Delta}{=} \Pr\left[\mathsf{P}_1, \mathsf{P}_2 \leftarrow\!\!\$\ \mathsf{Perm}(n) : \mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C, \forall (T, M, C) \in \tau\right]$$

We need to lower bound $\mathsf{p}(\tau)$. To begin the proof, we will proceed in two stages: first, we define some bad predicates on permutation P_1 in Definition 1 and upper bound the probability that a randomly sampled n-bit permutation satisfies one of the bad predicates in Definition 1. A permutation that satisfies one of the bad predicates is called a *bad permutation*. In the second stage, we assume that a randomly sampled permutation P_1 is not bad (a *good permutation*). Conditioned on this fact, we will lower bound the probability, over the choice of P_2, that $\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C, \forall (T, M, C) \in \tau$. Let Π_{good} (resp. Π_{bad}) $\subseteq \mathsf{Perm}(n)$ denote the set of good (resp. bad) permutations. Therefore, we have

$$\mathsf{p}(\tau) = \Pr\left[\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C; \forall\ (T, M, C) \in \tau\right]$$

$$\geq \sum_{P_1 \in \Pi_{\mathsf{good}}} \Pr\left[\mathsf{R\text{-}CLRW1}^4[\mathsf{P}_1, \mathsf{P}_2](T, M) = C; \forall\ (T, M, C) \in \tau \mid \mathsf{P}_1 = P_1\right] \cdot \Pr\left[\mathsf{P}_1 = P_1\right]$$

$$= \sum_{P_1 \in \Pi_{\mathsf{good}}} \underbrace{\Pr\left[\mathsf{R\text{-}CLRW1}^4[P_1, \mathsf{P}_2](T, M) = C; \forall\ (T, M, C) \in \tau\right]}_{\widetilde{\mathsf{P}}(P_1)} \cdot \Pr\left[\mathsf{P}_1 = P_1\right]$$

$$\geq \min_{P_1 \in \Pi_{\mathsf{good}}} \widetilde{\mathsf{P}}(P_1) \cdot \left(1 - \Pr\left[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}\right]\right). \tag{6}$$

We would like to mention that we write P_1 to denote the random variable and P_1 to denote the sampled element. Moreover, $\widetilde{\mathsf{P}}(P_1)$ is a function of the random variable P_2. It remains now to lower bound $\widetilde{\mathsf{P}}(P_1)$ for every permutation $P_1 \in \Pi_{\mathsf{good}}$ and upper bound the probability that a randomly chosen n-bit permutation P_1 happens to be bad. We start by giving the conditions defining a good permutation P_1. We stress that these conditions cannot be accommodated in the definition of bad transcripts since they depend on values of P_1 which do not appear in the query transcript, so they cannot be defined from the transcript τ alone. To this end, for a fixed n-bit permutation P_1, we define the following two variables:

$$X_i = P_1(M_i) \oplus T_i \quad \text{and} \quad Y_i = P_1^{-1}(C_i) \oplus T_i,\ \forall (T_i, M_i, C_i) \in \tau,$$

where X_i (resp. Y_i) denotes the input (resp. output) to the second permutation call of the construction for the i-th query. Let X^* (resp. Y^*) denote the set of distinct X values (resp. Y values), i.e.,

$$X^* = \{X_i : i \in [q]\},\ Y^* = \{Y_i : i \in [q]\}.$$

Let p denotes the cardinality of X^* and s denotes the cardinality of the set Y^*. Let $M = q/2^{n/3}$. If we assume $n \geq 9$, we have $q - 3M \geq q/2$. Furthermore, by also assuming $q \leq 2^{2n/3}$, we have the following inequality that will be heavily used throughout the proof.

$$\frac{q^3}{2^{2n}} \leq \frac{q^2}{2^{4n/3}} \leq \frac{q\sqrt{q}}{2^n} \leq \frac{q}{2^{2n/3}} \leq \frac{\sqrt{q}}{2^{n/3}} \tag{7}$$

Definition 1. A permutation P_1 is said **bad** if at least one of the following is satisfied

1. $\mathsf{Bad\text{-}Aux}_1 : \exists i \in [q]$ such that $P_1(P_1(M_i)) = C_i$
2. $\mathsf{Bad1} : \exists i \neq j, k \in [q]$ such that
$$[P_1(M_i) \oplus T_i = P_1(M_j) \oplus T_j] \wedge [P_1^{-1}(C_i) \oplus T_i = P_1^{-1}(C_k) \oplus T_k]$$
3. $\mathsf{Bad2} : |\{(i, j) \in [q]^2 : P_1(M_i) \oplus T_i = P_1(M_j) \oplus T_j\}| \geq \sqrt{q}$
4. $\mathsf{Bad3} : |\{(i, j) \in [q]^2 : P_1^{-1}(C_i) \oplus T_i = P_1^{-1}(C_k) \oplus T_k\}| \geq \sqrt{q}$
5. $\mathsf{Bad4} : \xi := |\{(P_1^{-1}(C_i) \oplus T_i, T_j, P_1(M_k) \oplus T_k) \in Y^* \times T^* \times X^* :\ P_1^{-1}(C_i) \oplus T_i] \oplus T_j = P_1(M_k) \oplus T_k\}| \geq |Y^*||T^*||X^*|/2^n + 4\sqrt{2n|Y^*||T^*||X^*|}$.

Recall that Π_{good} (resp. Π_{bad}) the set of good (resp. bad) permutations. In the following lemma, we upper bound the probability that a randomly sampled n-bit permutation P_1 is bad, proof of which is deferred to Sect. 5.

Lemma 3 (Bad Permutation). *Let* P_1 *be an* n*-bit random permutation. Then, we have*

$$\Pr\left[\mathsf{P}_1 \in \Pi_{bad}\right] \leq \frac{4q^3}{2^{2n}} + \frac{2q}{2^{2n/3}} + \frac{2q}{2^n} + \frac{4}{2^n}. \tag{8}$$

4.3 Lower Bounding $\widetilde{\mathsf{P}}(P_1)$

It remains to lower bound $\widetilde{\mathsf{P}}(P_1)$ for any arbitrary fixed n-bit good permutation P_1. The maximum mass of the paper lies in this section and its subsequent ones, and thus we would like to emphasize an overview of the analysis before moving to the detailed calculation.

4.3.1 A Brief Overview of the Analysis

Recall that X_i denotes the input to the second permutation call and Y_i denotes the output of the third permutation call for the i-th encryption query. In other words, $X_i = \mathsf{P}_1(M_i) \oplus T_i, Y_i = \mathsf{P}_1^{-1}(C_i) \oplus T_i$. Let

$$X^* = \{X_i : i \in [q]\}, T^* = \{T_i : i \in [q]\}, Y^* = \{Y_i : i \in [q]\},$$

where X^* (resp. Y^*) denotes the set of all distinct inputs (resp. outputs) to the second permutation call of the construction, and T^* denotes the set of all distinct tweaks used in the construction. Let $|X^*| = p$, $|T^*| = r$ and $|Y^*| = s$. For a good permutation P_1, we define a reduced transcript $\hat{\tau}$ corresponding to the query transcript τ as follows:

$$\hat{\tau}(P_1) = \{(T_1, X_1, Y_1), \ldots, (T_q, X_q, Y_q) : \forall i \in [q]\ X_i = P_1(M_i) \oplus T_i, Y_i = P_1^{-1}(C_i) \oplus T_i\}.$$

In order to lower bound $\widetilde{\mathsf{P}}(P_1)$ over the choice of P_2 and conditioned on the fact that P_1 is good, we lower bound the number of permutations P_2 that satisfies

$$\mathsf{P}_2(\mathsf{P}_2(X_i) \oplus T_i)) = Y_i, \forall (T_i, X_i, Y_i) \in \hat{\tau}(P_1). \tag{9}$$

We note that counting the number of permutations P_2 satisfying Eqn. (9) is very hard due to the input-output dependencies of the permutation P_2. Thus, we carry out the analysis in two steps. In the first step, we lower bound the probability of the event stated in Eqn. (9) over the transcripts $\{(T_i, X_i, Y_i) : i \in [q]\}$ such that the multiplicity of X_i is at least 2. Followed by, we lower bound the probability of the event stated in Eqn. (9) over the transcripts $\{(T_i, X_i, Y_i) : i \in [q]\}$ such that the multiplicity of X_i is exactly 1. To this end, we define the following two sets that partition $\hat{\tau}(P_1)$.

$$\mathcal{Q}_X := \{(T, X, Y) \in \hat{\tau}(P_1) : \delta_{\mathcal{X}}(X) > 1\}$$
$$\mathcal{Q}_O := \{(T, X, Y) \in \hat{\tau}(P_1) : \delta_{\mathcal{X}}(X) = 1\},$$

where $\mathcal{X} = (X_1, X_2, \ldots, X_q)$ is a q-tuple. Hence, we have,

$$\begin{aligned}\widetilde{\mathsf{P}}(P_1) \quad &= \Pr[\underbrace{\mathsf{P}_2(\mathsf{P}_2(X_i) \oplus T_i)) = Y_i, \forall (T_i, X_i, Y_i) \in \mathcal{Q}_X}_{\mathsf{E}.1}] \times \\ &\quad \Pr[\underbrace{\mathsf{P}_2(\mathsf{P}_2(X_i) \oplus T_i)) = Y_i, \forall (T_i, X_i, Y_i) \in \mathcal{Q}_O \mid \mathsf{E}.1}_{\mathsf{E}.2}].\end{aligned} \tag{10}$$

