Committing Authenticated Encryption

Sponges vs. Block-Ciphers in the case of the NIST LWC Finalists

Juliane Krämer[1], Patrick Struck[2], and Maximiliane Weishäupl[1]

[1] Universität Regensburg, Germany
{juliane.kraemer,maximiliane.weishaeupl}@ur.de
[2] Universität Konstanz, Germany
patrick.struck@uni-konstanz.de

Abstract. Committing security has gained considerable attention in the field of authenticated encryption (AE). This can be traced back to a line of recent attacks, which suggests that AE schemes used in practice should not only provide confidentiality and authenticity, but also committing security. Roughly speaking, a committing AE scheme guarantees that ciphertexts will decrypt only for one key. Despite the recent research effort in this area, the finalists of the NIST lightweight cryptography standardization process have not been put under consideration yet. We close this gap by providing an analysis of these schemes with respect to their committing security. Despite the structural similarities the finalists exhibit, our results are of a quite heterogeneous nature: We break four of the schemes with effectively no costs, while for two schemes our attacks are costlier, yet still efficient. For the remaining three schemes Isap, Ascon, and (a slightly modified version of) Schwaemm, we give formal security proofs. Our analysis reveals that sponges—due to their large states—are more favorable for committing security compared to block-ciphers.

* Work of Juliane Krämer was supported by the German Research Foundation (DFG) – SFB 1119 – 236615297. Patrick Struck acknowledges funding by the Bavarian State Ministry of Science and the Arts in the framework of the bidt Graduate Center for Postdocs (while working at University of Regensburg) and the Hector Foundation II. Work of Maximiliane Weishäupl was funded by the German Federal Ministry of Education and Research (BMBF) under the project Quant-ID (16KISQ111).

Table of Contents

1 Introduction

The most fundamental cryptographic concept is symmetric encryption, allowing two parties, Alice and Bob, which share some secret key, to securely exchange messages. The initial goal—and still a cornerstone—is confidentiality which prevents anyone but Alice and Bob from recovering the message from a ciphertext. In modern cryptography, security requirements have been enhanced to also incorporate authenticity[3] which ensures that no third party can produce a ciphertext that Bob would accept as one generated by Alice. On that account, *authenticated encryption* (AE), which encompasses both confidentiality and authenticity was introduced and has, since then, become the gold standard [11, 48]. While authenticated encryption has undergone some changes—from probabilistic over IV-based to nonce-based—nowadays, the research community agrees on *authenticated encryption with associated data* as the right approach. Such a scheme generates a ciphertext C by encrypting a message M under a *context* (K, N, A), consisting of a key K, a nonce N, and associated data A. Authenticity should hold for both the associated data and the message. In contrast, confidentiality is required only for the message.

The relevance of authenticated encryption is not only reflected by the conducted research, but also by the fact that AE schemes are deployed ubiquitously, e.g., in TLS 1.3 [51]. The CAESAR competition for authenticated encryption [11] and the recent NIST lightweight cryptography (LWC) standardization process [48], both called specifically for AE schemes which are deemed secure if they provide both confidentiality and authenticity.

However, a series of recent attacks [2,32,43] has shown that our understanding of what a secure AE scheme is, has to change once again. The *Facebook message franking attack* [32] enabled Alice, a malicious user, to send an offensive or even illegal content to Bob. If Bob tries to report this, it will fail as Facebook will see a harmless content—prepared by Alice as part of the attack—instead. Further examples are the *subscribe with Google attack* [2] and the *partitioning oracle attack* [43]. The latter allows for more efficient key recovery: The adversary crafts a ciphertext that decrypts validly under multiple keys (for instance taken from a leaked list of candidate keys) and sends it to the recipient; If the recipient rejects the ciphertext as invalid, the adversary can rule out all keys which are valid for the sent ciphertext.

In fact, these attacks can all be traced back to the same problem: The existence of ciphertexts that decrypt validly under more than one key. This is neither prevented by confidentiality nor by authenticity, baring the need for an additional security notion. To this end, committing security [9] was defined by requiring each ciphertext to be a commitment to the key ($\mathrm{CMT}_{\mathsf{K}}$) or even to the whole context (CMT). The latter notion is the strongest one and is formalized by the following security game: The adversary outputs two tuples $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$, each consisting of key, nonce, associated data,

[3] Otherwise attacks like the *padding oracle attack* [53], which exploits the absence of any authentication mechanism, are possible.

and message, and wins if their contexts differ, i.e., $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$, and $\textsc{Ae.Enc}(K, N, A, M) = \textsc{Ae.Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$ holds, for Ae the scheme under consideration.[4]

The aforementioned attacks demonstrate that the consequences of using non-committing authenticated encryption can be severe. Considering that there are most likely more attacks, which have yet to be discovered, it is important to deal with this problem. One possibility would be to design protocols in such a way that usage of non-committing authenticated encryption does not result in attacks.[5] However, this approach is ill-advised as it requires separate analysis for the different protocols and simply puts the burden on the designers of a protocol. A better approach is to design authenticated encryption schemes that are committing which can then be used in different protocols without worrying about committing attacks.

To this end, AE schemes used in practice need to be analyzed with respect to committing security. This process has already begun and a number of commonly used AE schemes (GCM, SIV, CCM, EAX, OCB3) have been examined [9,45]. A majority of them was shown to not achieve committing security. Arguably among the most important AE schemes are the finalists of the NIST LWC standardization process. While these schemes have received significant analysis with respect to confidentiality and authenticity [52], we are not aware of any research with respect to their committing security. Despite the announcement of the NIST that Ascon will be standardized, we consider all finalists relevant objects to study—especially considering that none of the finalists suffer from any weaknesses regarding their claimed security levels [52].

1.1 Contribution

In this paper we analyze the committing security of the NIST LWC finalists that are based on (tweakable) block-ciphers or sponges.[6] More precisely, we focus on the authenticated encryption mode of the schemes, while the underlying primitives, i.e., (tweakable) block-ciphers or permutations, are assumed to be ideal. We follow the example of [45], to define a boundary between committing insecure and secure schemes. The line is drawn at 64-bit security, i.e., a scheme providing at least 64-bit committing security is called secure, while all others are called insecure. This number stems from the fact that the cost of a committing attack is bounded below by the cost of finding colliding tags. As most present-day schemes employ 128-bit tags, the latter can be done with about 2^{64} queries using a birthday attack.

We divide the NIST LWC finalists into two groups: Firstly, Elephant [16] and Isap [29] follow the *Encrypt-then-MAC* (EtM) paradigm, as they start by

[4] For tidy schemes [47], an equivalent characterization of the notion requires the adversary to find $(C, (K, N, A), (\overline{K}, \overline{N}, \overline{A}))$ such that $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$ and $\textsc{Ae.Dec}(K, N, A, C), \textsc{Ae.Dec}(\overline{K}, \overline{N}, \overline{A}, C) \neq \perp$ [9].

[5] Note that the Facebook protocol was changed to prevent the message franking attack.

[6] This covers all finalists except Grain-128aead [37].

encrypting the message and then authenticate the resulting ciphertext alongside the context. Secondly, Romulus [40], Gift-Cofb [3], Photon-Beetle [5], Xoodyak [23], TinyJambu [56], Ascon [31], and Schwaemm [7] share a common structure, in the sense that all of them first process the context and then the message. We refer to schemes of the second type as *Context-pre-Processing* (CpP) schemes.

Surprisingly, even though the NIST finalists show strong structural similarities, our results regarding their committing security are of a very heterogeneous nature. As can be seen in Table 1, the results vary from attacks that require essentially no queries[7] to attacks that are costlier—while still using significantly less than 2^{64} queries—and proofs showing about 64-bit committing security. In summary, there are four schemes we break completely, two schemes we break efficiently, and three schemes[8] for which we show committing security.

While our attacks exploit different vulnerabilities, some of them share the same fundamental idea. This is the case for Romulus and Gift-Cofb, which are both block-cipher-based AE schemes. Further, both feature a state-update-function, which is invoked in an alternating manner with the block-cipher. The attacks boil down to the fact that for a fixed ciphertext, key, and nonce, one can find associated data such that the ciphertext decrypts validly under this context. For this, starting from the target ciphertext, the component that processes the message is inverted. Then the fact that associated data blocks are XORed onto the whole state is used to connect the initial state with the state obtained from the reverse computation. This attack strategy depends heavily on the invertibility of the state-update-function. For Romulus, we show that such an inversion is always possible, while for Gift-Cofb it works with a probability of $\frac{1}{2}$. This implies that the attack cost for Gift-Cofb depends on the length of the ciphertext, however, by choosing a short ciphertext we obtain a very efficient attack. The XORing of an input onto the whole state is a vulnerability that is also exploited in our attack on Elephant, a tweakable block-cipher-based AE scheme. In contrast to Romulus and Gift-Cofb, Elephant is an Encrypt-then-MAC scheme. This structure simplifies the committing attack as we only need to find two different contexts that verify the ciphertext correctly—in this case the decryption of Elephant will never return $\perp$.[9] Due to this, it suffices to concentrate on the MAC and, more precisely, finding a tag collision. The latter is easily achieved, as the associated data is XORed to the full state during the tag generation of Elephant.

Except for these three schemes, none of the other NIST finalists carry out a full-state XOR. Xoodyak arguably comes very close to this situation, as it is based on a full-state sponge, which reserves only a few bits for padding, which are not directly accessible via the inputs. Therefore, we are able to control most of the state by a direct XOR, while for the remaining bits a birthday attack

[7] More precisely, these attacks need only the minimal cost of computing the respective encryption algorithm twice (once for each of the output tuples).
[8] Note that we consider a slightly modified version of Schwaemm.
[9] Menda et al. [45] coin this property as *NoFailDecrypt*.

is applied. Similarly, the attack on TinyJambu, a block-cipher-based scheme, also boils down to a birthday attack. We exploit that TinyJambu uses a tag of just 64 bits (the shortest one among all finalists[10]), hence a tag collision can be produced with reasonable cost.

Our attack on Photon-Beetle, a sponge-based AE scheme, exploits the choice of the initial state. For most of the finalists, this state contains some fixed initialization vector, whereas for Photon-Beetle it consists exclusively of key and nonce. However, this implies that the initial state can be controlled completely by a committing adversary, which will turn out to be the key ingredient of our attack. Simply speaking, the attack allows to choose an intermediate state (outcome of the context-pre-processing) that results in the same ciphertext. We can invert this intermediate state for different associated data and take the outcome as the key-nonce pair.

None of these attacks are applicable to any of the sponge-based schemes Isap, Ascon, and Schwaemm. We give security proofs for these three schemes, showing that they achieve about 64-bit committing security. The high-level idea of all proofs is similar: we show that the schemes can be viewed as plain sponge constructions and give bounds for finding colliding tags, which directly translate to bounds on the committing security. Some extra care is necessary when dealing with the core features of the schemes—the re-keying mechanism deployed in Isap and the state-/output-blinding applied in both Ascon and Schwaemm.

Our analysis reveals that sponges are better suited for achieving committing authenticated encryption. Simultaneously, our attacks against Photon-Beetle and Xoodyak show that sponges are not always committing: Full-state sponges suffer from inherent weaknesses when it comes to committing security, regardless of how often the full-state property occurs.[11] To achieve security, a significant part of the state has to be kept "out-of-reach" of the committing adversary, meaning that it should not be affected by the inputs. Sponges are better suited for this due to their significantly larger states (≥ 256 bits) compared to block-ciphers that typically use smaller states (128 bits). Since a committing adversary knows the key, the general advantage of block-ciphers—namely, the permutation remains concealed from the adversary—vanishes.

1.2 Related Work

Committing security can be traced back to [1,33] where the focus was on public-key encryption. In [34]—using the name *key-robustness*—Farshim et al. gave first definitions of committing security for symmetric encryption. Recently, Bellare and Hoang [9] introduced different variants of committing security for authenticated encryption, covering the prior variants where a ciphertext is a commitment to the key, but also stronger forms where a ciphertext is a commitment to all inputs. Ultimately, Menda et al. [45] developed a framework for fine-grained committing security notions. Instead of just having a ciphertext being a commitment

[10] Elephant uses 64-bit tags as well, but also gives a parameter set with 128-bit tags.
[11] For Photon-Beetle only the initial state exhibits the full-state property.
[12] The costs correlate with the length of the ciphertext.

Table 1: Overview of results: a ✗ indicates a committing attack with essentially no queries; a ✦ indicates a committing (CMT) attack with significantly less than 2^{64} queries; and a ✓ indicates about 64-bit committing security.

Scheme	CMT	Theorem	Section	Illustration	Pseudocode
Elephant [16]	✗	Theorem 3	Section 3.1	Fig. 5	Fig. 20
Romulus [40]	✗	Theorem 4	Section 3.2	Fig. 6	Fig. 22
Gift-Cofb [3]	✗[12]	Theorem 5	Section 3.3	Fig. 7	Fig. 24
Photon-Beetle [5]	✗	Theorem 6	Section 3.4	Fig. 8	Fig. 26
Xoodyak [23]	✦	Theorem 7	Section 3.5	Fig. 9	Fig. 27
TinyJambu [56]	✦	Theorem 8	Section 3.6	Fig. 10	Fig. 28
Isap [29]	✓	Theorem 9	Section 3.7	Fig. 12	Fig. 30
Ascon [31]	✓	Theorem 10	Section 3.8	Fig. 13	Fig. 32
Schwaemm [7]	✓	Theorem 11	Section 3.9	Fig. 14	Fig. 35

to either the key or all inputs, it allows for variants where it is a commitment to, say, the key and the nonce. Along with these committing notions, they also coin the term *context discovery attacks*. In contrast to committing attacks, which require the adversary to find two contexts that decrypt the same ciphertext, context discovery attacks require finding a context that decrypts a given ciphertext. Concurrently to [45], Chan and Rogaway [20] also developed a more fine-grained definitional framework for committing security, for instance, allowing for variants where the adversary has to use honest keys, i.e., randomly sampled ones.

2 Authenticated Encryption and the NIST LWC Finalists

In this section, we first introduce some notation and recall the definitions of authenticated encryption schemes and committing security. We then provide a general classification of the NIST LWC finalists and high-level approaches for the committing attacks.

2.1 Notation

Throughout this work, we write $\{0,1\}^*$ for the set of bit strings with arbitrary length. By $\{0,1\}^{\leq r}$ ($\{0,1\}^{\geq r}$) we denote the set of bit string with length at most r (at least r). For a bit string S of length n, we write $\lceil S \rceil_r$, $\lfloor S \rfloor_c$, and $[S]_i^j$ for the first r bits, the last c bits, and the i-th to j-th bits of S, respectively. For bit strings X, Y, and Z, $|X|$ describes the length of X and $Y \parallel Z$ denotes the concatenation of Y and Z. For an integer k, the set $\{1,\ldots,k\}$ is abbreviated as $[k]$. We write $X_1,\ldots,X_l \xleftarrow{r} X$ to denote that X is split into bit strings X_1 to X_l s.t. $|X_i| = r$, for $i \in [l-1]$ and $|X_l| \leq r$. Bit rotation/shift of x by k bits to the left is written as $x \lll b$ / $x \ll b$ ($\ggg$/$\gg$ denote the same in the other direction). The encoding of x into one Byte is described by $\mathtt{enc}_8(x)$. For

sake of simplicity, we use ι as a generic value for domain separation in several schemes as our results are independent of it. Standard cryptographic background on sponges, block-ciphers (BC), and tweakable block-ciphers (TBC) as well as some results needed for our proofs are given in Appendix A.

2.2 Definitions

We recall the relevant definitions of authenticated encryption schemes and committing security.

Definition 1. *An authenticated encryption (AE) scheme with associated data is a pair of two algorithms* $(\textsc{Enc}, \textsc{Dec})$ *such that*

- $\textsc{Enc}\colon \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M} \to \mathcal{C}$ *takes a key* K*, a nonce* N*, associated data* A*, and a message* M *as input and outputs a ciphertext* (C, T)*.*
- $\textsc{Dec}\colon \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{C} \to \mathcal{M} \cup \{\perp\}$ *takes a key* K*, a nonce* N*, associated data* A*, and a ciphertext* (C, T) *as input and outputs a message* M *or a special symbol* $\perp$*.*

The sets $\mathcal{K}$, $\mathcal{N}$, $\mathcal{A}$, $\mathcal{M}$, and $\mathcal{C}$ denote the key space, nonce space, associated data space, message space, and ciphertext space, respectively. Throughout this work, we consider these sets to be bit strings of certain length, more precisely, $\mathcal{K} = \{0,1\}^\kappa$, $\mathcal{N} = \{0,1\}^\nu$, $\mathcal{A} = \{0,1\}^*$, $\mathcal{M} = \{0,1\}^*$, and $\mathcal{C} = \{0,1\}^* \times \{0,1\}^\tau$. An AE scheme is called *correct*, if $\textsc{Dec}(K, N, A, \textsc{Enc}(K, N, A, M)) = M$, for any (K, N, A, M). We note further that all considered schemes are *tidy* [47], i.e., $M = \textsc{Dec}(K, N, A, C)$ implies that $C = \textsc{Enc}(K, N, A, M)$. Following the nomenclature from [45], we call the triple (K, N, A) a *context*.

Simply speaking, committing security requires the adversary to find two context-message pairs that encrypt to the same ciphertext. We recall some weaker forms of committing security in Appendix A.

Definition 2. *Let* $\textsc{Ae} = (\textsc{Enc}, \textsc{Dec})$ *be an authenticated encryption scheme and the game* CMT *be defined as in Fig. 1. For any adversary* $\mathcal{A}$*, its* CMT *advantage is defined as*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Ae}}(\mathcal{A}) \coloneqq \Pr[\mathsf{CMT}(\mathcal{A}) \to 1]\,.$$

2.3 NIST LWC Finalists

The NIST LWC standardization process [48] required the submitted AE schemes to achieve the well-established notions of confidentiality and authenticity. For the former, the requirement was to maintain security as long as nonces are unique—security in case of repeating nonces can be mentioned as a special feature. Committing security is neither mentioned as a requirement nor a feature to be advertised. However, it is important to note that the call for algorithms was published the same year as the first attack [32] that exploited the absence

Game CMT (CMT-3 in [9])	Game CMT (CMT-4 in [9])
1 : $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$	1 : $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$
2 : **if** $(K, N, A) = (\overline{K}, \overline{N}, \overline{A})$	2 : **if** $(K, N, A, M) = (\overline{K}, \overline{N}, \overline{A}, \overline{M})$
3 : **return** 0	3 : **return** 0
4 : $(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$	4 : $(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
5 : $(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$	5 : $(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$
6 : **return** $((C, T) = (\overline{C}, \overline{T}))$	6 : **return** $((C, T) = (\overline{C}, \overline{T}))$

Fig. 1: Security game CMT. The version on the left requires the contexts to differ, the one on the right requires the context-message pair to differ. Bellare and Hoang [9] showed the notions to be equivalent due to correctness. We use the version on the left side.

of committing security. Due to the more recent research in this area, it can be expected that committing security will either become a requirement or at least a feature considered relevant for cryptographic standards.[13]

The AE schemes that we study in this work are the NIST LWC finalists that are based on (tweakable) block-ciphers or sponges.[14] In the following, we provide some more information about the schemes with regards to similarities and differences. Table 2 gives an overview and details will be given in this section. For each candidate we focus on the main parameter set as described in Table 3.

Classes of AE Schemes. The considered schemes can be divided into two categories. The first class encompasses AE schemes that follow the *Encrypt-then-MAC* (EtM) paradigm [10]. These schemes first encrypt the message and subsequently authenticate the resulting ciphertext alongside the nonce and the associated data. The second class comprises AE schemes that follow what we call *Context-pre-Processing* (CpP). These schemes first process the context (K, N, A) via a function $\textsc{Enc}_{\mathcal{C}}$. The result is then processed together with M and optionally K and N, yielding the ciphertext (C, T) via a function $\textsc{Enc}_{\mathcal{M}}$. The former (EtM) is illustrated in Fig. 2, the latter (CpP) in Fig. 3. Out of the schemes that we analyze in this work, Elephant and Isap follow the EtM paradigm, whereas the others—Romulus, Photon-Beetle, Gift-Cofb, Xoodyak, TinyJambu, Ascon, and Schwaemm—follow the CpP-approach.

Attacking Encrypt-then-MAC AE Schemes For the EtM schemes, we can focus on the underlying MAC. Once we have two contexts $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$

[13] This can be seen in other standardization processes by NIST. While additional security properties for signature schemes [21] have not been considered in the initial NIST post-quantum cryptography standardization process [49], they are named as desirable features in the very recent call for additional post-quantum signatures [50].

[14] This covers all finalists except Grain-128aead [37] which does not fall into one of these categories.

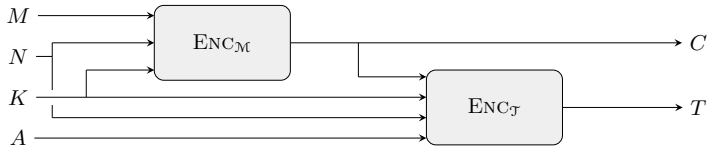

Fig. 2: Illustration of Encrypt-then-MAC AE schemes. Both Elephant and Isap follow this design.

that verify the same ciphertext (C, T), we can immediately derive a committing attack. This is the case because for the described contexts, the decryption algorithm will return some messages $M, \overline{M} \neq \perp$.[15] Using the tidyness property, we get $\textsc{Enc}(K, N, A, M) = (C, T) = \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$, which implies that we win the game CMT.

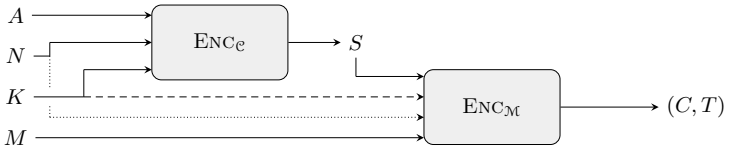

Fig. 3: Illustration of Context-pre-Processing AE schemes. The dotted/dashed arrows indicate that only some of the analyzed schemes exhibit these dependencies: Photon-Beetle and Xoodyak have neither of the two; Gift-Cofb, TinyJambu, Ascon, and Schwaemm have only the dashed line; and Romulus has both lines.

Attacking Context-pre-Processing AE Schemes For the CpP schemes, we focus on the state S that is outputted by $\textsc{Enc}_{\mathcal{E}}$ and then fed into $\textsc{Enc}_{\mathcal{M}}$. The general idea is to generate the first context (K, N, A) and ciphertext (C, T) at random. Then, we invert $\textsc{Enc}_{\mathcal{M}}$ for the same ciphertext (C, T) and a different key $\overline{K}$ and nonce $\overline{N}$, which yields the state $\overline{S}$ (along with the message $\overline{M}$). In the last step, we find associated data $\overline{A}$ such that $\textsc{Enc}_{\mathcal{E}}$ with input $(\overline{K}, \overline{N}, \overline{A})$ results in $\overline{S}$, which ultimately yields a committing attack as $\textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M}) = \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, \textsc{Enc}_{\mathcal{E}}(\overline{K}, \overline{N}, \overline{A}), \overline{M}) = \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, \overline{S}, \overline{M}) = (C, T)$. The step of finding $\overline{A}$ is essentially what was recently coined a *context discovery attack* [45]. This is a stronger attack that easily translates to a committing attack as shown in [45]. Indeed our committing attacks against Romulus, Gift-Cofb, Elephant and Photon-Beetle can easily be translated into context discovery attacks; for the other committing attacks this is not the case.

[15] The underlying decryption algorithm never returns $\perp$, thus the AE scheme returns $\perp$ iff the verification of the tag failed. For both Elephant and Isap, this is the case.

