# Efficient Variants of TNT with BBB Security

Ritam Bhaumik<sup>1</sup>, Wonseok Choi<sup>2</sup>, Avijit Dutta<sup>3,4</sup>, Cuauhtemoc Mancillas López<sup>5</sup>, Hrithik Nandi<sup>3,6</sup>, and Yaobin Shen<sup>7</sup>

<sup>1</sup> CRC, TII, Abu Dhabi bhaumik.ritam@gmail.com <sup>2</sup> Purdue University choi935@purdue.edu

<sup>3</sup> Institute for Advancing Intelligence (IAI), TCG CREST, Kolkata

avirocks.dutta130gmail.com, nandyhrithik0gmail.com

<sup>4</sup> Academy of Scientific and Innovative Research (AcSIR), Ghaziabad

<sup>5</sup> CINVESTAV-IPN cuauhtemoc.mancillas830gmail.com

<sup>6</sup> Ramakrishna Mission Vivekananda Educational and Research Institute, Belur <sup>7</sup> Xiamen University yaobins1800gmail.com

Abstract. At EUROCRYPT'20, Bao et al. have shown that three-round cascading of LRW1 construction, which they dubbed as TNT, is a strong tweakable pseudorandom permutation that provably achieves 2n/3-bit security bound. Jha et al. showed a birthday bound distinguishing attack on TNT and invalidated the proven security bound and proved a tight birthday bound security on the TNT construction in EUROCRYPT'24. In a recent work, Datta et al. have shown that four round cascading of the LRW1 construction, which they dubbed as CLRW1<sup>4</sup> is a strong tweakable pseudorandom permutation that provably achieves 3n/4-bit security. In this paper, we propose a variant of the TNT construction, called b-TNT1, and proved its security up to  $2^{3n/4}$  queries. However, unlike CLRW1<sup>4</sup>, b-TNT1 requires three block cipher calls along with a field multiplication. Besides, we also propose another variant of the TNT construction, called b-TNT2 and showed a similar security bound. Compared to b-TNT1, b-TNT2 requires four block cipher calls. Nevertheless, its execution of block cipher calls can be pipelined which makes it efficient over CLRW1<sup>4</sup>. We have also experimentally verified that both b-TNT1 and b-TNT2 outperform CLRW1<sup>4</sup>.

**Keywords:** Tweakable Block Cipher, Tweak-aNd-Tweak, Cascaded LRW1, Beyond Birthday Bound Security, Mirror Theory, Expectation Method.

# 1 Introduction

A tweakable block cipher is a rich cryptographic primitive that serves to introduce variability within the cipher's structure. A tweakable block cipher is defined as a family of permutations  $\tilde{\mathsf{E}} : \mathcal{K} \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$  indexed by secret key  $k \in \mathcal{K}$  and public tweak  $t \in \mathcal{T}$ . A prototypical design of a tweakable block cipher originally appeared in the Hasty Pudding Cipher [43], where an extra input, known as "spice" served the role of a tweak besides the key and the plaintext, to a block cipher. The actual intention of spice is to introduce randomization in the choice of the permutation family. Later in [32, 33], Liskov, Rivest, and Wagner formalized the design and referred to the primitive as a tweakable block cipher.

Tweakable block ciphers have received significant acceptance as a fundamental cryptographic object. Over the years, TBCs have found diverse applications, in designing of AE schemes, e.g., Deoxys [27], Romulus [36], and several other candidates of AE schemes [1,4,5,8,15,20,38,41]. TBC has also been extensively used in designing many AE candidates for NIST and CÆSAR competitions, including [16,22,25–27,44]. Besides, TBCs have also been used in designing wide block encryption modes [6,39], message authentication codes [8,9,11,24,35,37], hash functions [14,18,21], and pseudorandom functions [10].

LRW1 and LRW2, proposed by Liskov et al. [32], are the first examples of tweakable block ciphers, which are built from block ciphers assuming their strong pseudorandom permutation security. Over the years, a few variants of the LRW2 construction have been proposed in [7,34,42] which have been shown to be secure up to the birthday bound of the query complexity. Landecker et al. [31] showed that cascading two independent LRW2 constructions, called CLRW2, achieves security up to  $2^{2n/3}$  queries. Subsequent works [29] have improved the bound of Landecker et al. [31] from 2n/3 bits to 3n/4 bits. Lampe and Seurin [30] generalized CLRW2 construction to the cascading of  $r \geq 1$  LRW2 construction and proved that it achieves security up to  $2^{rn/(r+2)}$  queries for even r. Although the bound approaches the optimal security with increasing r, it comes at the cost of increasing the number of block cipher keys and primitive calls linearly with r. Bao et al. [2] showed that the three-round cascading of the LRW1 construction, called TNT (an abbreviation of "*The Tweak-aNd-Tweak*") achieves CCA security up to  $2^{2n/3}$  queries.

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

Guo et al. [17] showed a tight 3n/4-bit CPA security bound of the construction. Zhang et al. [45] studied the security analysis of the generalized *r*-round cascading of the LRW1 construction, called  $\mathsf{CLRW1}^r$  and showed that it achieves CCA security up to  $2^{(r-2)n/r}$  queries, with  $r \geq 2$ . Furthermore, when *r* is odd, the construction attains enhanced security for up to  $2^{(r-1)n/(r+1)}$  queries.

Jha et al. [28] showed a birthday bound CCA distinguishing attack on TNT, invalidated the previously asserted security claim of the construction, and proved a tight birthday bound security of the TNT construction. Recently, Datta et al. [13] showed that four round cascading of the LRW1 construction, called  $CLRW1^4$ , achieves CCA security up to  $2^{3n/4}$  queries. In fact, this result is the first one that provably shows the minimal number of rounds required for cascading LRW1 construction to ensure beyond-birthday-bound security against all CCA adversaries. We would like to mention that a parallel work of [13] also established a similar security bound of the construction [28].

#### 1.1 Our Contribution

Birthday bound security of the  $\mathsf{TNT}$  construction has rendered the designer to include one extra block cipher call in  $\mathsf{CLRW1}^4$  construction to achieve beyond-

birthday-bound security. However, this additional invocation of the block cipher comes at the cost of evaluating it for every query. Moreover, to accommodate the decryption query, one needs to invoke the decryption circuit of the extra block cipher for every distinct ciphertext. Besides, we cannot execute the additional block cipher call in parallel to the execution of the TNT construction, i.e., to evaluate  $E_{K_4}$ , one needs to wait for the output of TNT to become available.

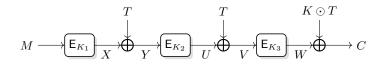


Fig. 1: **b-TNT1** construction based on three block cipher calls and a field multiplication.

To address the above issues, we propose a simple fix to the TNT construction that does not require any extra block cipher call. In particular, we blind the output of the TNT construction by multiplying an *n*-bit secret key with the tweak and call the resulting construction b-TNT1. A pictorial description of the construction is shown in Fig. 1. Since the field multiplication is less costly than evaluating a block cipher, our proposed construction outperforms CLRW1<sup>4</sup> in terms of throughput while retaining a similar level of security bound. Although b-TNT1 is better than CLRW1<sup>4</sup> in terms of throughput, it incurs a larger hardware area compared to CLRW1<sup>4</sup> due to the involvement of two different operations on the cipher. As a result, we propose b-TNT2, an another variant of the TNT construction, where we blind the output of the TNT construction with an encryption of the tweak. A pictorial description of the construction is shown in Fig. 2. Unlike CLRW1<sup>4</sup>, the last block cipher call of TNT can be executed parallel to that of the TNT construction. One can also pre-computes the last block cipher which becomes advantageous while making queries with same tweak.

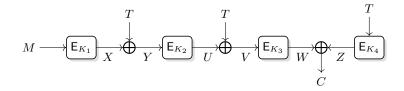


Fig. 2: b-TNT2 construction based on four block cipher calls.

In this paper, we have shown that both b-TNT1 and b-TNT2 provide security up to  $2^{3n/4}$  queries. In particular, we have the following security results, the proofs of which are deferred to Sect. 3.

**Theorem 1.** Let  $\mathsf{E} : \{0,1\}^{\kappa} \times \{0,1\}^n \to \{0,1\}^n$  be a block cipher. Then, for any (q,t) adversary  $\mathsf{A}^8$  against the strong tweakable pseudorandom permutation security of b-TNT1 with  $q \leq 2^{3n/4}$ , there exists a (q,t') adversary  $\mathsf{A}'$  against the strong pseudorandom permutation security of  $\mathsf{E}$ , where t' = t, such that

$$\mathbf{Adv}_{\mathsf{b-TNT1}}^{\text{STPRP}}(\mathsf{A}) \le 3\mathbf{Adv}_{\mathsf{E}}^{\text{SPRP}}(\mathsf{A}') + \frac{3q^2}{2^{2n}} + \frac{5q^{4/3}}{2^n} + \frac{45q^4}{2^{3n}} + \frac{1}{2^n}.$$

**Theorem 2.** Let  $\mathsf{E} : \{0,1\}^{\kappa} \times \{0,1\}^n \to \{0,1\}^n$  be a block cipher. Then, for any (q,t) adversary A against the strong tweakable pseudorandom permutation security of b-TNT2 with  $q \leq 2^{3n/4}$ , there exists a (q,t') adversary A' against the strong pseudorandom permutation security of E and a  $(\mu, t')$  adversary B against the pseudorandom permutation security of E, where  $\mu$  denotes the number of distinct tweaks queried and t' = t, such that

$$\mathbf{Adv}_{\mathsf{b}-\mathsf{TNT2}}^{\mathrm{STPRP}}(\mathsf{A}) \leq 3\mathbf{Adv}_{\mathsf{E}}^{\mathrm{SPRP}}(\mathsf{A}') + \mathbf{Adv}_{\mathsf{E}}^{\mathrm{PRP}}(\mathsf{B}) + \frac{4q^2}{2^{2n}} + \frac{6q^{4/3}}{2^n} + \frac{53q^4}{2^{3n}}$$

We have experimentally verified that both b-TNT1 and b-TNT2 perform better than the  $CLRW1^4$  in terms of throughput while achieving a similar level of security bound.

# 2 Preliminaries

4

**Notation.** 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 a natural number n and q, we write  $x^q$  to denote a q-tuple  $(x_1, x_2, \ldots, x_q)$  where each  $x_i \in \{0, 1\}^n$ . We write  $\hat{x}^q$  to denote the set  $\{x_i : i \in [q]\}$ . By an abuse of notation, we also write  $x^q$  to denote the multiset  $\{x_i : i \in [q]\}$  and  $\mu(x^q, x)$  denotes the multiplicity of  $x \in x^q$ . We also write  $\mu_x$  to denote the multiplicity of  $x \in x^q$ , when the multiset  $x^q$  is understood from the context. For a set  $\mathcal{I} \subseteq [q]$  and a q-tuple  $x^q$ , we write  $x^{\mathcal{I}}$  to denote the sub-tuple  $(x_i)_{i \in \mathcal{I}}$ . We write a 2-ary tuple  $(x^q, y^q)$  to denote the q tuple  $((x_1, y_1), (x_2, y_2), \ldots, (x_q, y_q))$ , where each  $x_i, y_i \in \{0, 1\}^n$ . We write  $x \leftarrow y$  to denote the assignment of the variable y to x.

For a random variable X,  $X \leftarrow \{0,1\}^n$  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 to all other previously sampled random variables. We

<sup>&</sup>lt;sup>8</sup> A (q, t) adversary A is one that makes a total of q queries to the oracle with running time of at most t steps.

