A Long Tweak Goes a Long Way: High Multi-user Security Authenticated Encryption from Tweakable Block Ciphers

Benoît Cogliati\textsuperscript{1}, Jérémy Jean\textsuperscript{2}, Thomas Peyrin\textsuperscript{3} and Yannick Seurin\textsuperscript{4}

\textsuperscript{1} Thales DIS France SAS, France
\textsuperscript{2} ANSSI, France
\textsuperscript{3} Nanyang Technological University, Singapore
\textsuperscript{4} Ledger, France

Abstract. We analyze the multi-user (mu) security of a family of nonce-based authentication encryption (nAE) schemes based on a tweakable block cipher (TBC). The starting point of our work is an analysis of the mu security of the SCT-2 mode which underlies the nAE scheme Deoxys-II, winner of the CAESAR competition for the defense-in-depth category. We extend this analysis in two directions, as we detail now.

First, we investigate the mu security of several TBC-based variants of the counter encryption mode (including CTRT, the encryption mode used within SCT-2) that differ by the way a nonce, a random value, and a counter are combined as tweak and plaintext inputs to the TBC to produce the keystream blocks that will mask the plaintext blocks. Then, we consider the authentication part of SCT-2 and study the mu security of the nonce-based MAC Nonce-as-Tweak (NaT) built from a TBC and an almost universal (AU) hash function. We also observe that the standard construction of an AU hash function from a (T)BC can be proven secure under the assumption that the underlying TBC is unpredictable rather than pseudorandom, allowing much better conjectures on the concrete AU advantage. This allows us to derive the mu security of the family of nAE modes obtained by combining these encryption/MAC building blocks through the NSIV composition method (including SCT-2).

These modes require an underlying TBC with a larger tweak length than what is usually available for existing ones. We then show the practicality of these modes by instantiating them with two new TBC constructions, Deoxys-TBC-512 and Deoxys-TBC-640, which can be seen as natural extensions of the Deoxys-TBC family to larger tweak lengths. Designing such TBCs with unusually large tweaks is prone to pitfalls: Indeed, we show that a large-tweak proposal for SKINNY published at EUROCRYPT 2020 presents an inherent construction flaw. We therefore provide a sound design strategy to construct large-tweak TBCs within the Superposition Tweakey (STK) framework, leading to new Deoxys-TBC and SKINNY variants. We provide software benchmarks indicating that while ensuring a very high security level, the performances of our proposals remain very competitive.

Keywords: authenticated encryption · tweakable block cipher · multi-user security

E-mail: benoit.cogliati@gmail.com (Benoît Cogliati), jeremy.jean@ssi.gouv.fr (Jérémy Jean), thomas.peyrin@ntu.edu.sg (Thomas Peyrin), yannick.seurin@m4x.org (Yannick Seurin)

This work is licensed under a “CC BY 4.0” license.

Date of this document: 2024-07-09.
1 Introduction

1.1 Background

Multi-user Security for AE. Authenticated Encryption (AE), providing confidentiality and integrity in a single primitive, is arguably among the most widely used components in applied cryptography. Its usage is widespread to protect all modern communication protocols, such as TLS, IPSec, SSH, etc. Traditionally, designers considered mainly single-user security, where an attacker is restrained to target only a single key. However, one can witness a growing concern towards large-scale adversaries, like state actors, which can potentially try to attack a large number of users at the same time, for example for mass-surveillance purpose. We observe more efforts conducted recently on multi-user security research and increased attention to these threats in standardization bodies discussions.

Multi-user (mu) security (sometimes called multi-key security) was first formally introduced for PRFs (as a technical tool) by Bellare, Canetti, and Krawczyk [BCK96] and later for public-key encryption (as a full-fledged security goal) by Bellare, Boldyreva, and Micali [BBM00]. In the mu setting, the attacker can distribute its resources to attack multiple users (all having independent keys) and is considered successful if it compromises at least one of them. In symmetric cryptography, multi-user attacks have been studied for block ciphers [Bih02,FJM14] and for stream ciphers [PPS15], and provable security results have been obtained for block ciphers [ML15,Tes15,HT16,HT17,GW18] and MACs [ADMA15,BBT16,SWGW21]. The study of the mu setting for AE has been first explored in [BT16] for the “randomized nonce” mechanism proposed for the Galois Counter-Mode with AES (AES-GCM) in TLS 1.3, an analysis later improved with tighter bounds in [LMP17,HTT18]. Results have also been obtained for the keyed duplex construction [DMA17] and AES-GCM-SIV [BHT18].

Single-user (su) security implies multi-user security, however the best generic reduction implies a security loss of $u$, where $u$ is the number of users that the adversary can simultaneously attack. Sometimes this loss is unavoidable, meaning there exists a matching attack, for example in the case of key-recovery attacks against block ciphers [Bih02]. However, for some schemes it can be shown that security does not degrade substantially when going from the su to the mu setting [ML15,Tes15,BBT16,HT16,HT17,LMP17]. In other words, in such a case, the adversary does not gain much from having the opportunity to allocate its attack resources (such as the total number of queries to oracles available in the security game) across multiple users.

Tweakable Block Ciphers. Tweakable block ciphers (TBC) are block ciphers with an extra public input, the tweak, that can be used to randomize the family of permutations defined. The first published TBC was the AES competition candidate Hasty Pudding Cipher [Sch98], but the first formal treatment of TBCs was due to Liskov et al. [LRW11]. We can generally observe two main strategies to build a TBC: either from an existing BC or directly with a dedicated design [Cro01,JNP14,BJK+16,Ava17,BLLS22]. Since their emergence, TBCs have proven to be very flexible and attractive primitives, leading to efficient AE operating modes with stronger security guarantees than classical BC-based ones [RBBK01,KR11,PS16,IMPS17,NS19,IKMP20,NS20a,NS20b]. Indeed, a recurrent issue with BC-based AE operating modes is that security is usually only guaranteed up to the birthday bound [Fer02], i.e., up to $2^{n/2}$ queries for a $n$-bit block cipher. While this might be sufficient for most practical cases using a 128-bit cipher such as AES, future applications might suffer from such limitation. This problem is even more present when multiple users are considered, which can further reduce the security guarantees. TBC-based AE modes, in contrary, would usually provide beyond-birthday bound (BBB) security, while retaining very good performances. It is therefore natural to consider the usage of TBCs to try to tackle efficiently the problem of multi-user security as well.
1.2 Our Contributions

Our contributions are two-fold. On one hand, we analyze the multi-user security of a family of nonce-based authenticated encryption (nAE) modes of operation for TBCs that follow the SCT-2 blueprint. On the other hand, we design new TBCs for instantiating these modes which require a large tweak length.

**Multi-user Analysis of nAE Modes for TBCs.** We consider a family of nAE modes called GNSIV that follow the NSIV composition paradigm [PS16] which combines a nonce-based PRF/MAC and a nonce and IV-based encryption scheme to obtain a nonce-misuse resistant nAE scheme [RS06]. Nonce-misuse resistance (meaning security does not break dramatically if a nonce is mistakenly repeated) has become an important security goal, as shown by a number of recent attacks [BZD+16,VP17,VP18]. The family of nAE modes we consider includes in particular the SCT-2 mode which underlies the nAE scheme Deoxys-II, selected as first choice in the portfolio of the CAESAR competition for the defense-in-depth category, and whose only existing security analysis is in the single-user setting [JNPS21,ABPV21]. For the encryption part, we consider a family of TBC-based variants of the counter encryption mode (including the CTRT [PS16] mode used in SCT-2) called GCTR combining a nonce, a random IV (to be generated pseudorandomly by the PRF applied to the inputs of the nAE mode once embedded in the NSIV construction), and a counter to define the tweak and plaintext inputs to the TBC to produce the keystream blocks. Our analysis can be seen as an extension of the recent work of Andreeva et al. [ABPV21], who carried out a similar investigation of TBC-based counter-like encryption modes, albeit limited to the single-user setting. Our proof is in the ideal (tweakable) cipher model, which is customary for multi-user analysis [BT16,BHT18,HTT18] as it allows to capture the offline computations performed by the adversary. For a TBC with key length $k$ and block length $n$, our security bounds show that some variants ensure security up to roughly $2^k$ ideal cipher queries and $2^n$ encryption queries in the nonce-respecting setting, with graceful security degradation as nonces are repeated (meaning security does not collapse at the first nonce repetition but instead deteriorates continuously as a function of the number of nonce repetitions), independently from the number of users. In other words, security does not suffer from collisions between user keys, something that can only be achieved for randomized cryptographic schemes.

Then, we consider the authentication part, for which we focus on the Nonce-as-Tweak (NaT) scheme [CLS17], a simple nonce-based PRF/MAC constructed from an almost universal (AU) hash function $H$ and a TBC $E$ and that follows the classical “Hash-then-PRF” paradigm. Again, our findings (in the ideal cipher model for $E$) indicate that NaT security does not degrade in the multi-user setting and that it provides security up to roughly $2^k$ ideal cipher queries and $2^n$ tag/verification queries in the nonce-respecting setting, with graceful security degradation as nonces are repeated. By combining NaT and any instance of the GCTR encryption mode through the NSIV composition method, we obtain an instance of the GNSIV family of nAE schemes (see Figure 1 for a high-level view of the mode). Note that this analysis assumes a generic AU hash function $H$. We then turn to the problem of instantiating $H$ from the same underlying TBC as the other components.

**Universal Hashing from Unpredictable TBCs.** Equipped with our results regarding NaT with a generic AU hash function, we consider the question of how to instantiate it from a TBC.\footnote{We could use statistical AU hash functions based for example on polynomial hashing, but we aim at a purely TBC-based design.} The standard construction consists in concatenating the input blocks with a counter, applying the TBC (or more generally any keyed function), and xoring the outputs,
something we refer informally to as the Xor-Hash construction, and whose origin can be traced back to Bellare et al. [BGR95]. Slight variants of this construction are the basis of multiple “Hash-then-PRF” MACs such as Protected Counter Sum [Ber99], PMAC [BR02], or LightMAC [LPTY16]. Usually, these constructions are proven secure under the assumption that the underlying primitive is pseudorandom. Indeed, it is not hard to prove that Xor-Hash is $\delta$-AU (i.e., for any pair of distinct inputs $X$ and $X'$, $H_K(X) = H_K(K)$ with probability at most $\delta$) for $\delta = 2^{-n} + \epsilon_{prp}$ (see for example [BS20], Section 7.2.3). However, we face the problem that the NaT security bound contains a term $\mu q \delta$, where $\mu$ is the maximal number of repetitions of nonces for any user and $q$ is the total number of tag/verification queries made by the adversary. Yet for a TBC (or any keyed deterministic function) with a $k$-bit key, there exist non-uniform attacks that can distinguish it from random with advantage $2^{-k/2}$ in constant time and queries [DTT10,BL13]. As a result, provable security for NaT caps at $q \approx 2^{k/2}$ even in the nonce-respecting setting, short of our objectives to deliver $k$-bit security. We overcome this hurdle by proving that Xor-Hash is almost universal assuming that the underlying TBC is only unpredictable, a much weaker assumption than pseudorandomness. In particular, no attacks better than key recovery or random guessing seems to be known against unpredictability, even in the non-uniform case. See [DS09, Section 7] for a discussion of unpredictability versus pseudorandomness.

A similar observation has previously been maid by Datta and Yasuda [DY15]: they showed that a two-key variant of PMAC [BR02] can be proven PRF-secure assuming the “internal” block cipher (used for hashing the message) is only a secure MAC (rather than a PRP/PRF as in previous work about PMAC), which for a block cipher is equivalent to unpredictability. The proof proceeds by showing that the block cipher-based hash function underlying PMAC is (computationally) almost universal assuming the underlying block cipher is unpredictable. While we consider a slightly different TBC-based hash function relying on a counter rather than masks as in PMAC, our proof is very similar. We observe though that computational almost universality might not be strong enough for a H-coefficients-based proof of security of NaT and extend the approach to statistical almost
Table 1: Summary of the nAE schemes from the GNSIV family considered in this paper. The second column shows the corresponding variant of the GCTR encryption mode family and the third column gives the corresponding tweak input for TBC calls in the GCTR variant, where $N$ is the nonce, $R$ is the random IV, and $\langle j \rangle_c$ is the encoding of the block counter $j$ over $c$ bits. The last column gives the required length $\ell$ for the tweakey (which combines the key and the tweak) for $|N| = |R| = c = 128$ (we neglect here the tweak prefix required for domain separation) depending on the size of the key $K_2$ used in the encryption mode ($\ell$ is simply the sum of $|K_2|$ and the tweak length as per the third column).

| GNSIV variant | GCTR variant | tweak input | tweakey length ($|K_2| = 128/256$) |
|---------------|--------------|-------------|----------------------------------|
| SCT-2         | CTRT         | $R + \langle j \rangle_c$ | 256/384                          |
| GNSIV-N       | GCTR-N       | $R\|\langle j \rangle_c$ | 384/512                          |
| GNSIV-R       | GCTR-R       | $N\|\langle j \rangle_c$ | 384/512                          |
| GNSIV-C       | GCTR-C       | $N\|R\|\langle j \rangle_c$ | 512/640                          |

universality. We also discuss concrete conjectures for the unpredictability of a secure TBC and derive a corresponding concrete bound for NaT.

**Large-Tweak TBC Proposals.** While our new nAE modes are clean and simple, they require a larger tweak length than what is usually available from existing TBCs such as Deoxys-TBC [JNP14, JNPS16, JNPS21] or SKINNY [BJK+16]. Our second contribution is therefore to propose new TBC designs achieving larger tweak lengths. This is not a trivial task: we first show that SKINNYe-64/256, a recent large-tweak variant of SKINNY proposed at EUROCRYPT 2020 [NSS20a] (64-bit block with 256-bit tweakey) presents some structural design flaws. Indeed, the STK construction extension proposed by the authors does not follow the actual STK design paradigm and allows many unwanted cancellations in the subtweakeys throughout the round when the attacker inserts well-chosen differences in the tweak input. As the number of difference cancellations is directly used in the corresponding MILP model, our finding implies that the proven bounds on the number of active Sboxes for this primitive are actually wrong.2

We therefore provide a sound design strategy to construct large-tweak TBCs using the STK paradigm [JNP14], leading to new Deoxys-TBC variants. More specifically, we propose Deoxys-TBC-512 and Deoxys-TBC-640, which can be seen as natural extensions of the Deoxys-TBC family to larger tweak lengths (512 and 640 bits, respectively). We also propose a patch for the SKINNYe-64/256 variant. Finally, we produced software benchmarks for our modes when instantiated with Deoxys-TBC-512 and Deoxys-TBC-640. They indicate that while ensuring an extremely high security level, even in the multi-user setting, the performance remains competitive. We believe these new large-tweak TBCs might find further applications such as TBC-based hashing.

We provide a summary of the modes of the GNSIV family and the corresponding instantiations of Deoxys-TBC in Table 1.

**Comparison with AES-GCM-SIV.** To finish, we compare the security bound of our new modes with the one proven for AES-GCM-SIV in [BHT18]. Their bound is dominated by

---

2This flaw was communicated to the authors of [NSS20a] in October 2020, who updated the ePrint version of their paper. Subsequently, Qin et al. [QDW+22] used this flaw to devise a full-fledged attack against SKINNYe-64/256.
\[ \sigma B/2^n + d(\sigma + q_{nc})/2^k, \]
where \( \sigma \) is the total number of blocks in encryption and decryption queries, \( B \) is an upper bound on the number of blocks encrypted for any user, \( d \) is an upper bound on the number of users that re-use a particular nonce value, and \( q_{nc} \) is the number of ideal cipher queries. In comparison, the bound for our new AE mode with the best variant of the counter encryption mode (namely GCTR-Z, see Table 2) is dominated by
\[
nq_{nc}/2^k + \mu\sigma/2^k + \mu q/2^n + \min\{\mu_{max}^{\text{enc}} q_{ver}/2^k, Bq_{ver}/2^k\},
\]
where \( q \) is the total number of encryption and decryption queries, \( \mu \) is the maximal number of repetitions of any nonce for any user, and \( \ell_{\text{enc}} \) is the maximal number of blocks in any encryption query. This calls for some comments. First, note that even in the nonce-respecting setting, \( d \) can be as large as the number of encryption queries \( q_{enc} \), e.g. if all users rely on a counter with the same initial value to generate nonces. Hence, AES-GCM-SIV can only achieve beyond-birthday security if some small upper bound on \( d \) can be enforced (e.g. when all users generate nonces randomly), or by increasing the key length of the underlying block cipher. Second, since one always has \( \mu \leq B \), assuming \( k \geq n \), the terms \( \mu\sigma/2^k, \mu q/2^n \), and \( \min\{\mu_{max}^{\text{enc}} q_{ver}/2^k, Bq_{ver}/2^k\} \) are always smaller than \( \sigma B/2^n \), and in fact much smaller when \( \mu \ll B \) and \( q_{ver} \ll \sigma \), a common situation when nonces are only “mildly” misused and a large number of blocks are encrypted/decrypted per user and in total.

2 Preliminaries

2.1 General Notation and Definitions

Given a finite non-empty set \( X \), we let \( X \leftarrow X \) denote the draw of an element \( X \) from \( X \) uniformly at random. Given a positive integer \( n \), we let \( \{0,1\}^n \) denote the set of all \( n \)-bit strings of length \( n \), \( \{0,1\}^* \) denote the set of all bit strings of length at most \( n \), and \( \{0,1\}^* \) denote the set of all bit strings. The empty string is denoted \( \epsilon \) and the all-zero string of length \( n \geq 1 \) is denoted \( 0^n \). The length of a bit string \( X \) is denoted \( |X| \). The concatenation of two bit strings \( X \) and \( Y \) is denoted \( X||Y \). Given a bit string \( X \) and an integer \( n > 0 \), we define \( |X|_n = |X|/n \) and for \( X \neq \epsilon \) we write the parsing operation of \( X \) into \( n \)-bit blocks as \( X_0||\cdots||X_{\ell-1} \leftarrow^{\text{ozp}} X \), where \( \ell = |X|_n, |X_i| = n \) for \( i = 0, \ldots, \ell-1 \) and \( |X_{\ell-1}| \leq |n| \). Given an integer \( n > 0 \), we let \( \text{ozp}_n \) (or \( \text{ozp} \) when the parameter \( n \) is implicit) denote the padding function defined for \( X \neq \epsilon \) as
\[
\text{ozp}(X) := \begin{cases} 
X & \text{if } |X| = n \\
X||0^{n-|X|} & \text{if } |X| < n.
\end{cases}
\]
Given a bit string \( X \) of length \( i \) or larger, the \( i \) leftmost bits of \( X \) are denoted \( [X]_i \), and the \( i \) rightmost bits of \( X \) are denoted \( [X]_i \). We let \( X \ll a \) denote the bit string \( X \) rotated by \( a \) positions to the left. Given two integers \( b > 0 \) and \( i \geq 0 \) such that \( i < 2^b \), we let \( (i)_b \) denote the \( b \)-bit binary representation of \( i \). We simply write \( (i) \) when the length \( b \) is clear from the context or unspecified. Given integers \( a \leq b \), we let \( [a,b] \) denote the set \( \{a,\ldots,b\} \). We say that a function \( F : X \to Y \) is regular if all \( Y \in Y \) have the same number of preimages by \( F \). We say that a function \( H : K \times X \to Y \) is \( \gamma \)-uniform if for every \( X \in X \) and every \( Y \in Y \), \( \Pr[K \leftarrow K : H(K,X) = Y] \leq \gamma \).

Almost Universal Hashing. Let \( K, X \) and \( Y \) be two non-empty sets with \( K \) and \( Y \) finite. A keyed hash function with key space \( K \), domain \( X \), and range \( Y \) is a function \( H : K \times X \to Y \). We write \( H_K(X) \) for \( H(K,X) \). We defined two notions of almost universal (AU) hashing, namely computational AU (cAU) hashing and statistical AU (sAU) hashing.
**Definition 1** (AU hash function). Let $\mathcal{K}$, $\mathcal{X}$ and $\mathcal{Y}$ be two non-empty sets with $\mathcal{K}$ and $\mathcal{Y}$ finite and let $\text{len}: \mathcal{X} \to \mathbb{N}$ be some function measuring the “length” of an element of $\mathcal{X}$ (e.g., the number of $n$-bit blocks when $\mathcal{X}$ consists of bit strings). Let $H: \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a keyed hash function and $\mathcal{A}$ be an adversary. The advantage of $\mathcal{A}$ against the AU security of $H$ is defined as

$$\text{Adv}_{\text{AU}}^H(\mathcal{A}) = \Pr[(X, X') \leftarrow \mathcal{A}(\cdot), K \leftarrow \mathcal{K}: X \neq X' \land H_K(X) = H_K(X')].$$

Let $\delta: \mathbb{N} \to [0,1]$ be some function. We say that $H$ is a $(\zeta, \delta)$-cAU hash function (w.r.t $\text{len}$) if for any adversary $\mathcal{A}$ running in time at most $\zeta$ and returning $X$ and $X'$ of length at most $\ell$ (i.e., $\max\{\text{len}(X), \text{len}(X')\} \leq \ell$), $\text{Adv}_{\text{AU}}^H(\mathcal{A}) \leq \delta(\ell)$. We say that $H$ is a $\delta$-sAU hash function if for any adversary $\mathcal{A}$ (even computationally unbounded) returning $X$ and $X'$ of length at most $\ell$, $\text{Adv}_{\text{AU}}^H(\mathcal{A}) \leq \delta(\ell)$. Equivalently, $H$ is $\delta$-sAU if and only if for every $(X, X') \in \mathcal{X}^2$ with $X \neq X'$ and $\max\{\text{len}(X), \text{len}(X')\} \leq \ell$, $\Pr[K \leftarrow \mathcal{K}: H_K(X) = H_K(X')] \leq \delta(\ell)$.

The last equivalence is easily proven: the if direction is trivial; for the only if direction, assume towards a contradiction that there exists $(X, X') \in \mathcal{X}^2$ with $X \neq X'$ and $\max\{\text{len}(X), \text{len}(X')\} \leq \ell$ such that $\Pr[K \leftarrow \mathcal{K}: H_K(X) = H_K(X')] > \delta(\ell)$ and consider the adversary $\mathcal{A}$ which computes the collision probability for every pair of messages of length at most $\ell$ and returns $(X, X')$; then $\mathcal{A}$ has AU advantage strictly larger than $\delta(\ell)$, contradicting the assumption that $H$ is $\delta$-sAU.

### 2.2 Tweakable Block Ciphers

A tweakable block cipher (TBC) with key space $\mathcal{K}$, tweak space $\mathcal{T}$, and domain $\mathcal{X}$ is a mapping $E: \mathcal{K} \times \mathcal{T} \times \mathcal{X} \to \mathcal{X}$ such that for any key $K \in \mathcal{K}$ and any tweak $T \in \mathcal{T}$, the mapping $E(K, T, \cdot)$ is a permutation of $\mathcal{X}$. The set of all such TBCs will be denoted $\text{TBC}(\mathcal{K}, \mathcal{T}, \mathcal{X})$. Slightly abusing the notation, given positive integers $k$, $t$, and $n$, we also let $\text{TBC}(k, t, n)$ denote $\text{TBC}([0,1]^k, [0,1]^t, [0,1]^n)$ and we often write $E_K(T, X)$ or $E^X_K(T)$ in place of $E(K, T, X)$. A tweakable permutation with tweak space $\mathcal{T}$ and domain $\mathcal{X}$ is a mapping $P: \mathcal{T} \times \mathcal{X} \to \mathcal{X}$ such that for any tweak $T \in \mathcal{T}$, the mapping $P(T, \cdot)$ is a permutation of $\mathcal{X}$. The set of all tweakable permutations with tweak space $\mathcal{T}$ and domain $\mathcal{X}$ will be written $\text{TP}(\mathcal{T}, \mathcal{X})$. We consider two security notions for TBCs, namely pseudorandomness and unpredictability.