State-Update-Function. Out of the nine finalists, Romulus, Gift-Cofb, Photon-Beetle, and Schwaemm deploy a so-called *state-update-function*[16]. This function takes as input a state S and some additional input data I, and outputs a new state Y and additional output data O (cf. Fig. 4). The state-update-functions work very similar for all four schemes: one of the outputs is the XOR of the inputs whereas the other is the XOR of the input data I and some underlying function—which depends on the respective scheme—applied to the input states.

Typically, the state-update-function is used to process the associated data and the message: The current state is used as the input state S while the associated data or the message—more precisely, a block of it—is used as the input data I. The output state Y is used as the new state while the output data O yields the ciphertext or is simply discarded when the associated data is processed. For decryption, the schemes use the inverse of ξ. Here it is important to note that, inverse is to be understood *only* in relation to the output data, i.e., for any (S, I), $\xi(S, I) = (Y, O) \Rightarrow \xi^{-1}(S, O) = (Y, I)$.

For our attacks against Romulus and Gift-Cofb, we need to invert the state-update-function with respect to *both* outputs, which is not obviously possible from the specifications. Our committing attack against Photon-Beetle is independent of the used state-update-function and for Schwaemm the state-update-function is incorporated into the committing security proof.

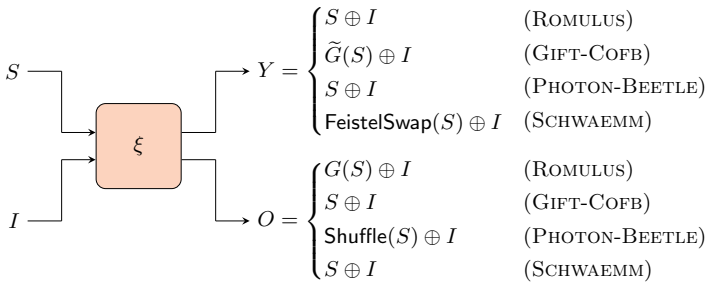

Fig. 4: Illustration of the state-update-function ξ for the different schemes. The components G, $\widetilde{G}$, Shuffle, and FeistelSwap—if relevant for our results—are described along with the schemes in the respective sections.

Achieving Committing Security via Transformations. There are several transformations that turn an arbitrary AE scheme into one that is committing. Clearly, such transformations can be applied to the NIST LWC finalists to make them committing. However, there are several reasons against this: Firstly, these transformations often do not achieve CMT security as we target here but weaker

[16] We adopt this name from [40], the other three finalists use different names.

Table 2: Overview of the NIST LWC finalists regarding similarities and differences in their design. Here, CpP stands for Context-pre-Processing, EtM for Encrypt-then-MAC, and (T)BC for (tweakable) block-cipher.

Scheme	Class of scheme	Underlying Primitive	State-update-function
Elephant [16]	EtM	TBC	No
Romulus [40]	CpP	TBC	Yes
Gift-Cofb [3]	CpP	BC	Yes
Photon-Beetle [5]	CpP	Sponge	Yes
Xoodyak [23]	CpP	Sponge	No
TinyJambu [56]	CpP	BC	No
Isap [29]	EtM	Sponge	No
Ascon [31]	CpP	Sponge	No
Schwaemm [7]	CpP	Sponge	Yes

Table 3: Parameters of the NIST LWC finalists. Values for rate and capacity are only given for the sponge-based schemes. For Isap, the parameters are for the version using Keccak-P, the version using Ascon-P has $n = 320$ and $r = 64$. Note that for Xoodyak and Isap, components of the schemes use rates deviating from the ones given above: The $\mathrm{ENC}_{\mathcal{C}}$ component in Xoodyak is a full-state sponge and in Isap's re-keying mechanism a minimal rate of 1 is used.

Scheme	Key κ	Nonce ν	Tag τ	State n	Rate r	Capacity c
Elephant [16]	128	96	64	160	-	-
Romulus [40]	128	128	128	128	-	-
Gift-Cofb [3]	128	128	128	128	-	-
Photon-Beetle [5]	128	128	128	256	128	128
Xoodyak [23]	128	128	128	384	192	192
TinyJambu [56]	128	96	64	128	-	-
Isap [29]	128	128	128	400	144	256
Ascon [31]	128	128	128	320	64	256
Schwaemm [7]	128	256	128	384	256	128

notions [45]. Secondly, these transformations impose some overhead which—especially considering the lightweight aspect of these schemes—might render them impractical. Thirdly, consider, say, ISAP, which comes with a formal security proof incorporating side-channel leakage. Since none of the transformations are analyzed w.r.t. side-channel leakage, applying them to ISAP can render the leakage security guarantees obsolete.

3 Security Analysis

Here, we analyze the CMT security of the NIST LWC finalists. For ELEPHANT (cf. Section 3.1), ROMULUS (cf. Section 3.2), GIFT-COFB (cf. Section 3.3), and PHOTON-BEETLE (cf. Section 3.4), we give attacks that break committing security with essentially no cost—requiring the bare minimum of two encryptions. For XOODYAK (cf. Section 3.5) and TINYJAMBU (cf. Section 3.6), we give committing attacks requiring significantly less than 2^{64} queries. For ISAP (cf. Section 3.7), ASCON (cf. Section 3.8), and SCHWAEMM (cf. Section 3.9), we give formal proofs showing that the schemes achieve committing security of about 64-bit. For all of the NIST schemes, we provide a figure illustrating the scheme as well as the pseudocode. The figures are in the respective subsections of the present section, while the pseudocodes can be found in Appendix B.

3.1 ELEPHANT

The AE scheme ELEPHANT [15, 16] is based on tweakable block-ciphers. More precisely, it relies on a cryptographic permutation which gets masked using linear feedback shift registers similar to the masked Even-Mansour construction [35].

Description of ELEPHANT. The pseudocode of ELEPHANT is given in Fig. 20 and further illustration is provided in Fig. 5. ELEPHANT follows the Encrypt-then-MAC paradigm, i.e., it first encrypts the message $C \leftarrow \text{ENC}_{\mathcal{M}}(K, N, M)$ and afterwards computes the tag $T \leftarrow \text{ENC}_{\mathcal{T}}(K, N, A, C)$. Note that in $\text{ENC}_{\mathcal{T}}$ the nonce and associated data are padded together, i.e., the first associated data block contains the nonce and the first bits of the associated data. This is in contrast to all other schemes, where the associated data blocks do not contain the nonce. Furthermore, note that the underlying encryption is an involution, i.e., to decrypt a ciphertext, we simply compute $\text{ENC}_{\mathcal{M}}(K, N, C)$.

Committing Attack against ELEPHANT. Since ELEPHANT follows the EtM paradigm, we only need to focus on the underlying MAC $\text{ENC}_{\mathcal{T}}$. If we can find two contexts that verify a ciphertext-tag pair, applying $\text{ENC}_{\mathcal{M}}$ to the ciphertext and each context, gives back two valid messages. The following attack, which is the simplest one in this work, shows that ELEPHANT [16] does not achieve committing security.

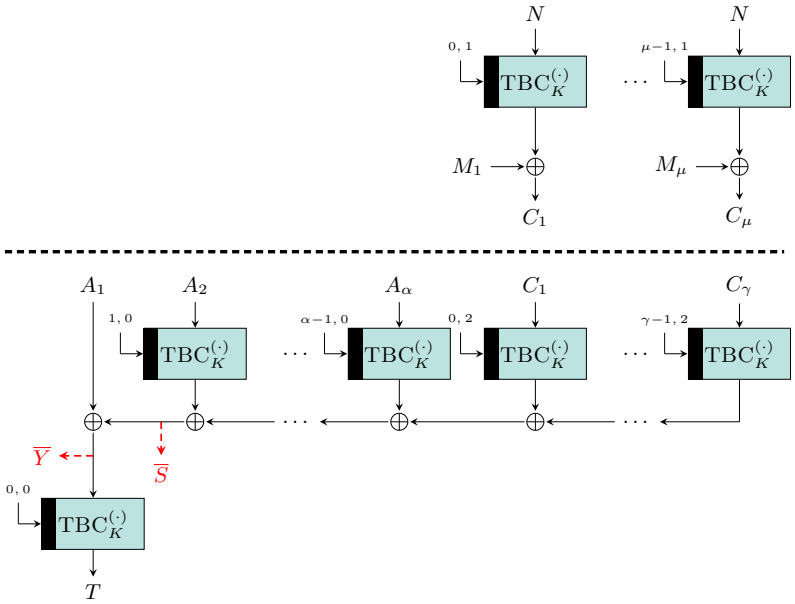

Fig. 5: Illustration of ELEPHANT in terms of $\text{ENC}_{\mathcal{M}}$ (top) and $\text{ENC}_{\mathcal{T}}$ (bottom). The states $\overline{S}$ and $\overline{Y}$, marked in red, are used in our CMT attack.

Theorem 3. *Consider* ELEPHANT *illustrated in Fig. 5. Let* TBC *be modeled as an ideal tweakable cipher* $\widetilde{\mathsf{E}}$*. Then there exists an adversary* $\mathcal{A}$*, making* q *queries to* $\widetilde{\mathsf{E}}$*, such that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{ELEPHANT}}(\mathcal{A}) = 1\,,$$

where $q = 2\mu + 2\gamma + \alpha + \overline{\alpha}$*. Here,* μ *is the number of message blocks while computing* $\text{ENC}_{\mathcal{M}}$ *and* γ *is the number of ciphertext blocks while computing* $\text{ENC}_{\mathcal{T}}$*.*[17] *Furthermore,* α *and* $\overline{\alpha}$ *are the number of associated data blocks for the two tuples that* $\mathcal{A}$ *outputs.*

Proof (sketch). The full proof is given in Appendix B.1, here we provide a sketch. Adversary $\mathcal{A}$ generates a ciphertext (C, T) by encrypting an arbitrary context-message pair $((K, N, A), M)$. Then, it computes the states $\overline{Y}$ and $\overline{S}$ shown in Fig. 5 for (C, T), a fresh key $\overline{K}$, and new associated data blocks $\overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}}$. It sets the remaining associated data block $\overline{A}_1$ to the XOR of $\overline{Y}$ and $\overline{S}$. Finally, $\mathcal{A}$ extracts $(\overline{N}, \overline{A})$ from the associated data blocks, computes $\overline{M} \leftarrow \text{ENC}_{\mathcal{M}}(\overline{K}, \overline{N}, C)$ and outputs $((K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}))$. □

[17] Note that μ and γ might not be the same.

3.2 Romulus

Romulus [39, 40] is an authenticated encryption scheme based on tweakable block-ciphers. For the concrete instantiation of the TBC, they use Skinny [8] and the authenticated encryption mode bears similarities with Cofb [19]. Romulus comes in three different variants Romulus-N, Romulus-M, and Romulus-T. The former is the main candidate while the other two are designed with additional security guarantees in mind: Romulus-M achieves security against nonce-misuse while Romulus-T is designed to maintain security even in the presence of side-channel leakage. Throughout this work we only consider the main variant Romulus-N, which we simply refer to as Romulus.

Description of Romulus. The pseudocode of Romulus is given in Fig. 22. The scheme is further illustrated in Fig. 6. It follows the CpP approach, i.e., it first computes $S \leftarrow \textsc{Enc}_{\mathcal{E}}(K, N, A)$ and afterwards $(C, T) \leftarrow \textsc{Enc}_{\mathcal{M}}(K, N, S, M)$. Both $\textsc{Enc}_{\mathcal{E}}$ and $\textsc{Enc}_{\mathcal{M}}$ apply the tweakable block-cipher and the state-update-function ξ in an alternating manner. The state-update-function

$$\xi : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n \times \{0,1\}^n,\ \xi(S, I) = (S \oplus I, G(S) \oplus I)$$

is an important component of Romulus and the matrix G it utilizes, is

$$G = \begin{pmatrix} G_s & 0 & 0 & \cdots & 0 \\ 0 & G_s & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & G_s & 0 \\ 0 & \cdots & 0 & 0 & G_s \end{pmatrix}, \text{ where } G_s = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Committing Attack Against Romulus. We show that Romulus is insecure with respect to committing security. The attack is stated in the following theorem.

Theorem 4. *Consider* Romulus *illustrated in Fig. 6. Let* TBC *be modeled as an ideal tweakable cipher* $\widetilde{\mathsf{E}}$*. Then there exists an adversary* $\mathcal{A}$*, making* q *queries to* $\widetilde{\mathsf{E}}$*, such that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Romulus}}(\mathcal{A}) = 1\,,$$

where $q = 2\mu + \left\lfloor \frac{\alpha}{2} \right\rfloor + \left\lfloor \frac{\overline{\alpha}}{2} \right\rfloor + 2$*. Here,* μ *is the number of blocks for the message,* α *is the number of blocks for the associated data of the first tuple, and* $\overline{\alpha}$ *is the number of blocks for the second tuple that* $\mathcal{A}$ *outputs.*

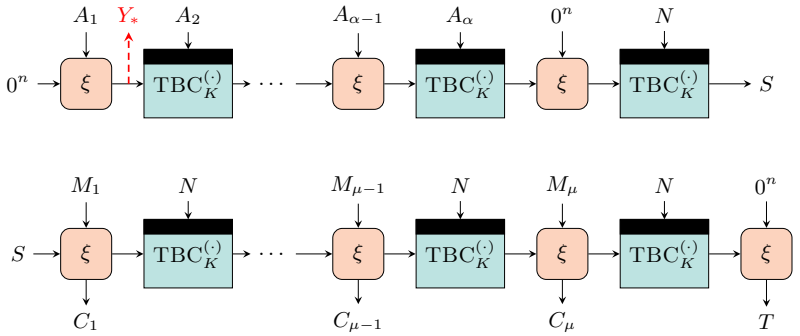

Fig. 6: Illustration of Romulus (for α an even number) in terms of $\mathrm{Enc}_{\mathcal{E}}$ (top) and $\mathrm{Enc}_{\mathcal{M}}$ (bottom). The values that are input from the top into $\mathrm{TBC}_K^{(\cdot)}$ are used as tweaks (we drop the counters making the tweaks unique for simplicity). The state Y_*, marked in red, is used in our CMT attack.

Proof (sketch). The full proof is given in Appendix B.2, here we provide a sketch. A key ingredient of the proof is to show that the state-update-function is invertible with respect to both the output and the state. Exploiting this, the following committing attack is possible. The adversary $\mathcal{A}$ generates a ciphertext (C, T) by encrypting an arbitrary context-message pair $((K, N, A), M)$. It then inverts $\mathrm{Enc}_{\mathcal{M}}$ for the ciphertext (C, T), a new key $\overline{K}$, and new nonce $\overline{N}$, by inverting the state-update-function and the block-cipher one by one, resulting in some state S. Finally, it inverts $\mathrm{Enc}_{\mathcal{E}}$ on S (using $\overline{K}$, $\overline{N}$, and new associated data blocks $\overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}}$) up to the state Y_* (cf. Fig. 6). Setting the first associated data block $\overline{A}_1$ to Y_* then guarantees that $\mathcal{A}$ wins the game CMT by outputting $((K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}))$. □

3.3 Gift-Cofb

The AE scheme Gift-Cofb uses the Cofb mode [19] for authenticated encryption and instantiates the block-cipher using Gift [4].

Description of Gift-Cofb. The pseudocode of Gift-Cofb is given in Fig. 24 while Fig. 7 provides an illustration of it. Gift-Cofb follows the CpP approach, i.e., it first computes $S \leftarrow \mathrm{Enc}_{\mathcal{E}}(K, N, A)$ and afterwards $(C, T) \leftarrow \mathrm{Enc}_{\mathcal{M}}(K, S, M)$. The scheme features the state-update-function

$$\xi : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n \times \{0,1\}^n,\ \xi(S, I) = (\widetilde{G}(S) \oplus I, S \oplus I)$$

that is invoked in an alternating manner with the block-cipher. This function makes use of the following matrix $\widetilde{G}$, which gets a bit string of length n, swaps the two halves, and additionally applies a bit rotation to the (new) second half:

$$\widetilde{G} \colon \{0,1\}^{n/2} \times \{0,1\}^{n/2} \to \{0,1\}^{n/2} \times \{0,1\}^{n/2},\ (S_1, S_2) \ \mapsto (S_2, S_1 \lll 1)\,.$$

In between the state-update-function and the block-cipher, GIFT-COFB applies some masking by XORing some value Δ to the state.

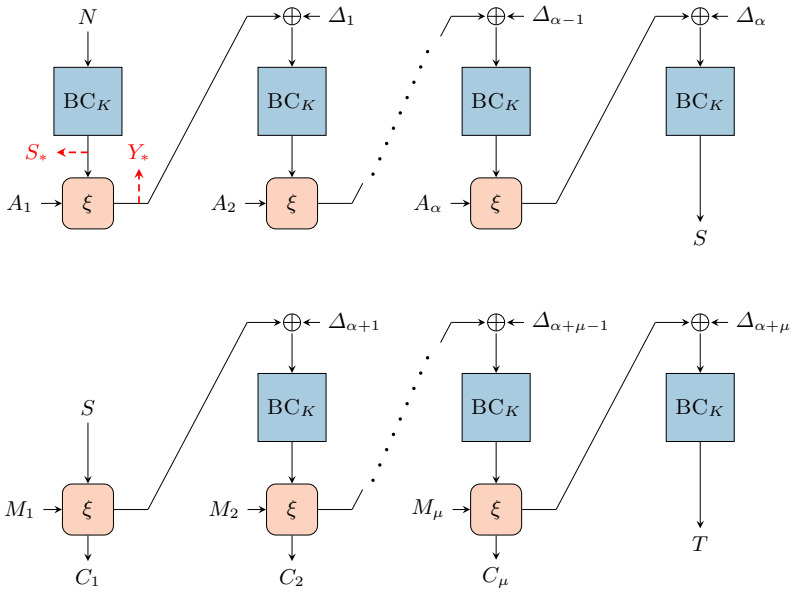

Fig. 7: Illustration of GIFT-COFB in terms of $\mathrm{ENC}_{\mathfrak{C}}$ (top) and $\mathrm{ENC}_{\mathfrak{M}}$ (bottom). The different indices for Δ only indicate that the values are different—in the pseudocode the value Δ is constantly updated. The states S_* and Y_*, marked in red, are used in our CMT attack.

Committing Attack Against GIFT-COFB. The AE scheme GIFT-COFB does not achieve committing security. The scheme is very similar to ROMULUS which allows to apply the same attack strategy. However, GIFT-COFB uses a different state-update-function ξ than ROMULUS. It turns out that we cannot always invert ξ—inversion only works with probability $\frac{1}{2}$. This is the reason, why the advantage depends on the number of ciphertext blocks.

Theorem 5. *Consider* GIFT-COFB *illustrated in Fig. 7. Let* BC *be modeled as an ideal cipher* E. *Then there exists an adversary* $\mathcal{A}$, *making* q *queries to* E, *such that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{GIFT-COFB}}(\mathcal{A}) = \frac{1}{2^\mu}\,,$$

where $q = 2\mu + \alpha + \overline{\alpha} + 2$. *Here,* μ *is the number of message blocks,* α *is the number of associated data blocks for the first tuple, and* $\overline{\alpha}$ *for the second tuple that* $\mathcal{A}$ *outputs.*

Proof (sketch). The full proof is given in Appendix B.3, here we provide a sketch. The gist is the same as for ROMULUS: generate a ciphertext, compute S_* and Y_* (cf. Fig. 7) for a new context, and set the associated data block $\overline{A}_1$ accordingly. An important difference is that inverting ξ in GIFT-COFB works only with probability $\frac{1}{2}$. Another detail that has to be considered is the usage of the correct masking values. □

3.4 PHOTON-BEETLE

The authenticated encryption scheme PHOTON-BEETLE [5] is a (duplex) sponge-based AE scheme. It uses the PHOTON permutation [36] as the underlying permutation and the BEETLE mode of operation [18]. In contrast to the plain duplex, the BEETLE mode uses a feedback function to determine the next input to the underlying permutation of the sponge.

Description of PHOTON-BEETLE. PHOTON-BEETLE is described in Fig. 26 and illustrated in Fig. 8. It is a CpP scheme, i.e., it processes the context (K, N, A) via the function $\text{ENC}_{\mathcal{E}}$ and subsequently processes the message together with the output of $\text{ENC}_{\mathcal{E}}$—an important property is that no part of the context is input to $\text{ENC}_{\mathcal{M}}$. In $\text{ENC}_{\mathcal{M}}$, the permutation and the state-update-function are applied in an alternating fashion. We omit the description of the latter, as our CMT attack is independent of it.

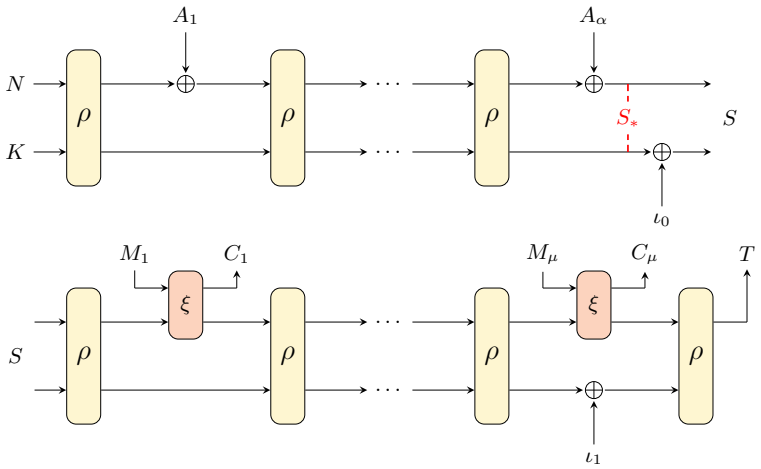

Fig. 8: Illustration of PHOTON-BEETLE in terms of $\text{ENC}_{\mathcal{E}}$ (top) and $\text{ENC}_{\mathcal{M}}$ (bottom). The state S_*, marked in red, is used in our CMT attack.

Committing Attack Against PHOTON-BEETLE**.** In this section, we show that PHOTON-BEETLE does not achieve committing security. The CMT attack is given in Theorem 6 below.

Theorem 6. *Consider* PHOTON-BEETLE *illustrated in Fig. 8 Let ρ be modeled as an ideal permutation. Then there exists an adversary $\mathcal{A}$, making q queries to ρ, such that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{PHOTON-BEETLE}}(\mathcal{A}) = 1\,,$$

where $q = \alpha + \overline{\alpha}$. Here, α is the number of blocks for the associated data of the first tuple, and $\overline{\alpha}$ is the number of blocks for the second tuple that $\mathcal{A}$ outputs.

Proof (sketch). The full proof is given in Appendix B.4, here we provide a sketch. The attack targets the state right before applying the domain separation at the end of $\text{ENC}_{\mathcal{E}}$ (the state S_* marked in red in Fig. 8). Since the initial state of PHOTON-BEETLE is entirely dependent on the context, we can take an arbitrary S_*, invert it for different associated data A and $\overline{A}$, and take the results as the key-nonce pairs (K, N) and $(\overline{K}, \overline{N})$. Since $\text{ENC}_{\mathcal{M}}$ only depends on the outcome of $\text{ENC}_{\mathcal{E}}$ and the message M, $\mathcal{A}$ wins by outputting $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, M)$ for an arbitrary message M. □

3.5 XOODYAK

The authenticated encryption scheme XOODYAK [23] is an AE scheme based on a full-state keyed duplex. XOODYAK uses the XOODOO permutation [22] as the underlying permutation and the CYCLIST mode of operation. The latter was introduced as part of the XOODYAK specification and is an adaption of the KEYAK mode [14] to the lightweight setting.

Description of XOODYAK**.** The pseudocode of XOODYAK is given in Fig. 27 and further illustration is provided in Fig. 9. XOODYAK is a CpP scheme, i.e., first $S \leftarrow \text{ENC}_{\mathcal{E}}(K, N, A)$ is computed, followed by the computation of ciphertext and tag as $(C, T) \leftarrow \text{ENC}_{\mathcal{M}}(S, M)$. In $\text{ENC}_{\mathcal{E}}$, the inputs are XORed onto the full-state (note that the last 32 bits are reserved for padding). In contrast to this, $\text{ENC}_{\mathcal{M}}$ uses a rate of 192 bits for the computation of ciphertext and tag.