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 Perm(n). We say that a 2-ary tuple  $(x^q, y^q)$  is permutation compatible, denoted as  $x^q \nleftrightarrow y^q$ , if there exists at least one permutation  $\mathsf{P} \in \text{Perm}(n)$  such that for all  $i \in [q], x_i = x_j \Leftrightarrow y_i = y_j, i \neq j \in [q]$ . Moreover, if  $(x^q, y^q)$  is not permutation compatible, then we denote it as  $x^q \nleftrightarrow y^q$ . For three tuples  $x^q = (x_1, x_2, \ldots, x_q), y^q = (y_1, y_2, \ldots, y_q), \text{ and } \lambda^q = (\lambda_1, \lambda_2, \ldots, \lambda_q)$  of q n-bit elements, we write  $x^q \oplus y^q = \lambda^q$ , if for all  $i \in [q]$ , it holds that  $x_i \oplus y_i = \lambda_i$ . 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.

### 2.1 (Tweakable) Block Cipher

Let  $n, \kappa, t \in \mathbb{N}$  be three 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  $k \in \{0,1\}^{\kappa}$  and an *n*-bit string  $x \in \{0,1\}^n$  and outputs an element  $y \in \{0,1\}^n$  such that for each  $k \in \{0,1\}^{\kappa}$ ,  $\mathsf{E}(k, \cdot)$  is a bijective function from  $\{0,1\}^n$  to  $\{0,1\}^n$ . A tweakable block cipher (TBC) is a mapping  $\widetilde{\mathsf{E}} : \{0,1\}^{\kappa} \times \{0,1\}^t \times \{0,1\}^n \to \{0,1\}^n$ , such that for all key  $k \in \{0,1\}^{\kappa}$  and for all tweak  $T \in \{0,1\}^t$ ,  $\widetilde{\mathsf{E}}(k,T,\cdot)$  is a permutation over  $\{0,1\}^n$ . 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  $T \in \{0,1\}^t$ ,  $\widetilde{\mathsf{P}}(T,\cdot)$  is a permutation over  $\{0,1\}^n$ . We write  $\mathsf{TP}(\{0,1\}^t,n)$  to denote the set of all tweakable permutations with tweak space  $\{0,1\}^t$  and *n*-bit messages. We fix positive even integers  $n, \kappa$  (resp. t) to denote the block size, key size (resp. tweak size) of the block cipher (resp. tweakable block cipher) respectively in terms of number of bits.

#### 2.2 Security Definition of (Tweakable) Block Cipher

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

$$\mathbf{Adv}_{\widetilde{\mathsf{E}}}^{\mathrm{STPRP}}(\mathsf{A}) \stackrel{\Delta}{=} |\Pr[\mathsf{A}^{\widetilde{\mathsf{E}}_{K}(\cdot,\cdot),\widetilde{\mathsf{E}}_{K}^{-1}(\cdot,\cdot)} = 1] - \Pr[\mathsf{A}^{\widetilde{\mathsf{P}}(\cdot,\cdot),\widetilde{\mathsf{P}}^{-1}(\cdot,\cdot)} = 1]|, \quad (1)$$

where the first probability is calculated over the randomness of  $\widetilde{\mathsf{P}} \leftarrow \{0,1\}^{\kappa}$  and the second probability is calculated over the randomness of  $\widetilde{\mathsf{P}} \leftarrow \mathsf{TP}(\{0,1\}^t,n)$ . When the adversary is given access only to the tweakable permutation and not its inverse, then we say the tweakable pseudorandom permutation (TPRP) advantage of A against  $\widetilde{\mathsf{E}}$ . We say that  $\widetilde{\mathsf{E}}$  is  $(q, t, \epsilon)$  secure if the maximum strong tweakable pseudorandom permutation advantage of  $\widetilde{\mathsf{E}}$  is  $\epsilon$  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 t. We assume throughout the paper the tweak size t 6

of the tweakable block cipher is equal to its block size n. When the tweak set is empty, then the notion of STPRP (resp. TPRP) boils down to the SPRP (resp. PRP) security.

#### 2.3 Mirror Theory For Tweakable Random Permutations

Mirror theory fundamentally works for bounding the pseudorandomness of the sum-of-permutations [3, 12, 19, 40] based constructions with respect to a random function. However, its traditional setup is not suited for bounding the pseudorandomness of tweakable block ciphers with respect to tweakable random permutation. Jha and Nandi [29] developed a variant of mirror theory result tailored for tweakable tweakable random permutations. We revisit their result below.

For a given system of linear equations  $\mathcal{L}$ , we associate an edge-labeled bipartite graph  $\mathcal{L}(G) = (\mathcal{X} \cup \mathcal{Y}, \mathcal{E})$  with the labeling function L, an edge (x, y)with label  $\lambda$  is called an *isolated-edge* if the degree of both x and y is 1. We say that a component  $\mathcal{C}$  is a *star* if  $\xi_{\mathcal{C}} \geq 3$ , where  $\xi_{\mathcal{C}}$  denotes the number of vertices in component  $\mathcal{C}$ , and there exists an unique vertex, called *center vertex*, with degree  $\xi_{\mathcal{C}} - 1$  and all the other vertices have degree exactly 1. A component  $\mathcal{C}$ is called  $\mathcal{X}$ -type (resp.  $\mathcal{Y}$ -type) if the center vertex of the component  $\mathcal{C}$  lies in  $\mathcal{X}$ (resp.  $\mathcal{Y}$ ).

For a given system of linear equations  $\mathcal{L}$  and its corresponding associated equation graph  $\mathcal{L}(G)$ , we write  $\alpha$  (resp.  $\beta$ ,  $\gamma$ ) to denote the number of isolated edges (resp. number of components of  $\mathcal{X}$ -type and number of components of  $\mathcal{Y}$ -type). Similarly,  $q_1$  denotes the number of equations such that none of its variables have collided with any other variables.  $q_2$  denotes the number of equations of  $\mathcal{X}$ -type and  $q_3$  denotes the number of equations of  $\mathcal{Y}$ -type. Note that  $\alpha = q_1$ . Following result from [29] has given a lower bound on the number of solutions for a given system of linear equations  $\mathcal{L}$  such that  $X'_i$  values are pairwise distinct and  $Y'_i$  values are pairwise distinct.

**Theorem 3.** Let  $\mathcal{L}$  be a system of the linear equation as defined above with  $q \leq 2^{n-2}$  and any component of  $\mathcal{L}(G)$  have at most  $2^{n-1}$  edge. Then the number of tuple of solution  $(x_1, x_2, \ldots, x_{q_X}, y_1, y_2, \ldots, y_{q_Y})$  of  $\mathcal{L}$ , denoted by h(q), where  $x_i \neq x_j$  and  $y_i \neq y_j$ , for all  $i \neq j$ , satisfies

$$h(q) \ge \left(1 - \frac{13q^4}{2^{3n}} - \frac{2q^2}{2^{2n}} - \left(\sum_{i=\alpha+1}^{\beta+\gamma} \zeta_i^2\right) \frac{4q^2}{2^{2n}}\right) \times \frac{(2^n)_{q_1+\beta+q_3} \times (2^n)_{q_1+q_2+\gamma}}{\prod_{\lambda \in \lambda^q} (2^n)_{\mu_\lambda}} (2)$$

where  $\zeta_i$  denote the number of edge in *i*-th component  $\forall i \in [\alpha + \beta + \gamma]$ .

### 3 Proof of Theorem 1 and Theorem 2

This section is devoted to establishing the security bound as demonstrated in Theorem 1 and Theorem 2. Due to the structural similarity of the proofs of Theorem 1 and Theorem 2, we present a combined proof of both the results.

However, we will explicitly highlight the differences between the proofs of the two constructions.

From now onwards we use the notation b-TNTd, where d = 1 stands for the construction b-TNT1 and d = 2 denotes the construction b-TNT2. Initially, we replace the three independently keyed block ciphers,  $E_{K_1}$ ,  $E_{K_2}$  and  $E_{K_3}$ , used in the constructions with three independently sampled *n*-bit random permutations, P<sub>1</sub>, P<sub>2</sub> and P<sub>3</sub> (for b-TNT2 fourth block cipher  $E_{K_4}$  will be replaced by another independently sampled *n*-bit random permutation comes at the cost of the strong pseudorandom permutation advantage of the underlying block cipher (replacement of  $E_{K_4}$  comes at the cost of pseudorandom permutation advantage). We denote the resulting construction as b-TNTd<sup>\*</sup>. Therefore, we have

$$\mathbf{Adv}_{b\text{-TNTd}}^{\mathrm{STPRP}}(\mathsf{A}) \leq \begin{cases} 3\mathbf{Adv}_{\mathsf{E}}^{\mathrm{SPRP}}(\mathsf{A}') + \overbrace{\mathbf{Adv}_{b\text{-TNTd}^{\star}}^{\mathrm{STPRP}}(\mathsf{A})}^{\delta^{*}}, \text{ for } \mathsf{d} = 1\\ 3\mathbf{Adv}_{\mathsf{E}}^{\mathrm{SPRP}}(\mathsf{A}') + \mathbf{Adv}_{\mathsf{E}}^{\mathrm{PRP}}(\mathsf{B}) + \underbrace{\mathbf{Adv}_{b\text{-TNTd}^{\star}}^{\mathrm{STPRP}}(\mathsf{A})}_{\delta^{*}}, \text{ for } \mathsf{d} = 2 \end{cases}$$

where A' is a (q, t') adversary such that t' = t. Our goal is now to upper bound  $\delta^*$ . Note that, we have

$$\delta^* \leq \max_{\mathsf{A}} \left| \Pr[\mathsf{A}^{\mathsf{b-TNTd}^*, (\mathsf{b-TNTd}^*)^{-1}} = 1] - \Pr[\mathsf{A}^{\widetilde{\mathsf{P}}, \widetilde{\mathsf{P}}^{-1}} = 1] \right|,$$

where  $\widetilde{\mathsf{P}} \leftarrow \mathsf{TP}(\{0,1\}^n, n)$ . This formulation of the problem now allows us to use the Expectation Method [23].