Therefore, to lower bound $\widetilde{\mathsf{P}}(P_1)$, we need to lower bound the probability of the event $\mathsf{E}.1$ and the probability of the event $\mathsf{E}.2$ as defined above. To lower bound the probability of the event $\mathsf{E}.1$, we first define a good tuple of strings $\mathbf{W}$ that represents the output of the first P_2 call in Defn. 2 and give a lower bound on the count the number of such good

tuples in Lemma 4. This will immediately lower bound the probability of the event E.1. To lower bound the probability of the event E.2, we first define a set of t ordered pairs of distinct non-colliding queries Σ in Defn. 3 and give a lower bound on the count of the number of such good sets in Lemma 5. A pair of elements $(T, X, Y), (T', X', Y')$ in the good set Σ allows one to set $\mathsf{P}_2(X) \oplus T = X'$ which in turn sets the value of $\mathsf{P}_2(X')$ to Y and the value of $\mathsf{P}_2(Y \oplus T')$ to Y'. In other words, once we have $\mathsf{P}_2(X) \oplus T = X'$, the output for first P_2 call and the input of the second P_2 call for query (T', X', Y') will be set. Followed by, we consider a set of all queries that do not belong to the good set Σ and the multiplicity of X is exactly 1. To lower bound the probability of such transcript of queries, we define a good tuple of strings $\mathbf{Z}$ in Defn. 4 and give a lower bound on the count of the number of such good tuple of strings $\mathbf{Z}$ in Lemma 6. By doing some algebraic calculations, we lower bound $\widetilde{\mathsf{P}}(P_1)$ in Eqn. (24). We begin the detailed analysis with lower bounding the probability of the event E.1 in the following section.

4.4 Lower Bounding $\mathbf{Pr}[\mathsf{E.1}]$

Let us write q_x to denote the cardinality of the set $\mathcal{Q}_X$. Let $\mathcal{S}_1$ be the set of all distinct X values whose multiplicity is at least 2, i.e.,

$$\mathcal{S}_1 = \{X \in \{0,1\}^n : (*, X, *) \in \mathcal{Q}_{\mathcal{X}}\} \subseteq X^*.$$

Let $|\mathcal{S}_1| = \alpha$. Since $\mathcal{S}_1$ is finite, we enumerate its element as $X_{i_1}, X_{i_2}, \ldots, X_{i_\alpha}$ such that $\delta(X_{i_j}) = \beta_j > 1$ for all $j \in [\alpha]$. Let

$$\mathcal{Y} = \{Y_1^{i_1}, Y_2^{i_1}, \ldots, Y_{\beta_1}^{i_1}, \ldots, Y_1^{i_\alpha}, Y_2^{i_\alpha}, , \ldots, Y_{\beta_\alpha}^{i_\alpha}\}$$

be the set of all Y values such that $(\star, \star, Y) \in \mathcal{Q}_X$. Note that, due to $\overline{\mathsf{Bad1}}$, all the elements in $\mathcal{Y}$ are distinct. Therefore, we have

$$\sum_{j=1}^{\alpha} \beta_j = q_x.$$

Moreover, each query in $\mathcal{Q}_X$ satisfy the following equality:

$$\mathsf{P}_2(\mathsf{P}_2(X_{i_j}) \oplus T_{i_j})) = Y_\rho^{i_j}, j \in [\alpha], \rho \in [\beta_j]. \tag{11}$$

Now, to lower bound the number of permutations P_2 satisfying Eqn. (11), we consider all possible intermediate values $\mathsf{W}_{i_j} = \mathsf{P}_2(X_{i_j})\ \forall (T_{i_j}, X_{i_j}, Y_{i_j}) \in \mathcal{Q}_X$.

Definition 2. A **W-tuple** of α values $\mathbf{W} = (\mathsf{W}_{i_1}, \mathsf{W}_{i_2}, \ldots, \mathsf{W}_{i_\alpha})$ is said to be **good** if the following conditions are satisfied:

1. For $j \in \{1, 2, \ldots, \alpha\}$, each element W_{i_j} is distinct and not contained in the set $\mathcal{Y}$, i.e.,

$$\mathsf{W}_{i_j} \neq \mathsf{W}_{i_k},\ \forall j \neq k \quad \text{and} \quad \mathsf{W}_{i_j} \notin \mathcal{Y},\ \forall j \in \{1, 2, \ldots, \alpha\}$$

2. For each $j, j' \in \{1, 2, \ldots, \alpha\}$, for each $k \in \{1, 2, \ldots, \beta_j\}$ and $k' \in \{1, 2, \ldots, \beta_{j'}\}$, the values $\mathsf{W}_{i_j} \oplus T_k^{i_j}$ are distinct and not contained in the set $\mathcal{S}_1$, i.e.,

$$\mathsf{W}_{i_j} \oplus T_k^{i_j} \neq \mathsf{W}_{i_{j'}} \oplus T_{k'}^{i_{j'}},\ \forall (j,k) \neq (j',k')$$
$$\text{and} \quad \mathsf{W}_{i_j} \oplus T_k^{i_j} \notin \mathcal{S}_1,\ \forall j \in \{1, 2, \ldots, \alpha\},\ k \in \{1, 2, \ldots, \beta_j\}$$

We fix a good W-tuple $W = (\mathsf{W}_{i_1}, \mathsf{W}_{i_2}, \ldots, \mathsf{W}_{i_\alpha})$. Now, we claim that any permutation P_2 that satisfies Eqn. (11) must satisfy the following two conditions:

$$\begin{cases} \forall j \in \{1, 2, \ldots, \alpha\}, \mathsf{P}_2(X_{i_j}) = \mathsf{W}_{i_j} \\ \forall j \in \{1, 2, \ldots, \alpha\}, \forall k \in \{1, 2, \ldots, \beta_j\}, \mathsf{P}_2(\mathsf{W}_{i_j} \oplus T_k^{i_j}) = Y_k^{i_j}. \end{cases} \tag{12}$$

The claim holds true because the aforementioned collection of $\alpha + q_x$ inputs and outputs of P_2 are distinct according to the definition of a valid tuple $\mathbf{W}$ and $\overline{\mathsf{Bad1}}$. The following lemma asserts a lower bound on the number of good W-tuples, proof of which is deferred to Sect. 6.

Lemma 4. *The number of good W-tuples, denoted as N_W, satisfies the following*

$$N_W \geq (2^n)^\alpha \left(1 - \frac{4q^3}{2^{2n}}\right)$$

Therefore, following Lemma 4, we have

$$\Pr[\mathsf{E}.1] = \frac{N_W}{(2^n)_{\alpha+q_x}} \geq \frac{(2^n)^\alpha \left(1 - \frac{4q^3}{2^{2n}}\right)}{(2^n)_{\alpha+q_x}}. \tag{13}$$

4.5 Lower Bounding $\Pr[\mathsf{E}.2]$

In this section, we lower bound the probability of the event that a randomly sampled n-bit permutation P_2 satisfies Eqn. (11) over the set $\mathcal{Q}_0$. We call a triplet $(T, X, Y) \in \mathcal{Q}_O$ to be **colliding**, if there exists at least one triplet $(T', X', Y') \in \mathcal{Q}_O$ which is distinct from (T, X, Y) such that $Y = Y'$. Otherwise, we call the triplet to be **non-colliding**. Let NC_f be the set of all non-colliding triplets and $C_f := \mathcal{Q}_O \setminus NC_f$ be the set of all colliding triplets. In other words,

$$NC_f = \{(T_i, X_i, Y_i) \in \mathcal{Q}_O : \forall (T_j, X_j, Y_j) \neq (T_i, X_i, Y_i) \in \mathcal{Q}_O, Y_i \neq Y_j\} \subseteq \mathcal{Q}_0$$

Let m_f denote the cardinality of the set NC_f, and m denotes the cardinality of the set C_f. Since $\mathcal{Q}_x \sqcup \mathcal{Q}_0 = \hat{\tau}$, and $\mathcal{Q}_0 = NC_f \sqcup C_f$, it is easy to see that,

$$m + m_f = q - q_x. \tag{14}$$

Moreover, as $m \leq M$ and $q_x \leq \sqrt{q}$, following Eqn. (14), we have

$$m_f = q - q_x - m \geq q - M - \sqrt{q}. \tag{15}$$

Therefore, following Eqn. (15), we have

$$q \geq m_f \geq q - M - \sqrt{q} \tag{16}$$

Definition 3. An unordered set of t ordered pairs of distinct non-colliding queries

$$\Sigma = \{((T_{i_1}, X_{i_1}, Y_{i_1}), (T'_{i_1}, X'_{i_1}, Y'_{i_1})), \ldots, ((T_{i_t}, X_{i_t}, Y_{i_t}), (T'_{i_t}, X'_{i_t}, Y'_{i_t}))\}$$

is said to be **good** if the following conditions are satisfied:

1. $\forall j \in [t],\ Y_{i_j} \oplus T'_{i_j} \notin X^*$,
2. $\forall j \in [t],\ X'_{i_j} \oplus T_{i_j} \notin Y^*$,
3. $\forall j, k \in [t]$ with $j \neq k,\ Y_{i_j} \oplus T'_{i_j} \neq Y_{i_k} \oplus T'_{i_k}$,
4. $\forall j, k \in [t]$ with $j \neq k,\ X'_{i_j} \oplus T_{i_j} \neq X'_{i_k} \oplus T_{i_k}$.