XOODYAK exhibits a form of padding, that is used by none of the other NIST candidates and hence will be described shortly in the following: For a bit string X of length at most 352 and $p \in \{0,1\}^8$ define

$$\mathtt{pad}_{\mathcal{E}}(X, p) = X \parallel 00000001 \parallel 0^{368-|X|} \parallel p\,,$$

which is used for padding the context blocks. Further, for $M \in \{0,1\}^{\leq 192}$, we define $\mathtt{pad}_{\mathcal{M}}(M) = M \parallel 00000001$, which will be used to pad the message blocks.

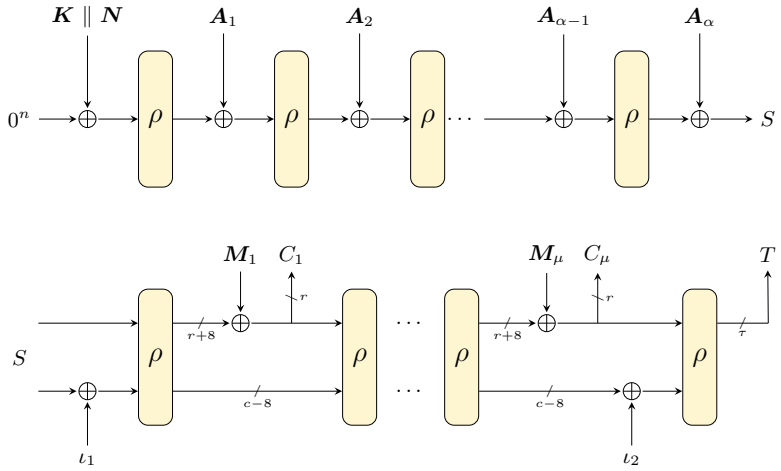

Fig. 9: Illustration of XOODYAK in terms of $\text{ENC}_{\mathcal{C}}$ (top) and $\text{ENC}_{\mathcal{M}}$ (bottom). Here, $\boldsymbol{K} \parallel \boldsymbol{N} = \texttt{pad}_{\mathcal{C}}((K \parallel N \parallel \texttt{enc}_8(|N|)), 00000010)$, $\boldsymbol{M}_i = \texttt{pad}_{\mathcal{M}}(M_i)$, $\boldsymbol{A}_1 = \texttt{pad}_{\mathcal{C}}(A_1, 00000011)$, and for $i = 2, \ldots, \alpha$, $\boldsymbol{A}_i = \texttt{pad}_{\mathcal{C}}(A_i, 0^8)$. In $\text{ENC}_{\mathcal{M}}$, the increased rate $(r+8)$ is required for the padding $\texttt{pad}_{\mathcal{M}}$ which appends 8 bits to the message blocks.

Committing Attack Against XOODYAK. We show that XOODYAK does not achieve committing security. The attack is stated in the following theorem.

Theorem 7. *Consider* XOODYAK *illustrated in Fig. 9 Let ρ be modeled as an ideal permutation. Then there exists an adversary $\mathcal{A}$ that makes $q = 2^{17} + 1$ queries to ρ and fulfills*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{XOODYAK}}(\mathcal{A}) \geq \frac{1}{2}\,.$$

Proof (sketch). The full proof is given in Appendix B.5, here we provide a sketch. The attack exploits the fact that $\text{ENC}_{\mathcal{M}}$ only depends on the output of $\text{ENC}_{\mathcal{C}}$ and the message, i.e., two different contexts $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$ with $\text{ENC}_{\mathcal{C}}(K, N, A) = \text{ENC}_{\mathcal{C}}(\overline{K}, \overline{N}, \overline{A})$ easily yield a committing attack by outputting $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, M)$ for an arbitrary M. Though XOODYAK is a full-state sponge, one cannot simply choose contexts yielding a collision, as the last 32 bits are reserved for padding. With a birthday attack, the adversary can find a collision on the reserved bits and then choose the context accordingly. □

3.6 TINYJAMBU

TINYJAMBU [56] is a block-cipher-based authenticated encryption scheme. The specification introduces the TINYJAMBU mode, which is a lightweight variant of the JAMBU mode [55]. The latter was part of the CAESAR competition [11]. For the permutation underlying TINYJAMBU, a keyed permutation based on non-linear feedback shift registers is defined.

Description of TinyJambu**.** The pseudocode of TinyJambu is given in Fig. 28 and an illustration of the scheme can be found in Fig. 10. TinyJambu follows the CpP approach, i.e., it first processes the context (K, N, A) via the function $\text{Enc}_{\mathcal{C}}$ and then passes the output on to $\text{Enc}_{\mathcal{M}}$, where it is processed together with the message. Note that, as for Photon-Beetle, no part of the context is directly given as input into $\text{Enc}_{\mathcal{M}}$. TinyJambu uses two keyed permutations BC_1 and BC_2, both based on the same keyed permutation that is applied 640 and 1024 times for BC_1 and BC_2, respectively.

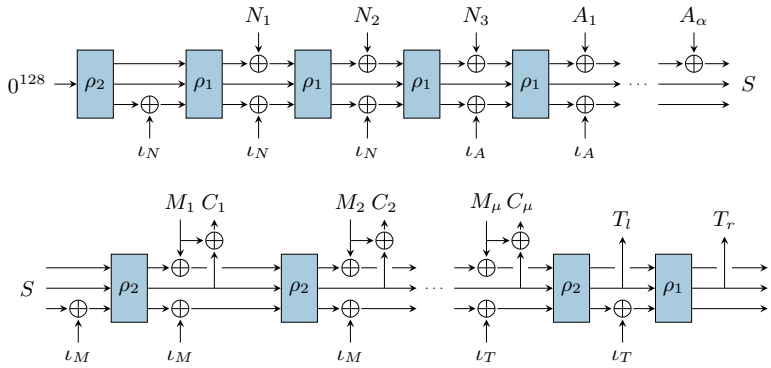

Fig. 10: Illustration of TinyJambu in terms of $\text{Enc}_{\mathcal{C}}$ (top) and $\text{Enc}_{\mathcal{M}}$ (bottom), where $\rho_1 = \text{BC}_1(K, \cdot)$ and $\rho_2 = \text{BC}_2(K, \cdot)$.

Committing Attack Against TinyJambu**.** In this section, we show that TinyJambu does not achieve CMT security. The attack exploits the short tag length of 64 bits in TinyJambu, which enables an efficient deployment of the birthday bound. In the security proof of TinyJambu (see [56, Section 6]), this setting is modeled with only one permutation ρ. We adopt the same for the TinyJambu attack given in the following.

Theorem 8. *Consider* TinyJambu *illustrated in Fig. 10. Let* BC_1 *and* BC_2 *be modeled as an ideal cipher* E*. Then there exists an adversary* $\mathcal{A}$ *that makes* q *queries to* E *such that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{TinyJambu}}(\mathcal{A}) \geq \frac{3}{8}\,.$$

Here, $q = 2(2^{32}+1)(6+\alpha+\mu)$ *for* α *and* μ *the number of associated data and message blocks, respectively, that* $\mathcal{A}$ *outputs.*

Proof (sketch). The full proof is given in Appendix B.6, here we provide a sketch. The core idea is to apply a birthday attack against the tag, which requires about 2^{32} queries due to the tag length of 64 bits. Some extra care is needed to ensure

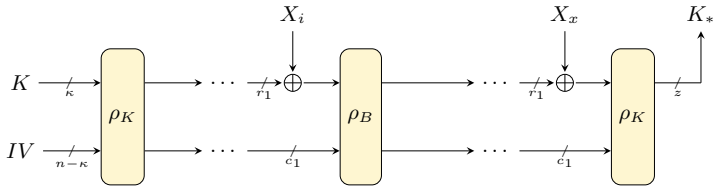

Fig. 11: Illustration of ISAP.RK.

that also the actual ciphertexts (and not just the tags) collide, as a tag collision is not sufficient for a CMT attack. Since the ciphertext is the XOR of the message and a keystream, this poses no hindrance to the attack, as the adversary can simply choose the messages in a way that yields the same ciphertext. □

3.7 ISAP

The authenticated encryption scheme ISAP [28–30] is a sponge-based scheme designed to withstand side-channel leakage. It features a re-keying approach that guarantees that for each input a different session key is used. The re-keying function uses a small rate to prevent adversaries from obtaining too much leakage.

Description of ISAP. The pseudocode of ISAP is given in Fig. 30 and further illustration is provided in Fig. 12. ISAP follows the EtM approach, i.e., first the message is encrypted via $\text{ENC}_{\mathcal{M}}$ resulting in a ciphertext C and afterwards, the tag T is computed using $\text{ENC}_{\mathcal{T}}$, which processes the context and the ciphertext. Both $\text{ENC}_{\mathcal{M}}$ and $\text{ENC}_{\mathcal{T}}$ internally uses the re-keying function ISAP.RK to derive the session key.

In ISAP, the underlying permutation is applied several times between two absorptions, the precise number depending on the position in the ISAP sponge. For ISAP.RK, for instance, more rounds are applied when the key is processed and when the tag is generated, while fewer rounds are used in between. In summary, ISAP uses four permutations ρ_K, ρ_H, ρ_B, and ρ_E, each based on the same permutation but applied a different number of times.

Committing Security of ISAP. We show that ISAP achieves CMT security. In the security proof for ISAP [28], the two permutations used in ISAP.RK (namely ρ_K and ρ_B) are modeled as one permutation. We adopt this for our proof and denote the permutation used in ISAP.RK by ρ_1 and the one used in $\text{ENC}_{\mathcal{T}}$ by ρ_2.[18] In conformity with this, the rate in ISAP.RK is denoted by r_1 and the rate in $\text{ENC}_{\mathcal{T}}$ by r_2.

Further, we consider a slightly different domain separation in $\text{ENC}_{\mathcal{T}}$, by XORing $1 \parallel 0^*$ instead of $0^* \parallel 1$. This neither influences the security of ISAP nor

[18] The proof is independent of $\text{ENC}_{\mathcal{M}}$ which is why we do not need a third permutation.

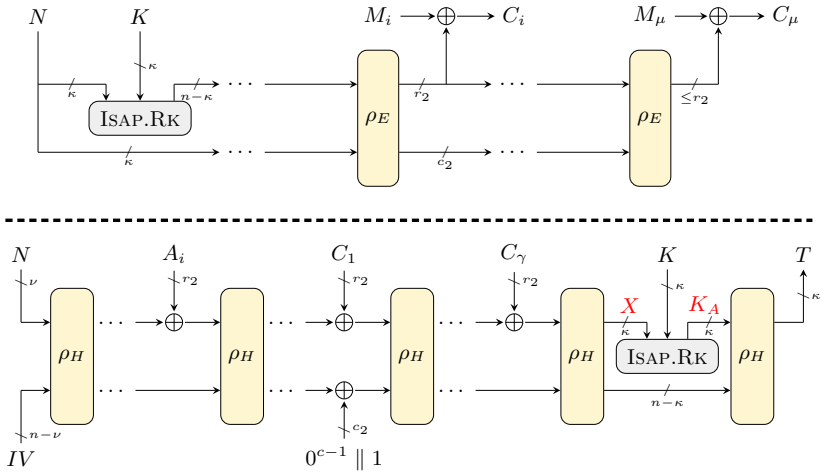

Fig. 12: Illustration of Isap in terms of $\text{Enc}_{\mathcal{M}}$ (top) and $\text{Enc}_{\mathcal{T}}$ (bottom), which both rely on the re-keying function IsapRk. The values X and K_A, marked in red, are used in our CMT proof.

interferes with the purpose of the domain separation but it allows us to view $\text{Enc}_{\mathcal{T}}$ as a sponge with increased rate of $r + 1$.[19]

Lastly, note that the scheme comes in two variants, using either Ascon-P or Keccak-P as a permutation. The theorem given below holds for both instances, considering for both of them the main parameter sets as described in Table 3. Since for all parameter sets under consideration the key-length equals the tag-length ($\kappa = \tau = 128$), we exclusively use the variable κ in the following proof.

Theorem 9. *Consider* Isap *illustrated in Fig. 12. Let ρ_1 and ρ_2 be modeled by ideal permutations ρ_1 and ρ_2, respectively. Then for any adversary $\mathcal{A}$ making q_1 and q_2 queries to ρ_1 and ρ_2, respectively, it holds that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{Isap}}(\mathcal{A}) \leq \frac{q_1(q_1-1)}{2^\kappa} + \frac{q_1(q_1+1)}{2^{n-\kappa}} + \frac{q_2(q_2-1)}{2^\kappa} + \frac{q_2(q_2+1)}{2^{n-\max\{\kappa, r_2+1\}}} .$$

Proof (sketch). The full proof is given in Appendix B.7, here we provide a sketch. The idea is to view $\text{Enc}_{\mathcal{T}}$ as a sponge-based hash function and leverage Theorem 15 to upper bound the advantage of finding two distinct contexts and a ciphertext that result in the same tag. For this, we adapt the rate in $\text{Enc}_{\mathcal{T}}$ in a way, that allows us to absorb the nonce as the first input. With this technical trick, we obtain sponges with constant initial states at the price of a decreased capacity. Further, we model the usage of IsapRk in $\text{Enc}_{\mathcal{T}}$ by a suitable XOR-operation with input denoted by Z. The proof is composed of two parts: Firstly,

[19] The same argument was also used for the sponge-based AE scheme Slae [26] and is also mentioned for Xoodyak in [23, Section 4.2.1].

we consider the case that at least one of the inputs (nonce, associated data, and Z—from the replacement of ISAP.RK) differs for the two tuples the adversary outputs. Then, Theorem 15 provides an upper bound for the probability that the hash values, i.e., the tags, agree. Secondly, we have to consider a special case, in which all of the inputs (nonce, associated data, and Z) are the same across both tuples.[20] Such inputs do not constitute a collision of $\text{ENC}_{\mathcal{T}}$ and hence the above argument does not apply. However, in this case we can show that the committing adversary has found a collision in ISAP.RK. The latter can also be viewed as a sponge-based hash function (by modifying the rate) and a second application of Theorem 15 finishes the proof. □

3.8 ASCON

ASCON [31] is a sponge-based authenticated encryption scheme. The scheme was chosen as the primary candidate for lightweight applications in the CAESAR competition. Furthermore, ASCON was selected to be standardized as part of the NIST LWC standardization process. As part of the CAESAR competition and the NIST LWC standardization process, ASCON enjoys a long line of research, in particular, with respect to the underlying permutation ASCON-P. For the authenticated encryption mode, no formal security analysis existed until recently, when Lefevre and Mennink [42] gave the first security proof for ASCON.[21]

Description of ASCON. The pseudocode of ASCON is given in Fig. 32 and further illustration is provided in Fig. 13. Similar to the other schemes, ASCON can be viewed as a CpP scheme which first processes the context using $\text{ENC}_{\mathcal{C}}$ before the message is processed using $\text{ENC}_{\mathcal{M}}$. A core feature of ASCON is that at the very start (first permutation of $\text{ENC}_{\mathcal{C}}$) and the very end (last permutation of $\text{ENC}_{\mathcal{M}}$), it uses more rounds of the underlying permutation for security (ρ^a and ρ^b for $a = 12$ and $b = 6$). Note that ASCON XORs the key three additional times: After the first permutation as well as before and after the last permutation. We call the former two instances state-blinding and the latter output-blinding.

Committing Security of ASCON. We show that ASCON achieves CMT security. We model the two permutations ρ^a and ρ^b by one ideal permutation ρ, as we did in the committing security proof of ISAP for the re-keying function. Further, we consider a slightly different order of inputs at two points in the ASCON encryption. Firstly, the initial state is changed by moving the initialization vector from the beginning of the state to the end. Secondly, the state-blinding is changed so that it affects the first bits of the inner state rather than the last bits.[22] Note that these are cosmetic changes. They do not influence the overall

[20] The careful reader might notice that in this case the CMT adversary outputs two tuples that differ solely in the keys but result in the same session keys.

[21] While an earlier work [41] argued that their proof covers ASCON, they actually only show security for a simplified version (namely without the state-/output-blinding).

[22] This only affects the first state-blinding; the second one is already of that form.

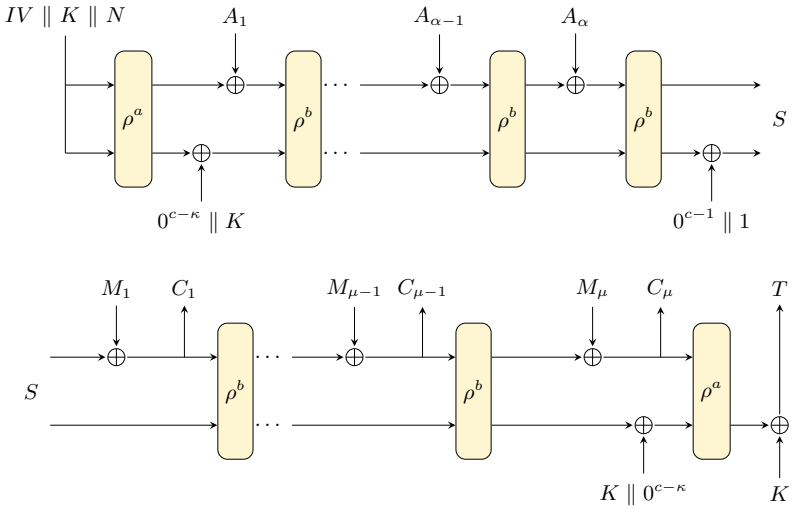

Fig. 13: Illustration of Ascon in terms of $\textsc{Enc}_{\mathcal{E}}$ (top) and $\textsc{Enc}_{\mathcal{M}}$ (bottom).

security of Ascon but greatly simplify our proof, as we can capture the state-blinding by considering a larger rate. Lastly, for sake of simplicity, we drop the domain separation of Ascon in the proof. It can, however, be easily incorporated by moving the domain separation to the first bits of the capacity, as we already did for Isap.

Similarly to Isap, we could show committing security by leveraging Theorem 15. However, for Ascon this would yield only about 32-bit committing security. This is due to the short initialization vector of 64 bits, which corresponds to a 64-bit capacity. By recasting the proof of Theorem 15, we can show that this limitation is not inherent, i.e., a smaller capacity in just the first round does not fundamentally affect the quality of the bound.

Theorem 10. *Consider* Ascon *illustrated in Fig. 13. Let ρ^a and ρ^b be modeled as a random permutation ρ. Then for any adversary $\mathcal{A}$ making $q \leq 2^{127}$ queries to ρ, it holds that*

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Ascon}}(\mathcal{A}) \leq 1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right) + \frac{q}{2^{63}} + \frac{q(q-1)}{2^{128}}\,.$$

Proof (sketch). The full proof is given in Appendix B.8, here we provide a sketch. The main idea is to adapt the proof of Theorem 15 to our particular setting. While we structurally follow the original proof, several changes are necessary and a more fine-grained analysis gives us a better bound.

The proof of Theorem 15 visualizes the attack via a graph, built from the adversary's queries. It defines colliding paths, which correspond to collisions in the hash function, and so-called problematic paths. The collision resistance is upper bounded by the probability of finding either (1) colliding paths that are not problematic or (2) finding problematic paths.

Our proof focuses on bounding the probability of finding colliding tags—which directly yields a bound on the CMT security of Ascon. We define paths in the adversary's graph that correspond to Ascon evaluations and model tag collisions by introducing the notion of colliding Ascon paths. Furthermore, we define problematic Ascon paths and, using both the state- and output-blinding deployed in Ascon, we give a bound on finding colliding Ascon paths that are not problematic. It is left to bound the probability of finding problematic paths. The gist of this part is to view Ascon paths as plain sponge paths, which comes at the cost of an increased rate—for the first round of absorption, the rate is increased to 256 while for the rest it is increased to 192. This allows us to understand the state blinding as part of the normal absorption into the rate part of the sponge. We then show that problematic Ascon paths translate to problematic plain sponge paths and continue by giving a bound on finding the latter. A careful analysis shows that the first absorption step is treated differently: The adversary has to hit the initialization vector with a query, whereas for the other absorption steps, it has to find a collision. Thus, despite the very large rate in the first round, we are still able to achieve about 64-bit of committing security. □

3.9 Schwaemm

Schwaemm [6,7] is a sponge-based AE scheme. The permutation used to instantiate Schwaemm is Sparkle which is inspired by the block-cipher Sparx [27]. The authentication mode is a variant of the Beetle mode [18].

Description of Schwaemm. The pseudocode of Schwaemm is given in Fig. 35 and further illustration is provided in Fig. 14. Schwaemm follows the CpP approach, i.e., first the context is processed resulting in $S \leftarrow \textsc{Enc}_{\mathcal{E}}(K, N, A)$ and afterwards the ciphertext is computed as $(C, T) \leftarrow \textsc{Enc}_{\mathcal{M}}(K, S, M)$. In Schwaemm the underlying permutation ρ is applied a varying number of times, depending on the position in the sponge (ρ^a and ρ^b for $a = 11$ and $b = 7$), similar to Isap and Ascon.

Like some of the other schemes, Schwaemm features a state-update-function that is defined as follows:

$$\xi\colon \{0,1\}^r \times \{0,1\}^r \to \{0,1\}^r \times \{0,1\}^r,$$
$$(S, I) \mapsto (\xi_1(S, I), \xi_2(S, I)) = (\mathsf{FeistelSwap}(S) \oplus I, S \oplus I)\,.$$

for $\mathsf{FeistelSwap} : \{0,1\}^r \to \{0,1\}^r$, $\mathsf{FeistelSwap}(S) = S_2 \parallel (S_2 \oplus S_1)$ with $S_1 = \lceil S \rceil_{\frac{r}{2}}$ and $S_2 = \lfloor S \rfloor_{\frac{r}{2}}$. Furthermore, Schwaemm deploys a so-called rate-whitening function, given by

$$\omega_{c,r}\colon \{0,1\}^c \to \{0,1\}^r, \quad \omega_{c,r}(I) = (I_1, I_2, I_1, I_2) \text{ for } I_1 = \lceil I \rceil_{\frac{c}{2}}\,,\ I_2 = \lfloor I \rfloor_{\frac{c}{2}}\,.$$

In each round, it is applied between the state-update-function and the permutation. After the final permutation, the last κ bits are XORed with the key to yield the tag. As for Ascon, we refer to this as output-blinding.

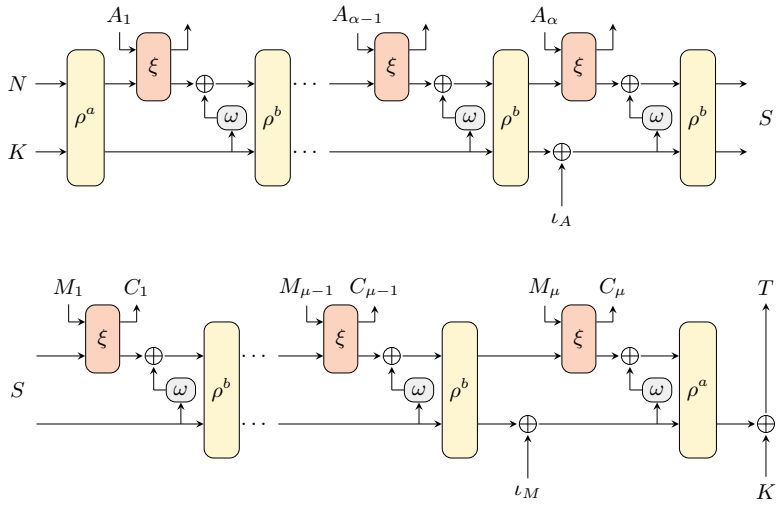

Fig. 14: Illustration of SCHWAEMM in terms of $\text{ENC}_{\mathcal{E}}$ (top) and $\text{ENC}_{\mathcal{M}}$ (bottom).