#### 3.1 Description of the Ideal World

The ideal world consists of two stages: in the first stage, which we call the *on*line stage, the ideal world simulates a random tweakable permutation  $\tilde{P}$ , i.e., for each encryption query (M,T), it returns  $\tilde{P}(M,T)$ . Similarly, for each decryption query (C,T), it returns  $\tilde{P}^{-1}(C,T)$ . Since the real world releases some additional information, the ideal world must generate these values as well. The ideal transcript random variable  $X_{id}$  is a 9-ary q-tuple

$$(M^{q}, T^{q}, C^{q}, X^{q}, Y^{q}, U^{q}, V^{q}, W^{q}, K(\text{for } d = 1)/Z^{q}(\text{for } d = 2))$$

defined below. However, the probability distribution of these additional random variables would be determined from their definitions. The initial transcript consists of  $(M^q, T^q, C^q)$ , where for all  $i \in [q]$ ,  $T_i$  is the *i*-th tweak value,  $M_i$  is the *i*-th plaintext value, and  $C_i$  is the *i*-th ciphertext value. Once the query-response phase is over, the next stage of the ideal world begins, which we call the *offline stage*. In the offline stage, the ideal world samples the intermediate random variables as follows: let us define the set

$$\mathbb{M}(M^q) = \{x : x = M_i, i \in [q]\}.$$

Let us assume that  $m := |\mathbb{M}(M^q)|$  be the number of distinct plaintexts. Then, it samples

$$X_{x_1}, X_{x_2}, \dots, X_{x_m} \xleftarrow{\operatorname{wor}} \{0, 1\}^n$$

where  $(x_1, x_2, \ldots, x_m)$  is an arbitrary ordering of the set  $\mathbb{M}(M^q)$ . For  $\mathsf{d} = 1$  the ideal world samples a key  $K \leftarrow \{0, 1\}^n$  independent over  $X_{x_i}$  and for  $\mathsf{d} = 2$  it samples the intermediate random variables  $Z^q$ ,

$$Z_{z_1}, Z_{z_2}, \dots, Z_{z_t} \xleftarrow{\operatorname{wor}} \{0, 1\}^n,$$

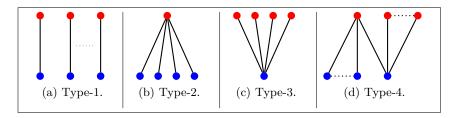
where  $(z_1, z_2, \ldots, z_t)$  is an arbitrary ordering of the set,  $\mathbb{T}(T^q) = \{z : z = T_i, i \in [q]\}$ . Let us assume that  $t := |\mathbb{T}(T^q)|$  is the distinct number of tweaks. Moreover,  $Z_{z_j}$  is independently sampled with  $X_{x_i}$ . From these sampled random variables  $(X_{x_1}, X_{x_2}, \ldots, X_{x_m})$ , we define q-tuple  $X^q$  as follows:  $X^q = (X_1, X_2, \ldots, X_q)$  such that  $X_i = X_{M_i}$  and  $Z^q = (Z_1, Z_2, \ldots, Z_q)$  such that  $Z_i = Z_{T_i}$ . Having defined q-tuple of random variables  $X^q$ , we define two q-tuples  $(Y^q, W^q)$  as follows: for each  $i \in [q]$ ,

$$Y_i = X_i \oplus T_i, \ W_i = \begin{cases} C_i \oplus (K \odot T_i), \text{ for } \mathsf{d} = 1, \\ Z_i \oplus C_i, \text{ for } \mathsf{d} = 2. \end{cases}$$

Given this partial transcript,  $X'_{id} = (M^q, T^q, C^q, X^q, Y^q, W^q, K \text{ or } Z^q)$ , we wish to define whether the sampled value  $X^q$  and  $(K \text{ or } Z^q)$  is good or bad. We say that a tuple  $(X^q, K \text{ or } Z^q)$  is **bad** if one of the following predicates hold:

- 1. Bad<sub>K</sub>:  $K = 0^n$  (This condition is only for d = 1).
- 2. Bad<sub>1</sub> (cycle of length 2):  $\exists i, j \in [q]$  such that the following holds:  $Y_i = Y_j, W_i = W_j$ .
- 3. Bad<sub>2</sub>:  $|\{(i, j) \in [q]^2 : i \neq j, Y_i = Y_j\}| \ge q^{2/3}$ .
- 4. Bad<sub>3</sub>:  $|\{(i,j) \in [q]^2 : i \neq j, W_i = W_j\}| \ge q^{2/3}$ .
- 5.  $\operatorname{Bad}_4(Y W Y \text{ path of length } 4): \exists i, j, k, l \in [q] \text{ such that the following holds: } Y_i = Y_j, W_j = W_k, Y_k = Y_l.$
- 6. Bad<sub>5</sub> (W-Y-W path of length 4):  $\exists i, j, k, l \in [q]$  such that the following holds:  $W_i = W_j, Y_j = Y_k, W_k = W_l$ .

If the sampled tuple  $(X^q, K \text{ or } Z^q)$  is bad, then  $U^q$  and  $V^q$  values are sampled degenerately, i.e.,  $U_i = V_i = 0$  for all  $i \in [q]$ . That is, we sample without maintaining any specific conditions, which may lead to inconsistencies. However, if the sampled tuple  $(X^q, K \text{ or } Z^q)$  is good, then we study a graph associated with  $(Y^q, W^q)$ . In particular, we consider the random transcript graph  $\mathcal{G}(Y^q, W^q)$ defined as follows: the set of vertices of the graph is  $Y^q \sqcup W^q$ . Moreover, we put a labeled edge between  $Y_i$  and  $W_i$  with label  $T_i$ . For two distinct indices  $i \neq j$ , if  $Y_i = Y_j$ , then we merge the corresponding vertices. Similarly, for two distinct indices, if  $W_i = W_j$ , then we merge the corresponding vertices. Therefore, the random transcript graph  $\mathcal{G}(Y^q, W^q)$  is a labeled bipartite graph. Now, we have the following lemma which asserts that the random transcript graph  $\mathcal{G}(Y^q, W^q)$ is **nice** if  $(X^q, K \text{ or } Z^q)$  is good. Fig. 3: Type-1 is a graph of isolated edges, and the maximum path length of a Type-1 graph is one. Type-2 is a star graph with Y being the centered vertex, and Type-3 is also a star graph with W being the centered vertex. The maximum path length of Type-2 and Type-3 graphs is two. Type-4 is a connected graph that is not an isolated edge or a star. It can have degree 2 vertices in both Y and W. The maximum path length of the Type-4 graph is three.



**Lemma 1.** The transcript graph  $\mathcal{G} := \mathcal{G}(Y^q, W^q)$  generated by a good tuple  $(X^q, K \text{ or } Z^q)$  is nice, i.e., it satisfies the following properties:

- G is simple, acyclic, and has no isolated vertices with no adjacent edges such that their labels are equal.
- maximum component size of  $\mathcal{G}$  is  $2q^{2/3}$  and every component of G is either a star graph, isolated edges, or contains a path of length 3.

Proof of this lemma is included in Appendix A. We depict the type of subgraphs generated from a good tuple  $(X^q, K \text{ or } Z^q)$  in Fig. 3. After describing the potential structure of random transcript graphs, we define the sampling of  $(U^q, V^q)$  when  $(X^q, K \text{ or } Z^q)$  is good. Referring to Fig. 3, we observe four types of possible random transcript graphs for a good tuple  $(X^q, K \text{ or } Z^q)$ , denoted as  $\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3$ , and  $\mathcal{G}_4$  respectively, where  $\mathcal{G}_i$  is a Type-i graph, for  $i \in [4]$ .

- $\mathcal{G}_1$  is the union of isolated edges.
- $\mathcal{G}_2$  is the union of star components containing Y as centered vertex.
- $\mathcal{G}_3$  is the union of star components containing W as centered vertex.
- $\mathcal{G}_4$  is the union of components containing at least one path of length three.

Therefore, we define for each  $b \in [4]$ ,

$$\mathcal{I}_b = \{ i \in [q] : (Y_i, W_i) \in \mathcal{G}_b \}.$$

Since, the collection of sets  $\mathcal{I}_b$  are disjoint, we have  $[q] = \mathcal{I}_1 \sqcup \mathcal{I}_2 \sqcup \mathcal{I}_3 \sqcup \mathcal{I}_4$ . We define  $\mathcal{I} = \mathcal{I}_1 \sqcup \mathcal{I}_2 \sqcup \mathcal{I}_3$ . Now, we consider the following system of equations

$$\mathcal{E} = \{ U_i \oplus V_i = T_i : i \in \mathcal{I} \},\$$

where  $U_i = U_j$  if and only if  $Y_i = Y_j$ . Similarly,  $V_i = V_j$  if and only of  $W_i = W_j$  for all  $i \neq j \in [q]$ . Thus, the solution set of  $\mathcal{E}$  is

$$\mathcal{S} = \{ (u^{\mathcal{I}}, v^{\mathcal{I}}) : u^{\mathcal{I}} \leadsto Y^{\mathcal{I}}, v^{\mathcal{I}} \leadsto W^{\mathcal{I}}, u^{\mathcal{I}} \oplus v^{\mathcal{I}} = T^{\mathcal{I}} \}.$$

Having defined the solution set for  $\mathcal{E}$ , we now define the sampling of the random variables  $(U^q, V^q)$  in the ideal world as follows:

- 10 R. Bhaumik, W. Choi, A. Dutta, C. M. López, H. Nandi, Y. Shen
- (i)  $(U^{\mathcal{I}}, V^{\mathcal{I}}) \leftarrow S$ , i.e., it uniformly samples one valid solution from the set of all valid solutions;
- (ii) For each component C of  $\mathcal{G}_4$ , let  $(Y_i, W_i) \in C$  corresponds to an edge in the component C such that the degree of both  $Y_i$  and  $W_i$  is at least 2. Then, we sample  $U_i \leftarrow \{0,1\}^n$  and set  $V_i = U_i \oplus T_i$ ;
- (iii) The final possibility is that for each edge  $(Y_i, W_i) \in \mathcal{C}$  such that  $(Y_i, W_i) \neq (Y_j, W_j)$ , where  $(Y_j, W_j) \in \mathcal{C}$ . Suppose,  $Y_i = Y_j$ , then  $U_i = U_j$  and  $V_i = U_i \oplus T_i$ . Similarly, if  $W_i = W_j$ , then  $V_i = V_j$  and  $U_i = V_i \oplus T_i$ .

Therefore, we completely define the random variable represents the ideal world transcript as follows:

 $X_{id} = (M^q, T^q, C^q, X^q, Y^q, U^q, V^q, W^q, K(\text{for } d = 1)/Z^q(\text{for } d = 2)).$ 

In this way, we achieve both the consistency of the equations in the form  $\{U_i \oplus V_i = T_i\}$  and the permutation compatibility within each component of the graph  $\mathcal{G}$  when the tuple  $(X^q, K \text{ or } Z^q)$  is good. However, we need to anticipate collisions among U values or V values across different components of the random transcript graph  $\mathcal{G}$ , which we will discuss in detail in the next section.

### 3.2 Definition and Probability of Bad Transcripts

Given the description of the transcript random variable in the ideal world, we define the set of all attainable transcripts  $\Omega$  as the set of all q tuples

$$\tau = (M^q, T^q, C^q, X^q, Y^q, U^q, V^q, W^q, K \text{ or } Z^q),$$

where  $M^q, T^q, C^q, X^q, Y^q, U^q, V^q, W^q, Z^q \in (\{0,1\}^n)^q, K \in \{0,1\}^n, Y^q = X^q \oplus T^q, W^q = C^q \oplus (K \odot T^q)$  (for d = 1) or  $W^q = Z^q \oplus C^q$  (for d = 2) and  $(M^q, T^q)$  is tweakable permutation compatible with  $(C^q, T^q)$ . Now, we will discuss what specific events constitute a bad condition.

- Consider the event  $Y^{\mathcal{I}} \stackrel{\times}{\longleftrightarrow} U^{\mathcal{I}}$  or  $W^{\mathcal{I}} \stackrel{\times}{\longleftrightarrow} V^{\mathcal{I}}$  that occurs while sampling  $(U^{\mathcal{I}}, V^{\mathcal{I}})$ , where  $\mathcal{I}$  encodes the edges that belongs to either Type-1 or Type-2 or Type-3 graphs. However, this condition cannot arise as we sample a valid solution from the set of all valid solutions  $\mathcal{S}$ ;
- Due to the sampling of  $(U^q, V^q)$ , it may so happen that  $Y^q \xleftarrow{\times} U^q$  or  $W^q \xleftarrow{\times} V^q$ .

We define transcripts to be bad depending upon the characterization of the pair of q-tuples  $(X^q, K \text{ or } Z^q)$ . Following the ideal world description, we say that a pair of q-tuples  $(X^q, K \text{ or } Z^q)$  is bad if and only if the following predicate is true:

 $\operatorname{Bad}_K \vee \operatorname{Bad}_1 \vee \operatorname{Bad}_2 \vee \operatorname{Bad}_3 \vee \operatorname{Bad}_4 \vee \operatorname{Bad}_5$  ( $\operatorname{Bad}_K$  is only for d = 1).