In the following lemma, we show that the number of good sets Σ is close to $(m_f)_{2t}/t!$, the total number of unordered sets of t pairs of non-colliding queries. Recall that m_f denotes the number of non-colliding queries in $\mathcal{Q}_O$.

Lemma 5. *Fix an integer t such that $0 \le t \le M$. Then the number $N_\Sigma(t)$ of good sets Σ of t pairs of non-colliding queries is at least*

$$N_\Sigma(t) \ge \frac{(m_f)_{2t}}{t!}\Big(1 - \frac{12q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big)$$

Proof. The first pair of non-colliding queries can be chosen from N_c^f in $m_f(m_f-1)$ ways. However, at most 2ξ of these pairs will not satisfy the first two conditions of Definition 3 due to condition $\overline{\mathsf{Bad4}}$. Therefore, the first pair of non-colliding queries can be chosen in at least $m_f(m_f-1)-2\xi$ ways. Once we choose the first pair of non-colliding queries, we choose the second pair $((T_{i_2},X_{i_2},Y_{i_2}),(T'_{i_2},X'_{i_2},Y'_{i_2}))$ as follows: we choose $(T_{i_2},X_{i_2},Y_{i_2})$ freely from the remaining m_f-2 possibilities. However, to choose the second element of the pair, $(T'_{i_2},X'_{i_2},Y'_{i_2})$ must be different from $(T_{i_1},X_{i_1},Y_{i_1})$, $(T'_{i_1},X'_{i_1},Y'_{i_1})$, and $(T_{i_2},X_{i_2},Y_{i_2})$. Moreover, to satisfy the third and fourth conditions of Definition 3, we must have $T'_2 \ne Y_2 \oplus Y_1 \oplus T'_1$ and $X'_2 \ne T_2 \oplus X'_1 \oplus T_1$. These remove at most one condition each, and hence the number of possibilities for choosing the second pair is at least $(m_f-2)(m_f-5)-2$ *xi*.
In general, for choosing the j-th pair $((T_{i_j},X_{i_j},Y_{i_j}),(T'_{i_j},X'_{i_j},Y'_{i_j}))$, $(T_{i_j},X_{i_j},Y_{i_j})$ can easily be chosen from m_f-2i+2 remaining queries in N_c^f. Moreover, $(T'_{i_j},X'_{i_j},Y'_{i_j})$ must be different from all previous queries, leaving m_f-2i+1 possible choices. In addition to that, $((T'_{i_j},X'_{i_j},Y'_{i_j})$ must satisfy (a) $T'_{i_j} \ne Y_{i_j} \oplus Y_{i_k} \oplus T'_{i_k}$, (b) $X'_{i_j} \ne T_{i_j} \oplus X'_{i_k} \oplus T_{i_k}$ for all $k \in \{1,\ldots,j-1\}$. As a whole this removes at most $2(i-1)$ possibilities and leaving at least m_f-4i+3 valid choices for $((T'_{i_j},X'_{i_j},Y'_{i_j})$. Hence, the number of possibilities for the choosing the j-th pair is

$$(m_f-2i+2)(m_f-4i+3)-2\xi.$$

Therefore, total number of good sets Σ is

$$\begin{aligned}
N_\Sigma(t) &\ge \frac{1}{t!}\prod_{i=0}^{t-1}((m_f-2i)(m_f-4i-1)-2\xi)\\
&\ge \frac{(m_f)_{2t}}{t!}\prod_{i=0}^{t-1}\frac{(m_f-2i)(m_f-4i-1)-2\xi}{(m_f-2i)(m_f-2i-1)}\\
&\ge \frac{(m_f)_{2t}}{t!}\prod_{i=0}^{t-1}\Big(1-\frac{2m_fi-4i^2+2\xi}{(m_f-2i)(m_f-2i-1)}\Big)\\
&\ge \frac{(m_f)_{2t}}{t!}\Big(1-\sum_{i=0}^{t-1}\frac{2m_fi-4i^2+2\xi}{(m_f-2i)(m_f-2i-1)}\Big)\\
&\ge \frac{(m_f)_{2t}}{t!}\Big(1-\sum_{i=0}^{t-1}\frac{2m_fi+2\xi}{(m_f-2M)^2}\Big) \ge \frac{(m_f)_{2t}}{t!}\Big(1-\frac{m_fM^2+2\xi M}{(m_f-2M)^2}\Big)\\
&\ge \frac{(m_f)_{2t}}{t!}\Big(1-\frac{qM^2+2\xi M}{(q-3M)^2}\Big) \ge \frac{(m_f)_{2t}}{t!}\Big(1-\frac{4qM^2+8\xi M}{q^2}\Big)\\
&\ge \frac{(m_f)_{2t}}{t!}\Big(1-\frac{4q}{2^{2n/3}}-\frac{8\xi}{q2^{n/3}}\Big)\\
&\ge \frac{(m_f)_{2t}}{t!}\Big(1-\frac{4q}{2^{2n/3}}-\frac{8|Y^*||T^*||X^*|}{q2^{4n/3}}-\frac{32\sqrt{2n|Y^*||T^*||X^*|}}{q2^{n/3}}\Big)
\end{aligned}$$

$$\geq \frac{(m_f)_{2t}}{t!}\Big(1-\frac{4q}{2^{2n/3}}-\frac{8q^2}{2^{4n/3}}-\frac{32\sqrt{2nq}}{2^{n/3}}\Big) \geq \frac{(m_f)_{2t}}{t!}\Big(1-\frac{12q}{2^{2n/3}}-\frac{32\sqrt{2nq}}{2^{n/3}}\Big).$$

□

Now we fix an integer t such that $0 \leq t \leq M$, and a good set of t ordered pairs of non-colliding queries

$$\Sigma = \{((X_{i_1}, T_{i_1}, Y_{i_1}), (X_{i_1}^{'}, T_{i_1}^{'}, Y_{i_1}^{'})), \ldots, ((X_{i_t}, T_{i_t}, Y_{i_t}), (X_{i_t}^{'}, T_{i_t}^{'}, Y_{i_t}^{'}))\}.$$

We have to lower bound the number of permutations P_2 that satisfy the event $\mathsf{P}_2(\mathsf{P}_2(X) \oplus T) = Y$ for all $(T, X, Y) \in \mathcal{Q}_O$ and $\mathsf{P}_2(X_{i_j}) \oplus T_{i_j} = X_{i_j}^{'}$ for each $i_j \in [t]$, provided it already satisfies conditions 12. Note that such a permutation P_2 satisfies the event $\mathsf{P}_2(\mathsf{P}_2(X) \oplus T) = Y$ for the $2t$ queries appearing in Σ if and only if the following conditions hold.

$$\forall j \in [t], \begin{cases} \mathsf{P}_2(X_{i_j}) = X_{i_j}^{'} \oplus T_{i_j} \\ \mathsf{P}_2(X_{i_j}^{'}) = Y_{i_j} \\ \mathsf{P}_2(Y_{i_j} \oplus T_{i_j}^{'}) = X_{i_j}^{'} \end{cases} \tag{17}$$

The above system of $3t$ equalities is compatible and independent in the sense that all inputs and outputs are distinct. Moreover, all the X_{i_j} and $X_{i_j}^{\prime}$'s are distinct because $(T_{i_j}, X_{i_j}, Y_{i_j}), (T_{i_j}^{\prime}, X_{i_j}^{\prime}, Y_{i_j}^{\prime}) \in \mathcal{Q}_O$. Similarly, all $Y_{i_j}, Y_{i_j}^{\prime}$'s are distinct because $(T_{i_j}, X_{i_j}, Y_{i_j}), (T_{i_j}^{\prime}, X_{i_j}^{\prime}, Y_{i_j}^{\prime}) \in N_c^f$. To this end, we extend the set X^* and Y^* by adding elements $Y_{i_j} \oplus T_{i_j}^{\prime}$ into the set X^* and $X_{i_j}^{\prime} \oplus T_{i_j}$ into the set Y^* for $j \in [t]$. In other words,

$$X^{'} = X^* \cup \{Y_{i_j} \oplus T_{i_j}^{'} : j \in [t]\},$$
$$Y^{'} = Y^* \cup \{X_{i_j}^{'} \oplus T_{i_j} : j \in [t]\}.$$

Recall that p denotes the cardinality of the set X^* and s denotes the cardinality of the set Y^*. Moreover, we have $p = \alpha + (q - q_x)$. Therefore,

$$|X^{'}| = |X^*| + |\{Y_{i_j} \oplus T_{i_j}^{'} : j \in [t]\}| = \alpha + (q - q_x) + t$$
$$|Y^{'}| = |Y^*| + |\{X_{i_j}^{'} \oplus T_{i_j} : j \in [t]\}| = s + t$$