Committing Security of SCHWAEMM. We show that SCHWAEMM achieves committing security. At the first glance, it looks like one can apply the same attack used against PHOTON-BEETLE: Invert $\text{ENC}_{\mathcal{E}}$ for some S and take the result as the concatenation of key and nonce. However, SCHWAEMM deploys output-blinding in $\text{ENC}_{\mathcal{M}}$ (the last XOR of the key in Fig. 14), that makes the attack unlikely to succeed. Output-blinding is a feature we have also encountered in ASCON, as one of the central properties making it committing secure. Despite that, we cannot show committing security in the same way, as SCHWAEMM lacks the state-blinding, that ASCON has, and SCHWAEMM's initial state does not contain a fixed component. However, we noticed that introducing an IV to SCHWAEMM's initial state suffices to obtain a committing secure scheme—despite the weaker blinding mechanism. More precisely, we decrease the length of the nonce from 256 to 128 bits[23] and instead incorporate a fixed IV (of length 128 bit) into the initial state. For the resulting modified version of SCHWAEMM, denoted by SCHWAEMM_{IV}, we can show about 64-bit committing security. For the proof, we model the two permutations ρ^a and ρ^b by one ideal permutation ρ, as it was done for ISAP and ASCON. We further drop the domain separation in our proof for sake of simplicity. This part can easily be incorporated at the cost of reducing the committing security by the number of bits required for the domain separation.

Theorem 11. *Consider* SCHWAEMM *illustrated in Fig. 14 and* SCHWAEMM_{IV}*, its modified version described above. Let* ρ^a *and* ρ^b *be modeled as a random permutation* ρ*. Then for any adversary* $\mathcal{A}$ *making* $q \leq 2^{127}$ *queries to* ρ*, it holds*

[23] Note that the modified scheme is still in accordance to the NIST requirements [48] that nonces are at least 96 bits long.

that

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Schwaemm}_{IV}}(\mathcal{A}) \leq 1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right) + \epsilon\,,$$

for $\epsilon > \frac{(1-2^{-256})q^2+(1+2^{-256})q}{2^{129}}$.

Proof (Sketch). The full proof is given in Appendix B.9, here we provide a sketch. The idea is to view $\textsc{Schwaemm}_{IV}$ as a plain sponge and show that it is hard to find colliding tags, which, in turn, yields that it is hard to break committing security. Due to the output-blinding, tag collisions are not equal to collisions of the plain sponge, as the keys—used to blind the output—can be different. To deal with that, we define shifted collisions, which encompass these tag collisions for the plain sponge. We conclude the proof by utilizing the indifferentiability of the plain sponge from a random function and providing a bound on finding shifted collisions for the latter. □

4 Conclusion

Out of the nine considered NIST finalists, we have shown that six do not achieve committing security while the remaining three do. For the former, we gave concrete attacks, while the others are backed up by formal security proofs. For Elephant, Gift-Cofb, and Romulus, the attacks can be traced back to the fact that inputs are XORed onto the entire state (we call this full-state XOR). Similarly to this, Xoodyak exhibits the same property as part of it ($\textsc{Enc}_{\mathcal{C}}$) is a full-state sponge: The whole state—except for a few bits reserved for padding—can be influenced by the adversary. Furthermore, also the attack against Photon-Beetle, which exploits the scheme's context-dependent initial state, can be regarded as a "full-state-attack". For this, note that an initial state that can be controlled entirely by the adversary, can easily be represented as a full-state XOR of some input onto the all-zero string. Our attack against TinyJambu is based on the scheme's short tag, which enables an efficient deployment of the birthday attack. Note that in order to obtain 64-bit committing security, a tag length of at least 128 bits is necessary.[24]

Overall, we observe that all block-cipher-based NIST finalists are broken with respect to committing security. At the same time, the three schemes for which we formally prove committing security are all sponge-based. This suggests that sponges are better suited to building committing authenticated encryption schemes, whereas block-cipher-based schemes seem to exhibit some inherent vulnerabilities. However, note that full-state sponges—which have proven to yield AE schemes that achieve confidentiality and authenticity [46]—are also vulnerable with respect to committing security. Our overall insight, i.e., that sponges are more favorable than block-ciphers when designing committing AE schemes, is also supported by prior results: While there is a number of committing attacks

[24] Though, longer tags are not sufficient to obtain CMT security for TinyJambu.

against block-cipher-based schemes [9,38,45], we are not aware of any committing attacks against sponge-based schemes. Moreover, very recently, positive results regarding committing AE from sponges emerged [24,25].

Looking at the positive results (Isap, Ascon, and Schwaemm), a common feature is that a significant part of the state is "out-of-reach" for the adversary, i.e., not manipulable by the input. All of these schemes have a capacity of at least 128 bits. Interestingly, TinyJambu features a similar design in the sense that a part of its state (96 bits) is unaffected by the inputs. However, due to the common, but comparably small, block size of $n = 128$, it is not possible for TinyJambu to achieve 64-bit committing security. This would require the entire 128-bit state to not be affected by the inputs throughout the whole scheme, which would render it pointless. In contrast to this, sponge constructions typically rely on permutations over much larger states: Ascon-P ($n = 320$), Keccak-P ($n = 400$), Sparkle ($n = 384$), Xoodoo ($n = 384$), and Photon ($n = 256$). These allow for efficient constructions while maintaining a large part of the state unaffected by the inputs, which is favorable for committing security.

References

1. Michel Abdalla, Mihir Bellare, and Gregory Neven. Robust encryption. In Daniele Micciancio, editor, *TCC 2010*, volume 5978 of *LNCS*, pages 480–497. Springer, Heidelberg, February 2010.
2. Ange Albertini, Thai Duong, Shay Gueron, Stefan Kölbl, Atul Luykx, and Sophie Schmieg. How to abuse and fix authenticated encryption without key commitment. In Kevin R. B. Butler and Kurt Thomas, editors, *USENIX Security 2022*, pages 3291–3308. USENIX Association, August 2022.
3. Subhadeep Banik, Avik Chakraborti, Tetsu Iwata, Kazuhiko Minematsu, Mridul Nandi, Thomas Peyrin, Yu Sasaki, Siang Meng Sim, and Yosuke Todo. GIFT-COFB. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
4. Subhadeep Banik, Sumit Kumar Pandey, Thomas Peyrin, Yu Sasaki, Siang Meng Sim, and Yosuke Todo. GIFT: A small present - towards reaching the limit of lightweight encryption. In Wieland Fischer and Naofumi Homma, editors, *CHES 2017*, volume 10529 of *LNCS*, pages 321–345. Springer, Heidelberg, September 2017.
5. Zhenzhen Bao, Avik Chakraborti, Nilanjan Datta, Jian Guo, Mridul Nandi, Thomas Peyrin, and Kan Yasuda. PHOTON-Beetle. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
6. Christof Beierle, Alex Biryukov, Luan Cardoso dos Santos, Johann Großschädl, Léo Perrin, Aleksei Udovenko, Vesselin Velichkov, and Qingju Wang. Lightweight AEAD and hashing using the Sparkle permutation family. *IACR Trans. Symm. Cryptol.*, 2020(S1):208–261, 2020.
7. Christof Beierle, Alex Biryukov, Luan Cardoso dos Santos, Johann Großschädl, Léo Perrin, Aleksei Udovenko, Vesselin Velichkov, Qingju Wang, Amir Moradi, and Aein Rezaei Shahmirzadi. Schwaemm and Esch. Technical report, National

Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
8. Christof Beierle, Jérémy Jean, Stefan Kölbl, Gregor Leander, Amir Moradi, Thomas Peyrin, Yu Sasaki, Pascal Sasdrich, and Siang Meng Sim. The SKINNY family of block ciphers and its low-latency variant MANTIS. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part II*, volume 9815 of *LNCS*, pages 123–153. Springer, Heidelberg, August 2016.
9. Mihir Bellare and Viet Tung Hoang. Efficient schemes for committing authenticated encryption. In Orr Dunkelman and Stefan Dziembowski, editors, *EUROCRYPT 2022, Part II*, volume 13276 of *LNCS*, pages 845–875. Springer, Heidelberg, May / June 2022.
10. Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In Tatsuaki Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 531–545. Springer, Heidelberg, December 2000.
11. Daniel J. Bernstein. CAESAR: Competition for authenticated encryption: Security, applicability, and robustness. https://competitions.cr.yp.to/caesar.html, 2014.
12. Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. On the indifferentiability of the sponge construction. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 181–197. Springer, Heidelberg, April 2008.
13. Guido Bertoni, Joan Daemen, Michaël Peeters, and Gilles Van Assche. Sponge functions. In *ECRYPT Hash Workshop*, 2007.
14. Guido Bertoni, Joan Daemen, Michaël Peeters, Gilles Van Assche, and Ronny Van Keer. Keyak v2. Technical report, Submission to the CAESAR Competition, 2016. Available at https://keccak.team/files/Keyakv2-doc2.2.pdf.
15. Tim Beyne, Yu Long Chen, Christoph Dobraunig, and Bart Mennink. Dumbo, Jumbo, and Delirium: Parallel authenticated encryption for the lightweight circus. *IACR Trans. Symm. Cryptol.*, 2020(S1):5–30, 2020.
16. Tim Beyne, Yu Long Chen, Christoph Dobraunig, and Bart Mennink. Elephant. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
17. Dan Boneh and Victor Shoup. *A Graduate Course in Applied Cryptography*. 2023. Draft 0.6, http://toc.cryptobook.us/.
18. Avik Chakraborti, Nilanjan Datta, Mridul Nandi, and Kan Yasuda. Beetle family of lightweight and secure authenticated encryption ciphers. *IACR TCHES*, 2018(2):218–241, 2018. https://tches.iacr.org/index.php/TCHES/article/view/881.
19. Avik Chakraborti, Tetsu Iwata, Kazuhiko Minematsu, and Mridul Nandi. Blockcipher-based authenticated encryption: How small can we go? In Wieland Fischer and Naofumi Homma, editors, *CHES 2017*, volume 10529 of *LNCS*, pages 277–298. Springer, Heidelberg, September 2017.
20. John Chan and Phillip Rogaway. On committing authenticated-encryption. In Vijayalakshmi Atluri, Roberto Di Pietro, Christian Damsgaard Jensen, and Weizhi Meng, editors, *ESORICS 2022, Part II*, volume 13555 of *LNCS*, pages 275–294. Springer, Heidelberg, September 2022.
21. Cas Cremers, Samed Düzlü, Rune Fiedler, Marc Fischlin, and Christian Janson. BUFFing signature schemes beyond unforgeability and the case of post-quantum signatures. In *2021 IEEE Symposium on Security and Privacy*, pages 1696–1714. IEEE Computer Society Press, May 2021.

22. Joan Daemen, Seth Hoffert, Gilles Van Assche, and Ronny Van Keer. The design of Xoodoo and Xoofff. *IACR Trans. Symm. Cryptol.*, 2018(4):1–38, 2018.
23. Joan Daemen, Seth Hoffert, Michaël Peeters, Gilles Van Assche, Ronny Van Keer, and Silvia Mella. Xoodyak. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
24. Joan Daemen, Silvia Mella, and Gilles Van Assche. Committing authenticated encryption based on SHAKE. *IACR Cryptol. ePrint Arch.*, 2023:1494, 2023.
25. Jean Paul Degabriele, Marc Fischlin, and Jérôme Govinden. The indifferentiability of the duplex and its practical applications. In *ASIACRYPT 2023*, 2023.
26. Jean Paul Degabriele, Christian Janson, and Patrick Struck. Sponges resist leakage: The case of authenticated encryption. In Steven D. Galbraith and Shiho Moriai, editors, *ASIACRYPT 2019, Part II*, volume 11922 of *LNCS*, pages 209–240. Springer, Heidelberg, December 2019.
27. Daniel Dinu, Léo Perrin, Aleksei Udovenko, Vesselin Velichkov, Johann Großschädl, and Alex Biryukov. Design strategies for ARX with provable bounds: Sparx and LAX. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 484–513. Springer, Heidelberg, December 2016.
28. Christoph Dobraunig, Maria Eichlseder, Stefan Mangard, Florian Mendel, Bart Mennink, Robert Primas, and Thomas Unterluggauer. ISAP v2.0. *IACR Trans. Symm. Cryptol.*, 2020(S1):390–416, 2020.
29. Christoph Dobraunig, Maria Eichlseder, Stefan Mangard, Florian Mendel, Bart Mennink, Robert Primas, and Thomas Unterluggauer. ISAP. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
30. Christoph Dobraunig, Maria Eichlseder, Stefan Mangard, Florian Mendel, and Thomas Unterluggauer. ISAP – towards side-channel secure authenticated encryption. *IACR Trans. Symm. Cryptol.*, 2017(1):80–105, 2017.
31. Christoph Dobraunig, Maria Eichlseder, Florian Mendel, and Martin Schläffer. Ascon. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
32. Yevgeniy Dodis, Paul Grubbs, Thomas Ristenpart, and Joanne Woodage. Fast message franking: From invisible salamanders to encryptment. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part I*, volume 10991 of *LNCS*, pages 155–186. Springer, Heidelberg, August 2018.
33. Pooya Farshim, Benoît Libert, Kenneth G. Paterson, and Elizabeth A. Quaglia. Robust encryption, revisited. In Kaoru Kurosawa and Goichiro Hanaoka, editors, *PKC 2013*, volume 7778 of *LNCS*, pages 352–368. Springer, Heidelberg, February / March 2013.
34. Pooya Farshim, Claudio Orlandi, and Răzvan Roşie. Security of symmetric primitives under incorrect usage of keys. *IACR Trans. Symm. Cryptol.*, 2017(1):449–473, 2017.
35. Robert Granger, Philipp Jovanovic, Bart Mennink, and Samuel Neves. Improved masking for tweakable blockciphers with applications to authenticated encryption. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part I*, volume 9665 of *LNCS*, pages 263–293. Springer, Heidelberg, May 2016.
36. Jian Guo, Thomas Peyrin, and Axel Poschmann. The PHOTON family of lightweight hash functions. In Phillip Rogaway, editor, *CRYPTO 2011*, volume 6841 of *LNCS*, pages 222–239. Springer, Heidelberg, August 2011.

37. Martin Hell, Thomas Johansson, Willi Meier, Jonathan Sönnerup, Hirotaka Yoshida, and Alexander Maximov. Grain-128AEAD. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
38. Takanori Isobe and Mostafizar Rahman. Key committing security analysis of AEGIS. *IACR Cryptol. ePrint Arch.*, 2023:1495, 2023.
39. Tetsu Iwata, Mustafa Khairallah, Kazuhiko Minematsu, and Thomas Peyrin. Duel of the titans: The Romulus and Remus families of lightweight AEAD algorithms. *IACR Trans. Symm. Cryptol.*, 2020(1):43–120, 2020.
40. Tetsu Iwata, Mustafa Khairallah, Kazuhiko Minematsu, Thomas Peyrin, and Chun Guo. Romulus. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.
41. Philipp Jovanovic, Atul Luykx, and Bart Mennink. Beyond $2^{c/2}$ security in sponge-based authenticated encryption modes. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT 2014, Part I*, volume 8873 of *LNCS*, pages 85–104. Springer, Heidelberg, December 2014.
42. Charlotte Lefevre and Bart Mennink. Generic security of the Ascon mode: On the power of key blinding. *IACR Cryptol. ePrint Arch.*, 2023:796, 2023.
43. Julia Len, Paul Grubbs, and Thomas Ristenpart. Partitioning oracle attacks. In Michael Bailey and Rachel Greenstadt, editors, *USENIX Security 2021*, pages 195–212. USENIX Association, August 2021.
44. Moses Liskov, Ronald L. Rivest, and David Wagner. Tweakable block ciphers. In Moti Yung, editor, *CRYPTO 2002*, volume 2442 of *LNCS*, pages 31–46. Springer, Heidelberg, August 2002.
45. Sanketh Menda, Julia Len, Paul Grubbs, and Thomas Ristenpart. Context discovery and commitment attacks - how to break CCM, EAX, SIV, and more. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023, Part IV*, volume 14007 of *LNCS*, pages 379–407. Springer, Heidelberg, April 2023.
46. Bart Mennink, Reza Reyhanitabar, and Damian Vizár. Security of full-state keyed sponge and duplex: Applications to authenticated encryption. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015, Part II*, volume 9453 of *LNCS*, pages 465–489. Springer, Heidelberg, November / December 2015.
47. Chanathip Namprempre, Phillip Rogaway, and Thomas Shrimpton. Reconsidering generic composition. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 257–274. Springer, Heidelberg, May 2014.
48. National Institute of Standards and Technology. Lightweight cryptography standardization process. https://csrc.nist.gov/projects/lightweight-cryptography, 2015.
49. National Institute of Standards and Technology. Post-quantum cryptography standardization process. https://csrc.nist.gov/projects/post-quantum-cryptography, 2017.
50. National Institute of Standards and Technology. Call for additional digital signature schemes for the post-quantum cryptography standardization process. https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/call-for-proposals-dig-sig-sept-2022.pdf, 2022.
51. Eric Rescorla. The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446, August 2018.

52. Meltem Sönmez Turan, Kerry McKay, Donghoon Chang, Lawrence E. Bassham, Jinkeon Kang, Noah D. Waller, John M. Kelsey, and Deukjo Hong. Status report on the final round of the NIST lightweight cryptography standardization process. https://nvlpubs.nist.gov/nistpubs/ir/2023/NIST.IR.8454.pdf, 2023.
53. Serge Vaudenay. Security flaws induced by CBC padding - applications to SSL, IPSEC, WTLS... In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 534–546. Springer, Heidelberg, April / May 2002.
54. David Wagner. A generalized birthday problem. In Moti Yung, editor, *CRYPTO 2002*, volume 2442 of *LNCS*, pages 288–303. Springer, Heidelberg, August 2002.
55. Hongjun Wu and Tao Huang. The JAMBU lightweight authentication encryption mode (v2.1). Technical report, Submission to the CAESAR Competition, 2016. Available at https://competitions.cr.yp.to/round3/jambuv21.pdf.
56. Hongjun Wu and Tao Huang. TinyJambu. Technical report, National Institute of Standards and Technology, 2021. Available at https://csrc.nist.gov/projects/lightweight-cryptography/finalists.

A Additional Preliminaries

In Appendix A.1 we describe the used paddings and define more security notions. We give some background on (tweakable) block-ciphers and sponges in Appendix A.2 and Appendix A.3, respectively. Finally, in Appendix A.4, we provide some results that are relevant for our attacks and proofs.

A.1 Paddings and Security Notions

The authenticated encryption schemes considered in this work, use common paddings which we recall below. The one-zero padding $\mathtt{pad}_{10^*}(\cdot, r)$, appends a 1, followed by 0s until the desired length r is reached. Simply padding with 0s to length r is denoted by $\mathtt{pad}_{0^*}(\cdot, r)$. By $\mathtt{pad}_{\mathsf{L}}(\cdot, r)$, we denote the padding which appends 0, followed by appending the length of the input.

Below we define collision resistance of a hash function.

Definition 12. *Let $\mathcal{H}\colon \{0,1\}^* \to \{0,1\}^w$ be a hash function with output length w. For any adversary $\mathcal{A}$, its* CR *advantage is defined as*

$$\mathbf{Adv}^{\mathsf{CR}}_{\mathcal{H}}(\mathcal{A}) \coloneqq \Pr[\mathcal{H}(X_1) = \mathcal{H}(X_2) \wedge X_1 \neq X_2 \mid (X_1, X_2) \leftarrow \mathcal{A}()]\,.$$

Menda et al. [45] defined several variants of committing security. These variants require different parts of the contexts to disagree (and sometimes also others to agree). Below we recall their security notions.[25]

Definition 13. *Let* $\textsc{Ae} = (\textsc{Enc}, \textsc{Dec})$ *be an authenticated encryption scheme and the games $\mathsf{CMT}_{\mathsf{X}}$ and $\mathsf{CMT}^{\star}_{\mathsf{X}}$ for $\mathsf{X} \in \{\mathsf{K}, \mathsf{N}, \mathsf{A}\}$ be defined as in Fig. 1. For any adversary $\mathcal{A}$, its* $\mathrm{CMT}_{\mathsf{X}}$ *and* $\mathrm{CMT}^{\star}_{\mathsf{X}}$ *advantages are defined as*

$$\mathbf{Adv}^{\mathsf{CMT}_{\mathsf{X}}}_{\textsc{Ae}}(\mathcal{A}) \coloneqq \Pr[\mathsf{CMT}_{\mathsf{X}}(\mathcal{A}) \to 1]\,, \quad \mathbf{Adv}^{\mathsf{CMT}^{\star}_{\mathsf{X}}}_{\textsc{Ae}}(\mathcal{A}) \coloneqq \Pr[\mathsf{CMT}^{\star}_{\mathsf{X}}(\mathcal{A}) \to 1]\,.$$

Next, we define the advantage of finding colliding tags. At its core, this is a weakened version of committing security as the ciphertexts are not required to agree. We use this to bound the committing security of Ascon and Schwaemm.

Definition 14. *Let* $\textsc{Ae} = (\textsc{Enc}, \textsc{Dec})$ *be an authenticated encryption scheme and the game* $\mathsf{TagColl}$ *be defined as in Fig. 16. For any adversary $\mathcal{A}$, its* TagColl *advantage is defined as*

$$\mathbf{Adv}^{\mathsf{TagColl}}_{\textsc{Ae}}(\mathcal{A}) \coloneqq \Pr[\mathsf{TagColl}(\mathcal{A}) \to 1]\,.$$

[25] Note, however, that $\mathrm{CMT}_{\mathsf{K}}$ originates from [9].

Game CMT_K

$(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$
if $K = \overline{K}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Game $\mathsf{CMT}^\star_\mathsf{K}$

$((K, \overline{K}), N, A, (M, \overline{M})) \leftarrow \mathcal{A}()$
if $K = \overline{K}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, N, A, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Game CMT_N

$(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$
if $N = \overline{N}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Game $\mathsf{CMT}^\star_\mathsf{N}$

$(K, (N, \overline{N}), A, (M, \overline{M})) \leftarrow \mathcal{A}()$
if $N = \overline{N}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(K, \overline{N}, A, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Game CMT_A

$(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$
if $A = \overline{A}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Game $\mathsf{CMT}^\star_\mathsf{A}$

$(K, N, (A, \overline{A}), (M, \overline{M})) \leftarrow \mathcal{A}()$
if $A = \overline{A}$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(K, N, \overline{A}, \overline{M})$
return $((C, T) = (\overline{C}, \overline{T}))$

Fig. 15: Security games CMT_K, CMT_N, CMT_A, $\mathsf{CMT}^\star_\mathsf{K}$, $\mathsf{CMT}^\star_\mathsf{N}$, and $\mathsf{CMT}^\star_\mathsf{A}$ for authenticated encryption schemes. Here, $((K, \overline{K}), N, A, (M, \overline{M}))$ is an abbreviation for $(K, N, A, M), (\overline{K}, N, A, \overline{M})$, likewise used for the other context components.

Game $\mathsf{TagColl}$

$(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M}) \leftarrow \mathcal{A}()$
if $(K, N, A) = (\overline{K}, \overline{N}, \overline{A})$
  return 0
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{C}, \overline{T}) \leftarrow \textsc{Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M})$
return $(T = \overline{T})$

Fig. 16: Security Game $\mathsf{TagColl}$ for authenticated encryption schemes used in the proof of Theorem 10.