**Definition 2** (TPRP and UNP security). Let $E \in \text{TBC}(\mathcal{K}, \mathcal{T}, \mathcal{X})$ and $\mathcal{A}$ be an adversary. Consider games $\text{Real}_{E}^{\text{tprp}}(\mathcal{A})$ and $\text{Ideal}_{E}^{\text{tprp}}(\mathcal{A})$ defined in Figure 2 (left). The advantage of $\mathcal{A}$ in breaking the TPRP security of $E$ is defined as

$$\text{Adv}_{E}^{\text{tprp}}(\mathcal{A}) = |\Pr[1 \leftarrow \text{Real}_{E}^{\text{tprp}}(\mathcal{A})] - \Pr[1 \leftarrow \text{Ideal}_{E}^{\text{tprp}}(\mathcal{A})]|.$$  

Consider game $\text{Game}_{E}^{\text{unp}}(\mathcal{A})$ defined in Figure 2 (right). The advantage of $\mathcal{A}$ in breaking the UNP security of $E$ is defined as

$$\text{Adv}_{E}^{\text{unp}}(\mathcal{A}) = \Pr[\text{true} \leftarrow \text{Game}_{E}^{\text{unp}}(\mathcal{A})].$$

**The Ideal Tweakable Cipher Model.** Our proofs will use the ideal (tweakable) cipher model, where the TBC underlying a specific construction is drawn uniformly at random from the set of all TBCs with the adequate key, tweak, and message spaces, and given as a black box that can be queried in the forward and backward direction by the adversary. More formally, given a cryptographic scheme $\Pi$ based on a TBC $E \in \text{TBC}(k, t, n)$, the game defining the security of $\Pi$ is modified as follows: at the beginning of the game, a random TBC $E_{\text{ic}}$ is drawn uniformly at random from $\text{TBC}(k, t, n)$ and the adversary is given access to encryption and decryption oracles IC and IC$^{-1}$ such that a query IC($K, T, X$) returns $E_{\text{ic}}(K, T, X)$ and a query IC$^{-1}$(K, T, Y) returns $E_{\text{ic}}^{-1}(K, T, Y)$. 
2.3 Multi-user Security Notions

Below we specify the multi-user (mu) security notions for encryption, authentication, and authenticated encryption used in this paper.

**Nonce and IV-based Encryption Scheme.** The notion of combined nonce and IV-based encryption (nivE) scheme was introduced in [PS16]. Syntactically, an nivE scheme is a tuple $\Pi = (K, N, R, M, C, Enc, Dec)$ where $K$ is the key space, $N$ is the nonce space, $R$ the random value space, $M$ the message space, and $C$ is the ciphertext space, all being non-empty sets of bit strings with $K$, $N$, and $R$ finite, and $Enc$ and $Dec$ are algorithms such that:

- the encryption algorithm $Enc$ takes as input a tuple $(K, N, R, M) \in K \times N \times R \times M$ and outputs a ciphertext $C \in C$;
- the decryption algorithm takes as input a tuple $(K, N, R, C) \in K \times N \times R \times C$ and outputs a plaintext $M \in M$.

We require that for all tuples $(K, N, R, M) \in K \times N \times R \times M$, one has $Dec(K, N, R, Enc(K, N, R, M)) = M$.

We also require that if $M$ contains a bit string of length $m$, then it contains all bit strings of length $m$, and that $|Enc(K, N, R, M)| = |M|$ for all $(K, N, R, M) \in K \times N \times R \times M$. We write $Enc_K(N, R, M)$ for $Enc(K, N, R, M)$ and similarly for $Dec$. The multi-user security of an nivE scheme is defined as follows.

**Definition 3** (mu-nivE security). Let $\Pi = (K, N, R, M, C, Enc, Dec)$ be an nivE scheme and let $\mathcal{A}$ be an adversary. Consider games $\text{Real}_\Pi^{\text{mu-nive}}(\mathcal{A})$ and $\text{Ideal}_\Pi^{\text{mu-nive}}(\mathcal{A})$ defined in Figure 3. The advantage of $\mathcal{A}$ in breaking the mu-nivE security of $\Pi$ is defined as

$$\text{Adv}_\Pi^{\text{mu-nive}}(\mathcal{A}) = \left| \Pr[1 \leftarrow \text{Real}_\Pi^{\text{mu-nive}}(\mathcal{A})] - \Pr[1 \leftarrow \text{Ideal}_\Pi^{\text{mu-nive}}(\mathcal{A})] \right|.$$
Figure 3: The mu-nivE security games for a nivE scheme $\Pi = (K,N,R,M,C,Enc,Dec)$.

Figure 4: The mu-nPRMAC security games for a nonce-based pseudorandom MAC $\Pi = (K,N,U,V,Tag)$.

Nonce-based Pseudorandom MACs. A nonce-based pseudorandom MAC (nPRMAC for short) is a tuple $\Pi = (K,N,U,V,Tag)$, where $K$ is the key space, $N$ is the nonce space, $U$ is the domain, and $V$ is the tag space. $Tag$ is a deterministic algorithm which takes as input a tuple $(K,N,U) \in K \times N \times U$ and returns a tag $V \in V$. We write $Tag_K(N,X)$ for $Tag(K,N,X)$. The following security notion introduced in [JNPS21] combines the (nonce-based) PRF and MAC security notions.

Definition 4 (mu-nPRMAC security). Let $\Pi = (K,N,U,V,Tag)$ be a nonce-based pseudorandom MAC and $A$ be an adversary. Let games $\text{Real}_{\Pi}^{\mu\text{-nprmac}}(A)$ and $\text{Ideal}_{\Pi}^{\mu\text{-nprmac}}(A)$ be as defined in Figure 4. The advantage of $A$ in breaking the mu-nPRMAC security of $\Pi$ is defined as

$$\text{Adv}_{\Pi}^{\mu\text{-nprmac}}(A) = |\text{Pr}[1 \leftarrow \text{Real}_{\Pi}^{\mu\text{-nprmac}}(A)] - \text{Pr}[1 \leftarrow \text{Ideal}_{\Pi}^{\mu\text{-nprmac}}(A)]|.$$  

Nonce-based Authenticated Encryption. A nonce-based authenticated encryption scheme with associated data (nAE scheme for short) is a tuple $\Pi = (K,N,A,M,C,Enc,Dec)$ where $K,N,A,M,C$, and $V$ are non-empty sets of bit strings with $K$, $N$, and $V$ finite and $Enc$ and $Dec$ are deterministic algorithms such that:
We also require that if \( i \notin [1, u] \vee (i, N, A, M) \in Q \) then
\[
\text{return } \bot
\]
\[
(C, V) \leftarrow \text{Enc}_K(N, A, M)
\]
\[
(C, V) \leftarrow \emptyset \{0, 1\}^{|M|} \times \mathcal{V}
\]
\[
Q \leftarrow Q \cup \{(i, N, A, M)\}
\]
\[
Q' \leftarrow Q' \cup \{(i, N, A, C; V)\}
\]
\[
\text{return } (C, V)
\]
\[
\text{Oracle Dec}(i, N, A, C, V)
\]
\[
\text{return } \bot \text{return Dec}_K(N, A, C, V)
\]

**Figure 5:** The mu-nAE security games for a nAE scheme \( \Pi = (\mathcal{K}, \mathcal{N}, \mathcal{A}, \mathcal{M}, \mathcal{C}, \mathcal{V}, \text{Enc}, \text{Dec}) \).

- the encryption algorithm \( \text{Enc} \) takes as input a key \( K \in \mathcal{K} \), a nonce \( N \in \mathcal{N} \), associated data \( A \in \mathcal{A} \), and a message \( M \in \mathcal{M} \) and outputs a ciphertext \( C \in \mathcal{C} \) and a tag \( V \in \mathcal{V} \);
- the decryption algorithm \( \text{Dec} \) takes as input a key \( K \in \mathcal{K} \), a nonce \( N \in \mathcal{N} \), associated data \( A \in \mathcal{A} \), a ciphertext \( C \in \mathcal{C} \) and a tag \( V \in \mathcal{V} \), and outputs either a message \( M \in \mathcal{M} \) or a special symbol \( \bot \) that indicates that decryption failed.

We require that for all tuples \( (K, N, A, M) \in \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M} \), one has
\[
\text{Dec}(K, N, A, \text{Enc}(K, N, A, M)) = M.
\]

We also require that if \( \mathcal{M} \) contains a bit string of length \( m \), then it contains all bit strings of length \( m \), and that for all \( (K, N, A, M) \in \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M} \), \(|C| = |M|\) where \( (C, V) = \text{Enc}(K, N, A, M) \). We write \( \text{Enc}_K(N, A, M) \) and \( \text{Dec}_K(N, A, C, V) \) in place of \( \text{Enc}(K, N, A, M) \) and \( \text{Dec}(K, N, A, C, V) \). The multi-user security of an nAE scheme is defined as follows.

**Definition 5** (mu-nAE security). Let \( \Pi = (\mathcal{K}, \mathcal{N}, \mathcal{A}, \mathcal{M}, \mathcal{C}, \mathcal{V}, \text{Enc}, \text{Dec}) \) be an nAE scheme and \( \mathcal{A} \) an adversary. Consider games \( \text{Real}_{\Pi}^{\text{mu-nae}}(\mathcal{A}) \) and \( \text{Ideal}_{\Pi}^{\text{mu-nae}}(\mathcal{A}) \) defined in Figure 5. The advantage of \( \mathcal{A} \) in breaking the mu-nAE security of \( \Pi \) is defined as
\[
\text{Adv}_{\Pi}^{\text{mu-nae}}(\mathcal{A}) = |\Pr[1 \leftarrow \text{Real}_{\Pi}^{\text{mu-nae}}(\mathcal{A})] - \Pr[1 \leftarrow \text{Ideal}_{\Pi}^{\text{mu-nae}}(\mathcal{A})]|.
\]

### 2.4 The H-coefficients Technique

Let us fix a deterministic adversary \( \mathcal{A} \) trying to break a cryptographic scheme \( \Pi \) whose security is defined via a distinguishing experiment specified by two games \( \text{Real}_{\Pi}(\mathcal{A}) \) and \( \text{Ideal}_{\Pi}(\mathcal{A}) \). We summarize the attack in a queries transcript \( \tau \) encoding all queries made by the adversary to the oracles available in the games together with their answer. Let \( \Lambda_{re} \) and \( \Lambda_{id} \) be two random variables sampled according to the probability distribution of \( \tau \) in the real and in the ideal world respectively. We let \( \Theta \) denote the set of all attainable queries transcripts, i.e. transcripts \( \tau \) such that \( \Pr[\Lambda_{id} = \tau] > 0 \). One has the following result, the proof of which can be found for example in [CS14].
Then one has

$$\exists \beta, \Theta$$

Theorem 1

We note that our results should easily be generalizable to forkciphers [ALP+19] and multi-forkciphers as done in [ABPV21].

---

### Figure 6: The encryption and decryption algorithms of the GCTR[$E, f_T, f_X$] nivE scheme based on a TBC $E \in \text{TBC}(k, t, n)$.

#### Theorem 1 (H-coefficients technique [Pat09]).

Fix a deterministic adversary $A$. Let $\Theta_{\text{bad}}$ and $\Theta_{\text{good}}$ be two disjoint subsets of $\Theta$ such that $\Theta = \Theta_{\text{bad}} \cup \Theta_{\text{good}}$. Assume that there exists $\beta \geq 0$ such that, for every $\tau \in \Theta_{\text{good}}$,

$$\frac{\Pr[\Lambda_{\text{rec}} = \tau]}{\Pr[\Lambda_{\text{id}} = \tau]} \geq 1 - \beta.$$

Then one has

$$\text{Adv}_\Pi(A) := |\Pr[1 \leftarrow \text{Real}_\Pi(A)] - \Pr[1 \leftarrow \text{Ideal}_\Pi(A)]| \leq \Pr[\Lambda_{\text{id}} \in \Theta_{\text{bad}}] + \beta.$$

### 3 Multi-user Security of GCTR Encryption

In this section, we focus on the encryption part of the AE scheme and consider several variants of the counter encryption mode. Our goal is to fit three strings (a nonce $N$, a random value $R$, and a counter $j$) into the tweak and the plaintext inputs of a tweakable block cipher. Following Andreeva et al. [ABPV21], we define a family of nivE schemes called generic CTR (GCTR) based on a TBC $E$.\(^4\) For the remainder of this section, we fix positive integers $k$, $t$, and $n$ and a TBC $E \in \text{TBC}(k, t, n)$. The schemes are parameterized by positive integers $\nu$, $r$, and $c$ that denote respectively the nonce length, the random value length, and the counter length, all expressed in bits, and two functions $fr$ and $fx$ taking as input a tuple $(N, R, j) \in \{0, 1\}^\nu \times \{0, 1\}^r \times \{0, 2^c - 1\}$ and returning respectively an element of the tweak space and of the plaintext space of $E$. Given $E$, $fr$, and $fx$, the nivE scheme GCTR[$E, fr, fx$] has key space $\{0, 1\}^k$, nonce space $\{0, 1\}^\nu$, random value space $\{0, 1\}^r$, message and ciphertext spaces $\{0, 1\}^{\leq 2^n}$, and algorithms Enc and Dec as defined in Figure 6. A specific instantiation of the GCTR family is obtained by picking up two functions $fr$ and $fx$.

Andreeva et al. [ABPV21] performed a systematic analysis of GCTR in the single user setting, identifying 22 secure variants named GCTR-1 to GCTR-22. Here, in addition to CTRT (GCTR-3), we focus on two simple and natural options:

\(^4\)We note that our results should easily be generalizable to forkciphers [ALP+19] and multi-forkciphers as done in [ABPV21].

---

<table>
<thead>
<tr>
<th>GCTR[$E, f_T, f_X$], Enc$_K(N, R, M)$</th>
<th>GCTR[$E, f_T, f_X$], Dec$_K(N, R, C)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{if } M = \epsilon \text{ then return } \epsilon$</td>
<td>$\text{if } C = \epsilon \text{ then return } \epsilon$</td>
</tr>
<tr>
<td>$M_0</td>
<td></td>
</tr>
<tr>
<td>for $j := 0 \ldots \ell - 1$ do</td>
<td>for $j := 0 \ldots \ell - 1$ do</td>
</tr>
<tr>
<td>$T_j := f_T(N, R, j)$</td>
<td>$T_j := f_T(N, R, j)$</td>
</tr>
<tr>
<td>$X_j := f_X(N, R, j)$</td>
<td>$X_j := f_X(N, R, j)$</td>
</tr>
<tr>
<td>for $j := 0 \ldots \ell - 2$ do</td>
<td>for $j := 0 \ldots \ell - 2$ do</td>
</tr>
<tr>
<td>$C_j := M_j \oplus E^T_{K_j}(X_j)$</td>
<td>$M_j := C_j \oplus E^T_{K_j}(X_j)$</td>
</tr>
<tr>
<td>// last block might be incomplete</td>
<td>// last block might be incomplete</td>
</tr>
<tr>
<td>$C_{\ell-1} := M_{\ell-1} \oplus \left[ E^{T_{\ell-1}}<em>{K</em>{\ell-1}}(X_{\ell-1}) \right]</td>
<td><em>{M</em>{\ell-1}}$</td>
</tr>
<tr>
<td>return $C_0</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Variants of the GCTR nivE schemes family considered in this paper. For CTRT, + is addition mod $2^r$.

<table>
<thead>
<tr>
<th>name in this paper</th>
<th>name in [ABPV21]</th>
<th>$f_T(N, R, j)$</th>
<th>$f_X(N, R, j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTRT</td>
<td>3</td>
<td>$R + \langle j \rangle_c$</td>
<td>$N$</td>
</tr>
<tr>
<td>GCTR-N</td>
<td>1</td>
<td>$R \parallel \langle j \rangle_c$</td>
<td>$N$</td>
</tr>
<tr>
<td>GCTR-R</td>
<td>7</td>
<td>$N \parallel \langle j \rangle_c$</td>
<td>$R$</td>
</tr>
<tr>
<td>GCTR-C</td>
<td>5</td>
<td>$N \parallel R$</td>
<td>$\langle j \rangle_c$</td>
</tr>
<tr>
<td>GCTR-Z</td>
<td>---</td>
<td>$N \parallel R \parallel \langle j \rangle_c$</td>
<td>$\langle 0 \rangle_n$</td>
</tr>
</tbody>
</table>

- two strings are concatenated and used as the tweak, while the third one is used as the plaintext block (GCTR-1, -5, and -7 in [ABPV21]);
- the three strings are concatenated and used as the tweak while the plaintext is a constant: in this case, the TBC is simply used as a PRF applied to the three strings (not analyzed in [ABPV21]).

This yields five variants in total which are listed in Table 2. Except for CTRT, we refer to a specific variant using a suffix reflecting the block input $f_X$ as indicated in the table (e.g. GCTR-N refers to the variant such that $f_T(N, R, j) = R \parallel \langle j \rangle_c$ and $f_X(N, R, j) = N$). We establish the multi-user security of GCTR in the ideal cipher model for $E$ in the following theorem. The proof is deferred to Appendix B.

**Theorem 2** (mu security of GCTR). Let $k$, $t$, and $n$ be positive integers and $E \in \text{TBC}(k, t, n)$ be a tweakable block cipher modeled as an ideal tweakable cipher (IC, IC$^{-1}$). Let $\nu$, $r$, and $c$ be positive integers and let $(f_T, f_X)$ be one of the pairs of functions from Table 2. Let $q_{ic}$, $q_{enc}$, $\ell_{\max}$, $\sigma$ and $\mu$ be positive integers such that $q_{enc} \leq \min\{2^k, 2^r\}$, $\ell_{\max} \leq 2^c$, and $\sigma \leq \min\{2^c, 2^r\}$. Then, for any adversary $A$ making at most $q_{ic}$ queries in total to IC or IC$^{-1}$ and $q_{enc}$ queries to ENC of maximal length (in number of $n$-bit blocks) at most $\ell_{\max}$ and of total length (in number of $n$-bit blocks) at most $\sigma$, and such that any (user, nonce) pair $(i, N)$ appears at most $\mu$ times in ENC queries, one has

$$\text{Adv}_{\text{mu-nive}}^{\text{GCTR}[E, f_T, f_X]}(A) \leq \frac{1}{2^\nu} + \frac{1}{2^n} + \frac{2(r + n)q_{ic}}{2^k} + g(q_{ic}, q_{enc}, \ell_{\max}, \sigma, \mu),$$

where

$$g(q_{ic}, q_{enc}, \ell_{\max}, \sigma, \mu) :=
\begin{cases}
\frac{(4\mu + 1)\sigma}{2^{\ell_{\max}}} & \text{for CTRT}, \\
\frac{q_{enc}}{2^{\ell_{\max}}} + \frac{\sigma}{2^r} & \text{for GCTR-N}, \\
\frac{q_{enc}}{2^{\ell_{\max}}} + \frac{\sigma}{2^c} & \text{for GCTR-R}, \\
\frac{(2\mu - 1)\sigma}{2^{\ell_{\max}}} + \frac{\sigma}{2^{\ell_{\max} + r}} & \text{for GCTR-C}, \\
\frac{q_{enc}}{2^{\ell_{\max}}} & \text{for GCTR-Z}.
\end{cases}$$

**Discussion.** The common part of the bound is negligible as long as $q_{ic}$ is small compared to $2^{k - \log_2 n}$ (assuming $r \approx n$). Omitting constant factors, one can note that the bound for all variants except CTRT have a $\mu q_{enc}/2^r$ term but differ in how $\sigma$ shows up: the security bound for GCTR-Z is always better than the one for GCTR-N, which is always better than the one for GCTR-R and GCTR-C, whose relation depend on $\mu$ and $\ell_{\max}$. Note also that the bound for GCTR-Z does not depend on $\sigma$. The bound for CTTR has a term $\mu \sigma/2^r$ and hence is always worse than the bound of GCTR-Z and GCTR-N (for $r = n$).
4 Multi-user Security of NaT Authentication

In this section, we focus on the authentication part of the AE scheme. In Section 4.1, we consider the nonce-based pseudorandom MAC called Nonce-as-Tweak (NaT) \[\text{\cite{CLS17}}\] based on a TBC \(E\) and a generic AU hash function \(H\) and study its mu-nPROMAC security. In Section 4.2, we study how to instantiate \(H\) from an unpredictable TBC.

4.1 NaT Based on a Generic AU Hash Function

Let \(k, t, \) and \(n\) be positive integers, \(\mathcal{K}_n\) and \(\mathcal{U}\) be non-empty sets with \(\mathcal{K}_n\) finite, and \(\text{len} : \mathcal{U} \to \mathbb{N}\) be some length function. Let \(E \in \text{TBC}(k, t, n)\) be a tweakable block cipher and \(H : \mathcal{K}_n \times \mathcal{U} \to \{0, 1\}^n\) be a keyed hash function. Let also \(\nu\) be an integer with \(\nu \leq t\). We define the nonce-based pseudorandom MAC \(\text{NaT}[H, E]\) with key space \(\mathcal{K}_n \times \{0, 1\}^k\), nonce space \(\{0, 1\}^\nu\), domain \(\mathcal{U}\), and tag space \(\{0, 1\}^\ell\) as

\[
\text{NaT}[H, E]_{\mathcal{K}_n, \mathcal{K}_n}(N, U) := E_{\mathcal{K}_n}^{\text{len}(N)}(H_{\mathcal{K}_n}(U)).
\]

We establish the multi-user security of \(\text{NaT}\) in the ideal cipher model for \(E\) and assuming that \(H\) is a \(\text{sAU}\) hash function in the following theorem using the \(H\)-coefficients technique. The proof is deferred to Appendix C. We conjecture that a similar result can be proven assuming \(H\) is a \(\text{cAU}\) hash function. However, as far as we can tell, the proof of this conjecture cannot use the \(H\)-coefficients technique if \(H\) is only assumed computationally secure as there does not seem to be a way to replace it with a statistically secure counterpart as a first step in the reasoning. Hence, a game-based proof would be required instead, allowing to construct a reduction to the \(\text{cAU}\) security of \(H\). We leave this for future work as we are able to prove that the TBC-based instantiation of \(H[E]\) defined in Section 4.2 is \(\text{sAU}\) under a plausible unpredictability assumption on the underlying TBC \(E\).

Theorem 3 (mu security of NaT). Let \(k, t, \) and \(n\) be positive integers, \(\mathcal{K}_n\) and \(\mathcal{U}\) be non-empty sets with \(\mathcal{K}_n\) finite, \(\text{len} : \mathcal{U} \to \mathbb{N}\) be some length function (below, the length of an element \(U \in \mathcal{U}\) refers to \(\text{len}(U)\)), \(E \in \text{TBC}(k, t, n)\) be a tweakable block cipher modeled as an ideal tweakable cipher (IC, IC\(^{-1}\)), and \(H : \mathcal{K}_n \times \mathcal{U} \to \{0, 1\}^n\) be a keyed hash function. Let \(\nu\) be an integer such that \(\nu \leq t\). Assume that \(H\) is \(\delta\)-\(\text{sAU}\) and \(\gamma\)-uniform (w.r.t \(\text{len}\)) for \(\delta(\ell) = \gamma(\ell) = \alpha \ell/2^k + \beta/2^n\) with \(\beta \geq 1\) (and hence \(\delta \geq 2^{-n}\) and \(\gamma \geq 2^{-n}\)). Let \(q_{\text{tag}}\), \(q_{\text{ver}}\), \(\sigma_{\text{tag}}\), \(\sigma_{\text{ver}}\), \(\ell_{\text{max}}\), and \(\mu\) be positive integers such that \(q_{\text{tag}} \leq 2^n\), \(q_{\text{ver}} + q_{\text{tag}} \leq 2^k\), and \(q_{\text{ver}} + q_{\text{tag}} \leq \mu \leq 2^{n/2}\). Then, for any (computationally unbounded) adversary \(A\) against the mu-nPROMAC security of NaT[H, E] making at most \(q_{\text{tag}}\) queries in total to IC or IC\(^{-1}\), \(q_{\text{tag}}\) queries to \(\text{TAG}\) of maximal length \(\ell_{\text{max}}\), and of total length at most \(\sigma_{\text{tag}}\), \(q_{\text{ver}}\) queries to \(\text{VER}\) of total length at most \(\sigma_{\text{ver}}\), and such that any (user, nonce) pair \((i, N)\) appears at most \(\mu\) times in its \(\text{TAG}\) queries, one has

\[
\text{Adv}_{\text{NaT}[H, E]}^{\text{mu-npMAC}}(A) \leq \frac{2nq_{\text{tag}}}{2^k} + \frac{\alpha \mu q_{\text{tag}}}{2^k} + \frac{\alpha (\mu + 1) \sigma_{\text{ver}}}{2^k} + \frac{\alpha \mu \ell_{\text{max}} q_{\text{ver}}}{2^k} + 1 + \frac{\beta \mu q_{\text{tag}}}{2^n} + \frac{\beta (\mu + 3) q_{\text{ver}}}{2^n}.
\]

Moreover, if the total length of all \(\text{TAG}\) queries for any user is at most \(B\), then the term \(\alpha \mu \ell_{\text{max}} q_{\text{ver}}/2^k\) can be replaced by \(\alpha Bq_{\text{ver}}/2^k\).

Discussion. We will justify in Section 4.2 that \(\alpha\) and \(\beta\) can be conjectured to be small absolute constant, hence we let \(\alpha = \beta = 1\) for the sake of discussion. Then as long as \(\mu\) remains small, all terms are negligible as long as \(\sigma_{\text{tag}}\) and \(\sigma_{\text{ver}}\) are small compared to \(2^k\), \(q_{\text{tag}}\) is small compared to \(2^n\), \(q_{\text{ver}}\ell_{\text{max}}\) is small compared to \(2^n\), and \(q_{\text{tag}}\) is small compared
to $2^{k - \log_2 n}$. Note that in many situations, $\ell_{\text{tag}}^{\text{max}}$ can be reasonably upper bounded to provide beyond-birthday security with respect to $q_{\text{err}}$ (e.g., for $n = 128$ and $\ell_{\text{tag}}^{\text{max}} = 2^{32}$, security is ensured up to $2^{96}$ verification queries). When $\mu$ is allowed to grow as large as $q_{\text{tag}}$, we hit the birthday bound and security vanishes at $2^{n/2}$ tag/verification queries.