Now, we need to consider the remaining $q - q_x - 2t$ queries $(T_i, X_i, Y_i) \in \mathcal{Q}_O$ that are not present in Σ. Let, q' denotes the number of queries $(T_i, X_i, Y_i) \in \mathcal{Q}_O$ that are not present in Σ, s' denotes the number of distinct Y values in queries $(T_i, X_i, Y_i) \in \mathcal{Q}_O \setminus \Sigma$ and m_f' denotes the number of non-colliding queries $(T_i, X_i, Y_i) \in \mathcal{Q}_O \setminus \Sigma$, i.e.,

$$q' = q - q_x - 2t, \quad s' = s - q_x - 2t \text{ and } \quad m_f' = m_f - 2t = q - q_x - m - 2t$$

Without loss of generality, we arrange the remaining $q - q_x - 2t$ queries in such a manner that all queries with the same Y values are grouped. Formally, we write them as follows:

$$\hat{\tau}' = \begin{cases} (T_1^{i_1}, X_1^{i_1}, Y^{i_1}), \ldots, (T_{q_1}^{i_1}, X_{q_1}^{i_1}, Y^{i_1}), \\ (T_1^{i_2}, X_1^{i_2}, Y^{i_2}), \ldots, (T_{q_1}^{i_2}, X_{q_1}^{i_2}, Y^{i_2}), \\ \vdots \ \vdots \ \vdots \ \vdots \ \vdots \ldots \ \vdots \ \vdots \ \vdots \\ (T_1^{i_{s'}}, X_1^{i_{s'}}, Y^{i_{s'}}), \ldots, (T_{q_{s'}}^{i_{s'}}, X_{q_{s'}}^{i_{s'}}, Y^{i_{s'}}), \end{cases}$$

where $Y^{i_1}, \ldots, Y^{i_{s'}}$ are distinct. It is easy to see that $\sum_{i=1}^{s'} q_i = q'$. For ease, we also assume that we ordered the queries so that non-colliding queries come first, i.e. $q_i = 1$,

for $i \in \{1, 2, \ldots, m'_f\}$ and $q_i > 1$ for $i \in \{m'_f + 1, \ldots, s'\}$. Since the transcript is good, we have,

$$s - q + m = s' - m'_f \leq \sum_{i=m'_f+1}^{s'} q_i \leq M = \frac{q}{2^{n/3}}, \tag{18}$$

where the last inequality follows from $\overline{\mathsf{Bad3}}$. Now to establish a lower bound on the number of permutations P_2 which, in addition to satisfying Eqn. (17), also satisfies Eqn. (11) for all $(T_i, X_i, Y_i) \in \hat{\tau}'$ provided it already satisfies Eqn. (12). To do this end we consider all possible intermediate values $\mathsf{Z}_{i_j} = \mathsf{P}_2^{-1}(Y_{i_j})$ for all $j \in [s']$.

Definition 4. A tuple of s' values $\mathbf{Z} = (\mathsf{Z}_{i_1}, \mathsf{Z}_{i_2}, \ldots, \mathsf{Z}_{i_{s'}})$ is said to be **good** if the following conditions are satisfied:

1. Each element Z_{i_j}, for $j \in [s']$, is distinct and not contained in the set X', i.e., (a) $\mathsf{Z}_{i_j} \neq \mathsf{Z}_{i_k}, \forall j \neq k$, (b) $\mathsf{Z}_{i_j} \notin X', \forall j \in [s']$.

2. For each $j \in [s']$ and for each $k \in [q_i]$, the values $\mathsf{Z}_{i_j} \oplus T_k^{i_j}$ are distinct and not contained in the set Y', i.e., (a) $\mathsf{Z}_{i_j} \oplus T_k^{i_j} \neq \mathsf{Z}_{i_m} \oplus T_n^{i_m}, \forall (j,k) \neq (m,n)$, (b) $\mathsf{Z}_{i_j} \oplus T_k^{i_j} \notin Y', \forall j \in [s'], k \in [q_i]$.

For a fixed good Z-tuple $Z = (\mathsf{Z}_{i_1}, \mathsf{Z}_{i_2}, \ldots, \mathsf{Z}_{i_{s'}})$, the permutation P_2 satisfies Eqn. (11) for the q' queries if and only if

$$\begin{cases} \forall j \in [s'], \mathsf{P}_2(\mathsf{Z}_{i_j}) = Y_{i_j} \\ \forall j \in [s'], \forall k \in [q_i], \mathsf{P}_2(X_{i_j}) = \mathsf{Z}_{i_j} \oplus T_k^{i_j}. \end{cases} \tag{19}$$

The above collection of $q' + s'$ conditions are compatible and independent with equalities in Eqn. (12) and Eqn. (17) in the sense that all inputs and all outputs that appear in the equalities of Eqn. (19) are distinct from those in Eqn. (12) and Eqn. (17). In the following lemma, we show that the number of good Z-tuples is approximately close to $(2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f}$.

Lemma 6. *For a fixed t such that $0 \leq t \leq M$ and for a fixed good set Σ, the number of valid good Z-tuples is*

$$N_Z(t) \geq (2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f} \left(1 - \frac{4q}{2^{2n/3}}\right) \prod_{l=0}^{m'_f-1} \left(1 - \frac{l}{2^n - 3q - l}\right)$$

Proof. To lower bound the number of possible choices for $N_Z(t)$, we start with number of ways choosing Z_{i_1}. To choose Z_{i_1}, we need to ensure that $\mathsf{Z}_{i_1} \notin X'$ and $\mathsf{Z}_{i_1} \oplus T_k^{i_1} \notin Y'$, for all $k \in \{1, 2, \ldots, q_1\}$. This condition leaves at least $2^n - (q - q_x + \alpha + t) - q_1(s + t)$ possibilities for Z_{i_1}. Once Z_{i_1} is fixed, Z_{i_2} must be different from Z_{i_1} and $\mathsf{Z}_{i_1} \oplus T_k^{i_1} \oplus T_{k'}^{i_2}$, for all $k \in \{1, 2, \ldots, q_1\}$ and $k' \in \{1, 2, \ldots, q_2\}$. Additionally, we need to ensure that $\mathsf{Z}_{i_2} \notin X$ and $\mathsf{Z}_{i_2} \oplus T_j^{i_2} \notin Y'$, for all $j \in \{1, 2, \ldots, q_2\}$. Thus, there are at least $2^n - (q - q_x + \alpha + t) - 1 - q_2(s + t + q_1)$ possibilities for Z_{i_2}.
In general, for the l th choice Z_{i_l}, we must ensure it is distinct from all previously chosen $\mathsf{Z}_{i_1}, \mathsf{Z}_{i_2}, \ldots, \mathsf{Z}_{i_{l-1}}$ and $\mathsf{Z}_{i_l} \oplus T_k^{i_l} \oplus T_{k'}^{i_j}$, for all $j < l$, $k \in \{1, 2, \ldots, q_l\}$ and $k' \in [q_j]$. Additionally, we also need to ensure $\mathsf{Z}_{i_l} \notin X'$ and $\mathsf{Z}_{i_l} \oplus T_j^{i_l} \notin Y'$, for all $j \in \{1, 2, \ldots, q_l\}$.

Hence, there are at least $2^n - (q - q_x + \alpha + t) - (l-1) - q_i(s + t + \sum_{j=1}^{l-1} q_j)$ valid choices for Z_{i_l}. Hence, the number of good Z-tuples is at least

$$\begin{aligned} N_Z(t) &\geq \prod_{l=0}^{s'-1} \Big(2^n - (q - q_x + \alpha + t) - l - q_{i+1}\Big(s + t + \sum_{j=1}^{l} q_j\Big)\Big) \\ &= \prod_{l=0}^{m'_f-1} \Big(2^n - (q - q_x + \alpha + t) - l - q_{i+1}\Big(s + t + \sum_{j=1}^{l} q_j\Big)\Big) \\ &\times \prod_{l=m'_f}^{s'-1} \Big(2^n - (q - q_x + \alpha + t) - l - q_{i+1}\Big(s + t + \sum_{j=1}^{l} q_j\Big)\Big) \\ &\overset{(1)}{\geq} \underbrace{\prod_{l=0}^{m'_f-1} (2^n - q - s - 2t - 2l)}_{N_{Z,1}(t)} \underbrace{\prod_{l=m'_f}^{s'-1} (2^n - 2q - 2qq_{l+1})}_{N_{Z,2}(t)} \end{aligned} \tag{20}$$

where (1) holds due to $(q - q_x + \alpha + t + l) \leq 2q$, $s + t + \sum_{j=1}^{l} q_j \leq 2q$ and $\alpha \leq q_x$.

4.5.1 Bounding $N_{Z,1}(t)$

$$\begin{aligned} \frac{N_{Z,1}(t)}{(2^n - q - s - 2t)_{m'_f}} &= \prod_{l=0}^{m'_f-1} \Big(\frac{2^n - q - s - 2t - 2l}{2^n - q - s - 2t - l}\Big) \\ &= \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - q - s - 2t - l}\Big) \geq \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big) \end{aligned} \tag{21}$$

where the last inequality follows from $s \leq q$ and $2t \leq 2M = \dfrac{2q}{2^{n/3}} \leq q$.