We say that a transcript  $\tau$  is *tuple-induced* bad transcript if  $(X^q, K \text{ or } Z^q)$  is bad, which we denote as

 $\operatorname{Bad} := \operatorname{Bad}_K \lor \operatorname{Bad}_1 \lor \operatorname{Bad}_2 \lor \operatorname{Bad}_3 \lor \operatorname{Bad}_4 \lor \operatorname{Bad}_5$  ( $\operatorname{Bad}_K$  is only for d = 1).

The other type of event that we need to discard, arises due to the bad sampling of  $(U^q, V^q)$  which causes permutation incompatibility, i.e.,  $Y^q \stackrel{\times}{\longleftrightarrow} U^q$  or  $W^q \stackrel{\times}{\longleftrightarrow} V^q$ . To bound such bad events, we need to enumerate all the conditions that results to the above inconsistencies. Note that, when the tuple  $(X^q, K \text{ or } Z^q)$  is bad, then the transcript is trivially inconsistent as we sample  $(U^q, V^q)$  degenerately. Therefore, for a good tuple  $(X^q, K \text{ or } Z^q)$ , if  $Y_i = Y_j$ or  $W_i = W_j$ , then we always have  $U_i = U_j$  or  $V_i = V_j$  respectively and hence in that case permutation inconsistencies won't arise. Therefore, we say that a transcript  $\tau$  is sampling induced bad transcript if one of the following conditions hold: for  $\alpha \in [4]$  and  $\beta \in [\alpha, 4]$ ,

-  $\operatorname{Ucoll}_{\alpha\beta}$ :  $\exists i \in \mathcal{I}_{\alpha}, j \in \mathcal{I}_{\beta}$  such that  $Y_i \neq Y_j$  and  $U_i = U_j$ ; -  $\operatorname{Vcoll}_{\alpha\beta}$ :  $\exists i \in \mathcal{I}_{\alpha}, j \in \mathcal{I}_{\beta}$  such that  $W_i \neq W_j$  and  $V_i = V_j$ .

Note that, by varying  $\alpha$  and  $\beta$  over all possible choices, we would have obtained 20 conditions, but due to the sampling mechanism of  $(U^q, V^q)$ , some of them could be immediately thrown out. For example,  $Ucoll_{11}$ ,  $Ucoll_{12}$ ,  $Ucoll_{23}$ ,  $Ucoll_{33}$  does not get satisfied. Similarly, for  $Vcoll_{\alpha\beta}$ , where  $\alpha \in [3]$  and  $\beta \in [\alpha, 3]$ . For the sake of completeness, we listed out all the 20 conditions and combine them into a single event as follows:

$$\mathsf{Bad-samp} := \bigcup_{\substack{\alpha \in [4]\\\beta \in [\alpha, 4]}} (\mathsf{Ucoll}_{\alpha, \beta} \cup \mathsf{Vcoll}_{\alpha, \beta}). \tag{3}$$

Finally, we consider a transcript  $\tau \in \Omega_{\text{bad}}$  if  $\tau$  is either *tuple-induced* bad or it is *sampling-induced* bad. All other transcripts  $\tau \in \Omega_{\text{good}} := \Omega \setminus \Omega_{\text{bad}}$  are good and it is easy to see that all good transcripts are attainable one.

**3.2.1** Bad Transcript Analysis. Now, we analyze the probability of realizing a bad transcript in the ideal world. Based on the preceding discussion, it is evident that analyzing the probability of realizing a bad transcript is only possible if either of the following two conditions, Bad or Bad-samp, occur. Therefore, we have

$$\epsilon_{\text{bad}} = \Pr[\mathsf{X}_{\text{id}} \in \Omega_{\text{bad}}] = \Pr[\mathsf{Bad} \lor \mathsf{Bad}\mathsf{-samp}] \le \Pr[\mathsf{Bad}] + \Pr[\mathsf{Bad}\mathsf{-samp}], (4)$$

where these two probabilities are calculated using the ideal world distribution of the random variables. The following two lemmas establish an upper bound on the probability of the event Bad and Bad-samp under the ideal world distribution.

**Lemma 2.** Let  $X_{id}$  and the event Bad be defined as above. Then, for any integer q such that  $q \leq 2^{n-2}$ , one has

$$\Pr[\mathsf{Bad}] \le \begin{cases} \frac{q^2}{2^{2n}} + \frac{5q^{4/3}}{2^n} + \frac{1}{2^n}, \text{ for } \mathsf{d} = 1, \\ \frac{2q^2}{2^{2n}} + \frac{6q^{4/3}}{2^n}, \text{ for } \mathsf{d} = 2. \end{cases}$$

**Lemma 3.** Let  $X_{id}$  and the event Bad-samp be defined as above. Then, for any integer q such that  $q \leq 2^{n-2}$ , one has

$$\Pr[\mathsf{Bad}\text{-samp}] \le \frac{8q^4}{2^{3n}}.$$

Following Lemma 2, Lemma 3 and Eqn. (4), we obtain the probability of bad transcripts as

$$\Pr[\mathsf{X}_{\mathrm{id}} \in \Omega_{\mathrm{bad}}] \le \begin{cases} \frac{q^2}{2^{2n}} + \frac{5q^{4/3}}{2^n} + \frac{8q^4}{2^{3n}} + \frac{1}{2^n}, \text{ for } \mathsf{d} = 1, \\ \frac{2q^2}{2^{2n}} + \frac{6q^{4/3}}{2^n} + \frac{8q^4}{2^{3n}}, \text{ for } \mathsf{d} = 2. \end{cases}$$
(5)

**3.2.2** Proof of Lemma 2. Recall that  $Bad = Bad_K \cup Bad_1 \cup Bad_2 \cup Bad_3 \cup Bad_4 \cup Bad_5$  (the condition  $Bad_K$  is only for d = 1). In this section, we bound the probability of the individual events, and then by virtue of the union bound, we sum up the individual bounds to obtain the overall bound of the probability of the event Bad.

 $\Box$  **Bounding** Bad<sub>K</sub>. Since K is sampled uniformly at random after the query response phase is over, the probability that it becomes equal to all zero string is exactly  $2^{-n}$ . Therefore, we have

$$\Pr[\operatorname{Bad}_K] = \frac{1}{2^n}.$$
(6)

 $\Box$  **Bounding Bad<sub>1</sub>.** Here we need to consider only the case when  $T_i \neq T_j$ . Note that if  $T_i = T_j$  then  $M_i \neq M_j$  and  $C_i \neq C_j$ , and hence the probability of the event is 0. Now, when  $T_i \neq T_j$ , using the randomness of  $X_i$  and  $K(\text{or } Z_i)$ , the probability of the above event can be bounded by  $1/((2^n - m) \cdot 2^n)$  (or  $1/(2^n - m)(2^n - t)$ ). Therefore, by varying over all possible choices of indices, and by assuming  $q \leq 2^{n-1}$ , we have

$$\Pr[\operatorname{Bad}_1] \le \begin{cases} q^2/2^{2n}, \text{ for } \mathsf{d} = 1, \\ 2q^2/2^{2n}, \text{ for } \mathsf{d} = 2. \end{cases}$$
(7)

 $\Box$  **Bounding Bad<sub>2</sub>.** We first bound the probability of the event **Bad<sub>2</sub>**. For a fixed choice of indices, we define an indicator random variable  $\mathbb{I}_{i,j}$  which takes the value 1 if  $Y_i = Y_j$ , and 0 otherwise. Let  $\mathbb{I} = \sum_{i \neq j} \mathbb{I}_{i,j}$ . By linearity of expectation,

$$\mathbf{E}[\mathbb{I}] = \sum_{i \neq j} \mathbf{E}[\mathbb{I}_{i,j}] = \sum_{i \neq j} \Pr[Y_i = Y_j] \le \frac{q^2}{2^n}.$$

Applying Markov's inequality, we have

$$\Pr[\operatorname{Bad}_2] = \Pr[|\{(i,j) \in [q]^2 : Y_i = Y_j\}| \ge q^{2/3}] \le \frac{q^2}{2^n} \times \frac{1}{q^{2/3}} = \frac{q^{4/3}}{2^n}.$$
 (8)

 $\Box$  Bounding Bad<sub>3</sub>. Using a similar argument as used in bounding Bad<sub>2</sub>, we have

$$\Pr[\mathsf{Bad}_3] \le \frac{q^{4/3}}{2^n}.\tag{9}$$

13

□ **Bounding** (Bad<sub>4</sub> | Bad<sub>2</sub>). Let us consider the event (Bad<sub>4</sub> | Bad<sub>2</sub>). Due to Bad<sub>2</sub>, the number of (i, j), (k, l) pairs such that  $Y_i = Y_j$  and  $Y_k = Y_l$  holds is at most  $q^{4/3}$ . For each such choices of i, j, k, l, the probability of the event  $W_j = W_k$ , i.e.,  $K \odot (T_j \oplus T_k) = C_j \oplus C_k$  (for d = 1) or  $Z_j \oplus Z_k = C_j \oplus C_k$  (for d = 2) holds with at most  $1/2^n$  (for d = 1) or  $1/(2^n - t)$  (for d = 2). This is due to the randomness of K or Z values. Therefore,

$$\Pr[\text{Bad}_4 \mid \overline{\text{Bad}_2}] \le \begin{cases} q^{4/3}/2^n, \text{ for } d = 1, \\ 2q^{4/3}/2^n, \text{ for } d = 2. \end{cases}$$
(10)

 $\Box$  Bounding (Bad<sub>5</sub> | Bad<sub>3</sub>). Using a similar argument as used above and using the randomness of X values, we can obtain

$$\Pr[\operatorname{Bad}_5 \mid \overline{\operatorname{Bad}_3}] \le \frac{2q^{4/3}}{2^n}.$$
(11)

Finally, by combining Eqn. (6), Eqn. (7), Eqn. (8), Eqn. (9), Eqn. (10) and Eqn. (11), we obtain the result.

**3.2.3** Proof of Lemma 3. Recall that from Eqn. (3) we have

$$\Pr[\mathsf{Bad-Samp}] \le \Pr\left[\bigcup_{\substack{\alpha \in [4]\\\beta \in [\alpha, 4]}} (\mathsf{Ucoll}_{\alpha, \beta} \cup \mathsf{Vcoll}_{\alpha, \beta})\right]$$
$$\le \sum_{\alpha \in [4]} \sum_{\beta \in \{\alpha, \dots, 4\}} \Pr[\mathsf{Ucoll}_{\alpha, \beta} \cup \mathsf{Vcoll}_{\alpha, \beta}].$$
(12)

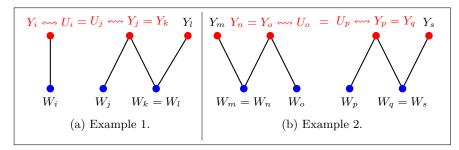
Now we will bound the probability for different values of  $(\alpha, \beta)$  as follows:  $\Box$  Case 1:  $\alpha \in [3], \beta \in [\alpha, 3]$ : In the ideal world we have done all the sampling of U and V consistently for all three  $\mathcal{I}_1, \mathcal{I}_2$  and  $\mathcal{I}_3$ . Recall that,  $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3$ . Now for any  $\alpha \in [3], \beta \in [\alpha, 3]$ , we have

$$\sum_{\alpha \in [3]} \sum_{\beta \in [\alpha, 3]} \Pr[\operatorname{Ucoll}_{\alpha, \beta} \cup \operatorname{Vcoll}_{\alpha, \beta}] = 0.$$
(13)

 $\Box$  Case 2:  $\alpha \in [3], \beta = 4$ : For this case we will analyze the probability for  $\alpha = 1 \wedge \overline{\beta} = 4$  and other five cases will attain the same bound by the same approach as bounding the probability of Vcoll<sub> $\alpha,\beta$ </sub> is similar to bounding that of Ucoll<sub> $\alpha,\beta$ </sub>. Hence we have to bound only Ucoll<sub>1,4</sub>. Example 1 in Fig. 4 illustrates the event Ucoll<sub>1,4</sub>. Recall that

$$Ucoll_{1,4} := \exists i \in \mathcal{I}_1, \ j \in \mathcal{I}_4$$
, such that  $Y_i \neq Y_j$  and  $U_i = U_j$ .