A.2 (Tweakable) Block-Ciphers

A block-cipher $\mathrm{BC}\colon \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$ takes as input a key K of length κ and a message M of length n, and outputs a ciphertext C of the same length as the message. For every $K \in \{0,1\}^\kappa$, $\mathrm{BC}(K,\cdot)$ is a permutation over $\{0,1\}^n$. A tweakable block-cipher [44] $\mathrm{TBC}\colon \{0,1\}^\kappa \times \mathcal{T} \times \{0,1\}^n \to \{0,1\}^n$ takes as input a key K of length κ, a tweak T (from some set of tweaks $\mathcal{T}$), and a message M of length n, and outputs a ciphertext C of the same length as the message. For every pair $(K,T) \in \{0,1\}^\kappa \times \mathcal{T}$, $\mathrm{TBC}(K,T,\cdot)$ is a permutation over $\{0,1\}^n$. We also use $\mathrm{TBC}^T(K,\cdot)$ as an alternative notation for $\mathrm{TBC}(K,T,\cdot)$.

For our results, we model the block-ciphers BC and tweakable block-ciphers TBC by an ideal cipher E and ideal tweakable cipher $\widetilde{\mathsf{E}}$, respectively.

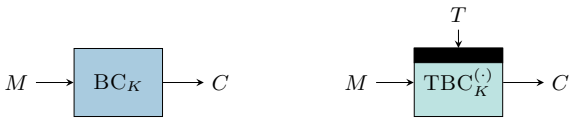

Fig. 17: Block-cipher (left) and tweakable block-cipher (right). For tweakable block-ciphers, the black bar indicates that the incoming arrow (T) is used as a tweak.

A.3 Sponges

Sponges [13] are a versatile tool for cryptographic primitives. Rather than just being relevant for cryptographic hash functions—as was their main design goal—they turned out to be more powerful as one can construct numerous cryptographic primitives from sponges.

The underlying component of a sponge is a permutation $\rho\colon \{0,1\}^n \to \{0,1\}^n$. Here, n is the size of the sponge state. The sponge operates in a round-wise fashion, where each round it absorbs a part of the input and applies ρ. The rate r describes how many bits of the input can be absorbed in each round by XORing them to the first r bits of the sponge state. The higher the rate the faster the sponge as fewer rounds, hence fewer invocations of ρ, are required to absorb the input. The part of the sponge state that is not affected by the input absorption is called the inner state and its size is denoted by the capacity c, thus we have $r + c = n$. The capacity is related to the security of the sponge, the higher the capacity the better the security of the sponge.

We refer to sponges of the form described above by *plain sponges* and provide an illustration in Fig. 18. It was shown that—especially in the context of AE schemes—one can also deploy *full-state sponges* and *duplex sponges*. The former XORs the input to the entire state, i.e., $r = n$ and $c = 0$. The latter absorbs and squeezes in each round, in contrast to the plain sponge which squeezes only after

the absorption is finished. XOODYAK uses a full-state sponge, while a duplex sponge is used, for instance, by ASCON.

Below we recall two results for the plain sponge construction that we will use later: First, a bound on the collision resistance of a simple sponge-based hash function and, second, the indifferentiability of sponges from a random function.

Theorem 15 ([17, Theorem 8.6]). *Let $\mathcal{H}$ be a hash function obtained from a permutation $\rho\colon \{0,1\}^n \to \{0,1\}^n$, with capacity c, rate r (so $n = r + c$), and output length $w \leq r$. For every adversary $\mathcal{A}$, if the number of ideal permutation queries plus the number of r-bit blocks in the output of $\mathcal{A}$ is bounded by q, it holds that*

$$\mathbf{Adv}^{\mathsf{CR}}_{H}(\mathcal{A}) \leq \frac{q(q-1)}{2^w} + \frac{q(q+1)}{2^c}\,.$$

Theorem 16 ([12, Theorem 2]). *Let $\mathcal{H}$ be a (padded) sponge construction obtained from a permutation $\rho\colon \{0,1\}^n \to \{0,1\}^n$, with capacity c and rate r (so $n = r + c$). Then, for any adversary $\mathcal{A}$, making significantly less than 2^c queries to ρ, $\mathcal{H}$ is indistinguishable from a random oracle F, except with probability at most $\frac{(1-2^{-256})q^2+(1+2^{-256})q}{2^{129}}$.*

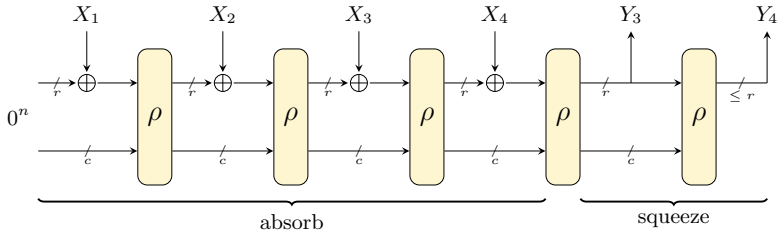

Fig. 18: Illustration of a plain sponge construction with four rounds of absorbing and two rounds of squeezing

A.4 Existing Results

The theorem below gives both upper and lower bounds on finding collisions for independent random variables.

Theorem 17 ([17, Theorem B.1]). *Let $\mathcal{M}$ be a set of size n and $X_1, \ldots, X_k$ be k independent random variables uniform in $\mathcal{M}$. Let C be the event that for some distinct $i, j \in \{1, \ldots, k\}$ we have that $X_i = X_j$. Then*

$$\Pr[\mathsf{C}] \geq 1 - \exp\left(\frac{-k(k-1)}{2n}\right) \geq \min\left\{\frac{-k(k-1)}{4n}, 0.63\right\} \text{ and}$$
$$\Pr[\mathsf{C}] \leq 1 - \exp\left(\frac{-k(k-1)}{n}\right) \text{ when } k < \frac{n}{2}\,.$$

We use the formulation of the birthday problem presented in [54], but provide a more formal description using Theorem 17.

Lemma 18. *Consider two lists L_1, L_2 of elements drawn uniformly and independently at random from $\{0,1\}^\tau$. We denote the size of L_1 and L_2 by l_1 and l_2, respectively, and define $l = l_1 + l_2$. Then one finds $x_k \in L_1$ and $x_j \in L_2$ such that $x_k \oplus x_j = 0$ with a probability of at least*

$$\left(1 - \exp\left(\frac{-l(l-1)}{2^{\tau+1}}\right)\right) \cdot \frac{2\,l_1 l_2}{l^2 - l}\,.$$

Proof. Consider the concatenation of the two lists $L = L_1 \parallel L_2$ and write $x_1, \ldots, x_l$ for its elements. Denote by C the event that $x_k = x_j$ holds for some $k \neq j$. Since the x_i are drawn uniformly and independently, Theorem 17 yields the bound $\Pr[\mathsf{C}] \geq 1 - \exp\left(\frac{-l(l-1)}{2^{\tau+1}}\right)$. However, this probability also counts internal collisions of L_1 and L_2, respectively. For two elements of L, the probability that they are not both from either L_1 or L_2 is $\frac{l_1}{l}\frac{l_2}{l-1} + \frac{l_2}{l}\frac{l_1}{l-1} = \frac{2\,l_1 l_2}{l^2 - l}$. Taking this into account, the probability that $x_k = x_j$ holds for some $x_k \in L_1$ and $x_j \in L_2$ is thus bounded above by $\left(1 - \exp\left(\frac{-l(l-1)}{2^{\tau+1}}\right)\right) \cdot \frac{2\,l_1 l_2}{l^2 - l}$. □

The following lemma contains two technical results needed for the committing security proof of Ascon (Theorem 10). While the computations are not hard, we give them here to provide a complete presentation.

Lemma 19. *Let $n, c \in \mathbb{N}$ such that $c \leq n$ and $IV \in \{0,1\}^c$. Let further ρ be a random permutation over $\{0,1\}^n$ and $\mathcal{A}$ be an adversary making queries to ρ. Consider the following events:*

1. *Event E_t (target hitting query):*
   $\mathcal{A}$ makes a query Y to ρ such that $\lfloor \rho(Y) \rfloor_{64} = IV$ or $\mathcal{A}$ makes a query S to ρ^{-1} such that $\lfloor \rho^{-1}(S) \rfloor_{64} = IV$.
2. *Event E_c (colliding queries):*
   $\mathcal{A}$ makes queries $Y \neq \overline{Y}$ to ρ such that $\lfloor \rho(Y) \rfloor_{128} = \lfloor \rho(\overline{Y}) \rfloor_{128}$ or $\mathcal{A}$ makes queries Y to ρ and $\overline{S}$ to ρ^{-1} such that $\lfloor \rho(Y) \rfloor_{128} = \lfloor \rho^{-1}(\overline{S}) \rfloor_{128}$.

If $\mathcal{A}$ makes $q \leq 2^{n-1}$ queries, then

$$\Pr[\mathsf{E}_t] \leq \frac{q}{2^{c-1}}\,.$$

and

$$\Pr[\mathsf{E}_c] \leq \frac{q(q-1)}{2^c}\,.$$

Proof. We start with the bound for E_t. Let X_i be the event that $\mathcal{A}$ triggers event E_t with its i-th query. It holds that

$$\Pr[\mathsf{E}_t] \leq \sum_{i=1}^{q} \Pr[\mathsf{X}_i] = \sum_{i=1}^{q} \frac{2^r}{2^n - i + 1} \leq \frac{2^r q}{2^n - q} \leq \frac{2^r q}{2^{n-1}} = \frac{q}{2^{c-1}}\,,$$

where $q \leq 2^{n-1}$ is used for the last inequality. Next, we bound event E_c. Let X_{ij} be the event that the j-th query by $\mathcal{A}$ forms a collision with the i-th query. Then it holds that

$$\Pr[\mathsf{E}_c] \leq \sum_{j=i}^{q} \sum_{i=1}^{j-1} \Pr[\mathsf{X}_{ij}] \leq \sum_{j=1}^{q} \frac{(j-1)2^r}{2^n - j + 1} \leq 2^r \left(\frac{q(q-1)}{2\left(2^n - q\right)} \right),$$

and using $q \leq 2^{n-1}$ again, we obtain

$$\Pr[\mathsf{E}_c] \leq 2^r \left(\frac{q(q-1)}{2^n} \right) = \frac{q(q-1)}{2^c},$$

which finishes the proof. □

B Deferred Proofs

B.1 Proof of Theorem 3 (Elephant)

Proof. We construct a CMT adversary $\mathcal{A}$ against Elephant as shown in Fig. 19. As a first step it samples a key K, a nonce N, associated data A, and a message M at random from the respective sets. It computes the ciphertext $C \leftarrow \textsc{Enc}_{\mathcal{M}}(K, N, M)$ and the tag $T \leftarrow \textsc{Enc}_{\mathcal{T}}(K, N, A, C)$. The ciphertext is parsed into blocks $C_1, \ldots, C_\gamma \xleftarrow{n} \mathtt{pad}_{10^*}(C, n)$. Next, the adversary samples a second, different key $\overline{K} \leftarrow\!\!\$\ \mathcal{K}\backslash\{K\}$ and associated data blocks $\overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}} \leftarrow\!\!\$\ \{0,1\}^n$.[26] The adversary then computes the state

$$\overline{S} \leftarrow \bigoplus_{i=2}^{\overline{\alpha}} (\widetilde{\mathsf{E}}(K, (i-1, 0), \overline{A}_i)) \oplus \bigoplus_{i=1}^{\gamma} (\widetilde{\mathsf{E}}(K, (i-1, 2), C_i)) \,,$$

shown in Fig. 5. The value $\overline{Y}$ is computed by querying $\widetilde{\mathsf{E}}^{-1}$ on $\overline{K}$, $(0,0)$, and T (padded with 0s to length n). Adversary $\mathcal{A}$ then computes $\overline{A}_1$ as the XOR of $\overline{S}$ and $\overline{Y}$. Together with the other associated data blocks, $\mathcal{A}$ computes $(\overline{N}, \overline{A}) \leftarrow \mathtt{pad}_{10^*}^{-1}(\overline{A}_1 \parallel \ldots \parallel \overline{A}_{\overline{\alpha}})$, i.e., removes the padding. It remains to compute the message $\overline{M}$ to which the ciphertext C decrypts under the context $(\overline{K}, \overline{N}, \overline{A})$. This can easily be achieved by setting $\overline{M} \leftarrow \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, C)$. Finally, $\mathcal{A}$ outputs $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. Observe that $\mathcal{A}$ wins the game CMT, as we have

$$\textsc{Elephant.Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M}) = (C, T) = \textsc{Elephant.Enc}(K, N, A, M) \,.$$

As for the queries to $\widetilde{\mathsf{E}}$, $\mathcal{A}$ makes μ queries to compute C and $\alpha + \gamma$ to compute T. Additionally, $\mathcal{A}$ makes μ queries to compute $\overline{M}$ and $\overline{\alpha} + \gamma$ queries to compute $\overline{S}$ and $\overline{Y}$, totalling up to $q = 2\mu + 2\gamma + \alpha + \overline{\alpha}$ queries. □

The gist of the attack, is finding a second associated data $\overline{A}$ which yields the target ciphertext. The attack easily extends to a context discovery attack ($\mathrm{CDY}^\star_A$) [45]. Hence, we can conclude that Elephant is also vulnerable with respect to the weaker security notions $\mathrm{CMT}_{\mathsf{K}}$ and $\mathrm{CMT}_{\mathsf{N}}$ by using [45, Corollar 3]. Furthermore, the attack can be translated to one against $\mathrm{CMT}_{\mathsf{A}}$ by observing that the adversary can choose $\overline{A}$ to differ from A at some point (note that $\mathcal{A}$ can freely choose all but one block). Finally, the attack is also extendable to the more restricted notion $\mathrm{CMT}^\star_{\mathsf{A}}$ by choosing the second key-nonce pair $(\overline{K}, \overline{N})$ not at random but equal to the first pair (K, N).[27]

[26] We assume that $\mathcal{A}$ chooses the last block to exhibit a valid padding.

[27] In this case, the adversary needs to target a different associated data block, which grants the freedom to choose the nonce $\overline{N}$.

ELEPHANT adversary $\mathcal{A}$

$K, N, A, M \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M}$
$C \leftarrow \text{ENC}_{\mathcal{M}}(K, N, M)$
$T \leftarrow \text{ENC}_{\mathcal{T}}(K, N, A, C)$
$C_1, \ldots, C_\gamma \xleftarrow{n} \mathtt{pad}_{10^*}(C, n)$
$\overline{K} \leftarrow\!\!\$\ \mathcal{K} \setminus \{K\}$
$\overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}} \leftarrow\!\!\$\ \{0,1\}^n$
$\overline{S} \leftarrow \mathcal{B}(C_1, \ldots, C_\gamma, \overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}})$
$\overline{Y} \leftarrow \widetilde{\mathsf{E}}^{-1}(\overline{K}, (0,0), \mathtt{pad}_{0^*}(T, n))$
$\overline{A}_1 \leftarrow \overline{Y} \oplus \overline{S}$
$(\overline{N}, \overline{A}) \leftarrow \mathtt{pad}_{10^*}^{-1}(\overline{A}_1, \ldots, \overline{A}_{\overline{\alpha}})$
$\overline{M} \leftarrow \text{ENC}_{\mathcal{M}}(\overline{K}, \overline{N}, C)$
return $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$

$\mathcal{B}(C_1, \ldots, C_\gamma, \overline{A}_2, \ldots, \overline{A}_{\overline{\alpha}})$

$\overline{S} \leftarrow 0^n$
for $i = 1, \ldots, \gamma$
    $\overline{S} \leftarrow \overline{S} \oplus \widetilde{\mathsf{E}}^{-1}(\overline{K}, (i-1, 2), C_i)$
for $i = 2, \ldots, \overline{\alpha}$
    $\overline{S} \leftarrow \overline{S} \oplus \widetilde{\mathsf{E}}^{-1}(\overline{K}, (i-1, 0), \overline{A}_i)$
return $\overline{S}$

Fig. 19: ELEPHANT adversary $\mathcal{A}$ from Theorem 3.

ELEPHANT.ENC(K, N, A, M)

$C \leftarrow \text{ENC}_{\mathcal{M}}(K, N, M)$
$T \leftarrow \text{ENC}_{\mathcal{T}}(K, N, A, C)$
return (C, T)

$\text{ENC}_{\mathcal{M}}(K, N, M)$

$M_1, \ldots, M_\mu \xleftarrow{n} \mathtt{pad}_{0^*}(M, n)$
for $i = 1, \ldots, \mu$
    $C_i \leftarrow M_i \oplus \text{TBC}^{(i-1,1)}(K, N)$
$C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
return C

$\text{ENC}_{\mathcal{T}}(K, N, A, C)$

$A_1, \ldots, A_\alpha \xleftarrow{n} \mathtt{pad}_{10^*}(N \parallel A, n)$
$C_1, \ldots, C_\gamma \xleftarrow{n} \mathtt{pad}_{10^*}(C, n)$
$T \leftarrow A_1$
for $i = 2, \ldots, \alpha$
    $T \leftarrow T \oplus \text{TBC}^{(i-1,0)}(K, A_i)$
for $i = 1, \ldots, \gamma$
    $T \leftarrow T \oplus \text{TBC}^{(i-1,2)}(K, C_i)$
$T \leftarrow \text{TBC}^{(0,0)}(K, T)$
return $\lceil T \rceil_\tau$

Fig. 20: Pseudocode of ELEPHANT [16] in terms of $\text{ENC}_{\mathcal{M}}$ and $\text{ENC}_{\mathcal{T}}$.

B.2 Proof of Theorem 4 (Romulus)

For the proof of Theorem 4, we formulate and prove two lemmas. Firstly, we show that the state-update-function ξ is invertible (Lemma 20). Secondly, we prove that we can invert both $\text{Enc}_{\mathcal{C}}$ and $\text{Enc}_{\mathcal{M}}$ (Lemma 21 and Lemma 22), where, for the latter, we make use of the invertibility of ξ.

Recall that the state-update-function ξ of Romulus maps a state S and an input I to a new state Y and an output O. In Romulus.Dec, the inverse of ξ is considered, however, inverse is understood only in relation to the output data. This means that the inverse function will *not* invert the state. When looking at Romulus, one can see that the output of ξ is discarded in $\text{Enc}_{\mathcal{C}}$—a fact that will be exploited later. For our attack against $\text{Enc}_{\mathcal{M}}$, this no longer works, as we have to invert the output of ξ while maintaining equal ciphertexts. The following lemma shows that we can invert ξ with respect to *both* its output and state. We write M for the input and C for the output of ξ (instead of I and O), which is the case for our scenario.

Lemma 20. *Let ξ be the state-update-function of* Romulus. *Let further $\mathcal{A}_\xi$ be the algorithm displayed in Fig. 21. Then for any $(Y, C) \in \{0,1\}^n \times \{0,1\}^n$, it holds that*

$$\xi(\mathcal{A}_\xi(Y, C)) = (Y, C)\,.$$

Proof. We first observe that the matrix G works block-wise on blocks of 8 bits. Since $\mathcal{A}_\xi$ splits the state and ciphertext into blocks of 8 bits each and runs $\mathcal{A}_G$ (cf. Fig. 21) on each of these blocks, we merely need to show that $\mathcal{A}_G$ correctly inverts an 8-bit block.

Let $(Y, C) \in \{0,1\}^8 \times \{0,1\}^8$ and $(S, M) \leftarrow \mathcal{A}_G(Y, C)$. For $i \in \{1, \ldots, 8\}$, denote the i-th bit of Y, C, S, and M by Y [i], C [i], S [i], and M [i], respectively. Algorithm $\mathcal{A}_G$ first computes $S[1] \leftarrow Y[8] \oplus C[8]$ and $M[1] \leftarrow Y[1] \oplus S[1]$. Subsequently, for $i \in \{2, \ldots, 8\}$, it computes $S[i] \leftarrow M[i-1] \oplus C[i-1]$ and $M[i] \leftarrow Y[i] \oplus S[i]$. Denote the output of ξ on input (S, M) by $(\overline{Y}, \overline{C})$. By construction we have:

$$\begin{aligned}
\overline{Y} &= M \oplus S \\
&= (M[1] \oplus S[1]) \parallel \ldots \parallel (M[8] \oplus S[8]) \\
&= (Y[1] \oplus S[1] \oplus S[1]) \parallel \ldots \parallel (Y[8] \oplus S[8] \oplus S[8]) \\
&= Y[1] \parallel \ldots \parallel Y[1] \\
&= Y
\end{aligned}$$

and

$$
\begin{aligned}
\overline{C} &= M \oplus G_s(S) \\
&= (M[1] \oplus S[2]) \parallel \ldots \parallel (M[7] \oplus S[8]) \parallel (M[8] \oplus S[8] \oplus S[1]) \\
&= (M[1] \oplus M[1] \oplus C[1]) \parallel \ldots \parallel (M[7] \oplus M[7] \oplus C[7]) \parallel (M[8] \oplus S[8] \oplus S[1]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel (M[8] \oplus S[8] \oplus S[1]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel (M[8] \oplus M[7] \oplus C[7] \oplus S[1]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel (M[8] \oplus M[7] \oplus C[7] \oplus Y[8] \oplus C[8]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel (Y[8] \oplus S[8] \oplus M[7] \oplus C[7] \oplus Y[8] \oplus C[8]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel (Y[8] \oplus M[7] \oplus C[7] \oplus M[7] \oplus C[7] \oplus Y[8] \oplus C[8]) \\
&= C[1] \parallel \ldots \parallel C[7] \parallel C[8] \\
&= C\,.
\end{aligned}
$$

Thus we obtain $\xi(\mathcal{A}_\xi(Y, C)) = \xi(S, M) = (\overline{Y}, \overline{C}) = (Y, C)$, which concludes the proof. □

Next, we give an adversary that inverts $\textsc{Enc}_{\mathcal{E}}$, i.e., for a given output S of $\textsc{Enc}_{\mathcal{E}}$ and a partial context (K, N), it finds matching associated data. We exploit the fact that the associated blocks are XORed to the full state in $\textsc{Enc}_{\mathcal{E}}$.

Lemma 21. *Consider* Romulus *described in Fig. 22. There exists an adversary $\mathcal{A}_{\mathcal{E}}$, making q queries to $\widetilde{\mathsf{E}}$ such that for any $(K, N, S) \in \mathcal{K} \times \mathcal{N} \times \{0,1\}^n$, it holds that*

$$\Pr[\textsc{Enc}_{\mathcal{E}}(K, N, A) = S \mid A \leftarrow \mathcal{A}_{\mathcal{E}}(K, N, S)] = 1\,.$$

The number of ideal tweakable cipher queries by $\mathcal{A}_{\mathcal{E}}$ is $q = \lfloor \frac{\alpha}{2} \rfloor + 1$ for α being the number of associated data blocks that $\mathcal{A}_{\mathcal{E}}$ outputs.

Proof. We construct $\mathcal{A}_{\mathcal{E}}$ as shown in Fig. 21. As input it receives (K, N, S). It chooses an arbitrary even number of associated data block α and chooses all except the first block at random, i.e., $A_2, \ldots, A_\alpha \leftarrow\!\!\$\ \{0,1\}^n$.[28] In addition, $\mathcal{A}$ sets $A_{\alpha+1} \leftarrow 0^n$.[29] Then $\mathcal{A}$ proceeds by inverting S using the ideal tweakable cipher to obtain Y. For $i \in \{1, \ldots, \frac{\alpha}{2}\}$, $\mathcal{A}$ first computes $S \leftarrow Y \oplus A_{2i+1}$ followed by computing $Y \leftarrow \widetilde{\mathsf{E}}^{-1}(K, A_{2i}, S)$, inverting the state-update-function ξ and the ideal tweakable cipher. Denote the resulting state by Y_* (see also the state marked in red in Fig. 6). Adversary $\mathcal{A}$ sets the first associated data as $A_1 \leftarrow Y_*$ which ensures that $\textsc{Enc}_{\mathcal{E}}(K, N, A) = S$.