4.5.2 Bounding $N_{Z,2}(t)$

$$\begin{aligned} \frac{N_{Z,2}(t)}{(2^n)^{s'-m'_f}} &= \prod_{l=m'_f}^{s'-1} \Big(\frac{2^n - q - 2qq_{l+1}}{2^n}\Big) \geq \prod_{l=m'_f}^{s'-1} \Big(1 - \frac{4qq_{l+1}}{2^n}\Big) \geq \Big(1 - \frac{4q \sum_{m'_f}^{s'-1} q_{i+1}}{2^n}\Big) \\ &\overset{(3)}{\geq} \Big(1 - \frac{4q^{3/2}}{2^n}\Big) \overset{(4)}{\geq} \Big(1 - \frac{4q}{2^{2n/3}}\Big). \end{aligned} \tag{22}$$

where (3) holds due to Eqn. (18) and (4) holds due to Eqn. (7). Finally, by combining Eqn. (20), Eqn. (21) and Eqn. (22), the result follows. □

Note that, there are exactly $3t$ equalities in Eqn. (17) and exactly $q' + s'$ equalities in Eqn. (19) on P_2. Note that these equalities are compatible by the very definition of a good set Σ and a good tuple Z. Summing over t, Σ and Z, we have

$$\begin{aligned} \Pr[\mathsf{E}.2] &\geq \sum_{0 \leq t \leq M} \frac{N_\Sigma(t) \cdot N_Z(t)}{(2^n - \alpha - q_x)_{3t+q'+s'}} \\ &\geq \sum_{0 \leq t \leq M} \frac{(m_f)_{2t}}{t!(2^n - \alpha - q_x)_{3t+q'+s'}} \Big(1 - \frac{12q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big) \end{aligned}$$

$$\times (2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f} \times \Big(1 - \frac{4q}{2^{2n/3}}\Big) \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big). \quad (23)$$

Finally, by combining Eqn. (10), Eqn. (13) and Eqn. (23), we have

$$\begin{aligned} \widetilde{\mathsf{P}}(P_1) \;\geq\; & \frac{(2^n)^{\alpha}\Big(1 - \frac{4q}{2^{2n/3}}\Big)}{(2^n)_{\alpha+q_x}} \times \sum_{0\leq t\leq M} \frac{(m_f)_{2t}}{t!(2^n - \alpha - q_x)_{3t+q'+s'}} \Big(1 - \frac{12q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big) \\ \times \;& (2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f} \times \Big(1 - \frac{4q}{2^{2n/3}}\Big) \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big). \quad (24) \end{aligned}$$

4.6 Ratio of Real to Ideal Interpolation Probability

In this section, we lower bound the ratio of the real to ideal interpolation probability for a transcript τ. Recall that the ideal interpolation probability for an attainable transcript τ is

$$\prod_{i=1}^{r} \frac{1}{(2^n)_{d_i}},$$

where d_i denotes the multiplicity of the i-th tweak. Therefore, following Eqn. (6), we have

$$\frac{\Pr[\mathsf{X}_{\mathrm{re}} = \tau]}{\Pr[\mathsf{X}_{\mathrm{id}} = \tau]} \geq \frac{\widetilde{\mathsf{P}}(P_1)}{\prod_{i=1}^{r} \frac{1}{(2^n)_{d_i}}} \cdot \sum_{P_1 \in \Pi_{\mathsf{good}}} \Pr[\mathsf{P}_1 = P_1] \geq \underbrace{\frac{\widetilde{\mathsf{P}}(P_1)}{\prod_{i=1}^{r} \frac{1}{(2^n)_{d_i}}}}_{R_1} \cdot \Big(1 - \Pr[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}]\Big).$$

4.7 Lower Bounding R_1

$$\begin{aligned} R_1 \overset{(5)}{\geq} & \prod_{i=1}^{r} (2^n)_{d_i} \times \frac{\Big(1 - \frac{4q}{2^{2n/3}}\Big)(2^n)^{\alpha}}{(2^n)_{\alpha+q_x}} \times \sum_{0\leq t\leq M} \frac{(m_f)_{2t}}{t!(2^n - \alpha - q_x)_{3t+q'+s'}} \\ & \times \Big(1 - \frac{12q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big) \times (2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f} \\ & \times \Big(1 - \frac{4q}{2^{2n/3}}\Big) \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big) \geq \Big(1 - \frac{12q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big)\Big(1 - \frac{4q}{2^{2n/3}}\Big)^2 \\ & \times \sum_{0\leq t\leq M} \frac{(m_f)_{2t} \prod_{i=1}^{r} (2^n)_{d_i} (2^n)^{\alpha} (2^n - q - s - 2t)_{m'_f} (2^n)^{s'-m'_f}}{t!(2^n)_{\alpha+q_x}(2^n - \alpha - q_x)_{3t+q'+s'}} \times \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big) \\ \overset{(6)}{\geq} & \Big(1 - \frac{20q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}}\Big) \times \underbrace{\frac{(2^n)^{\alpha}}{(2^n)_{\alpha}}}_{\geq 1} \times \underbrace{\frac{\prod_{i=1}^{r} (2^n)_{d_i}}{(2^n - m_f)_q} \prod_{l=0}^{m'_f-1} \Big(1 - \frac{l}{2^n - 3q - l}\Big)}_{A} \\ & \times \underbrace{\frac{(2^n)^{s'-m'_f}}{(2^n)_{s'-m'_f}}}_{\geq 1} \times \underbrace{\sum_{0\leq t\leq M} \frac{(m_f)_{2t}(2^n - q - s - 2t)_{m'_f}}{t!(2^n - s - m)_{m'_f+t}}}_{B} \end{aligned} \quad (25)$$

where (5) holds following Eqn. (24) and (6) holds due to the fact that $(2^n)_{\alpha+q_x}(2^n - \alpha - q_x)_{3t+q'+s'} \leq (2^n)_{\alpha}(2^n)_{s'-m'_f}(2^n - m_f)_q(2^n - s - m)_{m'_f+t}$.

Therefore, it remains to lower bound the expression (A) and (B) as defined above.

Lower Bounding A: To lower bound A, we have:

$$A = \frac{\prod_{i=1}^{r}(2^n)_{d_i}}{(2^n - m_f)_q} \prod_{l=0}^{m'_f-1}\Big(1 - \frac{l}{2^n - 3q - l}\Big) \tag{26}$$

Since, $d_i \leq m_f$ for all $i \in [r]$, we have $(2^n)_{d_i} \geq (2^n - m_f)^{d_i}$. Therefore,

$$\prod_{i=1}^{r}(2^n)_{d_i} \geq \prod_{i=1}^{r}(2^n - m_f)_{d_i} \geq (2^n - m_f)^{\sum_{i=1}^{r} d_i} = (2^n - m_f)^q \tag{27}$$

By combining Eqn. (26) and Eqn. (27), we have

$$\begin{aligned}
A &\geq \frac{(2^n - m_f)^q}{(2^n - m_f)_q} \prod_{l=0}^{m'_f-1}\Big(1 - \frac{l}{2^n - 3q - l}\Big) = \prod_{l=0}^{q-1}\Big(1 + \frac{l}{2^n - m_f - l}\Big) \prod_{l=0}^{m'_f-1}\Big(1 - \frac{l}{2^n - 3q - l}\Big) \\
&\geq \prod_{l=0}^{q-1}\Big(1 + \frac{l}{2^n - m_f - l}\Big)\Big(1 - \frac{l}{2^n - 3q - l}\Big) = \prod_{l=0}^{q-1}\Big(1 - \frac{3ql - m_f l + l^2}{(2^n - m_f - l)(2^n - 3q - l)}\Big) \\
&\geq \prod_{l=0}^{q-1}\Big(1 - \frac{4ql}{(2^n - 2q)(2^n - 4q)}\Big) \overset{(7)}{\geq} \Big(1 - \frac{4q \cdot q^2/2}{(\frac{3\cdot 2^n}{8})(\frac{2^n}{2})}\Big) \qquad (q \leq 2^n/8)) \\
&\geq \Big(1 - \frac{32q^3}{3 \cdot 2^{2n}}\Big) \overset{(8)}{\geq} \Big(1 - \frac{11q}{2^{2n/3}}\Big),
\end{aligned} \tag{28}$$

where (7) holds as we assume $q \leq 2^n/8$ and (8) holds following Eqn. (7).

Lower Bounding B: To lower bound the expression B, we recall a discrete probability distribution called the hypergeometric distribution. This distribution describes the probability of obtaining k successes (i.e., objects with a specific feature) in n random draws, without replacement, from a finite population of size N that contains exactly K objects with that specific feature, where each draw is either a success or a failure. The probability distribution function is therefore given by,

$$\mathbf{Hyp}_{N,K,n}(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

The mean of the distribution $\mathbf{Hyp}_{N,K,n}(k)$ is $\frac{nK}{N}$. Since,

$$\begin{aligned}
\mathbf{Hyp}_{2^n-q,m_f,m_f}(t) &= \frac{(m_f)_t (m_f)_t (2^n - q - m_f)_{m_f - t}}{t!(2^n - q)_{m_f}} \\
&= \frac{(m_f)_t (m_f)_t (2^n - q - m_f)_{m'_f + t}}{t!(2^n - q)_{m'_f + 2t}},
\end{aligned}$$

we rewrite B as:

$$B = \sum_{0 \leq t \leq M} \underbrace{\frac{(m_f)_{2t}}{(m_f)_t (m_f)_t}}_{C} \cdot \underbrace{\frac{(2^n - q)_{m'_f+2t}(2^n - q - s - 2t)_{m'_f}}{(2^n - q - m_f)_{m'_f+t}(2^n - s - m)_{m'_f+t}}}_{D} \cdot \mathbf{Hyp}_{2^n-q,m_f,m_f}(t) \tag{29}$$

Therefore, to lower bound B, we need to lower bound C and D separately.