Fig. 4: These are two events where Bad-samp occurs. Example 1 indicates the event  $Ucoll_{1,4}$  i.e.  $\exists i \in \mathcal{I}_1, j \in \mathcal{I}_4$ , such that  $Y_i \neq Y_j$  and  $U_i = U_j$ . Example 2 indicates the event  $Ucoll_{4,4}$  i.e.  $\exists o \& p \in \mathcal{I}_4$ , such that  $Y_o \neq Y_p$  and  $U_o = U_p$ .



Since  $j \in \mathcal{I}_4$ , so  $Y_j - W_j$  is an edge in some component of  $\mathcal{I}_4$  say C. This C is a connected component having a path of length 3. Hence, at least one of these  $Y_j$  and  $W_j$  have degree  $\geq 2$ . Let us consider following conditions:

- (i)  $\underline{\deg(Y_j) \geq 2}$  and  $\underline{\deg(W_j) \geq 2}$ : These two vertices of degree-2 clearly implies that there exist  $k, l \neq j$  such that  $W_k (Y_k = Y_j) (W_j = W_l) Y_l$  forms a path of length 3 in C. To satisfy this case, we need  $\mathbf{E}_1 := (Y_k = Y_j \wedge W_j = W_l)$ .
- (ii)  $\underline{\deg(Y_j) \geq 2}$  and  $\underline{\deg(W_j) = 1}$ : In this case having a 3-length path implies that there exists  $k, l \neq j$  such that  $Y_l - (W_l = W_k) - (Y_k = Y_j) - W_j$  path exists in C. Hence, we need  $\mathbb{E}_2 := (Y_j = Y_k \wedge W_k = W_l)$ .
- (iii)  $\frac{\deg(Y_j) = 1 \text{ and } \deg(W_j) \ge 2}{\text{existence of } k, l \ne j \text{ such that } W_l (Y_l = Y_k) (W_k = W_j) Y_j \text{ is path in } C.$  Hence, we need  $\mathbf{E}_3 := (Y_l = Y_k \land W_k = W_j).$

Clearly from random sampling of X's and K we have

$$\forall a, b, c \in [q], \ \Pr[Y_a = Y_b \land W_b = W_c] \le \frac{2}{2^{2n}}$$

Now clearly from the definition of  $Ucoll_{1,4}$  we have

$$\Pr[\operatorname{Ucoll}_{1,4}] = \Pr[\exists i \in \mathcal{I}_1, \exists j, k, l \in \mathcal{I}_4 : U_i = U_j \land (\mathsf{E}_1 \lor \mathsf{E}_2 \lor \mathsf{E}_3)]$$

$$\leq \sum_{i \in \mathcal{I}_1} \sum_{j \neq k \neq l \in \mathcal{I}_4} \Pr[U_i = U_j] \times \Pr[\mathsf{E}_1 \lor \mathsf{E}_2 \lor \mathsf{E}_3]$$

$$\leq q \times \binom{q}{3} \times \frac{1}{2^n} \times \frac{6}{2^{2n}} \leq \frac{q^4}{2^{3n}}.$$
(14)

As stated before following a similar approach we can achieve the same bound for other five cases  $Ucoll_{2,4}, Ucoll_{3,4}, Vcoll_{\alpha,4}$ , where  $\alpha \in [3]$ . Hence

$$\sum_{\alpha \in [3]} \sum_{\beta=4} \Pr[\operatorname{Ucoll}_{\alpha,\beta} \cup \operatorname{Vcoll}_{\alpha,\beta}] \le \frac{6q^4}{2^{3n}}.$$
(15)

 $\Box$  Case 3:  $\alpha = 4, \beta = 4$ : For this case we will follow the similar approach as the previous case. Here we will bound the probability of Ucoll<sub>4,4</sub> and other case will attain the same bound by a similar approach as bounding the probability of Vcoll<sub>4,4</sub> is similar to that of bounding Ucoll<sub>4,4</sub>. Hence, we have to bound only Ucoll<sub>4,4</sub>. Example 2 in Fig. 4 illustrates the event Ucoll<sub>4,4</sub>. Recall that

$$Ucoll_{4,4} := \exists i \& j \in \mathcal{I}_4$$
, such that  $Y_i \neq Y_j$  and  $U_i = U_j$ .

Since  $j \in \mathcal{I}_4$ , so  $Y_j - W_j$  is an edge in some component of  $\mathcal{I}_4$  say C. This C is a connected component having a path of length three. Hence at least one of these  $Y_j$  and  $W_j$  have degree  $\geq 2$ . Now, following the same approach as the previous case, we will have same  $E_1, E_2, E_3$  for some  $j \neq k \neq l \in \mathcal{I}_4$ . Then we will have the same final bound

$$\Pr[\operatorname{Ucoll}_{4,4}] \le \frac{q^4}{2^{3n}}.$$

Moreover, we will have same bound for other case  $Vcoll_{4,4}$ . Hence, we have

$$\Pr[\operatorname{Ucoll}_{4,4} \cup \operatorname{Vcoll}_{4,4}] \le \frac{2q^4}{2^{3n}}.$$
(16)

The result follows by combining Eqn. (13), Eqn. (15), and Eqn. (16).

#### 3.3 Analysis of Good Transcripts

We fix a good transcript  $\tau = (M^q, T^q, C^q, X^q, Y^q, U^q, V^q, W^q, K \text{ or } Z^q)$  and we have to lower bound the real interpolation probability and upper bound the ideal interpolation probability.

**Lemma 4.** Let  $X_{re}$  (resp.  $X_{id}$ ) be the transcript random variable induced by the interaction of adversary A with the real (resp. ideal) world. For any good transcript  $\tau$  and with the notations defined above, we have

$$\frac{\Pr[\mathsf{X}_{\rm re} = \tau]}{\Pr[\mathsf{X}_{\rm id} = \tau]} \ge \left(1 - \frac{13q^4}{2^{3n}} - \frac{2q^2}{2^{2n}} - \left(\sum_{i=e_1+1}^{\xi_2+\xi_3} \zeta_i^2\right) \frac{4q^2}{2^{2n}}\right).$$
(17)

Using Eqn. (5), Lemma 4 and the Expectation Method, both theorems follow. The proof of the above lemma and the subsequent analysis can be found in Appendix B.

# 4 Experimental Results

We have implemented CLRW1<sup>4</sup>, b-TNT1, and b-TNT2 using AES-NI instructions and school book multiplication with instruction PCLMULQDQ. The target processor is an Intel Core i9-9960X at 3.10 GHz. The results in cycles per byte are shown in Table 1. The source code was compiled with GCC 10.2.1 with 03 optimization. In the target processor, one AES round takes four clock cycles as it has a skylake architecture.

Construction	Cycles	Cycles per byte
CLRW1 <sup>4</sup>	184	11.5
b-TNT1	150	9.37
b-TNT2	164	10.25
CLRW1 <sup>4</sup> *	1719	107.44
b-TNT1*	1240	77.5
b-TNT2*	1645	102.81

Table 1: Cycles and cycles per byte for proposed constructions, constructions labeled with \* also include the key schedule cost.

If the proposed construction would be implemented in hardware, b-TNT1 would be the biggest one because it needs an additional multiplier besides one AES core. A sequential AES hardware implementation takes eleven clock cycles, while a multiplication can take two or four clock cycles, depending on the desirable speed. The best option for hardware implementation is to compute the round keys on the fly; this saves registers. So, for hardware implementation, b-TNT1 is not the best option as the throughput per area is less than for the other constructions based only on a block cipher.

Table 1 shows that reducing one permutation key has a notable impact as the key schedule is very costly; this is achieved for b-TNT1 construction. Both proposed constructions improve  $CLRW1^4$  as they need fewer clock cycles. It is important to note that all the AES calls are executed sequentially in all constructions.  $CLRW1^4$ \* and b-TNT2\* use four key schedules, but b-TNT2\* has two AES calls that can be performed in parallel or pipelined. It requires fewer clock cycles than  $CLRW1^4$ \*. The best performance is for b-TNT1, as it changes one block cipher call for one multiplication.

### 5 Conclusion

In this paper, we have proposed b-TNT1 and b-TNT2 and have shown that both of them provably achieve 3n/4-bit strong tweakable permutation security. We have experimentally verified that the throughput of b-TNT1 is better than CLRW1<sup>4</sup> in tens order of magnitude. We have also experimentally validated the fact that the evaluation of the last block cipher call for b-TNT2 can be made parallel to the execution of the TNT evaluation, whereas CLRW1<sup>4</sup> enforces the evaluation of the last block cipher until the output of the TNT is available. This phenomenon allows b-TNT2 to achieve a better throughput than CLRW1<sup>4</sup>.

Acknowledgements. This research problem originated in a research discussion at the workshop ASK 2023 in Guangzhou, China. Part of the work was done when Ritam Bhaumik was affiliated with EPFL, Lausanne. Wonseok Choi was supported in part by the National Research Foundation of Korea (NRF)

grant funded by the Korea government (MSIT) (No. RS-2023-00239620), by AnalytiXIN and by Sunday Group. Yaobin Shen was supported in part by Ant Research, Ant Group.

### References

- Elena Andreeva, Andrey Bogdanov, Atul Luykx, Bart Mennink, Elmar Tischhauser, and Kan Yasuda. Parallelizable and authenticated online ciphers. In Kazue Sako and Palash Sarkar, editors, Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part I, volume 8269 of Lecture Notes in Computer Science, pages 424–443. Springer, 2013. (Cited on p. 2.)
- 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. (Cited on p. 2.)
- 3. M. Bellare and R. Impagliazzo. A tool for obtaining tighter security analyses of pseudorandom function based constructions, with applications to prp to prf conversion. Cryptology ePrint Archive, Paper 1999/024, 1999. https://eprint.iacr.org/1999/024. (Cited on p. 6.)
- 4. Davide Bellizia, Francesco Berti, Olivier Bronchain, Gaëtan Cassiers, Sébastien Duval, Chun Guo, Gregor Leander, Gaëtan Leurent, Itamar Levi, Charles Momin, Olivier Pereira, Thomas Peters, François-Xavier Standaert, Balazs Udvarhelyi, and Friedrich Wiemer. Spook: Sponge-based leakage-resistant authenticated encryption with a masked tweakable block cipher. *IACR Trans. Symmetric Cryptol.*, 2020(S1):295–349, 2020. (Cited on p. 2.)
- Arghya Bhattacharjee, Ritam Bhaumik, and Mridul Nandi. Offset-based bbbsecure tweakable block-ciphers with updatable caches. In Takanori Isobe and Santanu Sarkar, editors, Progress in Cryptology - INDOCRYPT 2022 - 23rd International Conference on Cryptology in India, Kolkata, India, December 11-14, 2022, Proceedings, volume 13774 of Lecture Notes in Computer Science, pages 171–194. Springer, 2022. (Cited on p. 2.)
- Ritam Bhaumik, Eik List, and Mridul Nandi. ZCZ Achieving n-bit SPRP Security with a Minimal Number of Tweakable-Block-Cipher Calls, page 336–366. Springer International Publishing, 2018. (Cited on p. 2.)
- 7. Debrup Chakraborty and Palash Sarkar. A general construction of tweakable block ciphers and different modes of operations. In Helger Lipmaa, Moti Yung, and Dong-dai Lin, editors, Information Security and Cryptology, Second SKLOIS Conference, Inscrypt 2006, Beijing, China, November 29 December 1, 2006, Proceedings, volume 4318 of Lecture Notes in Computer Science, pages 88–102. Springer, 2006. (Cited on p. 2.)
- Debrup Chakraborty and Palash Sarkar. A general construction of tweakable block ciphers and different modes of operations. *IEEE Trans. Inf. Theory*, 54(5):1991– 2006, 2008. (Cited on p. 2.)
- Wonseok Choi, Akiko Inoue, ByeongHak Lee, Jooyoung Lee, Eik List, Kazuhiko Minematsu, and Yusuke Naito. Highly secure nonce-based macs from the sum of tweakable block ciphers. *IACR Trans. Symmetric Cryptol.*, 2020(4):39–70, 2020. (*Cited on p. 2.*)