The number of queries to the ideal tweakable cipher $\widetilde{\mathsf{E}}$ by $\mathcal{A}$ is $\lfloor \frac{\alpha}{2} \rfloor + 1$: the initial one plus $\lfloor \frac{\alpha}{2} \rfloor$ for the for loop. □

We now give an adversary that inverts $\textsc{Enc}_{\mathcal{M}}$, i.e., for a given ciphertext (C, T) and a partial input (K, N), it finds a matching pair of state S and message M. The attack relies heavily on the invertibility of ξ as shown in Lemma 20.

[28] We assume that these blocks are chosen to exhibit a valid padding.
[29] This corresponds to the input of the last application of ξ in the upper part of Fig. 6.

Lemma 22. *Consider* Romulus *as described in Fig. 22. There exists an adversary $\mathcal{A}_{\mathcal{M}}$, making q queries to $\widetilde{\mathsf{E}}$ such that for any $(K, N, (C, T)) \in \mathcal{K} \times \mathcal{N} \times \mathcal{C}$, it holds that*

$$\Pr[\textsc{Enc}_{\mathcal{M}}(K, N, S, M) = (C, T) \mid (S, M) \leftarrow \mathcal{A}_{\mathcal{M}}(K, N, (C, T))] = 1\,.$$

The number of ideal tweakable cipher queries by $\mathcal{A}_{\mathcal{M}}$ is $q = \mu$ for μ being the number of ciphertext blocks that $\mathcal{A}_{\mathcal{M}}$ receives as input.

Proof. We construct adversary $\mathcal{A}_{\mathcal{M}}$ as shown in Fig. 21. $\mathcal{A}_{\mathcal{M}}$ gets as input $(K, N, (C, T))$. For ease of exposition, we assume that the length of C is a multiple of the block size n.[30] Let $C_1, \ldots, C_\gamma \xleftarrow{n} C$. The adversary $\mathcal{A}_{\mathcal{M}}$ first sets $Y \leftarrow T$ and then computes $S \leftarrow G^{-1}(Y)$. For $i \in \{1, \ldots, \gamma\}$, $\mathcal{A}_{\mathcal{M}}$ computes $Y \leftarrow \widetilde{\mathsf{E}}^{-1}(K, N, S)$[31] followed by the computation of $(S, M_i) \leftarrow \mathcal{A}_{\xi}(Y, C_i)$ from Lemma 20. Finally, $\mathcal{A}_{\mathcal{M}}$ outputs (S, M), where $M = M_1 \parallel \ldots \parallel M_\gamma$. By construction, it holds that $\textsc{Enc}_{\mathcal{M}}(K, N, S, M) = (C, T)$ as $\mathcal{A}_{\mathcal{M}}$ inverted all invocations of $\widetilde{\mathsf{E}}$ and ξ during $\textsc{Enc}_{\mathcal{M}}$—using $\mathcal{A}_{\xi}$ from Lemma 20 for the latter.

$\mathcal{A}$ queries the ideal tweakable cipher $\widetilde{\mathsf{E}}$ a total of μ times. □

Having established Lemma 20, Lemma 21, and Lemma 22, we can now prove Theorem 4.

Proof (of Theorem 4). We construct the following adversary $\mathcal{A}$ against Romulus as shown in Fig. 21. It samples a context (K, N, A) together with a message M at random and computes the ciphertext $(C, T) \leftarrow \textsc{Romulus.Enc}(K, N, A, M)$. It then samples $(\overline{K}, \overline{N})$ at random, computes $(\overline{S}, \overline{M}) \leftarrow \mathcal{A}_{\mathcal{M}}(\overline{K}, \overline{N}, (C, T))$, and $\overline{A} \leftarrow \mathcal{A}_{\mathcal{C}}(\overline{K}, \overline{N}, \overline{S})$. Finally, $\mathcal{A}$ outputs $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. By using Lemma 21 and Lemma 22, we obtain

$$\begin{aligned}
\textsc{Romulus.Enc}(\overline{K}, \overline{N}, \overline{A}, \overline{M}) &= \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, \textsc{Enc}_{\mathcal{C}}(\overline{K}, \overline{N}, \overline{A}), \overline{M}) \\
&= \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, \overline{S}, \overline{M}) && \text{(Lemma 21)} \\
&= (C, T) && \text{(Lemma 22)} \\
&= \textsc{Romulus.Enc}(K, N, A, M)\,.
\end{aligned}$$

As for the number of queries to the ideal tweakable cipher $\widetilde{\mathsf{E}}$, $\mathcal{A}$ makes $\mu + \lfloor \frac{\alpha}{2} \rfloor + 1$ while computing the ciphertext (C, T) for the first tuple and additionally μ and $\lfloor \frac{\overline{\alpha}}{2} \rfloor + 1$ queries while running $\mathcal{A}_{\mathcal{M}}$ and $\mathcal{A}_{\mathcal{C}}$, respectively. This accumulates to $q = 2\mu + \lfloor \frac{\alpha}{2} \rfloor + \lfloor \frac{\overline{\alpha}}{2} \rfloor + 2$ queries in total and concludes the proof. □

The attack easily extends to $\mathrm{CMT}_{\mathsf{K}}$, $\mathrm{CMT}_{\mathsf{N}}$, $\mathrm{CMT}_{\mathsf{A}}$, and $\mathrm{CMT}^{\star}_{\mathsf{A}}$ attacks. The reasoning follows the one given for Elephant.

[30] This is justified by letting $\mathcal{A}$ choose a message satisfying this.
[31] Note that we drop the counter which is part of the tweak for simplicity.

ROMULUS adversary $\mathcal{A}()$

$(K, N, A, M) \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M}$
$(C, T) \leftarrow \text{ENC}(K, N, A, M)$
$(\overline{K}, \overline{N}) \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N}$
$(\overline{S}, \overline{M}) \leftarrow \mathcal{A}_{\mathcal{M}}(\overline{K}, \overline{N}, (C, T))$
$\overline{A} \leftarrow \mathcal{A}_{\mathcal{C}}(\overline{K}, \overline{N}, \overline{S})$
return $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$

$\mathcal{A}_{\xi}(Y, C)$

$Y_1, \ldots, Y_{\frac{n}{8}} \xleftarrow{8} S$
$C_1, \ldots, C_{\frac{n}{8}} \xleftarrow{8} C$
for $i = 1, \ldots, \frac{n}{8}$ **do**
    $(S_i, M_i) \leftarrow \mathcal{A}_G(Y_i, C_i)$
$S \leftarrow S_1 \parallel \ldots \parallel S_{\frac{n}{8}}$
$M \leftarrow M_1 \parallel \ldots \parallel M_{\frac{n}{8}}$
return (S, M)

$\mathcal{A}_G(Y, C)$

$Y[1], \ldots, Y[8] \xleftarrow{1} Y$
$C[1], \ldots, C[8] \xleftarrow{1} C$
$S[1] \leftarrow Y[8] \oplus C[8]$
$M[1] \leftarrow Y[1] \oplus S[1]$
for $i = 2, \ldots, 8$ **do**
    $S[i] \leftarrow M[i-1] \oplus C[i-1]$
    $M[i] \leftarrow Y[i] \oplus S[i]$
$S \leftarrow S_1 \parallel \ldots \parallel S_8$
$M \leftarrow M_1 \parallel \ldots \parallel M_8$
return (S, M)

$\text{ENC}_{\mathcal{C}}$ adversary $\mathcal{A}_{\mathcal{C}}(K, N, S)$

$\alpha \leftarrow\!\!\$\ 2\mathbb{N}$
$A_2, \ldots, A_{\alpha} \leftarrow\!\!\$\ \{0,1\}^n$
$A_{\alpha+1} \leftarrow 0^n$
$Y \leftarrow \widetilde{\mathsf{E}}^{-1}(K, N, S)$
for $i = \frac{\alpha}{2}, \ldots, 1$ **do**
    $S \leftarrow Y \oplus A_{2i+1}$
    $Y \leftarrow \widetilde{\mathsf{E}}^{-1}(K, A_{2i}, S)$
$A_1 \leftarrow Y$
$A \leftarrow A_1 \parallel \ldots \parallel A_{\alpha}$
return A

$\text{ENC}_{\mathcal{M}}$ adversary $\mathcal{A}_{\mathcal{M}}(K, N, (C, T))$

$C_1, \ldots, C_{\gamma} \xleftarrow{n} C$
$Y \leftarrow T$
$S \leftarrow G^{-1}(Y)$
for $i = \mu, \ldots, 1$ **do**
    $Y \leftarrow \widetilde{\mathsf{E}}^{-1}(K, N, S)$
    $(S, M_i) \leftarrow \mathcal{A}_{\xi}(Y, C_i)$
$M \leftarrow M_1 \parallel \ldots \parallel M_{\mu}$
return (S, M)

Fig. 21: ROMULUS adversary $\mathcal{A}$ from Theorem 4 and the state-update-function adversary $\mathcal{A}_{\xi}$ from Lemma 20.

ROMULUS.ENC(K, N, A, M)

1 : $S \leftarrow \text{ENC}_{\mathcal{E}}(K, N, A)$
2 : $(C, T) \leftarrow \text{ENC}_{\mathcal{M}}(K, N, S, M)$
3 : **return** (C, T)

ENC$_{\mathcal{E}}(K, N, A)$

4 : $A_1, \ldots, A_\alpha \xleftarrow{n} \texttt{pad}_{\text{L}}(A, n)$
5 : $S \leftarrow 0^n$
6 : **for** $i = 1, \ldots, \lfloor \frac{\alpha}{2} \rfloor$
7 : $\quad (Y, \cdot) \leftarrow \xi(S, A_{2i-1})$
8 : $\quad S \leftarrow \text{TBC}^{A_{2i}}(K, Y)$
9 : $V \leftarrow 0^n$
10 : **if** $\alpha \bmod 2 \neq 0$
11 : $\quad V \leftarrow A_\alpha$
12 : $(Y, \cdot) \leftarrow \xi(S, V)$
13 : $S \leftarrow \text{TBC}^N(K, Y)$
14 : **return** S

ENC$_{\mathcal{M}}(K, N, S, M)$

15 : $M_1, \ldots, M_\mu \xleftarrow{n} \texttt{pad}_{\text{L}}(M, n)$
16 : $S \leftarrow S$
17 : **for** $i = 1, \ldots, \mu - 1$
18 : $\quad (Y, C_i) \leftarrow \xi(S, M_i)$
19 : $\quad S \leftarrow \text{TBC}^N(K, Y)$
20 : $(Y, C_\mu) \leftarrow \xi(S, M_\mu)$
21 : $S \leftarrow \text{TBC}^N(K, Y)$
22 : $(\cdot, O) \leftarrow \xi(S, 0^n)$
23 : $T \leftarrow \lceil O \rceil_\tau$
24 : $C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
25 : **return** (C, T)

$\xi(S, I)$

26 : $Y \leftarrow S \oplus I$
27 : $O \leftarrow G(S) \oplus I$
28 : **return** (Y, O)

Fig. 22: Pseudocode of ROMULUS [40] in terms of ENC$_{\mathcal{E}}$ and ENC$_{\mathcal{M}}$. For sake of simplicity, we drop the counter that is part of the tweak.

B.3 Proof of Theorem 5 (GIFT-COFB)

For the proof of Theorem 5 we drop the XOR of the masking values. This avoids some very cumbersome and tedious bookkeeping. Subsequent to the proof, we discuss why the results also hold if the masking values are used.

Before giving the proof of Theorem 5, we give three lemmas that we will use to prove it. The first lemma, Lemma 23, shows that there is an algorithm that inverts the state-update-function of GIFT-COFB for random outputs with probability $\frac{1}{2}$. The second lemma, Lemma 24, shows that $\textsc{Enc}_{\mathcal{C}}$ can be inverted, i.e., given an arbitrary key K, a nonce N, and an output state S, there is an algorithm that outputs associated data A, such that $\textsc{Enc}_{\mathcal{C}}(K, N, A) = S$. The third lemma, Lemma 25, shows that there is an algorithm that inverts $\textsc{Enc}_{\mathcal{M}}$. The latter relies heavily on Lemma 23 which is why the success probability drops exponentially in the number of ciphertext blocks.

Lemma 23. *Let ξ be the state-update-function of* GIFT-COFB*. Let further $\mathcal{A}_\xi$ be the algorithm displayed in Fig. 23. Let C be an arbitrary bit string of length n. Then for $Y \leftarrow\!\!\$\ \{0,1\}^n$, it holds that*

$$\Pr[\xi(\mathcal{A}_\xi(Y, C)) = (Y, C)] = \frac{1}{2}\,.$$

Proof. We start by describing the algorithm $\mathcal{A}_\xi$ that is displayed in Fig. 23. We divide C and Y in two $\frac{n}{2}$-sized blocks, denoted by C_1, C_2 and Y_1, Y_2, respectively. Further, we denote the bits of $C_1 \oplus Y_1 \oplus C_2 \oplus Y_2$ by $z_1, \ldots, z_{\frac{n}{2}}$. Then, we randomly sample a bit s_1 and set $s_i \leftarrow s_{i-1} \oplus z_{i-1}$ for all $i = 2, \ldots, \frac{n}{2}$. By S_1 we denote the bit string resulting from concatenating the s_i's. Next, we define $S_2 \leftarrow C_1 \oplus Y_1 \oplus S_1$ and $S \leftarrow S_1 \parallel S_2$, which will be the first output of $\mathcal{A}_\xi$. Further, we set $M_1 \leftarrow Y_1 \oplus S_2$ and $M_2 \leftarrow Y_2 \oplus (S_1 \lll 1)$. The concatenated bit string $M \leftarrow M_1 \parallel M_2$ is the second output of $\mathcal{A}_\xi$, i.e., $\mathcal{A}_\xi(Y, C) = (S, M)$.

Next, we check that $\xi(\mathcal{A}_\xi(Y, C)) = (Y, C)$ holds with probability $\frac{1}{2}$. By definition of the state-update-function, the first component of $\xi(\mathcal{A}_\xi(Y, C)) = \xi(S, M)$ is given by

$$\begin{aligned}\widetilde{G}(S) \oplus M &= (S_2, S_1 \lll 1) \oplus M \\ &= (S_2 \oplus M_1, (S_1 \lll 1) \oplus M_2) \\ &= (Y_1, Y_2) = Y\,.\end{aligned}$$

The second component of $\xi(\mathcal{A}_\xi(Y, C)) = \xi(S, M)$ computes as

$$\begin{aligned}M \oplus S &= (M_1 \oplus S_1, M_2 \oplus S_2) \\ &= (Y_1 \oplus S_2 \oplus S_1, Y_2 \oplus (S_1 \lll 1) \oplus S_2)\,.\end{aligned}$$

Note that this expression is equal to C if and only if $S_1 \oplus (S_1 \lll 1) = C_1 \oplus Y_1 \oplus C_2 \oplus Y_2$, which results from solving $C_1 = Y_1 \oplus S_2 \oplus S_1$ for S_2 and plugging the result into $C_2 = Y_2 \oplus (S_1 \lll 1) \oplus S_2$. Breaking down this equation to the bit-level, yields

$$(s_1, s_2, \ldots, s_{\frac{n}{2}}) \oplus (s_2, s_3, \ldots, s_{\frac{n}{2}}, s_1) = (z_1, \ldots, z_{\frac{n}{2}})\,.$$

This gives the following equations

$$\begin{aligned} s_1 \oplus s_2 &= z_1 \\ s_2 \oplus s_3 &= z_2 \\ &\vdots \\ s_{\frac{n}{2}-1} \oplus s_{\frac{n}{2}} &= z_{\frac{n}{2}-1} \\ s_{\frac{n}{2}} \oplus s_1 &= z_{\frac{n}{2}} \end{aligned}$$

of which the first $\frac{n}{2} - 1$ equations hold by construction of the algorithm $\mathcal{A}_\xi$. Finally, we need to determine in which cases the last equation holds. For this, we replace $s_{\frac{n}{2}}$ by $s_{\frac{n}{2}-1} \oplus z_{\frac{n}{2}-1}$ which we get from the second to last equation. Then, in turn, we replace $s_{\frac{n}{2}-1}$ with $s_{\frac{n}{2}-2} \oplus z_{\frac{n}{2}-2}$ and continue this process until we get an equation depending only on the z_i's and s_1, namely

$$z_{\frac{n}{2}-1} \oplus \ldots z_2 \oplus z_1 \oplus s_1 \oplus s_1 = z_{\frac{n}{2}} \,,$$

which is equivalent to $z_{\frac{n}{2}} \oplus z_{\frac{n}{2}-1} \oplus \ldots z_2 \oplus z_1 = 0$. Thus we get $\xi(\mathcal{A}_\xi(Y, C)) = (\cdot, C)$ if and only if $C_1 \oplus Y_1 \oplus C_2 \oplus Y_2 = z_1 \parallel \ldots \parallel z_{\frac{n}{2}}$ contains an even number of 1s. Since Y is chosen uniformly at random, this is the case with probability $\frac{1}{2}$. This finishes the proof of the claim. □

Lemma 24. *Consider* GIFT-COFB *as described in Fig. 24. There exists an adversary $\mathcal{A}_\mathcal{C}$, making q queries to the ideal cipher E such that for any $(K, N, S) \in \mathcal{K} \times \mathcal{N} \times \{0,1\}^n$, it holds that*

$$\Pr[\textsc{Enc}_\mathcal{C}(K, N, A) = S \mid A \leftarrow \mathcal{A}_\mathcal{C}(K, N, S)] = 1 \,.$$

The number of ideal cipher queries by $\mathcal{A}_\mathcal{C}$ is $q = \alpha + 1$ for α being the number of associated data blocks that $\mathcal{A}_\mathcal{C}$ outputs.

Proof. We construct the following adversary $\mathcal{A}_\mathcal{C}$ against GIFT-COFB as shown in Fig. 23. Its input is (K, N, S). First, $\mathcal{A}_\mathcal{C}$ randomly picks associated data blocks $A_2, \ldots, A_\alpha$, i.e., all except the first one.[32] Next, $\mathcal{A}_\mathcal{C}$ computes both S_* and Y_* (cf. Fig. 7): For the former, $\mathcal{A}_\mathcal{C}$ computes $S_* \leftarrow \mathsf{E}(K, N)$ and for the latter, $\mathcal{A}_\mathcal{C}$ consecutively inverts the ideal cipher E (starting from the input S) and the state-update-function ξ (by XORing an associated data block and inverting $\widetilde{G}$—note that we merely need to invert $\widetilde{G}$ as the output of ξ is discarded). Finally, $\mathcal{A}_\mathcal{C}$ computes $A_1 \leftarrow S_* \oplus Y_*$ and outputs A. By construction, it holds that

$$\textsc{Enc}_\mathcal{C}(K, N, A) = S \,.$$

The number of ideal cipher queries by $\mathcal{A}_\mathcal{C}$ is $\alpha + 1$. □

[32] We assume that $\mathcal{A}_\mathcal{C}$ picks the blocks such that they exhibit a valid padding according to GIFT-COFB.

Lemma 25. *Consider* GIFT-COFB *described in Fig. 24. There exists an adversary* $\mathcal{A}_{\mathcal{M}}$*, making* q *queries to the ideal cipher* E *such that*

$$\Pr[\textsc{Enc}_{\mathcal{M}}(K,N,S,M) = (C,T) \mid (S,M) \leftarrow \mathcal{A}_{\mathcal{M}}(K,N,(C,T))] = \frac{1}{2^{\mu}}\,,$$

holds for any $(K,N,(C,T)) \in \mathcal{K} \times \mathcal{N} \times \mathcal{C}$*. The number of ideal cipher queries by* $\mathcal{A}_{\mathcal{M}}$ *is* $q = \mu$ *for* μ *being the number of ciphertext blocks that* $\mathcal{A}_{\mathcal{M}}$ *receives as input.*

Proof. We construct an adversary $\mathcal{A}_{\mathcal{M}}$ against GIFT-COFB as shown in Fig. 23. It gets $(K,N,(C,T))$ as input. For ease of exposition, we assume that the length of C is a multiple of the block size n, hence $C_1,\ldots,C_{\mu} \xleftarrow{n} C$ yields μ full blocks of length n. First, Adversary $\mathcal{A}_{\mathcal{M}}$ sets $S \leftarrow T$. Next, $\mathcal{A}_{\mathcal{M}}$ consecutively inverts the ideal cipher $Y \leftarrow \mathsf{E}^{-1}(K,S)$ and computes $(S,M_i) \leftarrow \mathcal{A}_{\xi}(Y,C_i)$. Finally, it sets $M \leftarrow M_1 \,\|\, \ldots \,\|\, M_{\mu}$ and outputs (S,M). Provided that $\mathcal{A}_{\xi}$ correctly inverts ξ, we obtain

$$\textsc{Enc}_{\mathcal{M}}(K,N,S,M) = (C,T)\,.$$

By Lemma 23—using Y is uniformly random due to E being an ideal cipher—every inversion of ξ succeeds with probability $\frac{1}{2}$. Since $\mathcal{A}_{\xi}$ is run μ times by $\mathcal{A}_{\mathcal{M}}$, we get

$$\Pr[\textsc{Enc}_{\mathcal{M}}(K,N,S,M) = (C,T) \mid (S,M) \leftarrow \mathcal{A}_{\mathcal{M}}(K,N,(C,T))] = \frac{1}{2^{\mu}}\,.$$

The number of queries to the ideal cipher E by $\mathcal{A}_{\mathcal{M}}$ is μ. □

We are now ready to give the proof of Theorem 5.

Proof (of Theorem 5). We construct $\mathcal{A}$ against GIFT-COFB as displayed in Fig. 23. It starts by sampling a context (K,N,A) together with a message M at random and computes $(C,T) \leftarrow \textsc{Gift-Cofb.Enc}(K,N,A,M)$. Next, it samples a fresh key-nonce pair $(\overline{K},\overline{N})$, computes $(\overline{S},\overline{M}) \leftarrow \mathcal{A}_{\mathcal{M}}(\overline{K},\overline{N},(C,T))$, and $\overline{A} \leftarrow \mathcal{A}_{\mathcal{C}}(\overline{K},\overline{N},\overline{S})$. Finally, $\mathcal{A}$ outputs $(K,N,A,M),(\overline{K},\overline{N},\overline{A},\overline{M})$.

By Lemma 24 and Lemma 25, we have the following equalities with probability $\frac{1}{2^{\mu}}$:

$$\begin{aligned}
\textsc{Gift-Cofb.Enc}(\overline{K},\overline{N},\overline{A},\overline{M}) &= \textsc{Enc}_{\mathcal{M}}(\overline{K},\overline{N},\textsc{Enc}_{\mathcal{C}}(\overline{K},\overline{N},\overline{A}),\overline{M}) \\
&= \textsc{Enc}_{\mathcal{M}}(\overline{K},\overline{N},\overline{S},\overline{M}) && \text{(Lemma 24)} \\
&= (C,T) && \text{(Lemma 25)} \\
&= \textsc{Gift-Cofb.Enc}(K,N,A,M)\,.
\end{aligned}$$

This yields

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Gift-Cofb}}(\mathcal{A}) = \frac{1}{2^{\mu}}\,.$$

Adversary $\mathcal{A}$ makes a total of $q = 2\mu + \alpha + \overline{\alpha} + 2$ queries to E: $\mu + \alpha + 1$ for computing the first tuple, μ for inverting $\textsc{Enc}_{\mathcal{M}}$, and $\overline{\alpha} + 1$ for inverting $\textsc{Enc}_{\mathcal{C}}$. □

When considering the masking values, inverting $\textsc{Enc}_{\mathcal{M}}$ might seem to be a problem, as the adversary needs to invert the function using the correct masking values. We observe that the masking values depend on the key, the nonce, and the length of associated data and message/ciphertext. In particular, they are independent of the exact values of A, M, and C. Thus, when inverting $\textsc{Enc}_{\mathcal{M}}$, the adversary merely has to choose how long the associated data will be, as this allows to use the correct masking values.