Lower Bounding C: For any fixed t with $0 \leq t \leq M$, we have

$$\begin{aligned} C &= \frac{(m_f)_{2t}}{(m_f)_t(m_f)_t} \geq \frac{(m_f)_{2t}}{(q-m)_t(q-m)_t} = \frac{(q-m-q_x)_{2t}}{(q-m)_t(q-m)_t} \geq \frac{(q-m-q_x-2M)^{2t}}{(q-m)^{2t}} \\ &\geq 1 - \frac{2t(q_x+2M)}{q-m} \overset{(9)}{\geq} 1 - \frac{2M(q_x+2M)}{q-M} \geq 1 - \frac{4M^2+2Mq_x}{q-M} \overset{(10)}{\geq} 1 - \frac{8M^2+4Mq_x}{q} \\ &\geq 1 - \frac{8M^2+4M\sqrt{q}}{q} \quad (\text{Since} \quad q_x \leq \sqrt{q}) \overset{(11)}{=} 1 - \left(\frac{8q}{2^{2n/3}} + \frac{4\sqrt{q}}{2^{n/3}}\right). \end{aligned} \tag{30}$$

where (9) holds following Eqn. (16), (10) holds due to the fact that $q - 3M \geq q/2$ and (11) holds as $M = q/2^{n/3}$.

Lower Bounding D: To lower bound D, we have:

$$\begin{aligned} D &= \frac{(2^n-q)_{m'_f+t}(2^n-q-m'_f-t)_t(2^n-q-s-2t)_{m'_f}}{(2^n-q-m_f)_t(2^n-q-m_f-t)_{m'_f}(2^n-s-m)_{m'_f+t}} \\ &= \frac{(2^n-q)_{m'_f+t}}{(2^n-s-m)_{m'_f+t}} \cdot \frac{(2^n-q-m_f+t)_t}{(2^n-q-m_f)_t} \cdot \frac{(2^n-q-s-2t)_{m'_f}}{(2^n-q-m_f-t)_{m'_f}} \\ &= \frac{(2^n-q)_{m'_f+t}}{(2^n-q-(s-q+m))_{m'_f+t}} \cdot \frac{(2^n-q-m_f+t)_t}{(2^n-q-m_f)_t} \cdot \frac{(2^n-q-s-2t)_{m'_f}}{(2^n-q-m_f-t)_{m'_f}}. \end{aligned}$$

Now from Eqn. (18), it is clear that $(s-q+m) \geq 0$. Also, since $2^n-q-m_f+t \geq 2^n-q-m_f$, the first two terms of D are ≥ 1 and hence,

$$\begin{aligned} D &\geq \prod_{i=0}^{m'_f-1} \frac{2^n-q-s-2t-i}{2^n-q-m_f-t-i} = \prod_{i=0}^{m'_f-1} \left(1 - \frac{s-m_f+t}{2^n-q-m_f-t-i}\right) \\ &\geq 1 - \frac{m'_f(s-m_f+M)}{2^n-3q} \overset{(12)}{\geq} 1 - \frac{q(q-(q-M-\sqrt{q})+M)}{2^n-3q} \\ &= 1 - \frac{2qM}{2^n-3q} - \frac{q\sqrt{q}}{2^n-3q} \overset{(13)}{\geq} 1 - \frac{4q^2}{2^{4n/3}} - \frac{2q\sqrt{q}}{2^n} \overset{(14)}{\geq} 1 - \frac{6q}{2^{2n/3}}. \end{aligned} \tag{31}$$

where (12) holds following Eqn. (16) and both $m'_f, s \leq q$, (13) holds by plugging-in $M = q/2^{n/3}$ and $q \leq 2^n/6$. Finally, (14) holds following Eqn. (7). Recall that the mean of the hypergeometric distribution $\mathbf{Hyp}_{2^n-q,m_f,m_f}$ is given by $\frac{m_f^2}{2^n-q}$. Therefore, we have:

$$\sum_{t>M} \mathbf{Hyp}_{2^n-q,m_f,m_f}(t) \leq \frac{m_f^2}{M(2^n-q)} \leq \frac{2q^2}{M2^n} = \frac{2q}{2^{2n/3}}.$$

By plugging-in the value of $M = q/2^{n/3}$ and by applying the Markov's inequality with the assumption $q \leq 2^n/2$, we have

$$\sum_{0 \leq t \leq M} \mathbf{Hyp}_{2^n-q,m_f,m_f}(t) \geq \left(1 - \frac{2q}{2^{2n/3}}\right). \tag{32}$$

Finally combining Eqn. (29), Eqn. (30), Eqn. (31), and Eqn. (32), we obtain:

$$\begin{aligned} B &\geq \left(1 - \frac{8q}{2^{2n/3}} - \frac{4\sqrt{q}}{2^{n/3}}\right)\left(1 - \frac{6q}{2^{2n/3}}\right) \sum_{0 \leq t \leq M} \mathbf{Hyp}_{2^n-q,m_f,m_f}(t) \\ &\geq 1 - \frac{16q}{2^{2n/3}} - \frac{4\sqrt{q}}{2^{n/3}}. \end{aligned} \tag{33}$$

Thus, by combining Eqn. (25), Eqn. (28), and Eqn. (33) together, we have

$$R_1 \geq 1 - \frac{47q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}} - \frac{4\sqrt{q}}{2^{n/3}}. \tag{34}$$

Moreover, following Lemma 8, we have

$$\Pr[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}] \leq \frac{4q^3}{2^{2n}} + \frac{2q}{2^{\frac{2n}{3}}} + \frac{2q}{2^n} + \frac{4}{2^n}. \tag{35}$$

Finally, by combining Eqn. (34) and Eqn. (35), we obtain the following

$$\begin{aligned}
\frac{\Pr[\mathsf{X}_{\mathrm{re}} = \tau]}{\Pr[\mathsf{X}_{\mathrm{id}} = \tau]} &\geq \frac{\widetilde{\mathsf{P}}(P_1)}{\prod_{i=1}^{r} \frac{1}{(2^n)_{d_i}}} \cdot \Big(1 - \Pr\left[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}\right]\Big) \\
&\geq \Big(1 - \frac{47q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}} - \frac{4\sqrt{q}}{2^{n/3}}\Big) \cdot \Big(1 - \frac{4q^3}{2^{2n}} - \frac{2q}{2^{\frac{2n}{3}}} - \frac{2q}{2^n} - \frac{4}{2^n}\Big) \\
&\geq 1 - \frac{49q}{2^{2n/3}} - \frac{32\sqrt{2nq}}{2^{n/3}} - \frac{4\sqrt{q}}{2^{n/3}} - \frac{4q^3}{2^{2n}} - \frac{2q}{2^n} - \frac{4}{2^n}.
\end{aligned}$$

This completes the proof of the theorem. □

4.8 Comparison in the Proof Techniques: Four-Keyed vs. Two-Keyed CLRW1

In [DDDM23], authors have proved $3n/4$-bit security of $\mathsf{CLRW1}^4$ construction invoked with four independent permutations. In the security proof, the input-output pairs for the outer permutations P_1 and P_4 have been released after the interaction. Hence, for each query, evaluating the central permutations P_2 and P_3 on the corresponding intermediate values are boiled down to the following system of equations:

$$P_2(P_1(M_i) \oplus T_i) \oplus P_3^{-1}(P_4^{-1}(C_i) \oplus T_i) = T_i.$$

The authors have defined appropriate bad conditions on the transcript and treated this part using the mirror theory result, as two central permutations P_2 and P_3 are independent, allowing us to get a lower bound on the number of solutions for the corresponding system of equations through mirror theory.
In contrast, our construction invokes the same keyed permutation P_2 twice in the central part and P_1 twice in the outer part of the construction. As before, we release the input-output pairs for the outer permutations P_1 after the interaction. Thus, for each query, evaluating the central permutations P_2 on the corresponding intermediate values boils down to the following system of equations:

$$P_2(P_1(M_i) \oplus T_i) \oplus P_2^{-1}(P_1^{-1}(C_i) \oplus T_i) = T_i$$

Note that, as P_2 is not independent of P_2^{-1}, it restricts us from using the mirror theory directly in our construction, as the authors did in [DDDM23]. One can also observe that our structure resembles the SUMPIP [GSWG19] constructions $\mathsf{P} \oplus \mathsf{P}^{-1}$, where handling the interdependencies between the input-output of the permutation is subtle and requires careful bounding of collision conditions compared to the mirror theory technique.