- 18 R. Bhaumik, W. Choi, A. Dutta, C. M. López, H. Nandi, Y. Shen
- Wonseok Choi, Jooyoung Lee, and Yeongmin Lee. Building PRFs from TPRPs: Beyond the Block and the Tweak Length Bounds. *IACR Cryptol. ePrint Arch.*, page 918, 2022. (Cited on p. 2.)
- Benoît Cogliati, Jooyoung Lee, and Yannick Seurin. New Constructions of MACs from (Tweakable) Block Ciphers. IACR Trans. Symmetric Cryptol., 2017(2):27–58, 2017. (Cited on p. 2.)
- 12. Wei Dai, Viet Tung Hoang, and Stefano Tessaro. Information-Theoretic Indistinguishability via the Chi-Squared Method. 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 497–523. Springer, 2017. (Cited on p. 6.)
- 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. (Cited on p. 2.)
- Niels Ferguson, Stefan Lucks, Bruce Schneier, Doug Whiting, Mihir Bellare, Tadayoshi Kohno, Jon Callas, and Jesse Walker. Skein. SHA-3 submission to NIST, 2006 - 2012. (Cited on p. 2.)
- 15. Robert Granger, Philipp Jovanovic, Bart Mennink, and Samuel Neves. Improved masking for tweakable blockciphers with applications to authenticated encryption. In Marc Fischlin and Jean-Sébastien Coron, editors, Advances in Cryptology EUROCRYPT 2016 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part I, volume 9665 of Lecture Notes in Computer Science, pages 263–293. Springer, 2016. (Cited on p. 2.)
- 16. Vincent Grosso, Gaëtan Leurent, François-Xavier Standaert, Kerem Varici, Anthony Journault, François Durvaux, Lubos Gaspar, and Stéphanie Kerckhof. Scream. Submission to CAESAR Competition, 2013 - 2019. https: //competitions.cr.yp.to/round2/screamv3.pdf. (Cited on p. 2.)
- 17. 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. (Cited on p. 2.)
- Chun Guo, Tetsu Iwata, Mustafa Khairallah, Kazuhiko Minematsu, and Thomas Peyrin. Romulus v1.3. Submission to NIST Lightweight Cryptography Standardization Process, 2018 - 2023. https://csrc.nist.gov/CSRC/media/Projects/ lightweight-cryptography/documents/finalist-round/updated-spec-doc/ romulus-spec-final.pdf. (Cited on p. 2.)
- Chris Hall, David Wagner, John Kelsey, and Bruce Schneier. Building PRFs from PRPs. In Hugo Krawczyk, editor, Advances in Cryptology — CRYPTO '98, pages 370–389, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg. (Cited on p. 6.)
- 20. Shoichi Hirose. Compactly committing authenticated encryption using tweakable block cipher. In Miroslaw Kutylowski, Jun Zhang, and Chao Chen, editors, Network and System Security - 14th International Conference, NSS 2020, Melbourne, VIC, Australia, November 25-27, 2020, Proceedings, volume 12570 of Lecture Notes in Computer Science, pages 187–206. Springer, 2020. (Cited on p. 2.)
- 21. Shoichi Hirose. Collision-resistant and pseudorandom hash function using tweakable block cipher. In Ilsun You and Taek-Young Youn, editors, *Information Security Applications - 23rd International Conference, WISA 2022, Jeju Island, South Ko*-

rea, August 24-26, 2022, Revised Selected Papers, volume 13720 of Lecture Notes in Computer Science, pages 3–15. Springer, 2022. (Cited on p. 2.)

- 22. Viet Tung Hoang, Ted Krovetz, and Phillip Rogaway. Robust authenticatedencryption 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. Springer, 2015. (Cited on p. 2.)
- Viet Tung Hoang and Stefano Tessaro. Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In Advances in Cryptology -CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part I, pages 3–32, 2016. (Cited on p. 7.)
- 24. 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. Springer, 2017. (Cited on p. 2.)
- 25. Jérémy Jean, Ivica Nikolic, and Thomas Peyrin. Joltik v1.3. Submission to CAESAR Competition, 2013 - 2019. https://competitions.cr.yp.to/round2/ joltikv13.pdf. (Cited on p. 2.)
- Jérémy Jean, Ivica Nikolic, and Thomas Peyrin. Kiasu v1. Submission to CAESAR Competition, 2013 - 2019. https://competitions.cr.yp.to/round1/kiasuv1. pdf. (Cited on p. 2.)
- 27. Jérémy Jean, Ivica Nikolic, Thomas Peyrin, and Yannick Seurin. The deoxys AEAD family. J. Cryptol., 34(3):31, 2021. (Cited on p. 2.)
- 28. Ashwin Jha, Mustafa Khairallah, Mridul Nandi, and Abishanka Saha. Tight security of TNT and beyond attacks, proofs and possibilities for the cascaded LRW paradigm. In Marc Joye and Gregor Leander, editors, Advances in Cryptology EUROCRYPT 2024 43rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zurich, Switzerland, May 26-30, 2024, Proceedings, Part I, volume 14651 of Lecture Notes in Computer Science, pages 249–279. Springer, 2024. (Cited on p. 2.)
- Ashwin Jha and Mridul Nandi. Tight security of cascaded LRW2. J. Cryptol., 33(3):1272–1317, 2020. (Cited on pp. 2, 6, 18, 19, 28, 29, and 30.)
- 30. 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. Springer, 2013. (Cited on p. 2.)
- 31. 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. Springer, 2012. (Cited on p. 2.)
- 32. Moses D. Liskov, Ronald L. Rivest, and David A. Wagner. Tweakable block ciphers. In Moti Yung, editor, Advances in Cryptology - CRYPTO 2002, 22nd Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-

22, 2002, Proceedings, volume 2442 of Lecture Notes in Computer Science, pages 31–46. Springer, 2002. (Cited on p. 2.)

- Moses D. Liskov, Ronald L. Rivest, and David A. Wagner. Tweakable block ciphers. J. Cryptol., 24(3):588–613, 2011. (Cited on p. 2.)
- 34. Kazuhiko Minematsu. Improved security analysis of XEX and LRW modes. In Eli Biham and Amr M. Youssef, editors, Selected Areas in Cryptography, 13th International Workshop, SAC 2006, Montreal, Canada, August 17-18, 2006 Revised Selected Papers, volume 4356 of Lecture Notes in Computer Science, pages 96–113. Springer, 2006. (Cited on p. 2.)
- 35. Yusuke Naito. Full PRF-Secure Message Authentication Code Based on Tweakable Block Cipher. In Man Ho Au and Atsuko Miyaji, editors, Provable Security - 9th International Conference, ProvSec 2015, Kanazawa, Japan, November 24-26, 2015, Proceedings, volume 9451 of Lecture Notes in Computer Science, pages 167–182. Springer, 2015. (Cited on p. 2.)
- Yusuke Naito. Tweakable blockciphers for efficient authenticated encryptions with beyond the birthday-bound security. *IACR Trans. Symmetric Cryptol.*, 2017(2):1– 26, 2017. (Cited on p. 2.)
- 37. Yusuke Naito. A highly secure MAC from tweakable blockciphers with support for short tweaks. In Julian Jang-Jaccard and Fuchun Guo, editors, Information Security and Privacy - 24th Australasian Conference, ACISP 2019, Christchurch, New Zealand, July 3-5, 2019, Proceedings, volume 11547 of Lecture Notes in Computer Science, pages 588–606. Springer, 2019. (Cited on p. 2.)
- Yusuke Naito and Takeshi Sugawara. Lightweight authenticated encryption mode of operation for tweakable block ciphers. *IACR Trans. Cryptogr. Hardw. Embed.* Syst., 2020(1):66–94, 2020. (Cited on p. 2.)
- Kazuki Nakaya and Tetsu Iwata. Generalized feistel structures based on tweakable block ciphers. *IACR Transactions on Symmetric Cryptology*, 2022(4):24–91, Dec. 2022. (Cited on p. 2.)
- 40. Jacques Patarin. Introduction to mirror theory: Analysis of systems of linear equalities and linear non equalities for cryptography. *IACR Cryptol. ePrint Arch.*, 2010:287, 2010. (Cited on p. 6.)
- 41. Thomas Peyrin and Yannick Seurin. Counter-in-tweak: Authenticated encryption modes for tweakable block ciphers. In Matthew Robshaw and Jonathan Katz, editors, Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part I, volume 9814 of Lecture Notes in Computer Science, pages 33–63. Springer, 2016. (Cited on p. 2.)
- 42. Phillip Rogaway. Efficient instantiations of tweakable blockciphers and refinements to modes OCB and PMAC. In Pil Joong Lee, editor, Advances in Cryptology - ASI-ACRYPT 2004, 10th International Conference on the Theory and Application of Cryptology and Information Security, Jeju Island, Korea, December 5-9, 2004, Proceedings, volume 3329 of Lecture Notes in Computer Science, pages 16–31. Springer, 2004. (Cited on p. 2.)
- Rich Schroeppel. Hasty pudding cipher specification. In First AES Candidate Workshop, 1998. (Cited on p. 1.)
- 44. Lei Wang. Shell v2.0. Submission to CAESAR Competition, 2013 2019. https: //competitions.cr.yp.to/round2/shellv20.pdf. (Cited on p. 2.)
- 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. (Cited on p. 2.)