### 4.2 TBC-based Almost Universal Hashing

**Warm-up.** Let $n$ and $c$ be positive integers and $\mathcal{K}$ and $\mathcal{X}$ be finite non-empty sets. Let $F : \mathcal{K} \times \mathcal{X} \times \{0, 1\}^c \rightarrow \{0, 1\}^n$ be a keyed function with key space $\mathcal{K}$, domain $\mathcal{X} \times \{0, 1\}^c$, and range $\{0, 1\}^n$. Consider the keyed hash function $F_{\oplus}$ with key space $\mathcal{K}$, domain $\mathcal{X}^\leq 2^c$ (the set of all sequences of elements of $\mathcal{X}$ of length at most $2^c$ equipped with the length function $\text{len}$ returning the length of a sequence), and range $\{0, 1\}^n$ defined for $K \in \mathcal{K}$ and $X = (X_0, \ldots, X_{a-1}) \in \mathcal{X}^\leq 2^c$ (with $a = \text{len}(X) = 0$ if $X$ is empty) as

$$F^\oplus_K(X_1, \ldots, X_a) := \begin{cases} 0^n & \text{if } a = 0, \\
\bigoplus_{j=0}^{a-1} F_K(X_j, \langle j \rangle_c) & \text{if } a \geq 1.
\end{cases}$$

Construction $F_{\oplus}$ is a simplified variant of the keyed hash function used for example in PMAC [BR02, Rog04]. The standard proof that $F_{\oplus}$ is an AU relies on the assumption that $F$ is a PRF (see for example [BS20], Section 7.2.3). Here we show that $F_{\oplus}$ is a cAU hash function assuming $F$ is unpredictable,\(^5\) as captured by the following theorem.

**Theorem 4.** Let $F : \mathcal{K} \times \mathcal{X} \times \{0, 1\}^c \rightarrow \{0, 1\}^n$ be a keyed function with key space $\mathcal{K}$, domain $\mathcal{X} \times \{0, 1\}^c$, and range $\{0, 1\}^n$. Then, for any adversary $A$ against the AU security of $F_{\oplus}$ running in time at most $\zeta$ and returning messages of length at most $\ell$, there exists an adversary $B$ against the UNP security of $F$ making at most $2\ell$ oracle queries and running in time at most $\zeta + \alpha$ for some small constant $\alpha$ that only depends on the computation model (but not on $F$), such that

$$\text{Adv}^\text{un}_{F_{\oplus}}(A) \leq \text{Adv}^\text{imp}_F(B).$$

**Proof.** Let $A$ be an adversary against the AU security of $F_{\oplus}$. We construct adversary $B$ against the UNP security of $F$ as follows. (Recall that $B$ has oracle access to $F_K$ for some random key $K$ and must return a pair $(x, F_K(x))$ for some $x$ that did not query its oracle). It runs $A$ (which returns two messages $X = (X_0, \ldots, X_{a-1})$ and $X' = (X_0', \ldots, X_{a'-1}')$ in $\mathcal{X}^\leq \ell$. Assuming $A$ is successful, then $X \neq X'$ and $F^\oplus_K(X) = F^\oplus_K(X')$, where $K$ is the key drawn at random by the UNP challenger. Assume wlog that $a \leq a'$, and let $J$ denote the set of integers $j \in [0, a-1]$ such that $X_j \neq X'_j$. Then $F^\oplus_K(X) = F^\oplus_K(X')$ is equivalent to

$$\left(\bigoplus_{j \in J} F_K(X_j, \langle j \rangle_c) \oplus F_K(X'_j, \langle j \rangle_c)\right) \oplus \bigoplus_{j=a}^{a'-1} F_K(X'_j, \langle j \rangle_c) = 0^n,$$

where the second summation is empty if $a = a'$. Note that all $F$ inputs appearing in (1) are distinct. Then $B$ proceeds as follows. If $a < a'$, then it queries all inputs appearing in (1) except $(X'_{a'-1}, (a'-1)_c)$ and returns

$$\left(\bigoplus_{j \in J} F_K(X_j, \langle j \rangle_c) \oplus F_K(X'_j, \langle j \rangle_c)\right) \oplus \bigoplus_{j=a}^{a'-2} F_K(X'_j, \langle j \rangle_c)$$

\(^5\)To avoid confusion, we warn that in his security proof of Protected Counter Sum, Bernstein uses “unpredictable” to mean pseudorandom, see beginning of [Ber99, Section 2].
as its guess for $F_K(X'_j, (j)_c)$. (Note that this expression equals $0^n$ if $J = \emptyset$ and $a' = a + 1$.) If $a = a'$, then $J \neq \emptyset$ as otherwise this would imply $X = X'$. Let $j_0 = \max(J)$. Then $B$ queries all inputs appearing in (1) except $(X_{j_0}, (j_0)_c)$ and returns

$$\bigoplus_{j \in J \setminus \{j_0\}} F_K(X_j, (j)_c) \oplus F_K(X'_j, (j)_c) \oplus F_K(X_{j_0}, (j_0)_c)$$

as its guess for $F_K(X'_{j_0}, (j_0)_c)$.

Clearly, $B$ succeeds when $A$ succeeds, makes at most $2\ell$ oracle queries, and runs in time at most $\zeta + \alpha \ell$ for some small constant $\alpha$. $B$’s additional computations besides running $A$ only consist in equality checking and xoring on $n$-bit blocks, hence $\alpha$ only depends on the computation model but not on $F$.

\[ \square \]

Theorem 4 establishes that $F^\oplus$ is a computational AU in a constructive way: given an AU adversary $A$ against $F^{\oplus}$, the proof describes an explicit UNP adversary $B$ against $F$ using $A$ as a black box. On the other hand, one can easily adapt the proof to show that $F^{\oplus}$ is a statistical AU in a non-constructive way, as shown in the following theorem. A similar constructive/non-constructive dichotomy for cAU/sAU security was proved by Bellare for the cascade construction applied to a PRF [Bel06, Lemma 3.1].

**Theorem 5.** Let $F : \mathcal{K} \times X \times \{0, 1\}^c \to \{0, 1\}^n$ be a keyed function with key space $\mathcal{K}$, domain $X \times \{0, 1\}^c$, and range $\{0, 1\}^n$. Let $\delta(\ell) := \max_{\mathcal{A}} \{ \text{Adv}^{\text{unp}}_{F^\oplus}(A) \}$, where the maximum is over all adversaries making at most $2\ell$ oracle queries and running in time at most $\alpha \ell$ for some small constant $\alpha$ that only depends on the computation model (but not on $F$). Then $F^{\oplus}$ is $\delta$-sAU.

**Proof.** Assume towards a contradiction that $F^{\oplus}$ is not $\delta$-sAU, i.e., there exists $\ell$ and $(X, X') \in (\mathcal{X}^{\leq \ell})^2$ such that $X \neq X'$ and

$$\delta' := \Pr[K \leftarrow \mathcal{K} : F^{\oplus}_K(X) = F^{\oplus}_K(X')] > \delta(\ell).$$

Define $B'$ as the UNP adversary having $X$ and $X'$ hardwired in its code\(^6\) and behaving exactly as $B$ from the proof of Theorem 4 (except it does not have to run any AU adversary $A$ to obtain $X$ and $X'$). Then $B'$ makes at most $2\ell$ oracle queries, runs in time $\alpha \ell$ for some small constant $\alpha$, and wins the UNP game with probability $\delta' > \delta(\ell)$, contradicting the definition of $\delta$.

\[ \square \]

**The Construction.** We now turn to the actual TBC-based hash function used to instantiate NaT in our nAE modes. Let $k$, $t$, and $n$ be positive integers and $E \in \text{TBC}(k, t, n)$ be a tweakable block cipher. Let $c$ be a positive integer such that $c \leq t - 3$ and let $m := t - c - 3$ and $L := 2^n \cdot (m + n)$.\(^7\) We define the keyed hash function $H[E]$ with key space $\{0, 1\}^k$, domain $\{0, 1\}^{\leq L} \times \{0, 1\}^{\leq L}$, and range $\{0, 1\}^n$ as

$$H[E]_K(A, M) := \begin{cases} E^{(c)}_{K} \oplus (0^n) & \text{if } (A, M) = (c, c), \\ H[E^0_K(A)] \oplus H[E^2_K(M)] & \text{otherwise}, \end{cases} \quad (2)$$

\(^6\)Note that one cannot construct adversary $B'$ above by running an sAU-adversary $A$ against $F^{\oplus}$ and waiting for it to output $(X, X')$ since it would be impossible to upper bound the running time of $B'$ (hence the non-constructiveness).

\(^7\)As usual, $c$ denotes the counter length, while $m$ denotes the number of message input bits that can fit in the tweak in addition to the counter and a 3-bit prefix.
where for $i \in \{0, 2\}$ and $X \in \{0, 1\}^{\leq L}$ parsed as $X_0||\cdots||X_{\ell-1} \leftarrow_{m+n} X$, 

$$H[E]^i_K(X) := \begin{cases} 
0^n & \text{if } X = \epsilon, \\
\bigoplus_{j=0}^{\ell-1} E_K^{(j)}[[j]\epsilon][X_j]_{m}(\left\lceil X_j \right\rceil_n) & \text{if } X \neq \epsilon, |X_{\ell-1}| = n, \\
\bigoplus_{j=0}^{\ell-2} E_K^{(j)}[[j]\epsilon][X_j]_{m}(\left\lceil X_j \right\rceil_n) \oplus E_K^{(\ell+1)}[[\ell+1]\epsilon][X^*_{\ell-1}]_{m}(\left\lceil X^*_{\ell-1} \right\rceil_n) & \text{if } X \neq \epsilon, |X_{\ell-1}| < n,
\end{cases}$$

where $X^*_{\ell-1} = \text{ozp}_{n+t-1}(X_{\ell-1})$. Define the length function for $(A, M) \in \{0, 1\}^{\leq L} \times \{0, 1\}^{\leq L}$ as $\text{len}(A, M) := |A|_{n+m} + |M|_{n+m}$.

We establish that $H$ is a sAU hash function in the following theorem. The proof follows the one of Theorem 4 and Theorem 5 and is deferred to Appendix D.

**Theorem 6** (sAU security of $H$). Let $k$, $t$, and $n$ be positive integers, $E \in \text{TBC}(k, t, n)$ be a tweakable block cipher, and $c$ and $m$ be integers such that $t = m + c + 3$. Let $\delta(t) := \max_{A} \{\text{Adv}^\text{sAU}_E(A)\}$, where the maximum is over all adversaries making at most $2t$ oracle queries and running in time at most $\alpha t$ for some small constant $\alpha$ that only depends on the computation model (but not on $E$). Then $H[E]$ is $\delta$-sAU.

**Conjectured Unpredictability of TBCs.** In view of Theorem 6, it is interesting to contrast unpredictability and pseudorandomness in order to see if anything has been gained compared to previous proofs of AU-security for $H[E]$ and variants based on (T)PRP assumptions. As mentioned in the introduction, there exist non-uniform attacks against any pseudorandom function with key length $k$ achieving advantage $2^{-k/2}$ in constant time and queries [DTT10, BL13]. On the other hand, no such attack is known for unpredictability. For adversaries running in time at most $\alpha t$ and making at most $2t$ queries, a reasonable conjecture is that $\max_{A} \{\text{Adv}^\text{sAU}_E(A)\}$ is close to $\alpha t / 2^k + 1/(2^n - 2t)$, where the first term captures the success probability of exhaustive key search and the second term the success probability of a random guess among the unqueried values. In our specific case, this can even be pushed to $\alpha t / 2^k + 1/(2^n - 2t)$ as it can be noted that adversaries constructed in the proof of Theorem 6 make at most 2 queries per tweak. Combined with Theorem 6, we obtain the following conjecture justifying the assumption regarding $\delta$ in Theorem 3.

**Conjecture 1.** Let $E \in \text{TBC}(k, t, n)$ be a “secure” tweakable block cipher. Then there exists small absolute constants $\alpha$ and $\beta$ such that $H[E]$ is $\delta$-sAU with $\delta(t) = \alpha t / 2^k + \beta / 2^n$.

**Uniformity of $H$.** We note that the proof of Theorem 6 can easily be adapted to prove that $H[E]$ restricted to inputs of length at most $\ell$ is $\gamma$-uniform for $\gamma = \max_{A} \{\text{Adv}^\text{sAU}_E(A)\}$ where the maximum is over all adversaries making at most $\ell$ oracle queries and running in time at most $\alpha t$ for some small constant $\alpha$. Note that we cannot simply let $H[E]_{\ell}(\epsilon, \epsilon) = 0^n$ as this would break the uniformity property required for the proof of Theorem 3. In particular, the adversary could provoke bad condition (C-3) with probability close to 1 by making $2^{\ell/2}$ queries $\text{Tag}(i, N, (\epsilon, \epsilon))$ for distinct users $i$ and some fixed nonce $N$ and $2^{\ell/2}$ queries $\text{IC}(K, 0^{\ell-t}\|N, 0^n)$.

## 5 Authenticated Encryption Modes

Given an nPRMAC and an nIV scheme, one can obtain an nAE scheme by combining them through the NSIV composition method [PS16], a variant of SIV [RS06]. Formally, given an AD space $\mathcal{A}$ and a message space $\mathcal{M}$, let $\Pi_{\text{mac}} = (\mathcal{K}_1, \mathcal{N}, \mathcal{U}, \mathcal{V}, \text{Tag})$ be an nPRMAC with $\mathcal{U} = \mathcal{A} \times \mathcal{M}$ and let $\Pi_{\text{enc}} = (\mathcal{K}_2, \mathcal{R}, \mathcal{M}, \mathcal{C}, \text{Enc}, \text{Dec})$ be an nIV scheme. Let $\text{Conv} : \mathcal{V} \rightarrow \mathcal{R}$ be a regular function. We define the nAE scheme $\Pi_{\text{AE}}[\Pi_{\text{mac}}, \Pi_{\text{enc}}]$ with key space $\mathcal{K}_1 \times \mathcal{K}_2$, \ldots
nonce space \( \mathcal{N} \), AD space \( \mathcal{A} \), message space \( \mathcal{M} \), ciphertext space \( \mathcal{C} \), and tag space \( \mathcal{V} \) as specified in Figure 7. As shown in [PS16,JNPS21] for the single user setting, assuming \( \Pi_{\text{mac}} \) is nPRMAC-secure and \( \Pi_{\text{enc}} \) is nivE-secure, then \( \Pi_{\Pi_{\text{mac}}, \Pi_{\text{enc}}} \) is nAE-secure.

Given three TBCs \( E_{\text{in}}, E_{\text{out}}, \) and \( E_{\text{enc}} \) with respective key spaces \( K_{\text{in}}, K_{\text{out}}, \) and \( K_{\text{enc}} \), any encryption scheme from the GCTR family (Section 3) based on \( E_{\text{enc}} \) can be combined through the NSIV composition method with the NaT MAC scheme (Section 4) based on \( E_{\text{in}} \) for hashing and \( E_{\text{out}} \) for encrypting the hash in order to obtain a TBC-based nAE scheme with key space \( K_{\text{in}} \times K_{\text{out}} \times K_{\text{enc}} \). However, using tweak domain separation, one can save on key material by letting \( E_{\text{out}} = E_{\text{enc}} \) and using the same key both for tag finalization (i.e., encrypting the hash) in NaT and in the encryption mode, separating the two types of calls by using a different tweak prefix. Moreover, it is possible to use a single TBC \( E = E_{\text{in}} = E_{\text{enc}} \), again by using tweak domain separation, as long as the hashing key and the encryption key are independent. This way, we obtain a family of nAE schemes that we call generic NSIV (GNSIV) based on a TBC \( E \) (or two TBCs \( E_{\text{in}} \) and \( E_{\text{enc}} \)) and parameterized by the pair of functions \((f_T, f_X)\) used to instantiate the encryption part as GCTR[\( E / E_{\text{enc}}, f_T, f_X \)].

We now describe the GNSIV family more formally, focusing on the single-TBC case for simplicity and study its nAE security. Let \( k, t, \) and \( n \) be positive integers and let \( E \in \text{TBC}(k, t, n) \). Let \( \nu \) and \( c \) be positive integers and let \( f_T \) and \( f_X \) be two functions from \( \{0,1\}^\nu \times \{0,1\}^c \) to \( \{0,1\}^{2^3} \) and \( \{0,1\}^n \) respectively.\(^8\) We use a 3-bit prefix for tweak domain separation, with prefixes 0 to 4 used for hashing (see definition of \( H[E] \) in Section 4.2), prefix 5 used to encrypt the hash, and prefix 6 used for message encryption. Given \( E, f_T, \) and \( f_X \), the nAE scheme GNSIV[\( E, f_T, f_X \)] has key space \( \{0,1\}^k \times \{0,1\}^k \), nonce space \( \{0,1\}^\nu \), AD, message, and ciphertext spaces \( \{0,1\}^{\leq 2^n} \), tag space \( \{0,1\}^n \), and algorithms Enc and Dec as defined in Figure 8 (see also Figure 1 in Section 1).

Unfortunately, the generic composition theorem from [PS16,JNPS21] no longer applies due to key reuse for both tag finalization and encryption and our use of the ideal cipher model. However, in Appendix E, we show that the security bound for GNSIV[\( E, f_T, f_X \)] (with a generic hash function in place of \( H[E] \)) is essentially the sum of the bounds of NaT and the specific GCTR variant.

### 6 TBC Instantiations

In this section, we propose new tweakable block ciphers with large tweak length that can be used in the modes described previously. Our analysis will focus on TBCs derived from the Superposition Tweakey (STK) framework [JNP14] that we recall in Section 6.1. In Section 6.2, we first discuss the SKINNYe-64/256 proposal from [NSS20a], a variant of the

\(^{8}\) We use the same counter length both for hashing and encryption for simplicity, but we could use two different lengths if needed.

\(^{9}\) The random value length \( r \) is equal to \( n \) here as \( R \) is the output of NaT[\( H[E_1], E_2 \)] which is \( n \)-bit.
We give security arguments for these new constructions in Section 6.4.

The TWEAKEY framework was put forward in 2014 by Jean [JNPS16], and discuss similar extensions for TBCs. The STK construction moreover specifies the type of update: in each block of data, the 16 elementary c-bit cells are permuted with the same permutation \(h_i\), and then, each cell from the \(i\)-th word is seen as an element of \(\mathbb{F}_{2^c}\) and gets multiplied by a constant \(a_i, i \in [0, p - 1]\) (see Figure 9). Finally, each \(n\)-bit sub-tweakey to be integrated in the key-alternating structure is simply taken as the XOR of the \(n\)-bit words.

The important consequence of this design that will matter for the rest of the section is that each fixed position \(i \in [0, 15]\) in the \(p\) tweakey words, seen as a \(p\)-element vector \(\mathbf{y}^i = (x_0^i, \ldots, x_{p-1}^i)\), is extended by the tweakey schedule to an \((r + 1)\)-element vector \(\mathbf{y}^i = (y_0^i, \ldots, y_r^i)\), containing the \(r + 1\) elements to be integrated in one position of the internal state by the \(r + 1\) sub-tweakey additions over \(r\) rounds.

With a small generalization, we can replace the multiplications by \(a_0, \ldots, a_{p-1} \in \mathbb{F}_{2^c}\).
by linear functions \( f_0, \ldots, f_{p-1} \) over \( \mathbb{F}_2 \) so that for \( k \in [0, r] \), one has \( y_k = \bigoplus_{j=0}^{n-1} f_j(x_j^k) \).

The number of differences cancellations on a nibble position \( i \) for \( r \) rounds is then defined as the number of zero-difference elements in \( y^i = (y_0^i, \ldots, y_{p-1}^i) \), given a non-zero difference is inserted in \( x^i \).

This simple construction allows to analyze the linear expansion \( x^i \rightarrow y^i \) to find a lower bound on the element-wise Hamming weight of the expanded vector \( y^i \). One can then feed this bound to MILP solvers to help them assess resistance against linear or differential cryptanalysis, regardless of the number of TK words.

### 6.2 SKINNYe-64/256: a Flawed TK4 Extension of SKINNY

Several tweakable block ciphers have been designed following the STK framework, notably SKINNY [BJK+16]. SKINNY is now considered a well-studied primitive, yet only TK1, TK2 and TK3 variants have been proposed by the designers. To address the need for a TK4 variant, Naito, Sasaki and Sugawara [NSS20a] proposed a TK4 variant of SKINNY-64 with lightweight applications in mind. They named their proposal SKINNYe-64/256.

Mimicking the TK3 design of SKINNY-64/192, the authors naturally extended the tweakey size by \( n = 64 \) bits by choosing a new 4-bit function \( f_3 \) and increasing the number of rounds to 44. While in SKINNY-64/192 \( f_0 \) is the identity function and \( f_1, f_2 \) are two 4-bit LFSRs, in SKINNYe-64/256 the chosen transformation\(^\text{10}\) \( f_3 \), although being a linear permutation over \( \mathbb{F}_2 \), is not an LFSR (as claimed in [NSS20a]).

However, we claim that this transformation fails to meet the assumptions required by the security analysis performed in [BJK+16] and reused in [NSS20a]. Namely, to allow a computer-aided security evaluation of STK-based ciphers like SKINNY with respect to differential cryptanalysis, its tweakey schedule must verify specific linear properties. More precisely, abstracting the inner workings of the tweakey schedule away, one must guarantee that the incoming round subtweakeys \( stk_i \) (see Figure 9) have a (lower) bounded number of active differences; or stated differently, that the number of differences cancellations made possible by the linearity of the tweakey schedule are (upper) bounded. In the original SKINNY paper, the MILP models correctly represent the cipher when the number of cancellations does not exceed \( p - 1 \) for each nibble position in the tweakey state for TK\( p \), \( p \in \{2, 3\} \), and when the number of rounds covered does not exceed 30.

By analyzing the effect of \( f_3 \) on that number of cancellations for SKINNYe-64/256 (TK4, \( p = 4 \)), we discover that it can be far higher than \( p - 1 = 3 \) for a well-chosen tweakey difference. For instance, the difference \((0x1, 0x4, 0x0, 0x5)\) applied to the same position in the four tweakey words yields subtweakeys where only 7 are nonzero across 14 rounds (see Table 3). More surprisingly, the joint effect of the linear transformations is such that each of the 7 active subtweakeys contain a single active nibble with difference equals to 1.

As a consequence, one can conclude that the security bounds computed in [NSS20a],

\[^{10}\text{The transformation is } f_3 : (x_3 \parallel x_2 \parallel x_1 \parallel x_0) \rightarrow (x_2 \parallel x_1 \parallel x_0 \parallel x_3 \oplus x_2 \oplus x_1).\]
Table 3: For starting difference (0x1, 0x4, 0x0, 0x5), the tweakey schedule of SKINNYe-64/256 outputs very sparse binary subtweakey differences.

<table>
<thead>
<tr>
<th>Round</th>
<th>TK1</th>
<th>TK2</th>
<th>TK3</th>
<th>TK4</th>
<th>stk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0x1</td>
<td>0x4</td>
<td>0x0</td>
<td>0x5</td>
<td>0x0</td>
</tr>
<tr>
<td>1</td>
<td>0x1</td>
<td>0x9</td>
<td>0x0</td>
<td>0x9</td>
<td>0x1</td>
</tr>
<tr>
<td>2</td>
<td>0x1</td>
<td>0x3</td>
<td>0x0</td>
<td>0x3</td>
<td>0x1</td>
</tr>
<tr>
<td>3</td>
<td>0x1</td>
<td>0x6</td>
<td>0x0</td>
<td>0x7</td>
<td>0x0</td>
</tr>
<tr>
<td>4</td>
<td>0x1</td>
<td>0xd</td>
<td>0x0</td>
<td>0xc</td>
<td>0x0</td>
</tr>
<tr>
<td>5</td>
<td>0x1</td>
<td>0xa</td>
<td>0x0</td>
<td>0xa</td>
<td>0x1</td>
</tr>
<tr>
<td>6</td>
<td>0x1</td>
<td>0x5</td>
<td>0x0</td>
<td>0x4</td>
<td>0x0</td>
</tr>
<tr>
<td>7</td>
<td>0x1</td>
<td>0xb</td>
<td>0x0</td>
<td>0xb</td>
<td>0x1</td>
</tr>
<tr>
<td>8</td>
<td>0x01</td>
<td>0x7</td>
<td>0x0</td>
<td>0x6</td>
<td>0x0</td>
</tr>
<tr>
<td>9</td>
<td>0x01</td>
<td>0xf</td>
<td>0x0</td>
<td>0xe</td>
<td>0x0</td>
</tr>
<tr>
<td>10</td>
<td>0x01</td>
<td>0xe</td>
<td>0x0</td>
<td>0xf</td>
<td>0x0</td>
</tr>
<tr>
<td>11</td>
<td>0x01</td>
<td>0xc</td>
<td>0x0</td>
<td>0xd</td>
<td>0x0</td>
</tr>
<tr>
<td>12</td>
<td>0x01</td>
<td>0x8</td>
<td>0x0</td>
<td>0x8</td>
<td>0x1</td>
</tr>
<tr>
<td>13</td>
<td>0x01</td>
<td>0x1</td>
<td>0x0</td>
<td>0x1</td>
<td>0x1</td>
</tr>
<tr>
<td>14</td>
<td>0x01</td>
<td>0x2</td>
<td>0x0</td>
<td>0x2</td>
<td>0x1</td>
</tr>
<tr>
<td>15</td>
<td>0x01</td>
<td>0x4</td>
<td>0x0</td>
<td>0x5</td>
<td>0x0</td>
</tr>
</tbody>
</table>

Table 2] from MILP optimization are incorrect: much sparser differential characteristics are likely to exist. We leave finding such characteristics and possible matching differential attacks out of the scope of this paper.