5 Proof of Lemma 3

Proof. We define the event $\mathsf{Bad} := \mathsf{Bad\text{-}Aux}_1 \vee \mathsf{Bad1} \vee \mathsf{Bad2} \vee \mathsf{Bad3} \vee \mathsf{Bad4}$. We first individually bound the probability of the event $\mathsf{Bad\text{-}Aux}_1$, $\mathsf{Bad1}$ to $\mathsf{Bad4}$ and then by the

virtue of the union bound, we sum up the individual bound to obtain the result, i.e., we have

$$\Pr\left[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}\right] = \Pr[\mathsf{Bad}] \leq \Pr[\mathsf{Bad\text{-}Aux}_1] + \Pr[\mathsf{Bad1}] + \Pr[\mathsf{Bad2}] + \Pr[\mathsf{Bad3}] + \Pr[\mathsf{Bad4}].$$

I. Bounding $\mathsf{Bad\text{-}Aux}_1$: For a fixed choice of indices, the probability of the event $P_1(P_1(M_i)) = C_i$ is upper bounded by $1/(2^n - q) \leq 2/2^n$, assuming $q \leq 2^{n-1}$. Therefore, by varying over all possible choices of indices, we have

$$\Pr[\mathsf{Bad\text{-}Aux}_1] \leq \frac{2q}{2^n}. \tag{36}$$

II. Bounding Bad1: For a fixed choice of indices i, j, and k, the event rank of the system of equations

$$\begin{cases} P_1(M_i) \oplus T_i = P_1(M_j) \oplus T_j \\ P_1^{-1}(C_i) \oplus T_i = P_1^{-1}(C_k) \oplus T_k \end{cases}$$

is at most 2. If the rank is 2, then we obtain at least two fresh variables and thus they can be used to bound the probability of the event. Therefore, in that case, the probability of the event is at most $1/(2^n - q)(2^n - q - 1) \leq 4/2^{2n}$ by assuming $q \leq 2^{n-1}$.

On the other hand, the rank of the above system of equations could be one and the only possible way we end up with rank 1 through the following equality:

$$P_1(P_1(M_i)) = C_k, P_1(P_1(M_j)) = C_i.$$

Note that the above equality holds with probability at most 2^{-n} using the randomness of P_1. Since the rank of the above system of equations is one, we get another 2^{-n} entropy to bound the event. Overall, the probability of the event when the rank of the system of equations is one is at most $4/2^{2n}$ by assuming $q \leq 2^{n-1}$. Therefore, by varying over all possible choices of indices, we have

$$\Pr[\mathsf{Bad1}] \leq \frac{4q^3}{2^{2n}}. \tag{37}$$

III. Bounding Bad2: For a fixed choice of indices, we define an indicator random variable $\mathbb{I}_{i,j}$ which takes the value 1, if $\mathsf{P}_1(M_i) \oplus T_i = \mathsf{P}_1(M_j) \oplus T_j$, and 0, otherwise. Let $\mathbb{I} = \sum_{i \neq j} \mathbb{I}_{i,j}$. By linearity of expectation,

$$\mathbf{E}[\mathbb{I}] = \sum_{i \neq j} \mathbf{E}[\mathbb{I}_{i,j}] = \sum_{i \neq j} \Pr[\mathsf{P}_1(M_i) \oplus T_i = \mathsf{P}_1(M_j) \oplus T_j] \leq \frac{q^2}{2^n}$$

Now,

$$\begin{aligned} \Pr[\mathsf{Bad2}] &\leq \Pr\left[|\{(i,j) \in [q]^2 : \mathsf{P}_1(M_i) \oplus T_i = \mathsf{P}_1(M_j) \oplus T_j\}| \geq \sqrt{q}\right] \\ &= \Pr[\mathbb{I} \geq \sqrt{q}] \overset{(1)}{\leq} \frac{q^2}{2^n\sqrt{q}} \leq \frac{q\sqrt{q}}{2^n} \overset{(2)}{\leq} \frac{q}{2^{2n/3}}, \end{aligned} \tag{38}$$

where (1) holds due to the Markov inequality and (2) holds due to 7.

IV. Bounding Bad3: Analysis of Bad3 is similar to the analysis of Bad2 except the fact that the bound here is given by M instead of $\sqrt{q}$. So we have,

$$\Pr[\mathsf{Bad3}] \leq \frac{q^2}{2^n M} = \frac{q^2 2^{\frac{n}{3}}}{2^n q} = \frac{q}{2^{\frac{2n}{3}}} \tag{39}$$

V. Bounding Bad4: Note that in our setting P_1 is a random permutation. So, to bound the event $\mathsf{Bad4}$ we apply a variant of Sum Capture lemma from Lemma 3 [CT21] with $A = Y^*$, $U = T^*$ and $V = X^*$ as follows:

$$\begin{aligned}\Pr[\mathsf{Bad4}] &\leq \Pr\left[\xi \geq \frac{|Y^*||T^*||X^*|}{2^n} + 4\sqrt{2n|Y^*||T^*||X^*|}\right] \\ &\leq \Pr\left[\mu(Y^*) \geq \frac{|Y^*||T^*||X^*|}{2^n} + 4\sqrt{2n|Y^*||T^*||X^*|}\right] \\ &\leq \frac{4}{2^n}\end{aligned} \tag{40}$$

Taking all together we have,

$$\Pr\left[\mathsf{P}_1 \in \Pi_{\mathsf{bad}}\right] \leq \frac{4q^3}{2^{2n}} + \frac{2q}{2^{2n/3}} + \frac{2q}{2^n} + \frac{4}{2^n}.$$

□

6 Proof of Lemma 4

To lower bound the number of possible choices for N_W, we count the the number of ways we can choose the elements $\mathsf{W}_{i_1}, \mathsf{W}_{i_2}, \mathsf{W}_{i_3}, \ldots \mathsf{W}_{i_\alpha}$ such that they satisfy the pre-specified conditions stated in Definition 2. To begin with, we count the number of ways we can choose a valid W_{i_1}. Note that, we need to ensure that

$$\mathsf{W}_{i_1} \notin Y_X, \text{ and } \mathsf{W}_{i_1} \oplus T_k^{i_1} \notin S_1, \forall k \in \{1, 2, \ldots, \beta_1\}.$$

This condition leaves at least $2^n - q_x - \alpha\beta_1$ possibilities for W_{i_1}. Once W_{i_1} is fixed, W_{i_2} must be different from W_{i_1} and $\mathsf{W}_{i_1} \oplus T_k^{i_1} \oplus T_{k'}^{i_2}$, for all $k \in \{1, 2, \ldots, \beta_1\}$ and $k' \in \{1, 2, \ldots, \beta_2\}$. Additionally, $\mathsf{W}_{i_2} \notin Y_X$ and $\mathsf{W}_{i_2} \oplus T_j^{i_2} \notin S_1$, for all $j \in \{1, 2, \ldots, \beta_2\}$. Thus, there are at least $2^n - q_x - 1 - \alpha\beta_2 - \beta_2\beta_1$ possibilities for W_{i_2}. In general, for the l th choice of W_{i_l}, we have to ensure it is distinct from all previously chosen $\mathsf{W}_{i_1}, \mathsf{W}_{i_2}, \ldots, \mathsf{W}_{i_{l-1}}$ and their respective $\mathsf{W}_{i_j} \oplus T_k^{i_j} \oplus T_{k'}^{i_l}$, for all $j < l$ and $k \in \{1, 2, \ldots, \beta_j\}$ and $k' \in \{1, 2, \ldots, \beta_l\}$. We also need to ensure that $\mathsf{W}_{i_l} \notin Y_X$ and $\mathsf{W}_{i_l} \oplus T_j^{i_l} \notin S_1$, for all $j \in \{1, 2, \ldots, \beta_l\}$. Hence, there are at least $2^n - q_x - (l-1) - \alpha\beta_l - \beta_l \sum_{j=1}^{l-1} \beta_j$ possibilities for W_{i_l}. Hence, the number of good W-tuples is at least,

$$\begin{aligned}N_W &\geq \prod_{l=1}^{\alpha} \left(2^n - q_x - (l-1) - \alpha\beta_l - \beta_l \sum_{j=1}^{l-1} \beta_j\right) = \prod_{l=0}^{\alpha-1} \left(2^n - q_x - l - \alpha\beta_{l+1} - \beta_{l+1} \sum_{j=1}^{l} \beta_j\right) \\ &= \prod_{l=0}^{\alpha-1} \left(2^n - q_x - l - \beta_{l+1}\left(\sum_{j=1}^{l} \alpha + \beta_j\right)\right) \geq \prod_{l=0}^{\alpha-1} \left(2^n - q_x - l - \beta_{l+1}(\alpha + q_x)\right) \\ &\geq \prod_{l=0}^{\alpha-1} \left(2^n - q_x - \alpha - \beta_{l+1}(\alpha + q_x)\right) \geq \prod_{l=0}^{\alpha-1} \left(2^n - 2q - 2q\beta_{l+1}\right).\end{aligned}$$

Hence, we have

$$\frac{N_W}{(2^n)^\alpha} = \prod_{l=0}^{\alpha-1} \left(\frac{2^n - 2q - 2q\beta_{l+1}}{2^n}\right) \geq \prod_{l=0}^{\alpha-1} \left(1 - \frac{4q\beta_{l+1}}{2^n}\right) \geq 1 - \left(\frac{4q \prod_{l=0}^{\alpha-1} \beta_{l+1}}{2^n}\right)$$

$$= 1 - \frac{4qq_x}{2^n} \geq 1 - \frac{4q\sqrt{q}}{2^n} \geq 1 - \frac{4q}{2^{2n/3}}.$$

7 Conclusion

We have shown that R-CLRW1[4] achieves $2n/3$-bit CCA security. However, our proven bound is not tight. It is interesting to see whether the bound can be improved to $3n/4$-bit and hence at par with the security level of the CLRW1[4] construction. In this regard, we believe that our proof approach may not yield an improved bound for the construction due to the application of the sum-capture lemma, which serves as a primary bottleneck in deriving the improved bound. Thus, one may need to come up with a different proof approach to manifest an improved security bound of the construction. Along with this, studying the security bound of a single-keyed variant of the CLRW1[4] construction will be another direction of future research. However, the same key across all rounds of the single-keyed variant of the CLRW1[4] construction yields a strong input-output dependency among the multiple invocations of the same permutation. We believe that such a strong inter-dependency complicates the security analysis. Nevertheless, a security proof might still be possible, but that would require a very careful handling of these dependencies and possibly a different set of technical tools.