The attack easily extends to $\mathrm{CMT}_{\mathsf{K}}$, $\mathrm{CMT}_{\mathsf{N}}$, $\mathrm{CMT}_{\mathsf{A}}$, and $\mathrm{CMT}^{\star}_{\mathsf{A}}$ attacks, following the argument we gave for Elephant.

GIFT-COFB adversary $\mathcal{A}()$

$(K, N, A, M) \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N} \times \mathcal{A} \times \mathcal{M}$
$(C, T) \leftarrow \textsc{Enc}(K, N, A, M)$
$(\overline{K}, \overline{N}) \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N}$
$(\overline{S}, \overline{M}) \leftarrow \mathcal{A}_{\mathcal{M}}(\overline{K}, \overline{N}, (C, T))$
$\overline{A} \leftarrow \mathcal{A}_{\mathcal{E}}(\overline{K}, \overline{N}, \overline{S})$
return $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$

$\textsc{Enc}_{\mathcal{E}}$ adversary $\mathcal{A}_{\mathcal{E}}(K, N, S)$

$A_2, \ldots, A_\alpha \leftarrow\!\!\$\ \{0,1\}^n$
$S_* \leftarrow \mathsf{E}(K, N)$
$Y \leftarrow \mathsf{E}^{-1}(K, S)$
for $i = \alpha, \ldots, 2$
    $S \leftarrow \widetilde{G}^{-1}(Y \oplus A_i)$
    $Y \leftarrow \mathsf{E}^{-1}(K, S)$
$Y_* \leftarrow Y$
$A_1 \leftarrow \widetilde{G}(S_*) \oplus Y_*$
$A \leftarrow A_1 \parallel \ldots \parallel A_\alpha$
return A

$\textsc{Enc}_{\mathcal{M}}$ adversary $\mathcal{A}_{\mathcal{M}}(K, N, (C, T))$

$S \leftarrow T$
for $i = \mu, \ldots, 1$
    $Y \leftarrow \mathsf{E}^{-1}(K, S)$
    $(S, M_i) \leftarrow \mathcal{A}_{\xi}(Y, C_i)$
$M \leftarrow M_1 \parallel \ldots \parallel M_\mu$
return (S, M)

$\mathcal{A}_{\xi}(Y, C)$

$C_1, C_2 \xleftarrow{\frac{n}{2}} C$
$Y_1, Y_2 \xleftarrow{\frac{n}{2}} Y$
$z_1, \ldots, z_{\frac{n}{2}} \xleftarrow{1} C_1 \oplus Y_1 \oplus C_2 \oplus Y_2$
$s_1 \leftarrow\!\!\$\ \{0,1\}$
for $i = 2, \ldots, \frac{n}{2}$ **do**
    $s_i \leftarrow s_{i-1} \oplus z_{i-1}$
$S_1 \leftarrow s_1 \parallel \ldots \parallel s_{\frac{n}{2}}$
$S_2 \leftarrow C_1 \oplus Y_1 \oplus S_1$
$S \leftarrow S_1 \parallel S_2$
$M_1 \leftarrow Y_1 \oplus S_2$
$M_2 \leftarrow Y_2 \oplus (S_1 \lll 1)$
$M \leftarrow M_1 \parallel M_2$
return (S, M)

Fig. 23: GIFT-COFB adversary $\mathcal{A}$ from Theorem 5 and the state-update-function adversary $\mathcal{A}_{\xi}$ from Lemma 23.

GIFT-COFB.ENC(K, N, A, M)

1 : $(S, \Delta) \leftarrow \text{ENC}_{\mathcal{E}}(K, N, A)$
2 : $(C, T) \leftarrow \text{ENC}_{\mathcal{M}}(K, (S, \Delta), M)$
3 : **return** (C, T)

$\text{ENC}_{\mathcal{M}}(K, (S, \Delta), M)$

4 : $M_1, \ldots, M_\mu \xleftarrow{n} \texttt{pad}_{10^*}(M, n)$
5 : **for** $i = 1, \ldots, \mu - 1$
6 : $\quad \Delta \leftarrow 2\Delta$
7 : $\quad (Y, C_i) \leftarrow \xi(S, M_i)$
8 : $\quad Y \leftarrow Y \oplus (\Delta \,\|\, 0^{n/2})$
9 : $\quad S \leftarrow \text{BC}(K, Y)$
10 : **if** $|M| \bmod n = 0$
11 : $\quad \Delta \leftarrow 3\Delta$
12 : **if** $|M| \bmod n \neq 0$
13 : $\quad \Delta \leftarrow 3^2\Delta$
14 : $(Y, C_\mu) \leftarrow \xi(S, M_\mu)$
15 : $Y \leftarrow Y \oplus (\Delta \,\|\, 0^{n/2})$
16 : $S \leftarrow \text{BC}(K, Y)$
17 : $T \leftarrow \lceil S \rceil_\tau$
18 : $C \leftarrow \lceil C_1 \,\|\, \ldots \,\|\, C_\mu \rceil_{|M|}$
19 : **return** (C, T)

$\text{ENC}_{\mathcal{E}}(K, N, A)$

20 : $Y \leftarrow N$
21 : $S \leftarrow \text{BC}(K, Y)$
22 : $\Delta \leftarrow \lceil S \rceil_{n/2}$
23 : $A_1, \ldots, A_\alpha \xleftarrow{n} \texttt{pad}_{10^*}(A, n)$
24 : **for** $i = 1, \ldots, \alpha - 1$
25 : $\quad \Delta \leftarrow 2\Delta$
26 : $\quad (Y, \cdot) \leftarrow \xi(S, A)$
27 : $\quad Y \leftarrow Y \oplus (\Delta \,\|\, 0^{n/2})$
28 : $\quad S \leftarrow \text{BC}(K, Y)$
29 : **if** $|A| \bmod n = 0$
30 : $\quad \Delta \leftarrow 3\Delta$
31 : $\quad j \leftarrow 1$
32 : **if** $|A| \bmod n \neq 0$
33 : $\quad \Delta \leftarrow 3^2\Delta$
34 : $\quad j \leftarrow 2$
35 : $(Y, \cdot) \leftarrow \xi(S, A_\alpha)$
36 : $Y \leftarrow Y \oplus (\Delta \,\|\, 0^{n/2})$
37 : $S \leftarrow \text{BC}(K, Y)$
38 : **return** (S, Δ)

$\xi(S, I)$

39 : $Y \leftarrow \widetilde{G}(S) \oplus I$
40 : $O \leftarrow S \oplus I$
41 : **return** (Y, O)

Fig. 24: Pseudocode of GIFT-COFB [3] in term of $\text{ENC}_{\mathcal{E}}$ and $\text{ENC}_{\mathcal{M}}$.

B.4 Proof of Theorem 6 (PHOTON-BEETLE)

Proof. We construct the following CMT adversary $\mathcal{A}$ against PHOTON-BEETLE as shown in Fig. 25. It chooses (K, N, A) uniformly at random from the respective sets and computes $S \leftarrow \textsc{Enc}_{\mathcal{C}}(K, N, A)$. Let S_* denote the state before the domain separation is applied (see Fig. 8). Adversary $\mathcal{A}$ then chooses different associated data $\overline{A}$ and inverts $\textsc{Enc}_{\mathcal{C}}$ starting from S_* up to the initial state. This initial state is then used as the concatenation of nonce N and key K.[33] As the final step, $\mathcal{A}$ picks a message M at random and outputs $((K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, M))$. By construction, we have $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$ and it holds that

$$\begin{aligned}
\textsc{Photon-Beetle.Enc}(K, N, A, M) &= \textsc{Enc}_{\mathcal{M}}(\textsc{Enc}_{\mathcal{C}}(K, N, A), M) \\
&= \textsc{Enc}_{\mathcal{M}}(S_*, M) \\
&= \textsc{Enc}_{\mathcal{M}}(\textsc{Enc}_{\mathcal{C}}(\overline{K}, \overline{N}, \overline{A}), M) \\
&= \textsc{Photon-Beetle.Enc}(\overline{K}, \overline{N}, \overline{A}, M)\,,
\end{aligned}$$

thus $\mathcal{A}$ wins the game CMT.

$\mathcal{A}$ makes α queries to compute S_* and additional $\overline{\alpha}$ queries to compute $(\overline{K}, \overline{N})$, resulting in $q = \alpha + \overline{\alpha}$ queries in total. □

The attack is by construction also a valid CMT_{A} attack as the associated data are chosen to be different. While the adversary does not choose the key $\overline{K}$ and nonce $\overline{N}$ for the second tuple, it is clear that the attack easily extends to CMT_{K} and CMT_{N}—if $(\overline{K}, \overline{N}) = (K, N)$, the adversary simply chooses different associated data $\overline{A}$ and repeats the attack until the keys and nonces differ.

PHOTON-BEETLE adversary $\mathcal{A}$	$\mathcal{B}(S_*, \overline{A})$
$(K, N) \leftarrow\!\!\$\ \mathcal{K} \times \mathcal{N}$	$\overline{A}_1, \ldots, \overline{A}_{\overline{\alpha}} \xleftarrow{r} \mathtt{pad}_{10^*}(\overline{A}, r)$
$S \leftarrow N \parallel K$	**for** $i = \overline{\alpha}, \ldots, 1$ **do**
$A \leftarrow\!\!\$\ \mathcal{A}$	$\quad S_* \leftarrow S_* \oplus (\overline{A}_i \parallel 0^c)$
$A_1, \ldots, A_\alpha \xleftarrow{r} \mathtt{pad}_{10^*}(A, r)$	$\quad S_* \leftarrow \rho^{-1}(S_*)$
for $i = 1, \ldots, \alpha$ **do**	$\overline{N} \parallel \overline{K} \leftarrow S_*$
$\quad S \leftarrow \rho(S)$	**return** $(\overline{K}, \overline{N})$
$\quad S \leftarrow S \oplus (A_i \parallel 0^c)$	
$S_* \leftarrow S$	
$\overline{A} \leftarrow\!\!\$\ \mathcal{A} \backslash \{A\}$	
$(\overline{K}, \overline{N}) \leftarrow \mathcal{B}(S_*, \overline{A})$	
$M \leftarrow\!\!\$\ \mathcal{M}$	
return $((K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, M))$	

Fig. 25: PHOTON-BEETLE adversary $\mathcal{A}$ from Theorem 6.

[33] Note that these are likely to be different than K and N but not guaranteed to be.

PHOTON-BEETLE.ENC(K, N, A, M)

1 : $S \leftarrow \text{ENC}_{\mathcal{E}}(K, N, A)$
2 : $(C, T) \leftarrow \text{ENC}_{\mathcal{M}}(S, M)$
3 : **return** (C, T)

$\text{ENC}_{\mathcal{M}}(S, M)$

4 : $M_1, \ldots, M_\mu \xleftarrow{r} \mathtt{pad}_{10^*}(M, r)$
5 : **for** $i = 1, \ldots, \mu$
6 : $\quad S \leftarrow \rho(S)$
7 : $\quad (\lceil S \rceil_r, C_i) \leftarrow \xi(\lceil S \rceil_r, M_i)$
8 : $S \leftarrow S \oplus (0^r \parallel \iota_1)$
9 : $S \leftarrow \rho(S)$
10 : $T \leftarrow \lceil S \rceil_\tau$
11 : $C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
12 : **return** (C, T)

$\text{ENC}_{\mathcal{E}}(K, N, A)$

13 : $S \leftarrow N \parallel K$
14 : $A_1, \ldots, A_\alpha \xleftarrow{r} \mathtt{pad}_{10^*}(A, r)$
15 : **for** $i = 1, \ldots, \alpha$
16 : $\quad S \leftarrow \rho(S)$
17 : $\quad S \leftarrow S \oplus (A_i \parallel 0^c)$
18 : $S \leftarrow S \oplus (0^r \parallel \iota_0)$
19 : **return** S

$\xi(S, I)$

20 : $O \leftarrow \mathsf{Shuffle}(S) \oplus I$
21 : $Y \leftarrow S \oplus I$
22 : **return** (Y, O)

Fig. 26: Pseudocode of PHOTON-BEETLE [5] in terms of $\text{ENC}_{\mathcal{E}}$ and $\text{ENC}_{\mathcal{M}}$. Here, $\mathsf{Shuffle}(S) = S_2 \parallel (S_1 \ggg 1)$ for $S_1, S_2 \xleftarrow{\frac{r}{2}} S$.

B.5 Proof of Theorem 7 (XOODYAK)

Proof. We construct a CMT adversary $\mathcal{A}$ against XOODYAK. It uses a birthday attack to find a collision in the last 32 bits of the sponge state after the first application of ρ. For this, $q = 2^{17} + 1$ different keys and q nonces are sampled randomly. For $i \in [q]$, we denote them by K_i and N_i and write $\boldsymbol{K}_i \parallel \boldsymbol{N}_i = \mathtt{pad}_{\mathcal{C}}((K_i \parallel N_i \parallel \mathtt{enc}_{\mathcal{S}}(N_i)), 00000010)$ for their padded concatenation. Further, we consider the following random function

$$f : \{0,1\}^{384} \to \{0,1\}^{32}, \quad f(X) = \lfloor \rho(X) \rfloor_{32} ,$$

and compute $f(\boldsymbol{K}_i \parallel \boldsymbol{N}_i)$ for each $i \in [q]$. By the birthday attack [17, Section 8.3][34], a collision of f is found with probability at least $\frac{1}{2}$. Assume that such a collision has been found and write (K, N) and $(\overline{K}, \overline{N})$ for the key-nonce pairs that lead to it, i.e., have $\lfloor \rho(\boldsymbol{K} \parallel \boldsymbol{N}) \rfloor_{32} = \lfloor \rho(\overline{\boldsymbol{K}} \parallel \overline{\boldsymbol{N}}) \rfloor_{32}$, where $\boldsymbol{K} \parallel \boldsymbol{N}$ and $\overline{\boldsymbol{K}} \parallel \overline{\boldsymbol{N}}$ denote the corresponding padded values. Next, adversary $\mathcal{A}$ picks $A \in \{0,1\}^{352}$ at random and computes $\overline{A} \leftarrow \lceil \rho(\overline{\boldsymbol{K}} \parallel \overline{\boldsymbol{N}}) \rceil_{352} \oplus \lceil \rho(\boldsymbol{K} \parallel \boldsymbol{N}) \rceil_{352} \oplus A$. Together with the collision on the last 32 bits, $\mathcal{A}$ has produced a collision on the whole state. Then, the adversary wins the game CMT against XOODYAK by outputting (K, N, A, M) and $(\overline{K}, \overline{N}, \overline{A}, M)$ for M some randomly sampled message. This is the case, as $(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$ and the states after the associated data is absorbed agree, after which point only the same input (namely M) is processed for both tuples. Hence, we obtain $\textsc{Xoodyak.Enc}(K, N, A, M) = \textsc{Xoodyak.Enc}(\overline{K}, \overline{N}, \overline{A}, M)$ and in total have shown that $\mathcal{A}$ wins with probability at least $\frac{1}{2}$ for $q = 2^{17} + 1$ queries. □

Note that, using the same strategy as presented above, we obtain an attacker that wins with probability 1 after making $2^{32} + 1$ queries. Further, the above attack is by construction also a valid $\mathsf{CMT}_{\mathsf{K}}$ and $\mathsf{CMT}_{\mathsf{N}}$ attack as keys and nonces, respectively, are chosen to be different. Moreover, it can be shown to be a $\mathsf{CMT}_{\mathsf{A}}$ attack: The same associated data blocks are only chosen if the states after the first application of ρ already coincide in their rate part. Since they agree in the last 32 bits by construction, this would constitute a full-state collision of ρ, which is impossible for a permutation.

[34] Note that the prerequisite, regarding the size of domain and codomain of f, is fulfilled as $2^{384} \geq 100 \cdot 2^{32}$.

XOODYAK.ENC(K, N, A, M)

$S \leftarrow \text{ENC}_{\mathcal{E}}(K, N, A)$
$(C, T) \leftarrow \text{ENC}_{\mathcal{M}}(S, M)$
return (C, T)

$\text{ENC}_{\mathcal{E}}(K, N, A)$

$S \leftarrow 0^n$
$X \leftarrow K \parallel N \parallel \texttt{enc}_8(|N|)$
$Y \leftarrow S \oplus \texttt{pad}_{\mathcal{E}}(X, 0^6 \parallel 10)$
$S \leftarrow \rho(Y)$
$A_1, \ldots, A_\alpha \xleftarrow{352} A$
$Y \leftarrow S \oplus \texttt{pad}_{\mathcal{E}}(A_1, 0^6 \parallel 11)$
for $i = 2, \ldots, \alpha$
    $S \leftarrow \rho(Y)$
    $Y \leftarrow S \oplus \texttt{pad}_{\mathcal{E}}(A_i, 0^8)$
$S \leftarrow Y$
return S

$\text{ENC}_{\mathcal{M}}(S, M)$

$Y \leftarrow S \oplus (0^{r+8} \parallel \iota_1)$
$M_1, \ldots, M_\mu \xleftarrow{r} M$
for $i = 1, \ldots, \mu$
    $\boldsymbol{M}_i \leftarrow \texttt{pad}_{\mathcal{M}}(M_i)$
    $S \leftarrow \rho(Y)$
    $Y \leftarrow S \oplus (\boldsymbol{M}_i \parallel 0^{c-8})$
    $C_i \leftarrow \lceil Y \rceil_r$
$Y \leftarrow Y \oplus (0^{r+8} \parallel \iota_2)$
$S \leftarrow \rho(Y)$
$T \leftarrow \lceil S \rceil_\tau$
$C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
return (C, T)

Fig. 27: Pseudocode of XOODYAK [23] in terms of $\text{ENC}_{\mathcal{E}}$ and $\text{ENC}_{\mathcal{M}}$.

B.6 Proof of Theorem 8 (TinyJambu)

Proof. We construct a CMT adversary $\mathcal{A}$ against TinyJambu as follows: First, we randomly choose two different keys $K \neq \overline{K}$ and a target ciphertext C. Note that, due to the structure of TinyJambu, the context produces a key stream which is XORed with the message to obtain the ciphertext. Hence, for a random context it is always possible to find a M such that the TinyJambu encryption results in the initially chosen target ciphertext C. In the following, we will implicitly consider this "matching" message for each context that occurs. Hence it suffices to find two different contexts that—together with their matching message—yield colliding tags.

We start by building two lists of tags, where for one we use K as key and in the other $\overline{K}$. For this, sample pairwise different (N_i, A_i) for $i \in \{1, \ldots, 2^{32}+1\}$ and compute the corresponding tags T_i using the key K. We then set $L = (T_i)_{i\in[2^{32}+1]}$. Analogously, we sample pairwise different $(\overline{N}_i, \overline{A}_i)$[35] for $i \in [2^{32}+1]$ and write the corresponding tags $\overline{T}_i$ (computed using $\overline{K}$) into the list $\bar{L} = (\overline{T}_i)_{i\in[2^{32}+1]}$. Building the two lists, takes a total of $q = 2(2^{32}+1)(6+\alpha+\mu)$ queries to ρ.

For the context (K, N_i, A_i), denote the states before the second to last permutation application (see Fig. 10) by S_i, and analogously for $(\overline{K}, \overline{N}_i, \overline{A}_i)$ by $\overline{S}_i$. Note that for a fixed key, TinyJambu can be considered a sponge-based function with rate 32 and capacity 96. Therefore, the event $S_i = S_j$ for $i \neq j$ (and analogously $\overline{S}_i \neq \overline{S}_j$ for $i \neq j$), constitutes an inner collision, which is—for a sponge with capacity 96—highly unlikely[36]. As we model both BC_1 and BC_2 by an ideal cipher E and the states S_i (and respectively $\overline{S}_i$) collide with negligible probability, we can assume the list elements to be distributed uniformly and independently.

This puts us in the situation of Lemma 18 (for $l_1 = l_2 = 2^{32}+1$ and $\tau = 64$), hence we obtain the following lower bound for finding a collision $T_i = \overline{T}_j$:

$$\left(1 - \exp\left(\frac{-(2^{33}+2)(2^{33}+1)}{2^{65}}\right)\right) \cdot \frac{2\cdot(2^{32}+1)^2}{(2^{33}+2)^2 - (2^{33}+2)} \tag{1}$$

Since

$$\frac{-(2^{33}+2)(2^{33}+1)}{2^{65}} \leq \frac{-2^{33}\cdot 2^{33}}{2^{65}} = -2\,,$$

the first factor in Eq. (1) can be bounded below by $1 - \exp\left(\frac{-(2^{33}+2)(2^{33}+1)}{2^{65}}\right) \geq 1 - e^{-2} \geq \frac{3}{4}$. The second factor in Eq. (1) simplifies to $\frac{2^{32}+1}{2^{33}+1}$ which is lower bound by $\frac{1}{2}$. In total, the probability for finding a tag collision (and hence winning the game CMT) is at least $\frac{3}{8}$. □

[35] For sake of simplicity, we assume that $\mathcal{A}$ chooses all associated data to have the same number of blocks α.

[36] More precisely, the probability is $\frac{q(q+1)}{2^{97}} - \frac{q(q-1)}{2^{129}}$ [13].

The attack exploits the fact that TinyJambu uses a very short tag (64 bits) compared to the other schemes—the only other scheme with a 64-bit tag is Elephant, though they also provide a parameter set with a larger tag. Increasing the tag length of TinyJambu would render our attack impractical. Note, however, that increasing the tag length to 128 does not make TinyJambu committing secure. For such a variant of TinyJambu, we can similarly apply a birthday attack to find a collision on the capacity part, while the associated data is processed. Such a 96-bit collision can be found with about 2^{48} queries and, by properly choosing the associated data, results in a full collision. One can modify the parameters such that a 127-bit collision has to be found—though this variant is impractical as the inputs would have to be processed bit by bit.

By construction, the above attack is a $\mathrm{CMT}_{\mathsf{K}}$ attack, as K and $\overline{K}$ were chosen to be different. Moreover, by requiring not only the tuples (N_i, A_i) to differ for all i, but the individual nonces and associated data, we also obtain a $\mathrm{CMT}_{\mathsf{N}}$ and a $\mathrm{CMT}_{\mathsf{A}}$ attack.

$\textsc{TinyJambu.Enc}(K, N, A, M)$

$S \leftarrow \textsc{Enc}_{\mathcal{E}}(K, N, A)$
$(C, T) \leftarrow \textsc{Enc}_{\mathcal{M}}(K, S, M)$
return (C, T)

$\textsc{Enc}_{\mathcal{E}}(K, N, A)$

$N_1, N_2, N_3 \xleftarrow{32} N$
$A_1, \ldots, A_\alpha \xleftarrow{32} \mathtt{pad}_{0^*}(A, 32)$
$S \leftarrow 0^{128}$
$S \leftarrow \mathrm{BC}_2(K, S)$
for $i = 1, \ldots, 3$
  $S \leftarrow S \oplus (0^{64} \,\|\, \iota_N)$
  $S \leftarrow \mathrm{BC}_1(K, S)$
  $S \leftarrow (\lceil S \rceil_{32} \oplus N_i) \,\|\, \lfloor S \rfloor_{96}$
for $i = 1, \ldots, \alpha$
  $S \leftarrow S \oplus (0^{64} \,\|\, \iota_A)$
  $S \leftarrow \mathrm{BC}_1(K, S)$
  $S \leftarrow (\lceil S \rceil_{32} \oplus A_i) \,\|\, \lfloor S \rfloor_{96}$
return S

$\textsc{Enc}_{\mathcal{M}}(K, S, M)$

$M_1, \ldots, M_\mu \xleftarrow{32} \mathtt{pad}_{0^*}(M, 32)$
for $i = 1, \ldots, \mu$
  $S \leftarrow S \oplus (0^{64} \,\|\, \iota_M)$
  $S \leftarrow \mathrm{BC}_2(K, S)$
  $S \leftarrow (\lceil S \rceil_{32} \oplus M_i) \,\|\, \lfloor S \rfloor_{96}$
  $C_i \leftarrow [S]_{33}^{64}$
$S \leftarrow S \oplus (0^{64} \,\|\, \iota_T)$
$S \leftarrow \mathrm{BC}_2(K, S)$
$T_l \leftarrow [S]_{33}^{64}$
$S \leftarrow S \oplus (0^{64} \,\|\, \iota_T)$
$S \leftarrow \mathrm{BC}_1(K, S)$
$T_r \leftarrow [S]_{33}^{64}$
$C \leftarrow \lceil C_1 \,\|\, \ldots \,\|\, C_\mu \rceil_{|M|}$
$T \leftarrow T_l \,\|\, T_r$
return (C, T)

Fig. 28: Pseudocode of TinyJambu [56] in terms of $\textsc{Enc}_{\mathcal{E}}$ and $\textsc{Enc}_{\mathcal{M}}$. If the last block of associated data or message is not of full length, TinyJambu XORs the respective lengths into the last bits (as part of ι_A and ι_M).