### 6.3 Extensions to TK4 and TK5

We now address the problem of designing a TBC with a large tweakey space within the STK framework. From a high-level point of view, what we want to design shares a lot of similarities with a linear code with bounded minimal distance. Indeed, an input vector of differences should expand to a vector containing as many nonzero differences as possible. Getting back to the original design, we show below that we can take advantage of results from coding theory to construct, under some assumptions, structured linear codes over $\mathbb{F}_2$ that can directly be used as tweakey schedules. Another option that we also consider below are transformations linear over $\mathbb{F}_2$.

**Linear code over $\mathbb{F}_2$.** In the original STK construction, the authors suggested to use linear functions $f_j : x \rightarrow \alpha_j x$ over $\mathbb{F}_{2^c}$ for constants $\alpha_i \in \mathbb{F}_{2^c}$. We revisit the approach and show that we can easily derive TBC with almost arbitrarily large tweakey space.

To build a TBC with $p$ words of tweakey, we can rely on a Reed-Solomon linear code constructed using a Vandermonde matrix

$$V(\beta_0, \ldots, \beta_{p-1}) = \begin{bmatrix}
1 & \cdots & 1 & \cdots & 1 \\
\beta_0 & \cdots & \beta_k & \cdots & \beta_r \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\beta_0^{p-1} & \cdots & \beta_k^{p-1} & \cdots & \beta_r^{p-1}
\end{bmatrix}$$

as generator matrix for distinct $\beta_0, \ldots, \beta_{p-1} \in \mathbb{F}_{2^c}$. Indeed, the code generated by the $p \times (r+1)$ generator matrix is known to be MDS. Hence, its minimal distance is $(r+1) - p + 1$, which means that the number of cancellations in the $(r+1)$ outputs elements $(r$ rounds) is upper bounded by $p - 1$.

Let $\alpha_j \in \mathbb{F}_{2^c}$ be the multiplication constants in the $j$-th tweakey word. To map the Vandermonde matrix structure to the STK construction, we can set $\alpha_j = \alpha^j$ for an element
\( \alpha \in \mathbb{F}_{2^c} \) with order at least \( r \). Indeed, we then have the \( p \times (r + 1) \) Vandermonde matrix

\[
V(\alpha^0, \ldots, \alpha^{p-1}) = \begin{bmatrix}
1 & \cdots & 1 & \cdots & 1 \\
1 & \cdots & \alpha^{k-1} & \cdots & \alpha^r \\
1 & \cdots & \alpha^{2(k-1)} & \cdots & \alpha^{2r} \\
\vdots & & \vdots & & \vdots \\
1 & \cdots & \alpha^{(p-1)(k-1)} & \cdots & \alpha^{(p-1)r}
\end{bmatrix}
\]

(3)

describing the tweakey schedule, with all \( \alpha^k \) being distinct, \( k \in [0, r] \), thanks to the condition on the order of \( \alpha \).

This construction therefore allows to construct a \( \text{TK}_p \) TBC with security arguments for \( \max_{e \in \mathbb{F}_{2^c}} (\text{ord}(e)) - 1 \) rounds, which can be as high as \( 2^c - 2 \) rounds by using any primitive element as \( \alpha \).

As concrete instantiations, we propose Deoxys-TBC-512 (TK4) with \( \alpha = x \in \mathbb{F}_{2^c} \) defined by the AES polynomial \( x^8 + x^7 + x^3 + x + 1 \). Informally, this results in multiplications by 1, 2, 4, and 8 for the four tweakey words. Similarly, we naturally propose Deoxys-TBC-640 (TK5) with multiplications by 1, 2, 4, 8 and 16 for the five words. We follow the rationale for the number of rounds of \( 10 + 2p \) used in Deoxys-TBC for \( \text{TK}_p \), so we use 18 rounds for Deoxys-TBC-512 and 20 rounds for Deoxys-TBC-640. The round function is left unchanged.

We note that the order of the chosen \( \alpha \in \mathbb{F}_{2^c} \) in both cases is not maximal and equals 51. We could have chosen \( x + 1 \) which is primitive, but first, we do not need security guarantees for more than 50 rounds (as Deoxys primitives stop at 20 rounds), and second, implementation-wise, it is more efficient to multiply by 2 than by 3.

About SKINNY, we do not formally propose new variants, but we note that the same reasoning would apply to extend the tweakey space.

**Remark 1.** When \( c = 4 \) (for instance SKINNY-64), we note that the maximal order of the field elements is \( 2^c - 1 = 15 \). Consequently, if ones relies on the Vandermonde structure with field multiplications to build the tweakey schedule, one cannot bound the number of cancellations for more than 14 rounds (i.e., 15 consecutive subtweakeys).

**Linear over \( \mathbb{F}_2 \).** To avoid the field multiplications described before, we can choose the functions \( f_i \) as being linear over \( \mathbb{F}_2 \). In practice, the original Deoxys variants as well as all SKINNY TBCs use this strategy, with LFSRs with cheap update functions (a few XORs).

Theoretically, one could describe the multiplications in \( \mathbb{F}_{2^c} \) as linear permutations over \( \mathbb{F}_2 \), but there are potentially more interesting \( \mathbb{F}_2 \)-linear mappings (LFSRs being a subset of them) than multiplications in \( \mathbb{F}_{2^c} \). However, one significant drawback is the absence of constructive method: one cannot rely anymore on a structured linear code to theoretically bound the number of cancellations. What has been done in Deoxys and SKINNY is an exhaustive search to derive this bound experimentally, but this can only work for low values of \( p \).

To extend SKINNY-64 to get a TK4 variant, we can for instance use the 4-bit linear permutation: \( (x_3 \parallel x_2 \parallel x_1 \parallel x_0) \rightarrow (x_1 \parallel x_0 \parallel x_2 \oplus x_3 \parallel x_1 \oplus x_2) \). Using this transformation as \( f_3 \) instead of the one used in SKINNYe-64/256 would patch the mistake from [NSS20a]. We indeed verified experimentally that the number of cancellations match the requirements of the STK approach applied to SKINNY.

### 6.4 Security Claims and Analysis

**Security claims.** For Deoxys-TBC-512 and Deoxys-TBC-640, our two new versions of Deoxys-TBC, even though the tweakey material has increased and could accommodate larger keys, we only claim security up to 256 bits with regards to key recovery in the...
single key model. Namely, for a usage of one of these two tweakable block ciphers with a $t$-bit tweak and $k$-bit key, we expect no key recovery attack with complexity lower than $2^{\min(256,k)}$ operations (up to a small constant). Thus, our security claims are actually the same as for Deoxys-TBC-256 (the smallest version of Deoxys-TBC) and lower than that of Deoxys-TBC-384. We believe this is justified by the fact that we do not foresee any application where a key larger than 256 bits would be needed, even in the post-quantum scenario. We emphasize that we do not claim any security in the related-, known-, or chosen-key models.

**Security rationale.** When proposing Deoxys-TBC versions with larger tweakey size, we naturally leverage the extensive security analysis previously performed on Deoxys-TBC-256 and Deoxys-TBC-384 by the original designers [JNPS16] as well as by third-party teams [CHP+17,Sas18,MMS18,ZDW19,WP19,ZDJ19,ZDJM19,ZDJ19,DQSW22,HSE23,SYC+24]. In other words, Deoxys-TBC-512 and Deoxys-TBC-640 are new instances of a generic design framework that has already been extensively analyzed and previous results generally apply to these new instances (ignoring that most of them are in the related-key model and that many have a complexity higher than $2^{256}$, thus outside our security claims). The only difference is the potential impact of the addition of more tweak material.

As our security claims do not surpass those of the small Deoxys-TBC version, potential issues that could arise from having very large keys and corresponding security claims can be safely discarded. Note that, on this aspect, Deoxys-TBC-640 is actually much harder to attack than Deoxys-TBC-384, implicitly leading to a few extra rounds of security margin (for example the best attack up to $2^{384}$ complexity on Deoxys-TBC-384 reaches 14 rounds [ZDJ19], while when limiting Deoxys-TBC-384 to 128/256-bit keys it can only reach 12/13 rounds respectively).

The crucial point to check with respect to larger tweaks is how it affects the probabilities of the best related-tweakey differential paths, which is exactly what we will study in details in this section. In particular, we would like to verify the trend for long paths, but also for short ones as all best attacks on Deoxys-TBC are boomerang/rectangle-type of attacks. We will observe that generally, in the Deoxys-TBC design framework, adding one tweakey word will lead to one extra free round for short differential trails and two extra free rounds for long differential trails (see Table 4). Thus, adding 2 rounds to maintain the security margin when increasing tweak size by one word seems adequate: classical differential cryptanalysis directly benefits from two free rounds, while boomerang/rectangle cryptanalysis benefits from a single round twice (once for each upper/bottom differential sub-trails). Therefore, it is natural that the number of rounds of Deoxys-TBC-512 and Deoxys-TBC-640 (18 and 20 respectively) has been chosen following the Deoxys-TBC rationale $10 + 2p$ (where $p$ is the number of 128-bit tweakey words).

Meet-in-the-middle attacks are currently far from being a threat to Deoxys-TBC ciphers [LJ19] (reaching only 8 and 10 rounds for Deoxys-TBC-256 and Deoxys-TBC-384 respectively), but they remain the best in the single-key scenario. We argue that, from a designer’s perspective, the same round increase reasoning applies. As full diffusion takes two rounds with AES, we estimate that it will be very difficult for the adversary to use one extra tweakey word to control more than two rounds within a full meet-in-the-middle attack. This is confirmed by observing that 2 rounds separate the best Deoxys-TBC-256 and Deoxys-TBC-384 meet-in-the-middle attacks.

**Differential analysis.** As noted previously, a MILP-based analysis of the STK construction abstracts the actual tweakey schedule away and simply relies on the number of cancellations in the incoming subtweakeys for $r$ rounds. Depending on the type of the update, $r$ can vary: it can for instance be as large at 50 for Deoxys for field multiplications in $F_{2^8}$, or lower when LFSRs are used.
Table 4: Lower bounds for the number of differentially active Sboxes for the structure used in Deoxys-TBC. We report the two original designs from [JNPS16] as well as our new TK4 and TK5 variants, with two types of linear update in the tweakey schedule: either an LFSR or a field multiplication in $F_{2^8}$ (“mul”).

<table>
<thead>
<tr>
<th>Type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK2</td>
<td>LFSR</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>12</td>
<td>16</td>
<td>19</td>
<td>23</td>
<td>26</td>
<td>29</td>
<td>32</td>
<td>35</td>
<td>38</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[JNPS16]</td>
</tr>
<tr>
<td>TK3</td>
<td>LFSR</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>14</td>
<td>18</td>
<td>21</td>
<td>24</td>
<td>28</td>
<td>31</td>
<td>35</td>
<td>37</td>
<td>41</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[JNPS16]</td>
</tr>
<tr>
<td>TK4</td>
<td>LFSR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>19</td>
<td>23</td>
<td>26</td>
<td>30</td>
<td>33</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TK4</td>
<td>mul</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>19</td>
<td>23</td>
<td>26</td>
<td>30</td>
<td>33</td>
<td>37</td>
<td>39</td>
<td>43</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TK5</td>
<td>LFSR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>12</td>
<td>14</td>
<td>17</td>
<td>21</td>
<td>25</td>
<td>28</td>
<td>31</td>
<td>33</td>
<td>34</td>
<td>36</td>
<td>38</td>
</tr>
<tr>
<td>TK5</td>
<td>mul</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>12</td>
<td>14</td>
<td>17</td>
<td>21</td>
<td>25</td>
<td>28</td>
<td>31</td>
<td>35</td>
<td>38</td>
<td>42</td>
<td>45</td>
</tr>
</tbody>
</table>

$^\dagger$ [CHP$^+$17] indicated 45.

By looking at the actual 8-bit LFSRs used in Deoxys, we remark that they have cycles of length 15 (even though most of the cycles are of length 30). Consequently, it is sufficient to study the element-wise Hamming weight of the expanded vector over one cycle, for $r - 1 = 14$ consecutive rounds so that there are $r = 15$ expanded elements (before they cycle).

However, in most of the security analyses of Deoxys we are aware of, e.g., [JNPS16, CHP$^+$17], the authors encode in the MILP model a lower bound of the number of cancellations for $r = 15$ rounds. Due to this short cycle, considering 15 rounds introduces an off-by-one error as one counts twice one element in the sequence of 16 elements. Consequently, the number of cancellations of $p - 1$ for TK$p$ over 15 rounds claimed in [JNPS16] should be $p - 1$ for 14 rounds (or the looser bound $p$ for 15 rounds). The MILP model from [CHP$^+$17] suffers from this off-by-one error, however, by chance, the Deoxys variants studied by the authors only uses up to 16 rounds (17 consecutive subtweakeys), which makes only the results on 16 rounds incorrect.

To accommodate this, and to be able to model more than 14 consecutive rounds, we slightly tweak the MILP model used for Deoxys based on LFSRs. In our model for TK$p$, for $r \leq 14$ rounds, if the Boolean variable $stk_j[i]$ represents the activity of the $i$-th byte of the $j$-th subtweakey, $i, j \in [0, 14] \times [0, r]$, the linear inequality $\sum_j stk_j[i] \geq \min(r+1, 15) - (p - 1)$ is added to the MILP model for each window of 14 consecutive rounds. For instance, to model 18 rounds, we introduce five inequalities, starting at Rounds 0, 1, 2, 3 and 4. This ensures that for any 15 consecutive subtweakeys, the linear property of the tweakey schedule is included in the MILP model.

In the case of the field multiplications in $F_{2^8}$, the length of the cycle depends on the order of the element used to construct the Vandermonde matrix (3). In our case, for Deoxys-TBC-512 and Deoxys-TBC-640, we use an element with order 51, so as long as we model $r \leq 50$ rounds, we can keep a single inequality per position $i$, namely: $\sum_{j=0}^{r} stk_j[i] \geq (r + 1) - (p - 1)$.

We report in Table 4 the results of the MILP optimizations for the different cases. Recall that Deoxys-TBC-512 corresponds to TK4 of “mul” type, and Deoxys-TBC-640 corresponds to TK5 of “mul” type. We include the other cases to emphasize the influence of the bounds on the modelization.

---

$^11$ We note that the situation is the same as described in the remark on the field multiplication in $F_{2^4}$ due to the size of its multiplicative subgroup.

$^{12}$ The inequality given is simplified, as all $stk_j[i]$ can be zero.
7 Implementations

In order to assess the performance of our new modes, we first make a general study on the expected computational cost when using Deoxys-TBC instances, focusing on GNSIV-N and GNSIV-Z. Then, we concentrate on the important case of Intel processors with AESNI instructions set. We compare our candidates with popular nonce-misuse resistant modes, such as AES-GCM-SIV [GLL17], ZAE [IMPS17] and SCT-2 [JNPS16, JNPS21]. For simplicity, we measure long messages, even though we note that GNSIV-N and GNSIV-Z, similarly to SCT-2, have a minimal overhead for small messages.

7.1 General Theoretical Performances

The encryption part of GNSIV-N and GNSIV-Z requires one TBC call per message block and this is independent of the Deoxys-TBC version used. Thus, we better use the smallest-tweak version of Deoxys-TBC (since it will use lesser AES rounds) as long as the tweak input is large enough to accommodate the mode inputs. Thus, considering an encryption key $K_2$, a nonce $N$, and a random value $R$ of 128 bits each and a counter of 125 bits (recall that we reserve 3 bits for tweak domain separation), the encryption part of GNSIV-N and GNSIV-Z can use Deoxys-TBC-384 and Deoxys-TBC-512 respectively. We note that we could consider increasing the encryption key length to 256 bits by using Deoxys-TBC-512 and Deoxys-TBC-640 instead.

Regarding the authentication part, using the same parameters, GNSIV-N and GNSIV-Z can rely on Deoxys-TBC-256 since only the counter is inserted in the tweak input of the TBC. However, using larger tweak instances of Deoxys-TBC could be useful as this extra input can directly take more associated data or message for a higher throughput. This is of course a trade-off as these Deoxys-TBC instances will use more AES rounds on the other hand.

As already observed in [IMPS17], in software with AESNI computing one Deoxys-TBC tweakey schedule line roughly represents a cost of 0.4 of an entire AES encryption, thus about 4 AES rounds. This is of course a very rough and relatively pessimistic estimation that is highly dependent on the actual platform on which the algorithm will be run, but it has the merit to take into account the extra cost coming from the tweakey schedule. Moreover, when measuring performances for long messages, if a tweak input in one TK position remains unchanged for many message blocks (typically the key, the nonce, etc.) the cost for this TK is amortized since all TK parts can be computed independently in Deoxys-TBC. The same applies to a counter since its evolution is predictable and thus updating the TK values from a counter increment is very cheap.

We summarize the theoretical cost in terms of AES rounds for various GNSIV-N and GNSIV-Z instantiations in Table 5. We note that GNSIV-N instantiated with Deoxys-TBC-384 for both authentication and encryption part would correspond to the recent Deoxys-AE2 mode [JNPS21] (proposed without security proof). We further note that while our modes require two passes on the plaintext input for encryption, the decryption can be done online with a single pass. Finally, we emphasize that our instances allow a counter with a size of almost 128 bits, in contrary to modes like AES-GCM-SIV that only handles inputs of size $2^{64}$ or less. Thus, for more accurate comparison, 64 bits of this counter could be traded for further authentication throughput improvement.

7.2 Software Performances with AESNI

As Deoxys-TBC is based on the AES round function and since our modes allow parallelization, they allow very efficient software implementations on the processors that support AES-accelerated instructions, and in particular AESNI.
Table 5: Theoretical performance in number of AES rounds (for one 128-bit AD block and one 128-bit M block) for various GNSIV-N and GNSIV-Z modes instantiations with Deoxys-TBC versions. Notation \( x \) (\( y \)) means that \( x \) rounds of AES are required when not considering potential extra tweakey schedule cost, while \( y \) takes into account the 1.4 factor that might occur depending on the platform. For AES-GCM-SIV, we consider that one POLYVAL evaluation is equivalent to 5 AES rounds [GLL17].

<table>
<thead>
<tr>
<th>Mode</th>
<th>Deoxys-TBC instance</th>
<th>AES round cost per AD per M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Hash</td>
</tr>
<tr>
<td>GNSIV-N ((</td>
<td>K_2</td>
<td>= 128))</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-384</td>
<td>Deoxys-TBC-384</td>
</tr>
<tr>
<td>GNSIV-N ((</td>
<td>K_2</td>
<td>= 256))</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-512</td>
<td>Deoxys-TBC-512</td>
</tr>
<tr>
<td>GNSIV-Z ((</td>
<td>K_2</td>
<td>= 128))</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-512</td>
<td>Deoxys-TBC-512</td>
</tr>
<tr>
<td>GNSIV-Z ((</td>
<td>K_2</td>
<td>= 256))</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-640</td>
<td>Deoxys-TBC-640</td>
</tr>
<tr>
<td>SCT-2 [JNPS21]</td>
<td>Deoxys-TBC-256</td>
<td>Deoxys-TBC-256</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-256</td>
<td>Deoxys-TBC-256</td>
</tr>
<tr>
<td>ZAE [IMPS17]</td>
<td>Deoxys-TBC-384</td>
<td>Deoxys-TBC-256</td>
</tr>
<tr>
<td></td>
<td>Deoxys-TBC-256</td>
<td>Deoxys-TBC-256</td>
</tr>
<tr>
<td>AES-GCM-SIV-128 [GLL17]</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

We have implemented GNSIV-N and GNSIV-Z using Intel Intrinsics and we provide in Table 6 performance measurements on Intel Haswell processor. For completeness, we also measured performances of AES-GCM-SIV implementation13 (“Intrinsics” ones), taking into account the key schedule.

Having to perform an update of one TK\( p \) word, when \( p \geq 2 \), would be too costly using AESNI, because the shifts/XORs involved in the LFSR would require too many operations in comparison to nicely pipelined AESNI instructions.

Our modes perform quite well: they are generally slower of a factor around 1.6 than AES-GCM-SIV for very long messages, but closer when the message gets smaller (actually on par for 64-byte messages). However, we recall that the Internet Mix is generally composed of mostly very small (smaller than 100 bytes) packets, then some medium size (around 500 bytes) packets and finally a low proportion of maximum-size (maximum Ethernet packet payload size is 1536 bytes). Considering the added security guarantees when compared to AES-GCM-SIV, we believe our modes represent a very competitive trade-off.

In addition, we remark that our benchmarks were performed in the most possible advantageous situation for AES-GCM-SIV: a processor that has the PCLMULQDQ hardware accelerated instruction that allows efficient GF(\( 2^{128} \)) multiplication on a different execution pipe than AESNI. Of course, our designs do not benefit from this advantage (for example having an hardware accelerated instruction to perform the tweakey schedule). On relatively cheap micro-controllers or, more importantly, in hardware, the situation will be very different. The GF(\( 2^{128} \)) multiplication requires a large area (or alternatively many cycles).

13Taken from https://github.com/Shay-Gueron/AES-GCM-SIV.
Table 6: Encryption benchmarks for GNSIV-N and GNSIV-Z (with Deoxys-TBC-384 in the authentication part), with comparison with AES-GCM-SIV-128, expressed in cycles per byte on AESNI enabled platforms (with Turbo Boost disabled) for increasing numbers of processed bytes. The key schedule is computed at each call and the loading of the bytes from the memory and storing them back to memory are included. The reported speed was taken as an average over multiple executions of the code with the same fixed message length. Naturally, smaller message or associated data sizes will lead to slower c/B performances, due to the various initialization overhead.

<table>
<thead>
<tr>
<th></th>
<th>Intel Haswell (gcc v9.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD bytes</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 0 0 64 576 1.5k 64</td>
</tr>
<tr>
<td>AD bytes</td>
<td>65k 65k</td>
</tr>
<tr>
<td>GNSIV-N (</td>
<td>K_2</td>
</tr>
<tr>
<td>GNSIV-N (</td>
<td>K_2</td>
</tr>
<tr>
<td>GNSIV-Z (</td>
<td>K_2</td>
</tr>
<tr>
<td>GNSIV-Z (</td>
<td>K_2</td>
</tr>
<tr>
<td>AES-GCM-SIV-128</td>
<td>10.00 2.62 1.77 8.87 1.90 1.14 6.29 1.68 1.23 1.34</td>
</tr>
</tbody>
</table>

In our modes, no specific operation is used in addition to the TBC calls; they would therefore largely outperform AES-GCM-SIV even when instantiated with Deoxys-TBC. If we consider using a lightweight cipher as internal primitive, the contrast would be even clearer: our modes with SKINNY-128/512 variant would remain lightweight, while GCM-SIV with a lightweight cipher would be very large (see for example Fig. 8 of [KMY18], where AES-GCM is the worst mode for area).

Acknowledgments.

We would like to thank the authors of [NSS20a] who provided us their MILP models generator for computing the number of active Sboxes in SKINNYe-64/256. We are grateful to the anonymous reviewers for their comments that improved the quality of this article.

Benoît Cogliati carried out this work in the framework of the French-German-Center for Cybersecurity, a collaboration of CISPA and LORIA. Thomas Peyrin was supported by the NRF-ANR project SELECT and the Singapore NRF Investigatorship research grant. Yannick Seurin carried out this work while at ANSSI.

References


[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, CRYPTO 2017,


A Balls-into-bins Lemmas

We will repeatedly appeal to the following classical result for the proof of Theorem 2 and Theorem 3.

**Lemma 1.** Assume we throw $M$ balls uniformly at random and independently into $N$ bins with $8 \leq M \leq N$. Then, with probability larger than $1 - 1/N$, the maximal number of balls in any bin is $2 \log_2(M)$.