# Supplementary Material

# A Proof of Lemma 1

It is easy to see that the random transcript graph  $\mathcal{G}(Y^q, W^q)$  is constructed in such a way that it contains no isolated vertices. Here we briefly justify the other properties of  $\mathcal{G}$  as follows:

- By virtue of  $\overline{\mathsf{Bad}_4} \wedge \overline{\mathsf{Bad}_5}$ , the maximum possible length of any path of  $\mathcal{G}$  is three.
- Due to  $\overline{\text{Bad}_1}$ ,  $\mathcal{G}$  contains no multiple edges or a cycle of length two. So,  $\mathcal{G}$  being a bipartite graph, the above conditions imply  $\mathcal{G}$  is simple and acyclic.
- The construction of the transcript graph  $\mathcal{G}$  implies it has no adjacent edges with equal labels.
- Owing to  $\overline{\text{Bad}_2} \wedge \overline{\text{Bad}_3}$ , the maximum component size of  $\mathcal{G}$  is  $2q^{2/3}$ . The maximum occurs when any component has a Y vertex and a W vertex linked by an edge and both of them have the maximum possible degree  $q^{2/3}$ .  $\Box$

# **B** Proof of Lemma 4 and Subsequent Analysis

We fix a good transcript  $\tau = (M^q, T^q, C^q, X^q, Y^q, U^q, V^q, W^q, K \text{ or } Z^q)$  and we have to lower bound the real interpolation probability and upper bound the ideal interpolation probability. Since the transcript is good, we know that the corresponding transcript graph  $\mathcal{G}$  is a nice graph and it is composed of the collection of components depicted in Fig. 3. From the definition of bad transcript in Sect. 3.2, we know that for a good transcript  $\tau$ , one must have

$$(M^q,T^q) \longleftrightarrow (C^q,T^q), Y^q \Longleftrightarrow U^q, W^q \Longleftrightarrow V^q, U^q \oplus V^q = T^q.$$

For  $i \in [4]$ ,  $\xi_i(\tau)$  and  $e_i(\tau)$  denotes the number of components and number of indices (corresponding to the edges), respectively, of Type-*i* graphs in  $\tau$ . Therefore, we have  $e_1(\tau) = \xi_1(\tau)$  and  $e_i(\tau) \ge 2\xi_i(\tau)$  for  $i \in \{2, 3\}$  and  $e_4(\tau) \ge 3\xi_4(\tau)$ . However, we have  $q = e_1(\tau) + e_2(\tau) + e_3(\tau) + e_4(\tau)$ . Let  $\zeta_i$  denote the number of edges in the *i*-th component. In our subsequent discussions, we will omit the parameter  $\tau$  whenever it is understood from the context. Recall that m, t denote the distinct number of plaintexts and tweaks respectively.

### **B.1** Real Interpolation Probability

In the real world,  $P_1$  is called exactly m times. Now, since the Type-1 graph is only isolated edges, so for each one of the isolated edges,  $P_2$ ,  $P_3$  is invoked once. The type-2 graph is a  $Y^*$ -star graph, which means that  $P_2$  is invoked once for every Type-2 component. However,  $P_3$  is invoked for each edge present in each of the Type-2 components. Similarly, for Type-3 graphs, which are  $W^*$ -star graph,  $P_3$  is invoked once for every Type-3 component. However,  $P_2$  is invoked for each edges present in each of the Type-3 components. Suppose, for Type-4 graph  $P_2$  is invoked  $t_1$  times. Since  $e_4$  is the number of indices (corresponding to the edges) of the Type-4 graph, therefore,  $P_3$  is invoked  $(e_4 - t_1 + \xi_4)$  times for the Type-4 graph. So,  $P_2$  is called exactly  $e_1 + \xi_2 + e_3 + t_1$  times and  $P_3$  is called exactly  $e_1 + \xi_3 + e_2 + e_4 - t_1 + \xi_4$  times.  $P_4$  is called exactly t times (for d = 2 only).

Therefore, the real interpolation probability is

$$\Pr[\mathsf{X}_{\rm re} = \tau] = \begin{cases} \frac{1}{(2^n)_m} \cdot \frac{1}{(2^n)_{e_1 + \xi_2 + e_3 + t_1}} \cdot \frac{1}{(2^n)_{e_1 + \xi_2 + e_3 + t_1}} \cdot \frac{1}{(2^n)_{e_1 + \xi_3 + e_2 + e_4 - t_1 + \xi_4}}, \text{ for } \mathsf{d} = 1\\ \frac{1}{(2^n)_m} \cdot \frac{1}{(2^n)_{e_1 + \xi_2 + e_3 + t_1}} \cdot \frac{1}{(2^n)_{e_1 + \xi_3 + e_2 + e_4 - t_1 + \xi_4}} \cdot \frac{1}{(2^n)_t}, \text{ for } \mathsf{d} = 2 \end{cases}$$
(18)

#### **B.2** Ideal Interpolation Probability

In the ideal world, the sampling of the random variables is done in three parts: in the first part, i.e., in the online stage of the sampling algorithm, it simulates a tweakable random permutation. Let  $(T_1, T_2, \ldots, T_t)$  denotes the tuple of distinct tweaks in  $T^q$  and for all  $i \in [t]$ , we have  $d_i = \mu(T^q, T_i)$ , i.e.,  $t \leq q$  and we have  $\sum_{i=1}^t d_i = q$ . Then, we have

$$\Pr[\widetilde{\mathsf{P}}(T^{q}, M^{q}) = C^{q}] = \prod_{i=1}^{t} \frac{1}{(2^{n})_{d_{i}}}$$
(19)

In the next stage of the sampling process, it samples the intermediate random variables. First, it samples the value  $X^q$  in without replacement manner, i.e.,  $X_i = X_j$  if and only if  $M_i = M_j$ . Then it samples K (for d = 1) or  $Z^q$  (for d = 2) independently. Since there are m distinct plaintexts and t distinct tweaks. Therefore, for any pair of q-tuples  $(x^q, k)$  or  $(x^q, z^q)$ , we have

$$\Pr[(X^{q}, K) = (x^{q}, k)] = \frac{1}{(2^{n})_{m}} \cdot \frac{1}{(2^{n})}, \text{ (for } d = 1)$$
  
or  
$$\Pr[(X^{q}, Z^{q}) = (x^{q}, z^{q})] = \frac{1}{(2^{n})_{m}} \cdot \frac{1}{(2^{n})_{t}}, \text{ (for } d = 2)$$
(20)

Now, we sample the intermediate random variables  $(U^q, V^q)$  in the following two stages:

- **Type-1, Type-2, Type-3 Sampling:** Recall that, we have defined three sets  $\mathcal{I}_1, \mathcal{I}_2$ , and  $\mathcal{I}_3$  such that  $i \in \mathcal{I}_b$  implies the edge  $(Y_i, W_i)$  belongs to Typeb graph, for  $b \in \{1, 2, 3\}$ . Recall that, we have defined the set  $\mathcal{I} = \mathcal{I}_1 \sqcup \mathcal{I}_2 \sqcup \mathcal{I}_3$ and the following system of equations

$$\mathcal{E} = \{ U_i \oplus V_i = \lambda_i : i \in \mathcal{I} \}.$$

Let  $(\lambda_1, \lambda_2, \dots, \lambda_s)$  denotes the tuple of distinct elements in  $\lambda^{\mathcal{I}}$ , and for all  $i \in [s]$ , we denote  $g_i = \mu(\lambda^{\mathcal{I}}, \lambda_i)$ . Note that, as the transcript is good, the

system of equations  $\mathcal{E}$  does not contain any cycle and is non-degenerate. Moreover, the maximum component size  $\xi_{\max}(\mathcal{E})$  is at most  $q^{2/3}$  due to  $\overline{\mathsf{Bad}_2}$ and  $\overline{\mathsf{Bad}_3}$ . Therefore, we apply Theorem 3 to lower bound on the number of valid solutions,  $|\mathcal{S}|$  for  $\mathcal{E}$ . Since, we sample  $(U^{\mathcal{I}}, V^{\mathcal{I}}) \leftarrow S$  and by virtue of Theorem 3, we have

$$\Pr[(U^{\mathcal{I}}, V^{\mathcal{I}}) = (u^{\mathcal{I}}, v^{\mathcal{I}})] \le \frac{\prod_{i=1}^{s} (2^{n})_{g_{i}}}{\Delta \cdot (2^{n})_{e_{1} + \xi_{2} + e_{3}} (2^{n})_{e_{1} + e_{2} + \xi_{3}}},$$
(21)

where

$$\Delta \stackrel{\Delta}{=} \left( 1 - \frac{13q^4}{2^{3n}} - \frac{2q^2}{2^{2n}} - \left( \sum_{i=e_1+1}^{\xi_2 + \xi_3} \zeta_i^2 \right) \frac{4q^2}{2^{2n}} \right). \tag{22}$$

- **Type-4 Sampling:** For the indices belongs to  $\mathcal{I}_4$ , a single value is sampled uniformly for each of the components, i.e., we have

$$\Pr[(U^{[q]\setminus\mathcal{I}}, V^{[q]\setminus\mathcal{I}}) = (u^{[q]\setminus\mathcal{I}}, v^{[q]\setminus\mathcal{I}})] = \frac{1}{(2^n)^{\xi_4}},$$
(23)

By combining Eqn. (27), Eqn. (28), Eqn. (29), and Eqn. (31), we have

$$\Pr[\mathsf{X}_{id} = \tau] \leq \begin{cases} \prod_{i=1}^{t} \frac{1}{(2^{n})_{d_{i}}} \cdot \frac{1}{(2^{n})_{m}} \cdot \frac{1}{(2^{n})} \cdot \frac{\prod_{i=1}^{s} (2^{n})_{g_{i}}}{\Delta \cdot (2^{n})_{e_{1} + \xi_{2} + e_{3}} (2^{n})_{e_{1} + e_{2} + \xi_{3}} (2^{n})^{\xi_{4}}}, \text{for } \mathsf{d} = 1\\ \prod_{i=1}^{t} \frac{1}{(2^{n})_{d_{i}}} \cdot \frac{1}{(2^{n})_{m}} \cdot \frac{1}{(2^{n})_{t}} \cdot \frac{\prod_{i=1}^{s} (2^{n})_{g_{i}}}{\Delta \cdot (2^{n})_{e_{1} + \xi_{2} + e_{3}} (2^{n})_{e_{1} + e_{2} + \xi_{3}} (2^{n})^{\xi_{4}}}, \text{for } \mathsf{d} = 2 \end{cases}$$

$$\tag{24}$$

#### B.3 Ratio of Real to Ideal Interpolation Probability

By taking the ratio of Eqn. (26) to Eqn. (32), we have the following:

$$\begin{split} \frac{\Pr[\mathsf{X}_{\mathrm{re}}=\tau]}{\Pr[\mathsf{X}_{\mathrm{id}}=\tau]} \geq \begin{cases} \prod_{i=1}^{t} (2^{n})_{d_{i}} \cdot \Delta \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} \cdot (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})^{\xi_{4}+1}}{\prod_{i=1}^{s} (2^{n})_{g_{i}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}+t_{1}} \cdot (2^{n})_{e_{1}+e_{2}+\xi_{3}+e_{4}-t_{1}+\xi_{4}}}, & \text{for } \mathsf{d}=1\\ \prod_{i=1}^{t} (2^{n})_{d_{i}} \cdot \Delta \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} \cdot (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})^{\xi_{4}}}{\prod_{i=1}^{s} (2^{n})_{g_{i}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} \cdot (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})_{e_{4}}}, & \text{for } \mathsf{d}=2\\ \geq \frac{\prod_{i=1}^{t} (2^{n})_{f_{i}} \cdot \prod_{i=1}^{t} (2^{n}-f_{i})_{d_{i}-f_{i}} \cdot \Delta \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})^{\xi_{4}}}{\prod_{i=1}^{s} (2^{n})_{g_{i}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}+t_{1}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} \cdot (2^{n})^{\xi_{4}}}, \\ \geq \frac{\prod_{i=1}^{t} (2^{n})_{f_{i}} \cdot \prod_{i=1}^{t} (2^{n}-f_{i})_{d_{i}-f_{i}} \cdot \Delta \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})^{\xi_{4}}}{\prod_{i=1}^{s} (2^{n})_{g_{i}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}+t_{1}} \cdot (2^{n})_{e_{1}+\xi_{2}+e_{3}} \cdot (2^{n})_{e_{1}+e_{2}+\xi_{3}} \cdot (2^{n})^{\xi_{4}}}, \\ \begin{pmatrix} (1) \\ \geq \Delta \cdot \underbrace{\prod_{i=1}^{s} (2^{n}-f_{i})_{d_{i}-f_{i}}} \\ (2^{n}-e_{1}-\xi_{2}-e_{3})_{t_{1}} (2^{n}-e_{1}-e_{2}-\xi_{3}-\xi_{4})_{e_{4}-t_{1}}}, \\ \rho \end{pmatrix} \end{cases}$$

where  $f_i = \mu(T^{\mathcal{I}}, T^i), i \in [t]$ . As the number of distinct internal masking values  $\lambda_i$  is at most the number of distinct tweaks  $T_i$  which implies that  $t \geq s$  and by the virtue of the Definition 2.1 of [29],  $\hat{T}^{\mathcal{I}}$  compresses <sup>9</sup> to  $\hat{\lambda}^{\mathcal{I}}$ . Hence, following Proposition 1 of [29], inequality (1) holds.