References

[BGGS20] Zhenzhen Bao, Chun Guo, Jian Guo, and Ling Song. Tnt: How to tweak a block cipher. In *Advances in Cryptology – EUROCRYPT 2020: 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10–14, 2020, Proceedings, Part II*, page 641–673, Berlin, Heidelberg, 2020. Springer-Verlag.

[CAE14] CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness, 2014. http://competitions.cr.yp.to/caesar.html.

[CT21] Yu Long Chen and Stefano Tessaro. Better security-efficiency trade-offs in permutation-based two-party computation. In Mehdi Tibouchi and Huaxiong Wang, editors, *Advances in Cryptology - ASIACRYPT 2021 - 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6-10, 2021, Proceedings, Part II*, volume 13091 of *Lecture Notes in Computer Science*, pages 275–304, Berlin, Heidelberg, 2021. Springer.

[DDDM23] Nilanjan Datta, Shreya Dey, Avijit Dutta, and Sougata Mandal. Cascading four round LRW1 is beyond birthday bound secure. *IACR Trans. Symmetric Cryptol.*, 2023(4):365–390, 2023.

[GGLS20] Chun Guo, Jian Guo, Eik List, and Ling Song. Towards closing the security gap of tweak-and-tweak (tnt). In *Advances in Cryptology – ASIACRYPT 2020: 26th International Conference on the Theory and Application of Cryptology and Information Security, Daejeon, South Korea, December 7–11, 2020, Proceedings, Part I*, page 567–597, Berlin, Heidelberg, 2020. Springer-Verlag.

[GSWG19] Chun Guo, Yaobin Shen, Lei Wang, and Dawu Gu. Beyond-birthday secure domain-preserving prfs from a single permutation. *Des. Codes Cryptography*, 87(6):1297–1322, June 2019.

[HKR15] Viet Tung Hoang, Ted Krovetz, and Phillip Rogaway. Robust authenticated-encryption AEZ and the problem that it solves. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, volume 9056 of *Lecture Notes in Computer Science*, pages 15–44, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[IMPS17] Tetsu Iwata, Kazuhiko Minematsu, Thomas Peyrin, and Yannick Seurin. ZMAC: A fast tweakable block cipher mode for highly secure message authentication. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part III*, volume 10403 of *Lecture Notes in Computer Science*, pages 34–65, Berlin, Heidelberg, 2017. Springer.

[JKNS23] Ashwin Jha, Mustafa Khairallah, Mridul Nandi, and Abishanka Saha. Tight security of tnt and beyond: Attacks, proofs and possibilities for the cascaded lrw paradigm. Cryptology ePrint Archive, Paper 2023/1272, 2023. https://eprint.iacr.org/2023/1272.

[JLM+17] Ashwin Jha, Eik List, Kazuhiko Minematsu, Sweta Mishra, and Mridul Nandi. XHX - A framework for optimally secure tweakable block ciphers from classical block ciphers and universal hashing. In *LATINCRYPT*, volume 11368 of *Lecture Notes in Computer Science*, pages 207–227, Berlin, Heidelberg, 2017. Springer.

[JN20] Ashwin Jha and Mridul Nandi. Tight security of cascaded LRW2. *J. Cryptol.*, 33(3):1272–1317, 2020.

[JNPS21] Jérémy Jean, Ivica Nikolic, Thomas Peyrin, and Yannick Seurin. The deoxys AEAD family. *J. Cryptol.*, 34(3):31, 2021.

[JNS23] Ashwin Jha, Mridul Nandi, and Abishanka Saha. Tight security of tnt: Reinforcing khairallah's birthday-bound attack. Cryptology ePrint Archive, Paper 2023/1233, 2023. https://eprint.iacr.org/2023/1233.

[Kha23] Mustafa Khairallah. Clrw1[3] is not secure beyond the birthday bound: Breaking tnt with $O(2^{n/2})$ queries. Cryptology ePrint Archive, Paper 2023/1212, 2023. https://eprint.iacr.org/2023/1212.

[LL18] ByeongHak Lee and Jooyoung Lee. Tweakable block ciphers secure beyond the birthday bound in the ideal cipher model. *IACR Cryptol. ePrint Arch.*, page 784, 2018.

[LPS12] Rodolphe Lampe, Jacques Patarin, and Yannick Seurin. An asymptotically tight security analysis of the iterated even-mansour cipher. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings*, volume 7658 of *Lecture Notes in Computer Science*, pages 278–295, Berlin, Heidelberg, 2012. Springer.

[LRW02] Moses Liskov, Ronald L. Rivest, and David Wagner. Tweakable block ciphers. In Moti Yung, editor, *Advances in Cryptology — CRYPTO 2002*, pages 31–46, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[LS13] Rodolphe Lampe and Yannick Seurin. Tweakable blockciphers with asymptotically optimal security. In Shiho Moriai, editor, *Fast Software Encryption - 20th International Workshop, FSE 2013, Singapore, March 11-13, 2013. Revised Selected Papers*, volume 8424 of *Lecture Notes in Computer Science*, pages 133–151, Berlin, Heidelberg, 2013. Springer.

[LST12] Will Landecker, Thomas Shrimpton, and R. Seth Terashima. Tweakable blockciphers with beyond birthday-bound security. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology - CRYPTO 2012 - 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings*, volume 7417 of *Lecture Notes in Computer Science*, pages 14–30, Berlin, Heidelberg, 2012. Springer.

[Mar10] Luther Martin. XTS: A mode of AES for encrypting hard disks. *IEEE Secur. Priv.*, 8(3):68–69, 2010.

[Men15] Bart Mennink. Optimally secure tweakable blockciphers. In Gregor Leander, editor, *Fast Software Encryption - 22nd International Workshop, FSE 2015, Istanbul, Turkey, March 8-11, 2015, Revised Selected Papers*, volume 9054 of *Lecture Notes in Computer Science*, pages 428–448, Berlin, Heidelberg, 2015. Springer.

[Men18] Bart Mennink. Towards tight security of cascaded lrw2. In *Theory of Cryptography: 16th International Conference, TCC 2018, Panaji, India, November 11–14, 2018, Proceedings, Part II*, page 192–222, Berlin, Heidelberg, 2018. Springer-Verlag.

[NIS18] NIST. Lightweight cryptography, 2018. Online: https://csrc.nist.gov/Projects/Lightweight-Cryptography. Accessed: August 01, 2019.

[Pat08] Jacques Patarin. The "coefficients h" technique. In *Selected Areas in Cryptography, 15th International Workshop, SAC 2008, Sackville, New Brunswick, Canada, August 14-15, Revised Selected Papers*, pages 328–345, 2008.

[Pat09] Jacques Patarin. The “coefficients h” technique. In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *Selected Areas in Cryptography*, pages 328–345, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[RBB03] Phillip Rogaway, Mihir Bellare, and John Black. OCB: A block-cipher mode of operation for efficient authenticated encryption. *ACM Trans. Inf. Syst. Secur.*, 6(3):365–403, 2003.

[Rog04] Phillip Rogaway. Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In *ASIACRYPT*, volume 3329 of *Lecture Notes in Computer Science*, pages 16–31, Berlin, Heidelberg, 2004. Springer.

[SS23] Yaobin Shen and François-Xavier Standaert. Optimally secure tweakable block ciphers with a large tweak from n-bit block ciphers. *IACR Trans. Symmetric Cryptol.*, 2023(2):47–68, 2023.

[Tes15] Stefano Tessaro. Optimally secure block ciphers from ideal primitives. In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part II*, volume 9453 of *Lecture Notes in Computer Science*, pages 437–462, Berlin, Heidelberg, 2015. Springer.

[WGZ+16] Lei Wang, Jian Guo, Guoyan Zhang, Jingyuan Zhao, and Dawu Gu. How to build fully secure tweakable blockciphers from classical blockciphers. In *ASIACRYPT*, pages 455–483, 2016.

[ZQG22] Zhongliang Zhang, Zhen Qin, and Chun Guo. Just tweak! asymptotically optimal security for the cascaded lrw1 tweakable blockcipher. *Des. Codes Cryptography*, 91(3):1035–1052, oct 2022.