**Proof.** Let $X_i$, $1 \leq i \leq N$, denote the load of the $i$-th bin and let $Y_{i,j}$, $1 \leq i \leq N$ and $1 \leq j \leq M$, be the indicator variable which is 1 if ball $j$ lands in bin $i$ and 0 otherwise. Then $X_i = \sum_{j=1}^{M} Y_{i,j}$ and $\Pr[Y_{i,j} = 1] = 1/N$. Let $\mu$ be the expected value of $X_i$. By
linearity of the expectation, \( \mu = M/N \). Since \( X_i \) is the sum of independent Bernoulli trials, by the multiplicative Chernoff bound, we have that for any \( \delta > 0 \),

\[
\Pr[X_i \geq (1 + \delta)\mu] \leq \left( \frac{e^{\delta}}{(1 + \delta)^{1+\gamma}} \right)^{\mu}.
\]

Letting \( \alpha = 1 + \delta \), we obtain

\[
\Pr[X_i \geq \alpha \mu] \leq \left( \frac{e^{\alpha-1}}{\alpha^\alpha} \right)^{\mu} < \left( \frac{e}{\alpha} \right)^{\alpha^\mu}.
\]

Taking \( \alpha = 2N \log_2(M)/M \) yields (note that \( \alpha > 1 \))

\[
\Pr[X_i \geq 2 \log_2(M)] \leq \left( \frac{eM}{2N \log_2(M)} \right)^{2 \log_2(M)} = \frac{M^2}{N^2} \left( \frac{e}{2 \log_2(M)} \right)^{2 \log_2(M)} \left( \frac{M}{N} \right)^{2 \log_2(M)-2} \leq \frac{1}{N^2} \left( \frac{M}{N} \right)^{2 \log_2(M)-2} \leq \frac{1}{N^2} \left( \frac{M}{N} \right)^{2 \log_2(M)-2}.
\]

By the union bound over all bins, the probability that there is a bin with \( 2 \log_2(M) \) balls or more is at most \( 1/N \).

We will also need the following variant.

**Lemma 2.** Consider \( N \) bins numbered from 0 to \( N-1 \). Fix some integer \( q, 1 \leq q \leq M \), and a sequence of integers \( (\ell_1, \ldots, \ell_q) \) with \( 1 \leq \ell_j \leq M \) and \( \sum_{j=1}^q \ell_j = M \). Consider the following random process: for \( j = 1, \ldots, q \), an initial bin \( x_j \) is drawn uniformly at random and a ball is thrown in each bin \( x_j, x_j+1, \ldots, x_j+\ell_j-1 \), where bin numbers are taken mod \( N \). (Hence, at the end of the process, \( M \) balls have been thrown in total.) Then, with probability larger than \( 1 - 1/N \), the maximal number of balls in any bin is \( 2 \log_2(M) \).

**Proof.** The proof is similar to the proof of Lemma 1. Let \( X_i, 1 \leq i \leq N \), denote the load of the \( i \)-th bin and let \( Y_{i,j}, 1 \leq i \leq N \) and \( 1 \leq j \leq q \), be the indicator variable which is 1 if some ball of the \( j \)-th throw lands in bin \( i \) and 0 otherwise. Then \( X_i = \sum_{j=1}^q Y_{i,j} \) (as for each throw, at most one ball can land in any bin) and \( \Pr[Y_{i,j} = 1] = \ell_j/N \). Let \( \mu \) be the expected value of \( X_i \). By linearity of the expectation, \( \mu = \sum_{j=1}^q \ell_j/N = M/N \). Since \( X_i \) is the sum of independent Poisson trials, the multiplicative Chernoff bound applies as in the proof of Lemma 1 with exactly the same parameters.

**B Proof of Theorem 2**

The proof of Theorem 2 relies on the H-coefficients method. Fix a deterministic adversary \( \mathcal{A} \) and assume without loss of generality that \( \mathcal{A} \) makes exactly \( q_{ic} \) queries in total to IC or IC\(^{-1}\) and exactly \( q_{enc} \) queries to Enc of total length (in number of \( n \)-bit blocks) exactly \( \sigma \) and that it never makes a pointless query, where a pointless query is either:

- a repeated query to IC, IC\(^{-1}\), or Enc,
• a query IC\((K, T, X)\) if there was a previous query IC\(^\{-1\}(K, T, Y)\) that returned X or a query IC\(^\{-1\}(K, T, Y)\) if there was a previous query IC\((K, T, X)\) that returned Y,

• a query ENC\((i, N, M)\) such that \(M = \epsilon\) or \(i > u\), where \(u\) is the current value of the counter keeping track of \(\text{New}\) queries.

The transcript consists of queries to the ideal cipher that we record as a list \(\tau_{\text{ic}}\) containing tuples \((K, T, X, Y)\) such that \(A\) made either a query IC\((K, T, X)\) that returned Y or a query IC\(^\{-1\}(K, T, Y)\) that returned X and queries to ENC that we record as a list \(\tau_{\text{enc}}\) containing tuples \((i, N, M, (R, C))\) such that \(A\) made a query ENC\((i, N, M)\) that returned \((R, C)\). If an ENC query is such that the message length is not a multiple of \(n\), we assume that the last block is padded with enough zeros so that it has length \(n\) before returning the oracle answer (this is \(\text{wlog}\) as this can only increase the adversary’s advantage). With this convention, the length of messages and ciphertexts in encryption query/answer pairs is always a multiple of \(n\), and we write such a pair \((i, N, M_0) \cdots \| M_{\ell - 1}, (R_0) \cdots \| C_{\ell - 1})\) where \(\ell = |M|\) and \(|M_j| = |C_j| = n\) for every \(j \in [0, \ell - 1]\). Queries to \(\text{New}\) are not explicitly recorded, but after the adversary has finished interacting with the oracles, we reveal the keys \(K = (K_1, \ldots, K_u)\) generated by \(\text{New}\) calls, where \(u\) is the final value of the counter keeping track of \(\text{New}\) queries (note that unlike \(q_{\text{enc}}\), \(u\) is not fixed and depends on the transcript) and add \(K\) to the transcript. (Again, this is \(\text{wlog}\) as this can only increase the adversary’s advantage; note that these keys are also generated in the ideal world, but never used afterwards.) We let \(\Lambda_{\text{re}}\), resp. \(\Lambda_{\text{id}}\) denote the probability distribution of the transcript when \(A\) interacts with game \(\text{Real}_{\text{GCTR}\{(E, f_{\text{IC}}), A\}}\), resp. \(\text{Ideal}_{\text{GCTR}\{(E, f_{\text{IC}}), A\}}\) (see Figure 3).

From the transcript of encryption queries \(\tau_{\text{enc}}\), we define the split transcript \(\tau_{\text{enc}}^\prime\) as follows. For a query \(Q = (i, N, M_0) \cdots \| M_{\ell - 1}, (R_0) \cdots \| C_{\ell - 1})\) \(\in \tau_{\text{enc}}\), we define the list
\[
S(Q) := \{(i, N, R, j, M_j, C_j)\}_{j \in [0, \ell - 1]}
\]
Then \(\tau_{\text{enc}}^\prime := \|Q \in \tau_{\text{enc}} S(Q)\), the concatenation of all lists \(S(Q)\) for \(Q \in \tau_{\text{enc}}\). In all the following, we call a tuple \((i, N, R, j, M, C)\) \(\in \tau_{\text{enc}}\) a split encryption query, or simply split query for short. Note that \(\tau_{\text{enc}}^\prime\) contains exactly \(\ell\) split queries and that each split query \((i, N, R, j, M, C) \in \tau_{\text{enc}}^\prime\) is associated with an evaluation of IC for which it allows to recover the corresponding key \(K_i\), tweak \(f_T(N, R, j)\), plaintext \(f_X(N, R, j)\), and output \(M \oplus C\).

We say that a transcript \(\tau = (\tau_{\text{ic}}, \tau_{\text{enc}}, K)\) is bad if one of the following conditions is satisfied:

(C-1) There exist two distinct\(^{14}\) split encryption queries \((i, N, R, j, M, C) \in \tau_{\text{enc}}\) and \((i', N', R', j', M', C') \in \tau_{\text{enc}}\) such that
\[
\begin{align*}
K_i = K_{i'}, \\
f_T(N, R, j) = f_T(N', R', j') \\
f_X(N, R, j) = f_X(N', R', j').
\end{align*}
\]

(C-2) There exist two distinct split encryption queries \((i, N, R, j, M, C) \in \tau_{\text{enc}}\) and \((i', N', R', j', M', C') \in \tau_{\text{enc}}\) such that
\[
\begin{align*}
K_i = K_{i'}, \\
f_T(N, R, j) = f_T(N', R', j') \\
M \oplus C = M' \oplus C'.
\end{align*}
\]

\(^{14}\)By distinct, we mean at different positions in the list \(\tau_{\text{enc}}\), but the two tuples might be equal.
(C-3) There exist \((i, N, R, j, M, C) \in \mathcal{E}_{\text{enc}}\) and \((K, T, X, Y) \in \tau_{\text{ic}}\) such that

\[
\begin{align*}
K_i &= K \\
R.T(N, R, j) &= T \\
R.X(N, R, j) &= X.
\end{align*}
\]

(C-4) There exist \((i, N, R, j, M, C) \in \mathcal{E}_{\text{enc}}\) and \((K, T, X, Y) \in \tau_{\text{ic}}\) such that

\[
\begin{align*}
K_i &= K \\
R.T(N, R, j) &= T \\
M \oplus C &= Y.
\end{align*}
\]

Otherwise, we say that \(\tau\) is good and let \(\Theta_{\text{bad}}\), resp. \(\Theta_{\text{good}}\) denote the set of bad, resp. good transcripts.

**Probability of bad transcripts.** First, we upper bound the probability of bad transcripts in the ideal world. This depends on the specific choice of functions \(R.T\) and \(R.X\). We consider each condition in turn, letting \(\Theta_i\) denote the set of transcripts satisfying condition (C-i), \(i \in [1, 4]\).

**Condition (C-1).** We must distinguish \(\text{CTRT}\) and the four other modes. Before that, we define two subsets \(\Theta_{1,1}\) and \(\Theta_{1,2}\) of \(\Theta_1\). Subset \(\Theta_{1,1}\) consists of transcripts such that there exist distinct split queries \((i, N, R, j, M, C)\) and \((i, N', R', j', M', C')\) in \(\mathcal{E}_{\text{enc}}\) such that \(R.T(N, R, j) = R.T(N', R', j')\) and \(R.X(N, R, j) = R.X(N', R', j')\). (Note the equality of the first component of the two split queries.) Subset \(\Theta_{1,2}\) consists of transcripts such that there exist distinct split queries \((i, N, R, j, M, C)\) in \(\mathcal{E}_{\text{enc}}\) and \((i', N', R', j', M', C')\) in \(\mathcal{E}_{\text{enc}}\) such that \(i \neq i', K_i = K_{i'}, R.T(N, R, j) = R.T(N', R', j')\), and \(R.X(N, R, j) = R.X(N', R', j')\).

**CTRT:** The condition is equivalent to \((K_i, R + \langle j \rangle_c, N) = (K_{i'}, R' + \langle j' \rangle_c, N')\). The two split queries cannot originate from the same encryption query as otherwise \(R = R'\) and hence \(\langle j \rangle_c = \langle j' \rangle_c\), which would imply that these two split queries are the same. First, we introduce some additional notation. For any (user, nonce) pair \((i, N)\) appearing in the encryption queries, let \(q_{\text{enc}}(i, N)\) denote the number of encryption queries for this pair \((i, N)\), let \(\ell_0(i, N), 1 \leq \ell \leq q_{\text{enc}}(i, N)\), denote the length of the \(\ell\)-th encryption query (for some arbitrary ordering) for pair \((i, N)\), and let \(\sigma(i, N) = \sum_{i=1}^{q_{\text{enc}}(i, N)} \ell(i, N)\). Note that \(q_{\text{enc}}(i, N) \leq \mu\) for any pair \((i, N)\) and \(\sum(i, N) \sigma(i, N) = \sigma\). We start with \(\Theta_{1,1}\) (i.e., the two split queries are for the same user, so that the equality of the keys is automatically satisfied). Fix some pair \((i, N)\) and consider two distinct encryption queries for this (user, nonce) pair \((i, N, M, (R, C))\) and \((i, N', M', (R', C'))\) of length respectively \(\ell\) and \(\ell'\). Then, the probability over the random draw of \(R\) and \(R'\) that \(R + \langle j \rangle_c = R' + \langle j' \rangle_c\) for some \(j \in [0, \ell - 1]\) and \(j' \in [0, \ell' - 1]\) is \((\ell + \ell' - 1)/2^\mu\). Summing over all pairs of queries...
and all pairs \((i, N)\), we have

\[
\Pr[\Lambda_{id} \in \Theta_{1,1}] \leq \sum_{(i, N)} \sum_{\ell_i = i + 1}^{q_{\text{enc}}(i, N) - 1} \sum_{\ell' = i + 1}^{q_{\text{enc}}(i, N) - 1} \frac{f_\ell(i, N) + \ell' \cdot (i, N) - 1}{2^r}
\]

\[
\leq \sum_{(i, N)} \sum_{\ell_i = 1}^{q_{\text{enc}}(i, N) - 1} \frac{(q_{\text{enc}}(i, N) - 1)\ell_i + \sigma(i, N)}{2^r}
\]

\[
\leq \sum_{(i, N)} 2(q_{\text{enc}}(i, N) - 1)\sigma(i, N)
\]

\[
\leq \frac{2(\mu - 1)\sigma(i, N)}{2^r}
\]

Consider now \(\Theta_{1,2}\). Let \(\ell_i, 1 \leq i \leq q_{\text{enc}}\), denote the length of the \(i\)-th encryption query (for some arbitrary ordering). Consider two distinct encryption queries \((i, N, M, (R, C))\) and \((i', N', M', (R', C'))\) with \(i \neq i'\) of length respectively \(\ell\) and \(\ell'\). Then \(K_i = K_{i'}\) with probability \(2^{-k}\) and, as before, the probability over the random draw of \(R\) and \(R'\) that \(R + \langle j \rangle_c = R' + \langle j' \rangle_c\) for some \(j \in [0, \ell - 1]\) and \(j' \in [0, \ell' - 1]\) is \((\ell + \ell' - 1)/2^r\). Summing over all pairs of queries, we have

\[
\Pr[\Lambda_{id} \in \Theta_{1,2}] \leq \sum_{i=1}^{q_{\text{enc}} - 1} \sum_{\ell_i = i + 1}^{q_{\text{enc}}} \frac{\ell_i + \ell_{i'} - 1}{2^{k+r}}
\]

\[
\leq \sum_{i=1}^{q_{\text{enc}} - 1} \frac{(q_{\text{enc}} - 1)\ell_i + \sigma}{2^{k+r}}
\]

\[
\leq \frac{2(q_{\text{enc}} - 1)\sigma}{2^{k+r}}
\]

All in all,

\[
\Pr[\Lambda_{id} \in \Theta_{1}] \leq \frac{2(\mu - 1)\sigma}{2^r} + \frac{2q_{\text{enc}}\sigma}{2^{k+r}} \leq \frac{2\mu\sigma}{2^r},
\]

where we used \(q_{\text{enc}}/2^k \leq 1\) for the last inequality.

**GCTR-N, R, C, Z:** One can easily check that for all four variants, condition (C-1) is equivalent to

\[
(K_i, N, R, \langle j \rangle_c) = (K_{i'}, N', R', \langle j' \rangle_c).
\]

Two split queries \((i, N, R, j, M, C) \in \tau_{\text{enc}}\) and \((i', N', R', j', M', C') \in \tau_{\text{enc}}\) satisfying (C-1) cannot originate from the same encryption query as otherwise \(\langle j \rangle_c = \langle j' \rangle_c\) would imply that these two split queries are the same. Hence, condition (C-1) is satisfied if there exist two distinct encryption queries \((i, N, M, (R, C)) \in \tau_{\text{enc}}\) and \((i', N', M', (R', C')) \in \tau_{\text{enc}}\) such that \((K_i, N, R) = (K_{i'}, N', R')\). We start with \(\Theta_{1,1}\). Since \(R\) and \(R'\) are independently distributed for two distinct encryption queries, \(R = R'\) is satisfied with probability \(2^{-r}\). Moreover, \(N = N'\) implies that we only need to apply the union bound on the at most \((\mu - 1)q_{\text{enc}}/2\) unordered pairs of encryption queries sharing the same nonce. (In details: fix an encryption query; then there are at most \(\mu - 1\) other encryption queries with the same nonce; summing over the \(q_{\text{enc}}\) queries and dividing by two as each pair is counted twice yields the result.) Hence, \(\Pr[\Lambda_{id} \in \Theta_{1,1}] \leq (\mu - 1)q_{\text{enc}}/2^{r+1}\). Consider now \(\Theta_{1,2}\). Then \(K_i = K_{i'}\) with probability \(2^{-k}\) and \(R = R'\) with probability \(2^{-r}\) (on the other hand, the adversary
can set \( N = N' \) at will. Summing over the \( q_{\text{enc}}(\frac{q_{\text{enc}} - 1}{2}) \) unordered pairs of encryption queries, we obtain \( \Pr[\Lambda_{id} \in \Theta_{1,2}] \leq \frac{q_{\text{enc}}^2}{2^{k+r+1}}. \) All in all,

\[
\Pr[\Lambda_{id} \in \Theta_1] \leq \frac{(\mu - 1)q_{\text{enc}}}{2^{r+1}} + \frac{q_{\text{enc}}}{2^{k+r+1}} \leq \frac{\mu q_{\text{enc}}}{2^{r+1}},
\]

where we used \( q_{\text{enc}}/2^{k} \leq 1 \) for the last inequality.

**Condition (C-2).** Here we must distinguish the five variants. Before that, we define two subsets \( \Theta_{2,1} \) and \( \Theta_{2,2} \) of \( \Theta_2 \). Subset \( \Theta_{2,1} \) consists of transcripts such that there exist distinct split queries \((i,N,R,j,M,C)\) and \((i,N',R',j',M',C')\) in \( \tau_{\text{enc}} \) such that \( f_{r}(N,R,j) = f_{r}(N',R',j') \) and \( M \oplus C = M' \oplus C' \). Subset \( \Theta_{2,2} \) consists of transcripts such that there exist distinct split queries \((i,N,R,j,M,C)\) \( \in \tau_{\text{enc}} \) and \((i',N',R',j',M',C')\) \( \in \tau_{\text{enc}} \) such that \( i \neq i' \), \( K_i = K_{i'} \), \( f_{r}(N,R,j) = f_{r}(N',R',j') \), and \( M \oplus C = M' \oplus C' \). In the following, we fix two split queries \((i,N,R,j,M,C)\) \( \in \tau_{\text{enc}} \) and \((i',N',R',j',M',C')\) \( \in \tau_{\text{enc}} \) (with either \( i = i' \) for \( \Theta_{2,1} \) or \( i \neq i' \) for \( \Theta_{2,2} \)), upper bound the probability that conditions are satisfied, and then apply the union bound.

**CTRT:** Then \( f_{r}(N,R,j) = f_{r}(N',R',j') \) translates to \( R + \langle j \rangle_c = R' + \langle j' \rangle_c \). We upper bound directly the probability that \( \Lambda_{id} \in \Theta_2 \) without distinguishing whether \( \Lambda_{id} \in \Theta_{2,1} \) or \( \Lambda_{id} \in \Theta_{2,2} \). The two split queries cannot originate from the same encryption query as otherwise \( R = R' \) and hence \( \langle j \rangle_c = \langle j' \rangle_c \), which would imply that these two split queries are the same. Hence \( R \) and \( R' \) are uniformly random and independent and \( R + \langle j \rangle_c = R' + \langle j' \rangle_c \) holds with probability \( 2^{-r} \). Moreover, \( C \) and \( C' \) are uniformly random and independent, hence \( M \oplus C = M' \oplus C' \) holds with probability \( 2^{-n} \). Summing over the \( \sigma (\sigma - 1)/2 \) unordered pairs of split encryption queries, one has

\[
\Pr[\Lambda_{id} \in \Theta_2] \leq \frac{\sigma^2}{2^{r+n+1}}.
\]

**GCTR-N:** Then \( f_{r}(N,R,j) = f_{r}(N',R',j') \) translates to \( (R,\langle j \rangle_c) = (R',\langle j' \rangle_c) \). We upper bound directly the probability that \( \Lambda_{id} \in \Theta_2 \) without distinguishing whether \( \Lambda_{id} \in \Theta_{2,1} \) or \( \Lambda_{id} \in \Theta_{2,2} \). The two split queries cannot originate from the same encryption query as otherwise \( \langle j \rangle_c = \langle j' \rangle_c \), which would imply that these two split queries are the same. Hence \( R \) and \( R' \) are uniformly random and independent and \( R + R' \) holds with probability \( 2^{-r} \). Moreover, \( C \) and \( C' \) are uniformly random and independent, hence \( M \oplus C = M' \oplus C' \) holds with probability \( 2^{-n} \). There are \( \sigma \) possible choices for the first split query \((i,N,R,j,M,C)\); then there are at most \( q_{\text{enc}} \) other split queries \((i',N',R',j',M',C')\) with \( \langle j' \rangle_c = \langle j \rangle_c \); dividing by two as each pair is counted twice, one has

\[
\Pr[\Lambda_{id} \in \Theta_2] \leq \frac{\sigma q_{\text{enc}}}{2^{r+n+1}}.
\]

**GCTR-R:** Then \( f_{r}(N,R,j) = f_{r}(N',R',j') \) translates to \( (N,\langle j \rangle_c) = (N',\langle j' \rangle_c) \). The two split queries cannot originate from the same encryption query as otherwise \( \langle j \rangle_c = \langle j' \rangle_c \), which would imply that these two split queries are the same. We start with \( \Theta_{2,1} \). One has \( M \oplus C = M' \oplus C' \) with probability \( 2^{-n} \). There are \( \sigma \) possible choices for the first split query \((i,N,R,j,M,C)\); then there are at most \( \mu - 1 \) other split queries \((i,N',R',j',M',C')\) with \( (N',\langle j' \rangle_c) = (N,\langle j \rangle_c) \); dividing by two as each pair is counted twice, one has \( \Pr[\Lambda_{id} \in \Theta_{2,1}] \leq (\mu - 1)\sigma / 2^n \). Consider now \( \Theta_{2,2} \). Then \( K_i = K_{i'} \) with probability \( 2^{-k} \) and \( M \oplus C = M' \oplus C' \) with probability \( 2^{-n} \). There are \( \sigma \) possible choices for the first split query \((i,N,R,j,M,C)\); then there are at most \( q_{\text{enc}} \) other split queries \((i',N',R',j',M',C')\) with \( \langle j' \rangle_c = \langle j \rangle_c \); dividing by two as each pair is counted twice, one has \( \Pr[\Lambda_{id} \in \Theta_{2,2}] \leq \sigma q_{\text{enc}} / 2^{k+n+1} \). All in all,
GCTR-C: Then $f_T(N, R, j) = f_T(N', R', j')$ translates to $(N, R) = (N', R')$. Note that here the two split queries may originate from the same encryption query. We start with $\Theta_{2.1}$. One has $M \oplus C = M' \oplus C'$ with probability $2^{-n}$. Let $F$ denote the event that there exist two distinct encryption queries for the same pair $(i, N)$ such that the resulting $R$ values collide. Then $\Pr[F] \leq (\mu - 1)q_{\text{enc}}/2^r + 1$. Conditioned on $\neg F$, we can count the number of pairs of split queries on which we must sum as follows: there are $\sigma$ possible choices for the first split query $(i, N, R, j, M, C)$; then there are at most $\ell_{\text{max}} - 1$ other split queries $(i, N', R', j', M', C')$ with $(N', R') = (N, R)$; dividing by two as each pair is counted twice, one has $\Pr[\Lambda_{\text{id}} \in \Theta_{2.1}] \leq (\mu - 1)q_{\text{enc}}/2^r + 1 + \sigma \ell_{\text{max}}/2^{n+1}$. Consider now $\Theta_{2.2}$. Then $K_i = K_{i'}$ with probability $2^{-k}$ and $M \oplus C = M' \oplus C'$ with probability $2^{-n}$. Summing over the $\sigma(\sigma - 1)$ unordered pairs of distinct split queries, one has $\Pr[\Lambda_{\text{id}} \in \Theta_{2.2}] \leq \sigma^2/2^{k+n+1}$. All in all,

$$\Pr[\Lambda_{\text{id}} \in \Theta_2] \leq \frac{(\mu - 1)q_{\text{enc}}}{2^r + 1} + \frac{\ell_{\text{max}}}{2^{n+1}} + \frac{\sigma^2}{2^{k+n+1}}.$$

GCTR-Z: Here we can observe that if $\tau \notin \Theta_1$, then necessarily $\tau \notin \Theta_2$. Indeed, since condition $f_X(N, R, j) = f_X(N', R', j')$ always vacuously holds, $\tau \notin \Theta_1$ means that for every distinct split encryption queries $(i, N, R, j, M, C) \in \tau_{\text{enc}}$ and $(i', N', R', j', M', C') \in \tau_{\text{enc}}$, either $K_i \neq K_{i'}$ or $(N, R, (j)_c) \neq (N', R', (j')_c)$, which implies that $\tau \notin \Theta_2$. Therefore,

$$\Pr[\Lambda_{\text{id}} \in \Theta_2 \mid \Lambda_{\text{id}} \notin \Theta_1] = 0.$$

Condition (C-3). We must distinguish CTRT and the four other modes.