**Proposition 1 ( [29]).** For  $r \geq s$ , let  $a = (a_i)_{i \in [r]}$  and  $b = (b_i)_{i \in [s]}$  be two sequences over  $\mathbb{N}$  such that a compresses to b. Then, for any n, such that,  $2^n \geq \sum_{i=1}^r a_i$  holds, we have  $\prod_{i=1}^r (2^n)_{a_i} \geq \prod_{j=1}^s (2^n)_{b_j}$ .

Moreover, from the following claim, we have  $\rho \geq 1$ . Finally, by plugging-in the value of  $\Delta$  from Eqn. (30), we have

$$\frac{\Pr[\mathsf{X}_{\rm re} = \tau]}{\Pr[\mathsf{X}_{\rm id} = \tau]} \ge \left(1 - \frac{13q^4}{2^{3n}} - \frac{2q^2}{2^{2n}} - \left(\sum_{i=e_1+1}^{\xi_2 + \xi_3} \zeta_i^2\right) \frac{4q^2}{2^{2n}}\right).$$
(25)

**Claim 1.** With the notations defined above,  $\rho \geq 1$ .

**Proof.** Note that,  $f_i$  denotes the multiplicity of the *i*-th tweak in the tuple  $T^{\mathcal{I}}$ . Hence, by definition, the multiplicity cannot be more than the number of components of the Type-1, Type-2, and Type-3 graphs as each of the components

<sup>&</sup>lt;sup>9</sup> Definition 2.1 of [29] says that a sequence  $(a_i)_{i \in [r]}$  compresses to an another sequence  $(b_i)_{i \in [s]}$ , where both the sequences are defined over  $\mathbb{N}$  if there exists a partition  $\mathcal{P}$  of [r] such that it contains exactly s classes  $\mathcal{P}_1, \ldots, \mathcal{P}_s$  and for all  $i \in [s]$ , we have  $b_i = \sum_{j \in \mathcal{P}_i} a_j$ 

of Type-2 and Type-3 graphs have distinct tweak values. Therefore,  $f_i \leq e_1 + \xi_2 + \xi_3 \leq e_1 + \xi_2 + e_3$ . Similarly,  $f_i \leq e_1 + \xi_2 + \xi_3 \leq e_1 + e_2 + \xi_3 + \xi_4$ . Note that,  $d_i$  denotes the multiplicity of the *i*-th tweak in the tuple  $T^q$ . Therefore,  $d_i$  cannot be more than the number of components of Type-1, Type-2, and Type-3 graph and twice that of the number of components of Type-4 graph. Therefore,

$$d_i \le \xi_1 + \xi_2 + \xi_3 + 2\xi_4 \le e_1 + \xi_2 + e_3 + t_1$$
  
$$d_i \le e_1 + e_2 + \xi_3 + e_4 - t_1 + \xi_4.$$

Moreover, it is easy to verify that  $\sum_{i=1}^{t} (d_i - f_i) = e_4$  as the total multiplicity of tweaks  $T \in T^{[q] \setminus \mathcal{I}}$  is exactly the number of edges in components of Type-4 graph. Therefore, we have the condition that

$$f_i \le e_1 + \xi_2 + e_3$$

$$f_i \le e_1 + e_2 + \xi_3 + \xi_4$$

$$d_i \le e_1 + \xi_2 + e_3 + t_1$$

$$d_i \le e_1 + e_2 + \xi_3 + e_4 - t_1 + \xi_4$$

$$\sum_{i=1}^{t} (d_i - f_i) = t_1 + (e_4 - t_1) = e_4.$$

The above conditions satisfy the conditions given in Proposition 2 of [29] and hence by virtue of Proposition 2 of [29], the result follows.  $\Box$ 

**Proposition 2** ([29]). For  $r \geq 2$ , let  $c = (c_i)_{i \in [r]}$  and  $d = (d_i)_{i \in [r]}$  be two sequences over  $\mathbb{N}$ . Let  $a_1, a_2, b_1, b_2 \in \mathbb{N}$  such that  $c_i \leq a_j, c_i + d_i \leq a_j + b_j$  for all  $i \in [r], j \in [2]$ , and  $\sum_{i=1}^r d_i = b_1 + b_2$ . Then, for any  $n \in \mathbb{N}$ , such that  $a_j + b_j \leq 2^n$  for  $j \in [2]$ , we have

$$\prod_{i=1}^{r} (2^n - c_i)_{d_i} \ge (2^n - a_1)_{b_1} (2^n - a_2)_{b_2}$$

Let  $\sim_Y$  be the equivalence relation over [q] defined as  $i \sim_Y j$  if and only if  $Y_i = Y_j$ . Similarly,  $\sim_W$  be the equivalence relation over [q] defined as  $i \sim_W j$  if and only if  $W_i = W_j$ . Note that, each  $\zeta_i$  is the random variable that corresponds to the cardinality of some non-singleton equivalence classes corresponding to the equivalence relation  $\sim_Y$  or  $\sim_W$ . Let  $\mathsf{E}_1, \mathsf{E}_2, \ldots, \mathsf{E}_y$  be the equivalence classes corresponding to the equivalence classes corresponding to the equivalence relation  $\sim_Y$ . Similarly,  $\mathsf{F}_1, \mathsf{F}_2, \ldots, \mathsf{F}_w$  be the equivalence classes corresponding to the equivalence relation  $\sim_Y$ . Similarly,  $\mathsf{F}_1, \mathsf{F}_2, \ldots, \mathsf{F}_w$  be the equivalence classes corresponding to the equivalence relation  $\sim_W$ . For every  $i \in [y]$ , let  $\nu_i = |\mathsf{E}_i|$  and for every  $i \in [w]$ , let  $\nu'_i = |\mathsf{F}_i|$ . In other words,  $\nu_i$  denotes the number of occurrences of  $Y_i$  and  $\nu'_i$  denotes the number of occurrences of  $W_i$ . We define  $\mathsf{coll}_Y$  to denote the number of colliding pairs in  $Y^q$ . Similarly, we define  $\mathsf{coll}_W$  to denote the number of colliding pairs in  $W^q$ . Then, we have the following lemma:

**Lemma 5** ([29]). Since  $\mathbf{E}[\operatorname{coll}_Y] \leq \frac{q^2}{2^n}$ ,  $\mathbf{E}[\operatorname{coll}_W] \leq \binom{q}{2}/2^n$  (for d = 1) or  $\mathbf{E}[\operatorname{coll}_W] \leq \frac{q^2}{2^n}$  (for d = 2), it holds

$$\begin{split} \mathbf{E}\left[\sum_{i=1}^{y}\nu_{i}^{2}\right] &= 2.\mathbf{E}[\operatorname{coll}_{Y}] + \sum_{i=1}^{y}\nu_{i} \leq 4\mathbf{E}[\operatorname{coll}_{Y}] \leq 4q^{2}/2^{n} \\ \mathbf{E}\left[\sum_{i=1}^{w}\nu_{i}^{\prime 2}\right] &= 2.\mathbf{E}[\operatorname{coll}_{W}] + \sum_{i=1}^{w}\nu_{i}^{\prime} \leq 4\mathbf{E}[\operatorname{coll}_{W}] \leq \begin{cases} 2q^{2}/2^{n}, \text{ for } \mathsf{d} = 1\\ 4q^{2}/2^{n}, \text{ for } \mathsf{d} = 2 \end{cases} \end{split}$$

It is easy to see that the expected number of colliding pairs in  $Y^q$  is  $2 \cdot {\binom{q}{2}}/{2^n}$ , as for a fixed choice of pairs (i, j), the probability that  $Y_i = Y_j$  holds with probability at most  $2/2^n$  due to the randomness of  $\mathsf{P}_1$ . Similarly, the expected number of colliding pairs in  $W^q$  is  ${\binom{q}{2}}/{2^n}$ , as for a fixed choice of pairs (i, j), the probability that  $W_i = W_j$  holds with probability at most  $2^{-n}$  (for  $\mathsf{d} = 1$ ) or  $2/2^n$ (for  $\mathsf{d} = 2$ ) due to the randomness of K (for  $\mathsf{d} = 1$ ) or  $\mathsf{P}_4$  (for  $\mathsf{d} = 2$ ). Therefore, due to the fact that  $X^q$  and K (or  $Z^q$ ) are independently sampled (justifies inequality (2)) and from Lemma 5 (justifies inequality (3)), the following holds.

$$\mathbf{E}\left[\sum_{i=e_1+1}^{\xi_2+\xi_3}\zeta_i^2\right] \stackrel{(2)}{\leq} \mathbf{E}\left[\sum_{i=1}^y \nu_i^2\right] + \mathbf{E}\left[\sum_{i=1}^w \nu_i'^2\right] \stackrel{(3)}{\leq} \begin{cases} \frac{6q^2}{2^n}, \text{ for } \mathsf{d} = 1\\ \frac{8q^2}{2^n}, \text{ for } \mathsf{d} = 2 \end{cases}$$
(26)

Finally, by combining Eqn. (5), Eqn. (17), Eqn. (34), and by following the Expectation Method, we have

$$\delta^{*} \leq \left(\frac{q^{2}}{2^{2n}} + \frac{5q^{4/3}}{2^{n}} + \frac{8q^{4}}{2^{3n}} + \frac{1}{2^{n}}\right) + \left(\frac{13q^{4}}{2^{3n}} + \frac{2q^{2}}{2^{2n}} + \mathbf{E}\left[\left(\sum_{i=e_{1}+1}^{\xi_{2}+\xi_{3}}\zeta_{i}^{2}\right)\right]\frac{4q^{2}}{2^{2n}}\right)$$

$$\leq \frac{3q^{2}}{2^{2n}} + \frac{5q^{4/3}}{2^{n}} + \frac{45q^{4}}{2^{3n}} + \frac{1}{2^{n}}. \quad \text{(for } \mathsf{d} = 1, \text{ i.e. for } \mathsf{b}\mathsf{-}\mathsf{TNT1}\text{)} \quad (27)$$
or
$$\delta^{*} \leq \left(\frac{2q^{2}}{2^{2n}} + \frac{6q^{4/3}}{2^{n}} + \frac{8q^{4}}{2^{3n}}\right) + \left(\frac{13q^{4}}{2^{3n}} + \frac{2q^{2}}{2^{2n}} + \mathbf{E}\left[\left(\sum_{i=e_{1}+1}^{\xi_{2}+\xi_{3}}\zeta_{i}^{2}\right)\right]\frac{4q^{2}}{2^{2n}}\right)$$

$$\leq \frac{4q^{2}}{2^{2n}} + \frac{6q^{4/3}}{2^{n}} + \frac{53q^{4}}{2^{3n}}. \quad \text{(for } \mathsf{d} = 2, \text{ i.e. for } \mathsf{b}\mathsf{-}\mathsf{TNT2}\text{)} \quad (28)$$