CTRT: We can view each encryption query $(i, N, M, (R, C)) \in \tau_{\text{enc}}$ as throwing $\ell$ balls $R + (j)_c$ into $2^r$ bins (where $\ell$ is the length of $M$ in $n$-bit blocks), the first bin being chosen uniformly at random and the balls are thrown in consecutive bins. In total, we throw $\sigma \leq 2^n$ balls. Hence, by Lemma 2, each bin contains at most $2 \log_2(\sigma) \leq 2r$ balls, except with probability at most $2^{-r}$. Assume this is the case and fix an ideal cipher query $(K, T, X, Y) \in \tau_{\text{ic}}$. Then there are at most $2\sigma$ split queries $(i, N, R, j, M, C) \in \tau_{\text{enc}}$ such that $R + (j)_c = T$ and the probability that $K_i = K$ for one of those split queries is at most $2r/2^k$. Summing over all $q_{\text{ic}}$ ideal cipher queries, one has

$$\Pr[\Lambda_{\text{id}} \in \Theta_3] \leq \frac{1}{2^r} + \frac{2rq_{\text{ic}}}{2^k}.$$

GCTR-N, R, C, Z: We can view each encryption query $(i, N, M, (R, C)) \in \tau_{\text{enc}}$ as throwing a ball $R$ uniformly at random into $2^r$ bins. In total, we throw $q_{\text{enc}} \leq 2^r$ balls and by Lemma 1 each bin contains at most $2 \log_2(q_{\text{enc}}) \leq 2r$ balls, except with probability at most $2^{-r}$. Assume this is the case and fix a query $(K, T, X, Y) \in \tau_{\text{ic}}$. Then, for all four modes, there are at most $2\sigma$ split queries $(i, N, R, j, M, C) \in \tau_{\text{enc}}$ such that $f_T(N, R, j) = T$ and $f_X(N, R, j) = X$ (since these two equations uniquely fix $R$ and $(j)_c$) and the probability that $K_i = K$ for one of those split queries is at most $2r/2^k$. Summing over all $q_{\text{ic}}$ ideal cipher queries, one has

$$\Pr[\Lambda_{\text{id}} \in \Theta_3] \leq \frac{1}{2^r} + \frac{2rq_{\text{ic}}}{2^k}.$$

Condition (C-4). The reasoning is the same for the five modes. We can view each split query $(i, N, R, j, M, C) \in \tau_{\text{enc}}$ as throwing a ball $M \oplus C$ uniformly at random into $2^r$ bins. In total, we throw $\sigma$ balls and by Lemma 1 each bin contains at most $2 \log_2(\sigma) \leq 2n$ balls, except with probability at most $2^{-n}$. Assume this is the case and fix a query $(K, T, X, Y) \in \tau_{\text{ic}}$. Then, there are at most $2n$ split queries $(i, N, R, j, M, C) \in \tau_{\text{enc}}$ such
that \( M \oplus C = Y \) and the probability that \( K_i = K \) for one of those split queries is at most \( 2n/2^k \). Summing over all \( q_{ic} \) ideal cipher queries, one has

\[
\Pr[A_{id} \in \Theta_k] \leq \frac{1}{2^n} + \frac{2nq_{ic}}{2^k}.
\]

We can obtain an upper bound on \( \Pr[A_{id} \in \Theta_{had}] \) by adding probabilities \( \Pr[A_{id} \in \Theta_i] \) above for \( i \in [1, 4] \).

**Good transcripts probability ratio.** We now prove that for every good transcript \( \tau \), \( \Pr[A_{re} = \tau] \geq \Pr[A_{id} = \tau] \) (meaning we can use Theorem 1 with \( \beta = 0 \)). Fix a good transcript \( \tau = (\tau_{ic}, \tau_{enc}, K) \) and let \( u \) be the length of \( K \). Let also \( \ell_i, i \in [1, q_{enc}] \), denote the length (in \( n \)-bit blocks) of the message (and hence the ciphertext) in the \( i \)-th encryption query. Note that \( \sum_{i \in [1, q_{enc}]} \ell_i = \sigma \). For any \( (K, T) \in \{0, 1\}^k \times \{0, 1\}^t \), let \( \mathcal{X}_{ic}(K, T) \) denote the set of inputs \( X \in \{0, 1\}^n \) such that there exists \( Y \) with \( (K, T, X, Y) \in \tau_{ic} \) and \( \mathcal{Y}_{ic}(K, T) \) denote the set of outputs \( Y \in \{0, 1\}^n \) such that there exists \( X \) with \( (K, T, X, Y) \in \tau_{ic} \). Note that since \( \tau \) is good, for any \( (K, T, X) \in \{0, 1\}^k \times \{0, 1\}^t \times \{0, 1\}^n \), there is at most one split query \( (i, N, R, j, M, C) \in \tau_{enc} \) such that \( K = K_i, T = f_T(N, R, j), \) and \( X = f_X(N, R, j) \) as otherwise condition (C-1) would be satisfied. With this in mind, let \( \mathcal{X}_{enc}(K, T) \) denote the set of \( X \in \{0, 1\}^n \) such that IC was evaluated on \( (K, T, X) \) in some encryption query, i.e.,

\[
\mathcal{X}_{enc}(K, T) := \{ X \in \{0, 1\}^n : \exists (i, N, R, j, M, C) \in \tau_{enc} : K = K_i \land T = f_T(N, R, j) \land X = f_X(N, R, j) \}.
\]

Note that by the previous observation, one has \( \sum_{(K, T)} |\mathcal{X}_{enc}(K, T)| = \sigma \). Similarly, for any \( (K, T, Y) \in \{0, 1\}^k \times \{0, 1\}^t \times \{0, 1\}^n \), there is at most one split query \( (i, N, R, j, M, C) \in \tau_{enc} \) such that \( K = K_i, T = f_T(N, R, j), \) and \( Y = M \oplus C \) (as otherwise condition (C-2) would be satisfied). Let \( \mathcal{Y}_{enc} \) be the set defined as

\[
\mathcal{Y}_{enc}(K, T) := \{ Y \in \{0, 1\}^n : \exists (i, N, R, j, M, C) \in \tau_{enc} : K = K_i \land T = f_T(N, R, j) \land Y = M \oplus C \}.
\]

Because \( \tau \) is good, for any \( (K, T) \), \( \mathcal{X}_{ic}(K, T) \cap \mathcal{X}_{enc}(K, T) = \emptyset \) (as otherwise condition (C-3) would be satisfied) and \( \mathcal{Y}_{ic}(K, T) \cap \mathcal{Y}_{enc}(K, T) = \emptyset \) (as otherwise condition (C-4) would be satisfied). Hence, for each pair \( (K, T), \tau_{ic} \) and \( \tau_{enc} \) together impose a set of \( |\mathcal{X}_{ic}(K, T)| + |\mathcal{X}_{enc}(K, T)| \) equations on the random permutation \( E_{ic}(K, T, \cdot) \) internally sampled by the ideal cipher of the form \( E_{ic}(K, T, X) = Y \) where all \( X \)'s are distinct and all \( Y \)'s are distinct. From this it follows that

\[
\Pr[A_{re} = \tau] = \frac{1}{2^{|K|}} \cdot \frac{1}{2^{|\mathcal{Y}_{enc}|}} \cdot \left( \prod_{K \in \{0, 1\}^k} \prod_{\tau \in \{0, 1\}^t} \frac{|\mathcal{X}_{ic}(K, T)|}{2^n} \cdot \frac{1}{2^{|\mathcal{Y}_{ic}|}} \right),
\]

where the first term accounts for the random choice of keys, the second term for the random choice of \( R \) values, and the third term for the probability that \( E_{ic} \) satisfies the constraints imposed by the transcripts \( \tau_{ic} \) and \( \tau_{enc} \). On the other hand, since in the ideal world the \( \sigma \) ciphertext blocks are uniformly random and independent, one has

\[
\Pr[A_{id} = \tau] = \frac{1}{2^{|K|}} \cdot \frac{1}{2^{|\mathcal{Y}_{enc}|}} \cdot \left( \prod_{K \in \{0, 1\}^k} \prod_{\tau \in \{0, 1\}^t} \frac{|\mathcal{X}_{ic}(K, T)|-1}{2^n} \cdot \frac{1}{2^{|\mathcal{Y}_{ic}|}} \right).
\]
Hence,
\[
\frac{\Pr[\Lambda_{ic} = \tau]}{\Pr[\Lambda_{id} = \tau]} = 2^{\sigma n} \left( \prod_{K \in \{0,1\}^k} \prod_{T \in \{0,1\}^t}^{\lvert \mathcal{X}_{enc}(K,T) \rvert - 1} \frac{1}{2^n - \lvert \mathcal{X}_{ic}(K,T) \rvert} \right) = \prod_{K \in \{0,1\}^k} \prod_{T \in \{0,1\}^t}^{\lvert \mathcal{X}_{enc}(K,T) \rvert - 1} \frac{2^n}{2^n - \lvert \mathcal{X}_{ic}(K,T) \rvert - 1} \geq 1,
\]
where we used that \( \sum_{(K,T)} \lvert \mathcal{X}_{enc}(K,T) \rvert = \sigma \).

**Concluding.** The theorem follows from Theorem 1 with \( \beta = 0 \) by collecting all cases from the bad transcripts analysis and simplifying the bounds using \( q_{enc}/2^r \leq 1, \ q_{enc}/2^k \leq 1, \) and \( \sigma/2^n \leq 1 \) as follows:

**CTRT:**
\[
\text{Adv}_{\text{CTRT}}^{\text{mu-nive}}(\mathcal{A}) \leq \frac{2\sigma^2}{2^r} \frac{\sigma q_{enc}}{2^{r+n+1}} + \frac{1}{2^r} + \frac{2r q_{ic}}{2^k} + \frac{1}{2^n} + \frac{2n q_{ic}}{2^k}
\]
\[
\leq \frac{1}{2^r} + \frac{1}{2^n} + \frac{2(r+n) q_{ic}}{2^k} + \frac{(4\mu + 1)\sigma}{2^{r+1}}
\]

**GCTR-N:**
\[
\text{Adv}_{\text{GCTR-N}}^{\text{mu-nive}}(\mathcal{A}) \leq \frac{\mu q_{enc}}{2^{r+1}} + \frac{\sigma q_{enc}}{2^{r+n+1}} + \frac{1}{2^r} + \frac{2r q_{ic}}{2^k} + \frac{1}{2^n} + \frac{2n q_{ic}}{2^k}
\]
\[
\leq \frac{1}{2^r} + \frac{1}{2^n} + \frac{2(r+n) q_{ic}}{2^k} + \frac{\mu q_{enc}}{2^{r+1}} + \frac{\sigma}{2^{r+1}}
\]

**GCTR-R:**
\[
\text{Adv}_{\text{GCTR-R}}^{\text{mu-nive}}(\mathcal{A}) \leq \frac{\mu q_{enc}}{2^{r+1}} + \frac{(\mu - 1)\sigma}{2^n} + \frac{\sigma q_{enc}}{2^{k+n+1}} + \frac{1}{2^r} + \frac{2r q_{ic}}{2^k} + \frac{1}{2^n} + \frac{2n q_{ic}}{2^k}
\]
\[
\leq \frac{1}{2^r} + \frac{1}{2^n} + \frac{2(r+n) q_{ic}}{2^k} + \frac{\mu q_{enc}}{2^{r+1}} + \frac{\mu \sigma}{2^n}
\]

**GCTR-C:**
\[
\text{Adv}_{\text{GCTR-C}}^{\text{mu-nive}}(\mathcal{A}) \leq \frac{\mu q_{enc}}{2^{r+1}} + \frac{(\mu - 1)q_{enc}}{2^{r+1}} + \frac{\sigma^2}{2^{k+n+1}} + \frac{1}{2^r} + \frac{2r q_{ic}}{2^k} + \frac{1}{2^n} + \frac{2n q_{ic}}{2^k}
\]
\[
\leq \frac{1}{2^r} + \frac{1}{2^n} + \frac{2(r+n) q_{ic}}{2^k} + \frac{(2\mu - 1)q_{enc}}{2^{r+1}} + \frac{\sigma^2}{2^{k+n+1}} + \frac{\sigma^2}{2^{k+n+1}}
\]

**GCTR-Z:**
\[
\text{Adv}_{\text{GCTR-Z}}^{\text{mu-nive}}(\mathcal{A}) \leq \frac{\mu q_{enc}}{2^{r+1}} + \frac{1}{2^r} + \frac{2r q_{ic}}{2^k} + \frac{1}{2^n} + \frac{2n q_{ic}}{2^k}
\]
\[
\leq \frac{1}{2^r} + \frac{1}{2^n} + \frac{2(r+n) q_{ic}}{2^k} + \frac{\mu q_{enc}}{2^{r+1}}
\]
Proof of Theorem 3

The proof uses the H-coefficients technique. Fix a deterministic adversary $A$ and assume without loss of generality that $A$ makes exactly $q_{\text{ic}}$ queries to IC or IC$^{-1}$, exactly $q_{\text{tag}}$ queries to Tag of total length at most $\sigma_{\text{tag}}$, exactly $q_{\text{ver}}$ queries to Ver of total length at most $\sigma_{\text{ver}}$, and that it never makes a pointless query, where a pointless query is either:

- a repeated query to IC, IC$^{-1}$, Tag, or Ver,
- a query IC$(K, T, X)$ if there was a previous query IC$^{-1}$(K, T, Y) that returned X or a query IC$^{-1}$(K, T, Y) if there was a previous query IC$(K, T, X)$ that returned Y,
- a query Tag$(i, N, U)$ or Ver$(i, N, U, V)$ such that $i > u$, where u is the current value of the counter keeping track of NEW queries,
- a query Ver$(i, N, U, V)$ if there was a previous query Tag$(i, N, U)$ that returned V.

The transcript consists of three types of queries:

- queries to the ideal cipher that we record as a list $\tau_{\text{ic}}$ containing tuples $(K, T, X, Y)$ such that $A$ made either a query IC$(K, T, X)$ that returned Y or IC$^{-1}$(K, T, Y) that returned X;
- queries to Tag that we record as a list $\tau_{\text{tag}}$ containing tuples $(i, N, U, V)$ such that $A$ made a query Tag$(i, N, U)$ that returned V;
- queries to Ver that we record as a list $\tau_{\text{ver}}$ containing all tuples $(i, N, U, V)$ that were queried to Ver (we do not keep track of the answers since we are interested in attainable transcripts, i.e., transcripts that can be obtained in the ideal world in which all queries to Ver return $\perp$).

We do not keep track explicitly of NEW queries, but when $A$ has finished interacting with the oracles, we reveal all keys $K = ((K^1_{\text{in}}, K^1_{\text{out}}), \ldots, (K^u_{\text{in}}, K^u_{\text{out}}))$ generated by calls to NEW, where $u$ is the final value of the counter keeping track of NEW queries (note that $u$ is transcript-dependent) and add them to the transcript (this is $\text{wlog}$ as this can only increase the adversary’s advantage; note that these keys are also generated in the ideal world but never used afterwards). We let $\Lambda_{\text{re}}$, resp. $\Lambda_{\text{id}}$ denote the probability distribution of the transcript when $A$ interacts with game $\text{Real}_{\text{Nat}[H, E]}(A)$, resp. $\text{Ideal}_{\text{Nat}[H, E]}(A)$ (see Figure 4).

We say that a transcript $\tau = (\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}}, K)$ is bad if one of the following conditions is satisfied:

(C-1) There exist two distinct queries $(i, N, U, V) \in \tau_{\text{tag}}$ and $(i', N', U', V') \in \tau_{\text{tag}}$ such that

\[
\begin{cases}
K^i_{\text{out}} = K'^{i'}_{\text{out}} \\
N = N' \\
H_{K^i_{\text{in}}}(U) = H_{K'^{i'}_{\text{in}}}(U').
\end{cases}
\]

(C-2) There exist two distinct queries $(i, N, U, V) \in \tau_{\text{tag}}$ and $(i', N', U', V') \in \tau_{\text{tag}}$ such that

\[
\begin{cases}
K^i_{\text{out}} = K'^{i'}_{\text{out}} \\
N = N' \\
V = V'.
\end{cases}
\]
(C-3) There exist queries \((i, N, U, V) \in \tau_{\text{tag}}\) and \((K, T, X, Y) \in \tau_{\text{ic}}\) such that
\[
\begin{align*}
K_i^i &= K \\
0^{-\nu} \| N &= T \\
H_{K_i^i}(U) &= X.
\end{align*}
\]

(C-4) There exist queries \((i, N, U, V) \in \tau_{\text{tag}}\) and \((K, T, X, Y) \in \tau_{\text{ic}}\) such that
\[
\begin{align*}
K_i^i &= K \\
0^{-\nu} \| N &= T \\
V &= Y.
\end{align*}
\]

(C-5) There exist queries \((i, N, U, V) \in \tau_{\text{tag}}\) and \((i', N', U', V') \in \tau_{\text{ver}}\) such that
\[
\begin{align*}
K_i^i &= K_i^{i'} \\
N &= N' \\
H_{K_i^i}(U) &= H_{K_i^{i'}}(U') \\
V &= V'.
\end{align*}
\]

(C-6) There exist queries \((i, N, U, V) \in \tau_{\text{ver}}\) and \((K, T, X, Y) \in \tau_{\text{ic}}\) such that
\[
\begin{align*}
K_i^i &= K \\
0^{-\nu} \| N &= T \\
H_{K_i^i}(U) &= X \\
V &= Y.
\end{align*}
\]

Otherwise, we say that \(\tau\) is good and let \(\Theta_{\text{bad}}, \text{ resp. } \Theta_{\text{good}}\) denote the set of bad, resp. good transcripts.

**Probability of bad transcripts.** First, we upper bound the probability of bad transcripts in the ideal world. We consider each condition in turn, letting \(\Theta_i\) denote the set of transcripts satisfying condition (C-\(i\)), \(i \in [1, 6]\). For conditions (C-1), (C-3), (C-5), and (C-6), we assume \(\text{wlog}\) that the queries transcript \((\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}})\) is fixed as we can upper bound the probability using only the randomness of keys \(K\). We assume that it involves \(u\) users, and for every \(i \in [1, u]\) and \(N \in \{0, 1\}^\nu\) we let \(q_{i,N}\) denote the number of \(\text{Tag}\) queries involving user \(i\) and nonce \(N\) (with \(q_{i,N} = 0\) if there were no such queries). When \(q_{i,N} \geq 1\), we also let \(\ell_{i,N,j}, 1 \leq j \leq q_{i,N}\), denote the length of the \(j\)-th \(\text{Tag}\) query for user \(i\) with nonce \(N\) and we assume that queries are reordered such that \(\ell_{i,N,1} \leq \ell_{i,N,2} \leq \cdots \leq \ell_{i,N,q_{i,N}}\). Note that by our assumptions, we have
\[
q_{i,N} \leq \mu \text{ for every } (i, N), \quad (4)
\]
\[
\sum_{i=1}^u \sum_{N \in \{0,1\}^\nu} q_{i,N} = q_{\text{tag}}, \quad (5)
\]
\[
\text{and } \sum_{i=1}^u \sum_{N \in \{0,1\}^\nu} \sum_{j=1}^{q_{i,N}} \ell_{i,N,j} \leq \sigma_{\text{tag}}. \quad (6)
\]

Similarly, for every \(i \in [1, u]\) and \(N \in \{0, 1\}^\nu\) we let \(q'_{i,N}\) denote the number of \(\text{Ver}\) queries involving user \(i\) and nonce \(N\) (with \(q'_{i,N} = 0\) if there were no such queries) and when
\(q'_{i,N} \geq 1\), we let \(\ell'_{i,N,j}, 1 \leq j \leq q'_{i,N}\), denote the length of the \(j\)-th VER query for user \(i\) with nonce \(N\) and we assume that queries are reordered such that \(\ell'_{i,N,1} \leq \ell'_{i,N,2} \leq \cdots \leq \ell'_{i,N,q'_{i,N}}\). Then

\[
\sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} q'_{i,N} = q_{\text{ver}} \tag{7}
\]

and

\[
\sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} \sum_{j=1}^{q'_{i,N}} \ell'_{i,N,j} \leq \sigma_{\text{ver}}. \tag{8}
\]

In the following, for \((U, U') \in U^2\), we let

\[
\delta_{U,U'} := \max\{\delta(\text{len}(U)), \delta(\text{len}(U'))\} = \alpha \max\{\text{len}(U), \text{len}(U')\}/2^k + \beta/2^n
\]

\[
\gamma_{U,U'} := \min\{\gamma(\text{len}(U)), \gamma(\text{len}(U'))\} = \alpha \min\{\text{len}(U), \text{len}(U')\}/2^k + \beta/2^n.
\]

**CONDITION (C-1).** We define two subsets \(\Theta_{1,1}\) and \(\Theta_{1,2}\) of \(\Theta_1\). Subset \(\Theta_{1,1}\) consists of transcripts \(\tau\) such that there exist distinct queries \((i, N, U, V)\) and \((i', N', U', V')\) in \(\tau_{\text{tag}}\) such that \(N = N'\) and \(H_{K_{\text{in}}}^i(U) = H_{K_{\text{in}}}^i(U')\). Subset \(\Theta_{1,2}\) consists of transcripts \(\tau\) such that there exists \((i, N, U, V)\) and \((i', N', U', V')\) in \(\tau_{\text{tag}}\) such that \(i \neq i'\), \(K_{\text{out}}^i = K_{\text{out}}^{i'}\), \(N = N'\), and \(H_{K_{\text{in}}}^i(U) = H_{K_{\text{in}}}^{i'}(U')\).

We start with \(\Theta_{1,1}\). For any two distinct \((i, N, U, V)\) and \((i', N', U', V')\) in \(\tau_{\text{tag}}\) such that \(N = N'\), one has \(U \neq U'\) by the assumption that \(A\) never repeats queries. Hence, by the assumption that \(H\) is \(\delta\)-sAAU, \(H_{K_{\text{in}}}^i(U) = H_{K_{\text{in}}}^{i'}(U')\) with probability at most \(\delta_{U,U'}\). By the union bound over users \(i \in [1, \nu]\), nonces \(N \in \{0, 1\}^{N}\), and pairs of queries involving \((i, N)\) and using our assumption that queries are ordered according to their lengths, one has

\[
\Pr[A_{id} \in \Theta_{1,1}] \leq \frac{\alpha}{2^k} \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} \sum_{j=1}^{q_{i,N}} (\alpha \ell_{i,N,j} + \beta/2^n)
\]

\[
= \frac{\alpha}{2^k} \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} \sum_{j=1}^{q_{i,N}} (\alpha \ell_{i,N,j} + \beta/2^n)
\]

\[
\leq \frac{\alpha}{2^k} \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} \sum_{j=1}^{q_{i,N}} (q_{i,N} - 1) \ell_{i,N,j} + \beta/2^n \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} q_{i,N} (q_{i,N} - 1)/2
\]

\[
\leq \frac{\alpha}{2^k} \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} \sum_{j=1}^{q_{i,N}} (\mu - 1) \ell_{i,N,j} + \beta/2^n \sum_{i=1}^{\nu} \sum_{N \in \{0, 1\}^{N}} q_{i,N} (\mu - 1)/2
\]

\[
\leq \frac{\alpha}{2^k} (\mu - 1) \sigma_{\text{tag}} + \frac{\beta}{2^n+1} \sigma_{\text{tag}},
\]

where we used (4) for the penultimate inequality and (5) and (6) for the last inequality.

Consider now \(\Theta_{1,2}\). For any two distinct \((i, N, U, V)\) and \((i', N', U', V')\) in \(\tau_{\text{tag}}\) such that \(i \neq i'\), \(K_{\text{out}}^i = K_{\text{out}}^{i'}\) with probability \(2^{-k}\) and \(H_{K_{\text{in}}}^i(U) = H_{K_{\text{in}}}^{i'}(U')\) with probability at most \(\gamma_{U,U'}\) by \(\gamma\)-uniformity of \(H\). (Indeed, assume \(\text{wlog}\) that \(\text{len}(U) \leq \text{len}(U')\); then for any \(K_{\text{out}}^i, H_{K_{\text{in}}}^i(U) = H_{K_{\text{in}}}^{i'}(U')\) with probability at most \(\gamma(\text{len}(U))\) over the draw of \(K_{\text{in}}\).) Consider the \(j\)-th query for user \(i\) and nonce \(N\). Then, by the union bound, the
probability that this query satisfies the condition with any other TAG query is at most \( q_{\text{tag}}(\alpha\ell_i,N,j)/2^{2k} + \beta/2^{k+n} \). By the union bound over all queries, we have
\[
\Pr[\Lambda_{id} \in \Theta_{1,2}] \leq \sum_{i=1}^{u} \sum_{N \in \{0,1\}^v} \sum_{j=1}^{q_{i,N}} \frac{q_{\text{tag}}(\alpha\ell_i,N,j)/2^{2k} + \beta/2^{k+n}}{2^{2k}}
\]
All in all,
\[
\Pr[\Lambda_{id} \in \Theta_1] \leq \frac{\alpha(q_{\text{tag}}/2^{2k}) + \beta(q_{\text{tag}}/2^{k+n})}{2^{2k}} + \frac{\alpha\ell_iN\gamma(\text{len}(U))/2^k + \beta}{2^{k+n}}.
\]

**CONDITION (C-2).** We define two subsets \( \Theta_{2,1} \) and \( \Theta_{2,2} \) of \( \Theta_2 \). Subset \( \Theta_{2,1} \) consists of transcripts \( \tau \) such that there exist distinct queries \((i,N,U,V)\) and \((i,N',U',V')\) in \( \tau_{\text{tag}} \) such that \( N = N' \) and \( V = V' \). Subset \( \Theta_{2,2} \) consists of transcripts \( \tau \) such that there exists \((i,N,U,V)\) and \((i',N',U',V')\) in \( \tau_{\text{tag}} \) such that \( i \neq i' \). \( K_{\text{out}}' = K_{\text{out}} \), \( N = N' \), and \( V = V' \). We start with \( \Theta_{2,1} \). For any two distinct \((i,N,U,V)\) and \((i',N',U',V')\) in \( \tau_{\text{tag}} \), one has \( V = V' \) with probability \( 2^{-n} \) since tags are uniformly random and independent in the ideal world. There are at most \((\mu - 1)q_{\text{tag}}/2\) unordered pairs of queries sharing the same nonce, hence by the union bound \( \Pr[\Lambda_{id} \in \Theta_{2,1}] \leq \frac{(\mu - 1)q_{\text{tag}}/2^{n+1}}{2^{k+n+1}} \). Consider now \( \Theta_{2,2} \). For any two distinct \((i,N,U,V)\) and \((i',N',U',V')\) in \( \tau_{\text{tag}} \) such that \( i \neq i' \), \( K_{\text{out}}' = K_{\text{out}} \) with probability \( 2^{-k} \) and \( V = V' \) with probability \( 2^{-n} \). Summing over the \( q_{\text{tag}}(q_{\text{tag}}-1)/2 \) unordered pairs of queries, one has \( \Pr[\Lambda_{id} \in \Theta_{2,2}] \leq \frac{q_{\text{tag}}/2}{2^{k+n+1}} \). All in all,
\[
\Pr[\Lambda_{id} \in \Theta_2] \leq \frac{(\mu - 1)q_{\text{tag}}}{2^{n+1}} + \frac{q_{\text{tag}}^2}{2^{k+n+1}}.
\]

**CONDITION (C-3).** For any \((i,N,U,V)\) in \( \tau_{\text{tag}} \) and \((K,T,X,Y)\) in \( \tau_{\text{ic}} \), the probability that \( K_{\text{out}}' = K_{\text{out}} \) is \( 2^{-k} \) and the probability that \( H_{K_{\text{in}}}'(U) = X \) is at most \( \gamma(\text{len}(U))/2^k + \beta/2^n \). Summing over all pairs of queries, one has
\[
\Pr[\Lambda_{id} \in \Theta_3] \leq q_{\text{ic}} \sum_{i=1}^{u} \sum_{N \in \{0,1\}^v} \sum_{j=1}^{q_{i,N}} \left( \frac{\alpha\ell_iN,j/2^{2k} + \beta/2^{k+n}}{2^{2k}} \right) \leq \frac{\alpha\ell_iN,\gamma(\text{len}(U))/2^k + \beta}{2^{k+n}} + \frac{q_{\text{ic}}q_{\text{tag}}}{2^{k+n}}.
\]

**CONDITION (C-4).** We can view each query \((i,N,U,V)\) in \( \tau_{\text{tag}} \) as throwing a ball \( V \) uniformly at random into \( 2^n \) bins. In total, we throw \( q_{\text{tag}} \leq 2^n \) balls and by Lemma 1 each bin contains at most \( 2n \) balls, except with probability at most \( 2^{-n} \). Then, for each query \((K,T,X,Y)\) in \( \tau_{\text{ic}} \), there are at most \( 2n \) queries \((i,N,U,V)\) in \( \tau_{\text{tag}} \) such that \( V = Y \) and the probability that \( K_{\text{out}}' = K_{\text{out}} \) for one of those \( 2n \) queries is at most \( 2n/2^k \). Summing over all \( q_{\text{ic}} \) ideal cipher queries yields
\[
\Pr[\Lambda_{id} \in \Theta_4] \leq \frac{1}{2^n} + \frac{2nq_{\text{ic}}}{2^k}.
\]

**CONDITION (C-5).** We define two subsets \( \Theta_{5,1} \) and \( \Theta_{5,2} \) of \( \Theta_5 \). Subset \( \Theta_{5,1} \) consists of transcripts \( \tau \) such that there exist \((i,N,U,V)\) in \( \tau_{\text{tag}} \) and \((i,N',U',V')\) in \( \tau_{\text{ver}} \) such that \( N = N' \), \( H_{K_{\text{in}}}'(U) = H_{K_{\text{in}}}'(U') \), and \( V = V' \). Subset \( \Theta_{5,2} \) consists of transcripts \( \tau \) such that there exist \((i,N,U,V)\) in \( \tau_{\text{tag}} \) and \((i',N',U',V')\) in \( \tau_{\text{ver}} \) such that \( i \neq i' \), \( K_{\text{out}}' = K_{\text{out}}' \), \( N = N' \), \( H_{K_{\text{in}}}'(U) = H_{K_{\text{in}}}'(U') \), and \( V = V' \).

We start with \( \Theta_{5,1} \). We fix any queries \((i,N,U,V)\) in \( \tau_{\text{tag}} \) and \((i,N',U',V')\) in \( \tau_{\text{ver}} \) and distinguish two cases. If the verification query \( \text{VER}(i,N',U',V') \) was made after the
tag query $\text{Tag}(i, N, U)$, then since we assumed that $\mathcal{A}$ does not make pointless queries, either $U \neq U'$ or $V \neq V'$ (otherwise we would have $(i, N', U', V') = (i, N, U, V)$). In the former case, $H_{K^\text{in}}(U) = H_{K^\text{in}}(U')$ with probability at most $\delta_{U, U'}$, while in the latter case the condition cannot be satisfied. If the verification query $\text{Ver}(i, N', U', V')$ was made before the tag query $\text{Tag}(i, N, U)$, then $V$ is uniformly random and independent from $V'$ and hence $V = V'$ with probability $2^{-n}$. In both cases, the condition is satisfied with probability at most $\max(\delta_{U, U'}, 2^{-n}) = \delta_{U, U'}$ by our assumption that $\delta \geq 2^{-n}$. Fix some user $i$ and nonce $N$ and consider the $j$-th tag query and the $j'$-th verification query for $(i, N)$. Then the condition is satisfied for this pair of queries with probability at most

$$\frac{\alpha \max \{\ell_i^{\text{tag}}, \ell_i'^{\text{tag}}\}}{2^k} + \frac{\beta}{2^n} \leq \frac{\alpha (\ell_i^{\text{tag}} + \ell_i'^{\text{tag}})}{2^k} + \frac{\beta}{2^n} \leq \frac{\alpha (\ell_i^{\text{tag}} + \ell_i'^{\text{tag}})}{2^k} + \frac{\beta}{2^n}.$$ 

By the union bound over all users $i$, nonce $N$, and pairs of tag and verification queries for $(i, N)$, we have

$$\Pr[\Lambda_{\text{id}} \in \Theta_{5,1}] \leq \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} \sum_{j=1}^{q_{i,N}'} \left( \frac{\alpha (\ell_i^{\text{tag}} + \ell_i'^{\text{tag}})}{2^k} + \frac{\beta}{2^n} \right).$$

$$\leq \frac{\alpha \mu \ell_i^{\text{tag}}}{2^k} \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} \frac{q_{i,N}'}{2^k} + \frac{\alpha \mu}{2^k} \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} \sum_{j=1}^{q_{i,N}'} \ell_i'^{\text{tag}}$$

$$+ \frac{\beta \mu}{2^n} \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} q_{i,N}'.$$

Consider now $\Theta_{5,2}$. For any $(i, N, U, V) \in \tau_{\text{tag}}$ and $(i', N', U', V') \in \tau_{\text{ver}}$ such that $i \neq i'$, $K^{\text{out}} = K^{\text{out}}$ with probability $2^{-k}$ and $H_{K^\text{in}}(U) = H_{K'_{\text{in}}}(U')$ with probability at most $\gamma_{U, U'}$. Summing over all pairs of queries, we have

$$\Pr[\Lambda_{\text{id}} \in \Theta_{5,2}] \leq \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} \sum_{i'=1}^{u} \sum_{N' \in \{0, 1\}^v} \sum_{j=1}^{q_{i',N',}} \left( \frac{\alpha \ell_i'^{\text{tag}}}{2^k} + \frac{\beta}{2^n} \right)$$

$$\leq \frac{\alpha \mu q_{\text{tag}}}{2^k} + \frac{\beta q_{\text{tag}}}{2^n}.$$ 

All in all,

$$\Pr[\Lambda_{\text{id}} \in \Theta_{5}] \leq \frac{\alpha \mu \ell_{\text{max}}^{\text{tag}}}{2^k} + \frac{\alpha \mu q_{\text{tag}}}{2^k} + \frac{\beta \mu q_{\text{ver}}}{2^n} + \frac{\alpha q_{\text{tag}} q_{\text{ver}}}{2^k} + \frac{\beta q_{\text{tag}} q_{\text{ver}}}{2^n}.$$ 

**CONDITION (C-6).** For any $(i, N, U, V) \in \tau_{\text{ver}}$ and $(K, T, X, Y) \in \tau_{\text{ic}}$, $K^{\text{out}} = K$ with probability $2^{-k}$ and $H_{K_{\text{in}}^i}(U) = X$ with probability at most $\gamma(\text{len}(U))$. By summing over all pairs, one has

$$\Pr[\Lambda_{\text{id}} \in \Theta_{6}] \leq q_{\text{ic}} \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^v} \sum_{j=1}^{q_{i,N}'} \left( \frac{\alpha \ell_i'^{\text{tag}}}{2^k} + \frac{\beta}{2^n} \right) \leq \frac{\alpha q_{\text{ic}} q_{\text{ver}}}{2^k} + \frac{\beta q_{\text{ic}} q_{\text{ver}}}{2^n}.$$
Collecting all probabilities, we obtain
\[
\Pr[\Lambda_{id} \in \Theta] \leq \frac{\alpha(\mu - 1)\sigma_{\tag} + \beta(\mu - 1)q_{\tag}}{2^k} + \frac{\alpha q_{\tag}\sigma_{\tag}}{2^{k+n}} + \frac{\beta q_{\tag}^2}{2^{k+n+1}}
\]
\[
+ \frac{\mu q_{\max} q_{\ver}}{2^k} + \frac{\alpha \mu \sigma_{\ver}}{2^n} + \frac{\beta \mu q_{\ver}}{2^n} + \frac{\alpha q_{\tag}\sigma_{\ver}}{2^{k+1}} + \frac{\beta q_{\tag} q_{\ver}}{2^{k+n}}
\]
\[
\leq \frac{2nq_{\tag}}{2^k} + \frac{\alpha(\mu - 1)\sigma_{\tag}}{2^k} + \frac{\alpha \mu \sigma_{\ver}}{2^n} + \frac{\alpha q_{\tag}\sigma_{\ver}}{2^{k+1}} + \frac{\beta q_{\tag} q_{\ver}}{2^{k+n}}
\]
\[
+ \frac{\mu q_{\max} q_{\ver}}{2^k} + \frac{\beta \mu q_{\ver}}{2^n} + \frac{\alpha q_{\tag} + q_{\ver}}{2^n} - \frac{(2q_{\tag} + q_{\ver})(q_{\tag} + q_{\ver})}{2^n}.
\]

where for the first transition we used \( \beta \geq 1 \) and for the second one we used \( 2q_{\tag} + q_{\ver} \leq 2^k \).

**Good transcripts probability ratio.** Let \( \tau = (\tau_{ic}, \tau_{\tag}; \tau_{\ver}; K) \) be a good transcript and let \( n \) be the length of \( K \). For any \((K, T) \in \{0,1\}^k \times \{0,1\}^t \), let \( \mathcal{X}_{ic}(K, T) \) denote the set of inputs \( X \in \{0,1\}^n \) such that there exists \( Y \) with \((K, T, X, Y) \in \tau_{ic} \) and \( \mathcal{Y}_{ic}(K, T) \) denote the set of outputs \( Y \in \{0,1\}^n \) such that there exists \( X \) with \((K, T, X, Y) \in \tau_{ic} \). Note that since \( \tau \) is good, for any \((K, T, X) \in \{0,1\}^k \times \{0,1\}^t \times \{0,1\}^n \), there is at most one query \((i, N, U, V) \in \tau_{ic} \) such that \( K = K_{\text{out}}^{i}, T = 0^t \wedge N, \) and \( X = H_{K_{\text{in}}^i}(U) \) as otherwise condition (C-1) would be satisfied. For any \((K, T) \in \{0,1\}^k \times \{0,1\}^t \), define
\[
\mathcal{X}_{\tag}(K, T) := \{X \in \{0,1\}^n : \exists(i, N, U, V) \in \tau_{\tag} : K = K_{\text{out}}^{i} \wedge T = 0^t \wedge N \wedge X = H_{K_{\text{in}}^i}(U)\}.
\]
Note that by the previous observation, one has \( \sum_{(K, T)} |\mathcal{X}_{\tag}(K, T)| = q_{\tag} \). Similarly, for any \((K, T, Y) \in \{0,1\}^k \times \{0,1\}^t \times \{0,1\}^n \), there is at most one query \((i, N, U, V) \in \tau_{\tag} \) such that \( K = K_{\text{out}}^{i}, T = 0^t \wedge N, \) and \( Y = V \) as otherwise condition (C-2) would be satisfied. For any \((K, T) \in \{0,1\}^k \times \{0,1\}^t \), define
\[
\mathcal{Y}_{\tag}(K, T) := \{Y \in \{0,1\}^n : \exists(i, N, U, V) \in \tau_{\tag} : K = K_{\text{out}}^{i} \wedge T = 0^t \wedge N \wedge Y = V\}.
\]
Because \( \tau \) is good, for any \((K, T), \mathcal{X}_{ic}(K, T) \cap \mathcal{X}_{\tag}(K, T) = \emptyset \) (as otherwise condition (C-3) would be satisfied) and \( \mathcal{Y}_{ic}(K, T) \cap \mathcal{Y}_{\tag}(K, T) = \emptyset \) (as otherwise condition (C-4) would be satisfied). Hence, for each pair \((K, T), \tau_{ic} \) and \( \tau_{\tag} \) together impose a set of \(|\mathcal{X}_{ic}(K, T)| + |\mathcal{X}_{\text{out}}(K, T)|\) equations on the random permutation \( E_{ic}(K, T, \cdot) \) internally.
where the first term accounts for the random choice of keys

\[ q \]

would be satisfied, depending on whether the equality stems from a IC query.

Then, according to Lemma 3 from [CLS17], one has

\[ \tau \]

\[ \tag \]

\[ \tau \]

\[ \tau \]

\[ \tau \] term for the probability that \( q(K,T) \) = \( Y \) in the sense that for any equality \( E \) (namely, every query \( (i,N,U,V) \) \( \in \tau \) translate to inequality \( E(K_{\text{out}},0^{\ell-\nu}|N,H_{K_{\text{in}}}(U)) \neq V \) that are “consistent” with previous equalities in the sense that for any equality \( E_{\text{ic}}(K,T,X,Y) \) and inequality \( E_{\text{ic}}(K',T',X') \neq Y' \), if \( (K,T) = (K',T') \), then either \( X \neq X' \) or \( Y \neq Y' \) (otherwise condition (C-5) or (C-6) would be satisfied, depending on whether the equality stems from a Tag query or an IC query).

We say that a TBC \( E \in \text{TBC}(k,t,n) \) is compatible with \( \tau_{\text{ic}}, \tau_{\text{tag}}, \) and \( \tau_{\text{ver}} \) if

- for every \( (K,T,X,Y) \in \tau_{\text{ic}}, E(K,T,X) = Y \),
- for every \( (I,N,U,V) \in \tau_{\text{tag}}, E(K_{\text{out}}^i,0^{\ell-\nu}|N,H_{K_{\text{in}}}(U)) = V \),
- for every \( (i,N,U,V) \in \tau_{\text{ver}}, E(K_{\text{out}}^i,0^{\ell-\nu}|N,H_{K_{\text{in}}}(U)) \neq V \),

and we let \( \text{Comp}(\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}}) \) denote the set of such TBCs and

\[ p(\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}}) := \text{Pr}[E \leftarrow \text{TBC}(k,t,n) : E \in \text{Comp}(\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}})]. \]

Then, according to Lemma 3 from [CLS17], one has

\[
p(\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}}) \geq \left( \prod_{K \in \{0,1\}^k} q(K,T)^{-1} \prod_{i=0}^{q(K,T)-1} \frac{1}{2^{n-i}} \right) \cdot \left( 1 - \frac{q_{\text{ver}}}{2^n - \max\{q(K,T)\}} \right) \]

\[
\geq \left( \prod_{K \in \{0,1\}^k} q(K,T)^{-1} \prod_{i=0}^{q(K,T)-1} \frac{1}{2^{n-i}} \right) \cdot \left( 1 - \frac{q_{\text{ver}}}{2^n - q_{\text{ic}} - \mu} \right)
\]

where \( q(K,T) := |X_{\text{ic}}(K,T)| + |X_{\text{tag}}(K,T)| \) and for the second inequality we used that \( \max\{q(K,T)\} \leq q_{\text{ic}} + \mu \). From this it follows that

\[
\text{Pr}[\Lambda_{\text{ic}} = \tau] \geq \frac{1}{|K_{\text{in}}|^n} \cdot \frac{1}{2^{n-k}} \cdot p(\tau_{\text{ic}}, \tau_{\text{tag}}, \tau_{\text{ver}}),
\]

where the first term accounts for the random choice of keys \( (K_{\text{in}}, K_{\text{out}}) \) and the second term for the probability that \( E_{\text{ic}} \) satisfies the constraints imposed by the transcripts \( \tau_{\text{ic}}, \tau_{\text{tag}}, \) and \( \tau_{\text{ver}} \). On the other hand, since in the ideal world the \( q_{\text{tag}} \) tags are uniformly random and independent, one has

\[
\text{Pr}[\Lambda_{\text{id}} = \tau] = \frac{1}{|K_{\text{in}}|^n} \cdot \frac{1}{2^{\mu+k}} \cdot \frac{1}{2^{n-k}} \cdot \left( \prod_{K \in \{0,1\}^k} |X_{\text{ic}}(K,T)|^{-1} \prod_{i=0}^{q(K,T)-1} \frac{1}{2^{n-i}} \right)
\]
Hence,
\[
\frac{\Pr[\Lambda_{\text{id}} = \tau]}{\Pr[\Lambda_{\text{id}} = \tau]} = 2^{n \cdot q_{\text{c}} n}, \quad \prod_{K \in \{0,1\}^k} \prod_{T \in \{0,1\}^t} \left( \frac{1}{2^{n - |X_{\text{ic}}(K,T)| - i}} \right) \left( 1 - \frac{q_{\text{ver}}}{2^n - q_{\text{c}} - \mu} \right)
\]
\[
= \left( 1 - \frac{q_{\text{ver}}}{2^n - q_{\text{c}} - \mu} \right) \prod_{K \in \{0,1\}^k} \prod_{T \in \{0,1\}^t} \left( \frac{1}{2^{n - |X_{\text{ic}}(K,T)| - i}} \right) \geq 1 - \frac{q_{\text{ver}}}{2^n - q_{\text{c}} - \mu},
\]
(10)
where for the last inequality we used that \(\sum_{(K,T)} |X_{\text{tag}}(K,T)| = q_{\text{tag}}\).

**Concluding.** The theorem follows by combining Theorem 1 with Equations (9) and (10) and using \(q_{\text{c}} + \mu \leq 2^n/2\) and \(\beta \geq 1\) which implies \(q_{\text{ver}}/(2^n - q_{\text{c}} - \mu) \leq 2\beta q_{\text{ver}}/2^n\).

**Improved Bound when Bounding the per-user data complexity.** Finally, we prove the “Moreover” part of the theorem. Assume that the total length of all TAG queries for any user is at most \(B\). We modify how we upper bound \(\Pr[\Lambda_{\text{id}} \in \Theta_{5,1}]\). Indeed, since for any \((i,N)\), \(\sum_{j=1}^{\nu_{i,N}} \ell_{i,N,j} \leq B\), one has

\[
\Pr[\Lambda_{\text{id}} \in \Theta_{5,1}] \leq \sum_{i=1}^{u} \sum_{N \in \{0,1\}^\nu} \sum_{j=1}^{\nu_{i,N}} \left( \frac{\alpha (\ell_{i,N,j} + \ell_{i,N,j}')} {2^n} + \frac{\beta} {2^n} \right)
\]
\[
\leq \frac{\alpha B} {2^n} \sum_{i=1}^{u} \sum_{N \in \{0,1\}^\nu} q_{i,N} + \frac{\alpha \mu} {2^n} \sum_{i=1}^{u} \sum_{N \in \{0,1\}^\nu} \sum_{j=1}^{\nu_{i,N}} \ell_{i,N,j}
\]
\[
+ \frac{\beta \mu} {2^n} \sum_{i=1}^{u} \sum_{N \in \{0,1\}^\nu} q_{i,N}
\]
\[
\leq \frac{\alpha B q_{\text{ver}}} {2^n} \frac{\alpha \mu q_{\text{ver}}} {2^n} + \frac{\beta \mu q_{\text{ver}}} {2^n}.
\]

### D Proof of Theorem 6

Assume towards a contradiction that \(H[E]\) is not \(\delta\)-sAU, i.e., there exists \(\ell\) and \((A,M)\) and \((A',M')\) in \(\{0,1\}^\ell \times \{0,1\}^\ell\) with \((A,M) \neq (A',M')\) and \(\max\{\text{len}(A,M),\text{len}(A',M')\} \leq \ell\) such that

\[
\delta' := \Pr[K \leftarrow K: H[E]_K(A,M) = H[E]_K(A',M')] > \delta.
\]

Let us show that there exists a non-empty set \(S\) of pairs \((T,X)\) in \(\{0,1\}^\ell \times \{0,1\}^n\) of size at most \(2\ell\) such that \(H[E]_K(A,M) = H[E]_K(A',M')\) is equivalent to

\[
\sum_{(T,X) \in S} E^i_{K}(X) = 0^n.
\]
The theorem will follow by defining $A'$ as the adversary having this set $S$ hardwired in its code and making oracle queries $\text{Enc}(T, X)$ for all pairs $(T, X) \in S$ except an arbitrary one $(T_0, X_0)$, computing $Y = \sum_{(T, X) \in S \setminus \{T_0, X_0\}} \text{Enc}(T, X)$, and returning $((T_0, X_0), Y)$ as “forgery”. Then $A'$ makes at most $2\ell$ oracle queries, runs in time at most $\alpha E$ for some small constant $\alpha$ independent of $E$, and wins the UNP game with advantage $\delta' > \delta(\ell)$, a contradiction.

Let us denote $S_0$ and $S'_0$ the set of pairs $(T, X)$ such that:

$$H[E]_K(A, M) = \sum_{(T, X) \in S_0} E^T_K(X), \quad \text{and} \quad H[E]_K(A', M') = \sum_{(T, X) \in S'_0} E^T_K(X).$$

By definition of $H[E]$, the tweaks involved in the computation of both hash value are pairwise distinct, and both $S_0$ and $S'_0$ are non-empty. Thus, the only way for the set $S$ to be empty is that one has $S_0 = S'_0$, since the pairs that belong to $S_0 \cap S'_0$ will be the only one that will cancel each other. What remains is to be shown is that at least one pair $(T, X)$ such that either $(T, X) \in S_0 \setminus S'_0$, or $(T, X) \in S'_0 \setminus S_0$. To this end, we distinguish several cases. If $(A, M) = \epsilon$, then $(A', M') \neq \epsilon$, and $H[E]_K(A, M) = E^{(4)}\{0^\ell-3, 0^n\}$. Moreover, no tweak with the prefix $(A')_3$ can appear in $S_0$, which means that $\{(A')_3\{0^\ell-3, 0^n\} \in S_0 \setminus S'_0\}$; the case where $(A', M') = \epsilon$ can be treated similarly. We now assume that $(A, M), (A', M') \neq \epsilon$. Consider the case where $A \neq A'$. Let us denote $\ell_A$ (resp. $\ell_{A'}$) the length of $A$ (resp. $A'$) in $m + n$-bit block, and denote $A_i$ (resp. $A'_i$) the $i$-th $n + m$-bit block of $A$ (resp. $A'$). Several subcases can occur.

- $\ell_A > \ell_{A'}$: (the case where $\ell_A < \ell_{A'}$ can be treated similarly): the pair $(T, X) = ((i_3)||A_{i-1}c||A_{i-1}m, |A_{i-1}|n)$, where $i \in \{0, 1\}$ depending whether $A$ was padded or not, is the only one involving a tweak with the prefix $(i_3)||A_{i-1}c|$, which means that $(T, X) \in S_0 \setminus S'_0$.

- $\ell_A = \ell_{A'}$, and there exists $i \in \{0, \ldots, \ell_A - 2\}$ such that $A_i \neq A'_i$: in that case, one has $((0_3)||i_3c||A_{i}m, |A_{i}|n) \in S_0 \setminus S'_0$, since otherwise we would have $A_i = A'_i$.

- $\ell_A = \ell_{A'}$, and $A_i = A'_i$ for $i$ in $\{0, \ldots, \ell_A - 2\}$: either $|A_{\ell_A-1}| \neq |A'_{\ell_{A'}-1}|$, or $|A_{\ell_A-1}| = |A'_{\ell_{A'}-1}|$ and $A_{\ell_A-1} \neq A'_{\ell_{A'}-1}$. In both cases, one necessarily has $(T, X) = ((i_3)||A_{i-1}c, |A_{i-1}m, |A_{i-1}|n) \in S_0 \setminus S'_0$. Indeed, in the former case, either both blocks are padded, in which case the padding rule implies that $\text{ozp}(A_{\ell_A-1}) \neq \text{ozp}(A'_{\ell_{A'}-1})$, or one block is padded and the other is not, which means that different tweak prefixes are used. In the latter case, the same tweak prefixes are used, but $(T, X) \in S'_0$ would imply $A_{\ell_A-1} = A'_{\ell_{A'}-1}$.

Finally, if $A = A'$, then one necessarily has $M \neq M'$, and we can apply the same argument to this new case, which ends the proof of Theorem 6.

## E Security of GNSIV

In this section, we study the security of GNSIV in the case where two independent keys are used (one for hashing, and a second one for tag finalization and encryption). Slightly abusing our notation, we write $\text{GNSIV}[E, f_T, f_X, H]$ for the mode of operation presented in Figure 8, where the hash function $H[E]$ defined in Section 4.2 is replaced by a generic hash function $H$. We prove the following result.

**Theorem 7** (mu security of GNSIV). Let $k$, $t$, and $n$ be positive integers, $K_{\text{in}}$, $K_{\text{out}}$, $A$ and $M$ be non-empty sets with $K_{\text{in}}$, $K_{\text{out}}$ finite, $\text{len} : A \times M \to \mathbb{N}$ be some length function, $E \in \text{TBC}(k, t, n)$ be a tweakable block cipher modeled as an ideal tweakable cipher $(IC, IC^{-1})$, and $H : K_{\text{in}} \times A \times M \to \{0, 1\}^n$ be a keyed hash function. Let $\nu$
be an integer such that $\nu \leq t$. Assume that $H$ is $\delta$-sAU and $\gamma$-uniform (w.r.t len) for $\delta(\ell) = \gamma(\ell) = \alpha/2^k + \beta/2^n$ with $\beta \geq 1$ (and hence $\delta \geq 2^{-n}$ and $\gamma \geq 2^{-n}$). Let $q_{\text{enc}}, q_{\text{dec}}, \sigma_{\text{enc}}, \sigma_{\text{dec}}, \ell_{\text{max}}$, and $\mu$ be positive integers such that $q_{\text{enc}} \leq 2^n$, $2q_{\text{enc}} + q_{\text{dec}} \leq 2^k$, and $q_{\text{enc}} + q_{\text{dec}} + \ell_{\text{dec}} \leq 2^n/2$. Then, for any (computationally unbounded) adversary $A$ against the mu-nAE security of GNSV1 in $E, f_T, f_X, H$ making at most $q_{\text{enc}}$ queries in total to IC or IC$^{-1}$, $q_{\text{dec}}$ queries to ENC of total length (as measured by len) at most $\ell_{\text{enc}}$, and no query longer than $\ell_{\text{enc}}$, $q_{\text{dec}}$ queries to DEC of total length (as measured by len) at most $\ell_{\text{dec}}$, and such that any (user, nonce) pair $(i, N)$ appears at most $\mu$ times in its ENC queries, one has

$$\text{Adv}_{\text{mu-nae}}^{\text{GNSV1}}[E, f_T, f_X, H](A) \leq \text{Adv}_{\text{mu-nive}}^{\text{GCTR}[E, f_T, f_X]}(n) (q_{\text{enc}}, q_{\text{dec}}, \sigma_{\text{enc}}, \ell_{\text{max}}, \mu) + \text{Adv}_{\text{n-mpmac}}^{\text{NaT}[H, E]}(n) (q_{\text{enc}}, q_{\text{dec}}, \sigma_{\text{enc}}, q_{\text{dec}}, \sigma_{\text{dec}}, \ell_{\text{max}}, \mu),$$

where the first term is the bound from Theorem 2 with $\sigma_{\text{enc}}$ and $\ell_{\text{max}}$ substituted respectively to $\sigma$ and $\ell_{\text{max}}$ and the second term is the bound from Theorem 3 with $q_{\text{enc}}, \sigma_{\text{enc}}, q_{\text{dec}}, \sigma_{\text{dec}}, \ell_{\text{max}}$ substituted respectively to $q_{\text{tag}}, \sigma_{\text{tag}}, q_{\text{ver}}, \sigma_{\text{ver}}, \ell_{\text{max}}$.

The proof of Theorem 7 uses the H-coefficients technique. Fix a deterministic adversary $A$ and assume without loss of generality that $A$ makes exactly $q_{\text{enc}}$ queries to IC or IC$^{-1}$, $q_{\text{dec}}$ queries to ENC and $q_{\text{dec}}$ queries to DEC, and that it never makes a pointless query, where a pointless query is either:

- a repeated query to IC, IC$^{-1}$, ENC or DEC
- a query IC$(K, T, X)$ if there was a previous query IC$^{-1}$(K, T, Y) that returned X or a query IC$^{-1}$(K, T, Y) if there was a previous query IC$(K, T, X)$ that returned Y,
- a query ENC$(i, N, A, M)$ or DEC$(i, N, A, V, C)$ such that $i > u$, where $u$ is the current value of the counter keeping track of NEW queries,
- a query DEC$(i, N, A, V, C)$ if there was a previous query ENC$(i, N, A, M)$ that returned $(V, C)$.

The transcript consists of three types of queries:

- queries to the ideal cipher that we record as a list $\tau_{\text{enc}}$ containing tuples $(K, T, X, Y)$ such that $A$ made either a query IC$(K, T, X)$ that returned Y or IC$^{-1}$(K, T, Y) that returned X;
- queries to ENC that we record as a list $\tau_{\text{enc}}$ containing tuples $(i, N, A, M, V, C)$ such that $A$ made a query ENC$(i, N, A, M)$ that returned the tag V and the ciphertext C;
- queries to DEC that we record as a list $\tau_{\text{dec}}$ containing all tuples $(i, N, A, V, C)$ such that were queried to DEC (we do not keep track of the answers since we are interested in attainable transcripts, i.e., transcripts that can be obtained in the ideal world in which all queries to DEC return ⊥).

We do not keep track explicitly of NEW queries, but when $A$ has finished interacting with the oracles, we reveal all keys $K = ((K_{1}^{\text{out}}, K_{2}^{\text{out}}), \ldots, (K_{\mu}^{\text{out}}, K_{\text{out}}^{\text{out}}))$ generated by calls to NEW, where $u$ is the final value of the counter keeping track of NEW queries. Besides, in the case where the length of an encryption query is not a multiple of $n$, we assume that, during the encryption pass, the last block is padded with enough zeros so that it has length $n$ before returning the oracle answer. Before defining bad transcript, we can remark that, for any transcript $\tau$, we can derive a transcript $\tau_{\text{tag}}$ of the interaction of $A$ with the authentication pass of nAE as follows: $\tau_{\text{tag}}$ contains the tuples $(i, N, A, M, V)$ for every query $(i, N, A, M, V, C)$ in $\tau_{\text{enc}}$. We now define the following events:
• we say that $\tau$ is NaT-bad if $(\tau_{\text{tag}}, \tau_{\text{ic}}, K)$ satisfies one of the conditions (C-1) to (C-4) from Section 4;

• we say that $\tau$ is GCTR-bad if $(\tau_{\text{enc}}, \tau_{\text{ic}}, K)$ satisfies any of the conditions (C-1) to (C-4) from Section 3.

Now, assuming that $\tau$ is neither NaT-bad nor GCTR-bad, we are going to release additional information to the attacker as follows:

• in the real world, for every $(i, N, A, V, C)$ in $\tau_{\text{dec}}$, we are going to release the corresponding plaintext $M$ by running the decryption algorithm of GNSIV;

• in the ideal world, we will release a dummy plaintext as follows: for every ciphertext block, if the corresponding $E(K, (6)_3\|f_T(N, V, j), (f_X(N, V, j))$ is already known (either from a query in $\tau_{\text{ic}}$ or $\tau_{\text{enc}}$), then the existing value is used; otherwise, we simulate what happens in the real world by drawing, uniformly at random and without replacement, an output in the set of values $\mathcal{Y}$ that do not appear as outputs in $\tau_{\text{ic}}$ or $\tau_{\text{enc}}$ for the considered (key,tweak) pair (this is possible as $\tau$ is neither NaT-bad nor GCTR-bad)\(^{15}\).

Note that this sampling is possible due to the fact that $\tau$ is neither NaT-bad nor GCTR-bad. After this step, queries from $\tau_{\text{dec}}$ will appear as tuples $(i, N, A, M, V, C)$, where $M$ corresponds to the computed plaintext. If $\tau$ is NaT-bad or GCTR-bad, we simply assume that $M = \epsilon$ for all decryption queries.

A transcript $\tau$ will be said bad if one of the following condition is satisfied:

(C-1) $\tau$ is NaT-bad;

(C-2) $\tau$ is GCTR-bad;

(C-3) There exist queries $(i, N, A, M, V, C) \in \tau_{\text{enc}}$ and $(i', N', A', M', V', C') \in \tau_{\text{dec}}$ such that

\[
\begin{align*}
K'_{\text{out}} &= K_{\text{out}}' \\
N &= N' \\
H_{K'_n}(A, M) &= H_{K'_n'}(A', M') \\
V &= V'.
\end{align*}
\]

(C-4) There exist queries $(i, N, A, M, V, C) \in \tau_{\text{dec}}$ and $(K, T, X, Y) \in \tau_{\text{ic}}$ such that

\[
\begin{align*}
K'_{\text{out}} &= K \\
0^{\nu} \| N &= T \\
H_{K'_n}(A, M) &= X \\
V &= Y.
\end{align*}
\]

Otherwise, we say that $\tau$ is good and let $\Theta_{\text{bad}}$, resp. $\Theta_{\text{good}}$ denote the set of bad, resp. good transcripts. Note that the last two conditions are well-defined thanks to the fact that we release the plaintexts that correspond to every decryption query.

\(^{15}\)Note that this is equivalent to sampling the output of an ideal cipher, conditioned on it being compatible with the one from the real world.
Probability of Bad Transcripts. First, we upper bound the probability of bad transcripts in the ideal world. We consider each condition in turn, letting Θi denote the set of transcripts satisfying condition (C-i), i ∈ [1, 4].

We assume that the queries transcript involves u users, and for every i ∈ [1, u] and N ∈ \{0, 1\}ν we let qisi denote the number of ENC queries involving user i and nonce N (with qisi = 0 if there were no such queries). When qisi ≥ 1, we also let τi,N,j, 1 ≤ j ≤ qisi, denote the length of the j-th ENC query for user i with nonce N and we assume that queries are reordered such that τi,N,1 ≤ τi,N,2 ≤ ⋯ ≤ τi,N,qisi. Note that by our assumptions, we have

\[ qisi ≤ \mu \text{ for every } (i, N), \]
\[ \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^\nu} qisi = q_{\text{enc}}, \]
\[ \text{and } \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^\nu} \sum_{j=1}^{qisi} \taui,N,j ≤ \sigma_{\text{enc}}. \]

Similarly, for every i ∈ [1, u] and N ∈ \{0, 1\}ν we let qdqi denote the number of DEC queries involving user i and nonce N (with qdqi = 0 if there were no such queries) and when qdqi ≥ 1, we let ℓi,N,j, 1 ≤ j ≤ qdqi, denote the length of the j-th DEC query for user i with nonce N and we assume that queries are reordered such that ℓi,N,1 ≤ ℓi,N,2 ≤ ⋯ ≤ ℓi,N,qdqi. Then

\[ \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^\nu} qdqi = q_{\text{dec}}, \]
\[ \text{and } \sum_{i=1}^{u} \sum_{N \in \{0, 1\}^\nu} \sum_{j=1}^{qdqi} \elli,N,j ≤ \sigma_{\text{dec}}. \]

In the following, for ((A, M), (A’, M’)) ∈ (A × M)2, we let

\[ \delta_{(A,M),(A’,M’)} := \max \{ \delta(\text{len}(A, M)), \delta(\text{len}(A’, M’)) \} \]
\[ = \alpha \max \{ \text{len}(A, M), \text{len}(A’, M’) \} / 2^k + \beta / 2^n \]
\[ \gamma_{(A,M),(A’,M’)} := \min \{ \gamma(\text{len}(A, M)), \gamma(\text{len}(A’, M’)) \} \]
\[ = \alpha \min \{ \text{len}(A, M), \text{len}(A’, M’) \} / 2^k + \beta / 2^n. \]

Conditions (C-1) and (C-2). As the random variable involved in these event have the same probability distribution in the ideal world as in Section 4 for (C-1) and Section 3 for (C-2), we can directly reuse the same upper bounds.

Condition (C-3). We define two subsets ΘA,1 and ΘA,2 of ΘA. Subset ΘA,1 consists of transcripts τ such that there exist queries (i, N, A, M, V, C) in τenc and (i, N’, A’, M’, V’, C’) in τdec such that N = N’, V = V’ and H_{K_{\text{in}}}(A, M) = H_{K_{\text{in}}}(A’, M’). Subset ΘA,2 consists of transcripts τ such that there exist queries (i, N, A, M, V, C) in τenc and (i’, N’, A’, M’, V’, C’) in τdec such that i ≠ i’, K_{\text{out}} = K’_{\text{out}}, N = N’, V = V’ and H_{K_{\text{in}}}(A, M) = H_{K_{\text{in}}}(A’, M’).

We start with ΘA,1. Fix (i, N, A, M, V, C) in τenc and (i, N’, A’, M’, V’, C’) in τdec such that N = N’. Assume that the decryption query occurred after the encryption query, and that V’ = V. Since A never makes pointless queries, then one has (A’, C’) ≠ (A, C). If A ≠ A’, the probability that H_{K_{\text{in}}}(A, M) = H_{K_{\text{in}}}(A’, M’) is smaller than δ_{(A,M),(A’,M’)}. The
If $A = A'$, then one necessarily has $C' \neq C$. Since $N = N'$ and $V = V'$, this implies that $M \neq M'$, as $M$ and $M'$ are the plaintexts of respectively $C$ and $C'$ under the same IV and nonce value. Thus, the probability that $H_{K^{\text{in}}_{\text{in}}}(A, M) = H_{K^{\text{in}}_{\text{in}}}(A', M')$ is also smaller than $\delta_{(A,M),(A',M')}$. Now assume that the encryption query occurred after the decryption query. Then, the probability that $V = V'$ is exactly $2^{-n}$. Then the condition is satisfied for this pair of queries with probability at most

$$\frac{\alpha \max(\ell, \ell_1, \ell_2)}{2^k} + \frac{\beta}{2^n} \leq \frac{\alpha(\ell + \ell_1, \ell_2)}{2^k} + \frac{\beta}{2^n} \leq \frac{\alpha(\ell + \ell_1, \ell_2)}{2^k} + \frac{\beta}{2^n}.$$ 

By the union bound over all users $i$, nonces $N$, and pairs of encryption and decryption queries for $(i, N)$, we have, like in the study of the set $\Theta_{5,1}$ in the proof of Theorem 3,

$$\Pr[\Lambda_{id} \in \Theta_{3,1}] \leq \sum_{i=1}^{u} \sum_{N \in \{0,1\}^*} \sum_{j=1}^{q_{enc}} \sum_{j'=1}^{q_{dec}} \left( \frac{\alpha(\ell + \ell_1, \ell_2)}{2^k} + \frac{\beta}{2^n} \right).$$

We now consider $\Theta_{3,2}$. Fix $(i, N, A, M, V, C)$ in $\tau_{enc}$ and $(i', N', A', M', V', C')$ in $\tau_{dec}$ such that $i \neq i'$. Then one clearly has $K^{\text{out}}_{\text{out}} = K^{\text{out}}_{\text{out}}$ with a probability smaller than $2^{-k_2}$. Moreover, the probability that $H_{K^{\text{in}}_{\text{in}}}(A, M) = H_{K^{\text{in}}_{\text{in}}}(A', M')$ is at most $\gamma_{(A,M),(A',M')}$. Summing over the at most $q_{enc}q_{dec}$ pairs of queries, we get

$$\Pr[\Lambda_{id} \in \Theta_{3,2}] \leq \frac{\alpha q_{enc} q_{dec}}{2^k} + \frac{\beta q_{enc} q_{dec}}{2^{k+n}}.$$ 

All in all,

$$\Pr[\Lambda_{id} \in \Theta_3] \leq \frac{\alpha \mu q_{enc} q_{dec}}{2^k} + \frac{\alpha(\ell + \ell_1, \ell_2)}{2^k} + \frac{\beta q_{enc} q_{dec}}{2^{k+n}} + \frac{\beta q_{enc} q_{dec}}{2^k} + \frac{\beta q_{enc} q_{dec}}{2^k}.$$ 

**CONDITION (C-4).** For any $(i, N, A, M, V, C) \in \tau_{dec}$ and $(K, T, X, Y) \in \tau_{enc}$, $K^{\text{out}}_{\text{out}} = K$ with probability $2^{-k_2}$ and $H_{K^{\text{in}}_{\text{in}}}(A, M) = X$ with probability at most $\gamma(\text{len}, (A, M))$. By summing over all pairs, one has

$$\Pr[\Lambda_{id} \in \Theta_4] \leq q_{dec} \sum_{i=1}^{u} \sum_{N \in \{0,1\}^*} \sum_{j=1}^{q_{enc}} \left( \frac{\alpha(\ell + \ell_1, \ell_2)}{2^k} + \frac{\beta}{2^k} \right) \leq \frac{\alpha q_{enc} q_{dec}}{2^k} + \frac{\beta q_{enc} q_{dec}}{2^{k+n}}.$$ 

Collecting all probabilities, we obtain that $\Pr[\Lambda_{id} \in \Theta]$ is the sum of the probabilities of bad transcripts in the proof of Theorem 2 and Theorem 3 (with substitutions as indicated in the theorem statement).

**GOOD TRANSCRIPT PROBABILITY RATIO.** Fix a good transcript $\tau$. We are going to divide the information about the ideal cipher that are contained in the transcript into several multisets:

- for each query to the ideal cipher, add a tuple $(K, T, X, Y)$ to $S_1$;
- for each encryption query $(i, N, A, M, V, C)$, add to $S_2$ the tuples $(K^{\text{out}}_{\text{out}}(i), 6i || f_{X}(N, V, j), f_{X}(N, V, j), M_i \oplus C_j)$, where $M = M_1 || \cdots || M_\ell$, $C = C_1 || \cdots || C_\ell$, $j = 1, \ldots, \ell_i$, and $|M_k| = |C_k| = n$ for all $k = 1, \ldots, \ell_i$;
• for each encryption query \((i, N, A, M, V, C)\), add to \(S_3\) the tuple
\[
(K^i_{out}, \langle 5 \rangle_3 || 0^{t-3-\nu} || N, H_{K^i_{in}}(A, M), V);
\]

• for each decryption query \((i, N, A, M, V, C)\), add to \(S_4\) the tuples
\[
(K^i_{out}, \langle 6 \rangle_3 || f_T(N, V, j), f_X(N, V, j), M_j \oplus C_j),
\]
where \(M = M_1 || \ldots || M_{\ell_i}, C = C_1 || \ldots || C_{\ell_i}, j = 1, \ldots, \ell_i,\) and \(|M_k| = |C_k| = n\) for all \(k = 1, \ldots, \ell_i\), and as long as the tuple did not already belong to \(S_1 \cup S_2 \cup S_4\);

• for each decryption query \((i, N, A, M, V, C)\), add to \(S_5\) the tuple
\[
(K^i_{out}, \langle 5 \rangle_3 || 0^{t-3-\nu} || N, H_{K^i_{in}}(A, M), V),
\]
if it did not already belong to \(S_5\).

Thanks to the fact that \(\tau\) is a good transcript, the domain separation, and the way \(S_4\) and \(S_5\) are generated, then the multisets \(S_1, \ldots, S_5\) do not contain any duplicated entries and are pairwise disjoint. Besides, the set \(S_1 \cup S_2 \cup S_4\) does not contain any distinct triples \((K, T, X, Y)\) and \((K', T', X', Y')\) such that \(K = K', T = T',\) and either \(X = X'\) or \(Y = Y'\).

Let us now compute the ratio \(\Pr[\Lambda_{re} = \tau]/\Pr[\Lambda_{id} = \tau]\). The event \(\Lambda_{id} = \tau\) can be broken down into several events:

- **Key** corresponds to the event where the key vector \(K\) has the correct value;
- **IdealPrim** corresponds to the ideal cipher that are compatible with \(S_1\) (i.e., tuple \((K, T, X, Y)\) is in \(S_1\) if and only if \(IC(K, T, X) = Y\));
- **IdealEnc** corresponds to the ciphertexts and authentication tags indicated by \(\tau\) being the actual values revealed (note that this corresponds to \(n(|S_2| + |S_3|)\) bits);
- **IdealDec** corresponds to the sampled queries agreeing with \(S_4\).

Similarly, the event \(\Lambda_{re} = \tau\) can be broken down into several events:

- **Key** corresponds to the event where the key vector \(K\) has the correct value;
- **RealPrim** corresponds to the ideal cipher being compatible with \(S_1\);
- **RealEnc** corresponds to the ideal cipher being compatible with \(S_2 \cup S_3\);
- **RealDec** corresponds to the ideal cipher being compatible with \(S_4\);
- **RealVer** corresponds to the ideal cipher not being compatible with any tuple in \(S_5\) (i.e. \(IC(K, T, X) \neq Y\) for every tuple \((K, T, X, Y)\) in \(S_5\)).

Note that the first event is common between both worlds (and independent from the second one),

\[
\Pr[RealPrim] = \Pr[IdealPrim]
\]
and

\[
\Pr[RealDec|Key \cap RealPrim \cap RealEnc] = \Pr[IdealDec|Key \cap IdealPrim \cap IdealEnc].
\]
Moreover, after ordering $S_2 \cup S_3$ (with an arbitrary ordering), one has

$$\frac{\Pr[\text{RealEnc}\mid \text{Key} \cap \text{RealPrim}]}{\Pr[\text{IdealEnc}\mid \text{Key} \cap \text{IdealPrim}]} = \frac{2^{n(|S_2| + |S_3|)}}{\prod_{i=1}^{|S_2 \cup S_3|} (2^n - n_i)} \geq 1,$$

where $n_i$ denotes the number of occurrences of $(K_i, T_i)$ in the first $i - 1$ elements of $S_2 \cup S_3$. Hence, one has

$$\frac{\Pr[\Lambda_{\text{re}} = \tau]}{\Pr[\Lambda_{\text{id}} = \tau]} \geq \Pr[\text{RealVer}\cap \text{RealDec} \cap \text{RealEnc} \cap \text{RealPrim} \cap \text{Key}]$$

$$\geq 1 - \frac{q_{\text{dec}}}{2^n - q_{\text{enc}} - q_{\text{dec}}}$$

$$\geq 1 - \frac{2\beta q_{\text{dec}}}{2^n} \quad (16)$$

$$\geq 1 - \frac{2\beta q_{\text{dec}}}{2^n} \quad (17)$$

where for the last inequality we used $q_{\text{enc}} + q_{\text{dec}} \leq 2^n / 2$ and $\beta \geq 1$. Combining the probability of bad transcripts and Equation 17 with Theorem 1 concludes the proof of Theorem 7.