B.7 Proof of Theorem 9 (Isap)

Proof (of Theorem 9). Let $\mathcal{A}$ be a CMT adversary against Isap with output denoted by $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. Further note that IV denotes the initialization vector used in Isap. We assume that $\mathcal{A}$ makes queries to ρ_1 and ρ_2 that correspond to its output, i.e., querying all states that occur during the evaluation of Isap for the two output tuples of $\mathcal{A}$. This assumption is without loss of generality, as we can easily transform any adversary into one that runs $\mathcal{A}$ to obtain $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$ and—before outputting the same—makes all queries to ρ corresponding to $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$.

The authentication component $\textsc{Enc}_{\mathcal{T}}$ uses a session key denoted by K_A (resp. $\overline{K}_A$), which results from an application of Isap.Rk to the key K (resp. $\overline{K}$) and the intermediate state X (resp. $\overline{X}$) computed during $\textsc{Enc}_{\mathcal{T}}$. We consider the event E that $(N, A) = (\overline{N}, \overline{A})$ and $K_A = \overline{K}_A$. Using this, the CMT advantage can be divided up as follows

$$\begin{aligned}\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Isap}}(\mathcal{A}) &= \Pr[\mathsf{CMT}(\mathcal{A}) \to 1] \\ &= \Pr[\mathsf{E} \wedge \mathsf{CMT}(\mathcal{A}) \to 1] + \Pr[\neg\mathsf{E} \wedge \mathsf{CMT}(\mathcal{A}) \to 1]\,.\end{aligned}$$

We start by giving an upper bound for the second summand. For this, we construct a CR (see Definition 12) adversary $\mathcal{B}$ against a sponge hash function $\mathcal{H}_2$ obtained from the permutation ρ_2 with rate $\overline{r}_2 = \max\{\kappa, r_2 + 1\}$[37], capacity $\overline{c}_2 = n - \overline{r}_2$, and output length κ. Further $0^\kappa \parallel IV$ is chosen as the initial state of $\mathcal{H}_2$. First, $\mathcal{B}$ runs $\mathcal{A}$, which outputs $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. For every query that $\mathcal{A}$ makes to ρ_2, the adversary $\mathcal{B}$ makes the same query to its own permutation and sends the response back to $\mathcal{A}$. Further, Adversary $\mathcal{B}$ simulates ρ_1 for $\mathcal{A}$. Using this, $\mathcal{B}$ computes for both output tuples of $\mathcal{A}$, the state in $\textsc{Enc}_{\mathcal{T}}$ after the associated data and the ciphertext blocks are absorbed. We denote the states for the first tuple and second tuple by X and $\overline{X}$, respectively and the session keys for $\textsc{Enc}_{\mathcal{T}}$ by K_A and $\overline{K}_A$, respectively (cf. Fig. 29). The states obtained after XORing these together are denoted by $Z = X \oplus K_A$ and $\overline{Z} = \overline{X} \oplus \overline{K}_A$.[38]

Let $A_1, \ldots, A_\alpha \xleftarrow{r_2} \mathtt{pad}_{10^*}(A, r_2)$ and $\overline{A}_1, \ldots, \overline{A}_{\overline{\alpha}} \xleftarrow{r_2} \mathtt{pad}_{10^*}(\overline{A}, r_2)$ be the division of A and $\overline{A}$ into blocks of length r_2. Analogously, the ciphertext $C = \textsc{Enc}_{\mathcal{M}}(K, N, M) = \textsc{Enc}_{\mathcal{M}}(\overline{K}, \overline{N}, \overline{M})$ is parsed as $C_1, \ldots, C_\gamma \xleftarrow{r_2} \mathtt{pad}_{10^*}(C, r_2)$. The adversary $\mathcal{B}$ then outputs

$$\begin{aligned}O &= (N \parallel 0^*, A_1 \parallel 0^*, \ldots, A_\alpha \parallel 0^*, C_1 \parallel 1 \parallel 0^*, \ldots, C_\gamma \parallel 0^*, Z \parallel 0^*) \\ \overline{O} &= (\overline{N} \parallel 0^*, \overline{A}_1 \parallel 0^*, \ldots, \overline{A}_{\overline{\alpha}} \parallel 0^*, C_1 \parallel 1 \parallel 0^*, \ldots, C_\gamma \parallel 0^*, \overline{Z} \parallel 0^*)\,,\end{aligned}$$

[37] The definition of the rate over the maximum ensures that the argument works for both Isap variants (Isap-K and Isap-A). More precisely, it guarantees that all inputs can still be fully absorbed after the adjustment of the rate.

[38] Note that $\mathcal{B}$ can compute these values by looking up the queries and responses from $\mathcal{A}$'s queries—using the assumption that it makes permutation queries corresponding to its output. Thus, this step does not require any additional permutation queries by $\mathcal{B}$.

where $\| 0^*$ denotes the padding with 0s up to length $\overline{r}$. A visualization for this is provided in Fig. 29. We show that if $\mathcal{A}$ wins the game CMT against Isap and the event $\neg\mathsf{E}$ holds, then the constructed adversary $\mathcal{B}$ wins the game CR against $\mathcal{H}_2$. Note that $\mathcal{A}$ winning the game CMT implies that $(K,N,A) \neq (\overline{K},\overline{N},\overline{A})$ and $\textsc{Enc}_{\mathcal{T}}(K,N,A,C) = T = \textsc{Enc}_{\mathcal{T}}(\overline{K},\overline{N},\overline{A},C)$. Hence, the output tuples of $\mathcal{B}$ are mapped to the same result under $\mathcal{H}_2$ (namely T) and it is only left to check that $O \neq \overline{O}$ to guarantee a collision. As event $\neg\mathsf{E}$ holds, we know that $(N,A) \neq (\overline{N},\overline{A})$ or $K_A \neq \overline{K}_A$. In the case that $(N,A) \neq (\overline{N},\overline{A})$, we have $O \neq \overline{O}$. Hence, we can assume from now on that $(N,A) = (\overline{N},\overline{A})$. Next, consider the case that $K_A \neq \overline{K}_A$. As $(N,A) = (\overline{N},\overline{A})$, we know that $X = \overline{X}$ and hence $Z = \lceil X \rceil_\kappa \oplus K_A \neq \lceil \overline{X} \rceil_\kappa \oplus \overline{K}_A = \overline{Z}$. Thus, $O \neq \overline{O}$ is also given in this case. We have shown that

$$\Pr[\neg\mathsf{E} \wedge \mathsf{CMT}(\mathcal{A}) \to 1] \leq \Pr[\mathsf{CR}(\mathcal{B}) \to 1] \leq \frac{q_2(q_2-1)}{2^\kappa} + \frac{q_2(q_2+1)}{2^{n-\overline{r}_2}}\,,$$

where the last inequality holds by Theorem 15, which bounds the probability of finding a collision in a general sponge hash function. Here, we exploit the fact that $\mathcal{B}$ makes the same number of queries to ρ_2 as $\mathcal{A}$.

Next, we give a bound for the first summand $\Pr[\mathsf{E} \wedge \mathsf{CMT}(\mathcal{A}) \to 1]$. For this, construct a CR adversary $\mathcal{C}$ against a sponge hash function $\mathcal{H}_1$ obtained from the permutation ρ_1 with rate $\overline{r}_1 = \kappa$, capacity $\overline{c}_1 = n - \overline{r}_2$ and output length κ. Further its initial state is given by $0^\kappa \| IV$. The adversary $\mathcal{C}$ starts by running $\mathcal{A}$, which outputs $(K,N,A,M),(\overline{K},\overline{N},\overline{A},\overline{M})$. For every query that $\mathcal{A}$ makes to ρ_1, the adversary $\mathcal{C}$ makes the same query to its own permutation and sends the response back to $\mathcal{A}$. Further, adversary $\mathcal{C}$ simulates ρ_2 for $\mathcal{A}$. Then, it computes X and $\overline{X}$ (analoguously to adversary $\mathcal{B}$), and we write $X_1,\ldots,X_\kappa \xleftarrow{1} X$ and $\overline{X}_1,\ldots,\overline{X}_\kappa \xleftarrow{1} \overline{X}$. Lastly, adversary $\mathcal{C}$ outputs $(K, X_1 \| 0^{\kappa-1},\ldots,X_\kappa \| 0^{\kappa-1})$ and $(\overline{K},\overline{X}_1 \| 0^{\kappa-1},\ldots,\overline{X}_\kappa \| 0^{\kappa-1})$. Next, we show that if $\mathcal{A}$ wins the game CMT against Isap and event E holds, then the constructed adversary $\mathcal{C}$ wins the game CR against $\mathcal{H}_1$. First observe that $\mathcal{A}$ winning the game CMT, implies that $(K,N,A) \neq (\overline{K},\overline{N},\overline{A})$, and $\textsc{Isap.Enc}(K,N,A,M) = \textsc{Isap.Enc}(\overline{K},\overline{N},\overline{A},\overline{M})$. If at the same time event E holds, i.e., $(N,A) = (\overline{N},\overline{A})$ and $K_A = \overline{K}_A$ hold, then $K \neq \overline{K}$, as otherwise $\mathcal{A}$ would not be a valid CMT adversary. Further note that $X = \overline{X}$ as $(N,A) = (\overline{N},\overline{A})$. Hence, $\mathcal{C}$ wins the game CR, because the tuples he outputs are different, but their image under $\mathcal{H}_1$ agrees (as $K_A = \overline{K}_A$). This implies that

$$\Pr[\mathsf{E} \wedge \mathsf{CMT}(\mathcal{A}) \to 1] \leq \Pr[\mathsf{CR}(\mathcal{C}) \to 1] \leq \frac{q_1(q_1-1)}{2^\kappa} + \frac{q_1(q_1+1)}{2^{n-\kappa}}\,,$$

where the last inequality holds by Theorem 15. Using $\overline{r}_2 = \max\{\kappa, r_2+1\}$, we obtain in total

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Isap}}(\mathcal{A}) \leq \frac{q_1(q_1-1)}{2^\kappa} + \frac{q_1(q_1+1)}{2^{n-\kappa}} + \frac{q_2(q_2-1)}{2^\kappa} + \frac{q_2(q_2+1)}{2^{n-\max\{\kappa,r_2+1\}}}\,,$$

which finishes the proof of the theorem. □

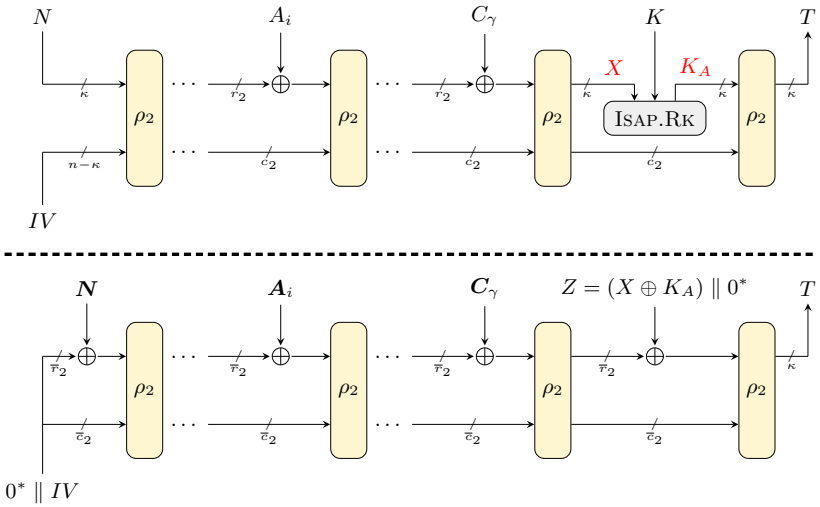

Fig. 29: Illustration of a proof step for ISAP (Theorem 9).. Here, $\overline{r}_2 = \max\{\kappa, r_H + 1\}$ (i.e., 129 for ISAP-A and 145 for ISAP-K) and $\overline{c}_2 = n - \overline{r}_2$; further write $\boldsymbol{N} = N \parallel 0^*$, $\boldsymbol{A}_i = A_i \parallel 0^*$, $\boldsymbol{C}_1 = C_1 \parallel 1 \parallel 0^*$ and $\boldsymbol{C}_i = C_i \parallel 0^*$ for $i \in \{2, \ldots, \gamma\}$.

The dominant term in the bound from Theorem 9 is $\frac{q_1(q_1-1)}{2^\kappa} + \frac{q_2(q_2-1)}{2^\kappa}$, thus by increasing κ (i.e., the tag and key length), we can increase the committing security. Note however, that—for ISAP-A—we can only increase κ up to 160 as for larger values the other term $\frac{q_1(q_1+1)}{2^{n-\kappa}} + \frac{q_2(q_2+1)}{2^{n-\max\{\kappa, r_2+1\}}}$ becomes the dominant term. This would result in about 80-bit committing security. A similar argument applies for ISAP-K, which deploys KECCAK-P as the underlying permutation. For this variant, we have $n = 400$ which allows to increase κ up to 200, allowing for about 100-bit committing security.

ISAP.ENC(K, N, A, M)

$C \leftarrow \text{ENC}_{\mathcal{M}}(K, N, M)$
$T \leftarrow \text{ENC}_{\mathcal{T}}(K, N, A, C)$
return (C, T)

$\text{ENC}_{\mathcal{T}}(K, N, A, C)$

$A_1, \ldots, A_\alpha \xleftarrow{r_2} \texttt{pad}_{10^*}(A, r_2)$
$C_1, \ldots, C_\gamma \xleftarrow{r_2} \texttt{pad}_{10^*}(C, r_2)$
$Y \leftarrow N \parallel IV$
$S \leftarrow \rho_H(Y)$
for $i = 1, \ldots, \alpha$
    $Y \leftarrow S \oplus (A_i \parallel 0^{c_2})$
    $S \leftarrow \rho_H(Y)$
$S \leftarrow S \oplus 0^{n-1} \parallel 1$
for $i = 1, \ldots, \gamma$
    $Y \leftarrow S \oplus (C_i \parallel 0^{c_2})$
    $S \leftarrow \rho_H(Y)$
$K_A \leftarrow \text{ISAP.RK}(K, \lceil S \rceil_\kappa)$
$S \leftarrow \rho_H(K_A, \lceil S \rceil_\kappa)$
$T \leftarrow \lceil S \rceil_\tau$
return T

$\text{ENC}_{\mathcal{M}}(K, N, M)$

$M_1, \ldots, M_\mu \xleftarrow{r_2} \texttt{pad}_{0^*}(M, r_2)$
$K_E \leftarrow \text{ISAP.RK}(K, N)$
$S \leftarrow K_E \parallel N$
for $i = 1, \ldots, \mu$
    $S \leftarrow \rho_E(S)$
    $C_i \leftarrow \lceil S \rceil_{r_2} \oplus M_i$
$C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
return C

ISAP.RK(K, X)

$X_1, \ldots, X_z \xleftarrow{r_1} X$
$Y \leftarrow K \parallel IV$
$S \leftarrow \rho_K(Y)$
for $i = 1, \ldots, z-1$
    $Y \leftarrow S \oplus (X_i \parallel 0^{n-r_1})$
    $S \leftarrow \rho_B(Y)$
$Y \leftarrow S \oplus (X_z \parallel 0^{n-r_1})$
$S \leftarrow \rho_K(Y)$
return $\lceil S \rceil_\kappa$

Fig. 30: Pseudocode of ISAP [29] in terms of $\text{ENC}_{\mathcal{M}}$ and $\text{ENC}_{\mathcal{T}}$.

B.8 Proof of Theorem 10 (Ascon)

Proof (of Theorem 10). Let $\mathcal{A}$ be a CMT adversary against Ascon with output denoted by $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. Further note that IV denotes the initialization vector used in Ascon. As a first step, we observe that finding different inputs to the Ascon encryption that give the same ciphertext is easy due to the duplex construction used in the sponge. The difficulty in breaking CMT security for Ascon lies in finding a tag collision, which is why we focus our attention on this task. An adversary that wins the game CMT against Ascon, in particular finds a tag collision, i.e., it wins the game TagColl (see Fig. 16), which allows the following reduction step

$$\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Ascon}}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{TagColl}}_{\textsc{Ascon}}(\mathcal{A})\,.$$

To bound the advantage on the right, we adapt the proof of Theorem 15 to fit our particular needs.

We build a directed graph G from the ideal permutation queries the adversary makes, in the following way: The nodes in G are the 2^{320} bit strings of length 320 and an edge from Y to S is added if $\mathcal{A}$ makes a query of the form $\rho(Y) = S$ or $\rho^{-1}(S) = Y$ (the graph starts with no edges). The edges resulting from ρ queries are called *forward edges* and the ones resulting from ρ^{-1} queries are referred to as *backward edges*.

We assume $\mathcal{A}$ to make queries to ρ that correspond to its output, i.e., querying all states that occur during the evaluation of Ascon for the two output tuples of $\mathcal{A}$. This assumption is without loss of generality, as we can easily transform any adversary into one that runs $\mathcal{A}$ to obtain $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$ and—before outputting the same—makes all queries to ρ corresponding to (K, N, A, M) and $(\overline{K}, \overline{N}, \overline{A}, \overline{M})$. Additionally, we assume $\mathcal{A}$ to make no redundant queries, i.e., once two values (Y, S) are known to be connected via an edge, no further ρ queries are made on Y and no further ρ^{-1} queries are made on S.

For this graph, we define a special kind of path, the *A-path*, which models an Ascon evaluation. An A-path P_{A} of length l[39] is a sequence of $2l$ nodes

$$Y_0, S_1, Y_1, S_2, \ldots, S_{l-1}, Y_{l-1}, S_l$$

with

1. $Y_0 = K \parallel N \parallel IV$,
2. $\lfloor Y_1 \rfloor_{256} = \lfloor S_1 \rfloor_{256} \oplus (K \parallel 0^{128})$ and $\lfloor Y_{l-1} \rfloor_{256} = \lfloor S_{l-1} \rfloor_{256} \oplus (K \parallel 0^{128})$,
3. $\lfloor Y_i \rfloor_{256} = \lfloor S_i \rfloor_{256}$ for all $i \in \{2, \ldots, l-2\}$, and
4. G contains edges from Y_{i-1} to S_i for all $i \in \{1, \ldots, l\}$,

for some $K, N \in \{0,1\}^{128}$. We define the *input* of an A-path P_{A} as $I := (K, N, X_1, \ldots, X_{l-1}) \in \{0,1\}^{128} \times \{0,1\}^{128} \times \{0,1\}^{64} \cdots \times \{0,1\}^{64}$ for $K \parallel N :=$

[39] Note that $l \geq 3$, as Ascon involves at least three applications of ρ (in case both associated data and message consist of a single block).

$\lceil Y_0 \rceil_{256}$ and $X_i = \lceil S_i \rceil_{64} \oplus \lceil Y_i \rceil_{64}$ for all $i \in \{1, \dots l-1\}$. The *result* of P_{A} is defined as $R = \lfloor S_l \rfloor_{128} \oplus K$. By construction, this models the tag generation of Ascon for key K, nonce N, and the tuple of associated data and message $(A, M) = (X_1, \dots, X_{l-1})$. As a notation for A-paths that incorporates the input, we write

$$(K \parallel N) | Y_0 \to \cdots \to S_{l-2} | X_{l-2} | Y_{l-2} \to S_{l-1} | X_{l-1} | Y_{l-1} \to S_l \,.^{40}$$

Next, we define two properties a pair $(P_{\text{A}}, \overline{P}_{\text{A}})$ of A-paths can have. For this, denote the nodes in $\overline{P}_{\text{A}}$ by $\overline{Y}_0, \overline{S}_1, \overline{Y}_1, \dots, \overline{Y}_{\bar{l}-1}, \overline{S}_{\bar{l}}$. Firstly, the paths P_{A} and $\overline{P}_{\text{A}}$ are *colliding* if their inputs differ but their results agree. Secondly, the paths P_{A} and $\overline{P}_{\text{A}}$ are *problematic* if their inputs differ and

1. $Y_{l-1} = \overline{Y}_{\bar{l}-1}$ or
2. at least one of the edges in P or $\overline{P}$ is a backward edge.

We are interested in the event CP that $\mathcal{A}$ finds a pair of colliding paths. Note that finding such paths means that $\mathcal{A}$ wins the game $\mathsf{TagColl}$. In order to compute the probability of CP, we define the auxiliary event PP_{A} that $\mathcal{A}$ finds a pair of problematic A-paths. Using this, we obtain

$$\mathbf{Adv}^{\mathsf{TagColl}}_{\text{Ascon}}(\mathcal{A}) = \Pr[\mathsf{CP}] \leq \Pr[\mathsf{CP} \wedge \neg \mathsf{PP}_{\text{A}}] + \Pr[\mathsf{PP}_{\text{A}}] \,,$$

and proceed by deriving upper bounds for both of the above summands.

We start with the easier case, which is giving an upper bound for the probability that $\mathsf{CP} \wedge \neg \mathsf{PP}_{\text{A}}$ holds, i.e., that $\mathcal{A}$ finds a pair of colliding A-paths that is not problematic. Hence, $\mathcal{A}$ finds two different inputs $I = (K, N, X_1, \dots, X_{l-1})$ and $\overline{I} = (\overline{K}, \overline{N}, \overline{X}_1, \dots, \overline{X}_{l-1})$ such that the corresponding A-paths

$$\begin{aligned} &Y_0, S_1, Y_1, S_2, \dots, S_{l-1}, Y_{l-1}, S_l \\ &\overline{Y}_0, \overline{S}_1, \overline{Y}_1, \overline{S}_2, \dots, \overline{S}_{\bar{l}-1}, \overline{Y}_{\bar{l}-1}, \overline{S}_{\bar{l}} \end{aligned}$$

fulfill $Y_{l-1} \neq \overline{Y}_{\bar{l}-1}$ and have equal results, i.e., $\lfloor \rho(Y_{l-1}) \rfloor_{128} \oplus K = R = \overline{R} = \left\lfloor \rho(\overline{Y}_{\bar{l}-1}) \right\rfloor_{128} \oplus \overline{K}$. By definition of an A-path, this implies

$$\left\lfloor \rho(S_{l-1} \oplus (X_{l-1} \| K \| 0^{128})) \right\rfloor_{128} \oplus K = \left\lfloor \rho(\overline{S}_{\bar{l}-1} \oplus (\overline{X}_{\bar{l}-1} \| \overline{K} \| 0^{128})) \right\rfloor_{128} \oplus \overline{K} \,. \tag{2}$$

Since ρ is a random permutation and $\mathcal{A}$ only used forward queries (as $\neg \mathsf{PP}_{\text{A}}$ holds), finding such a collision is unlikely. We assume—to the benefit of the adversary $\mathcal{A}$—that it can choose $S_{l-1}, \overline{S}_{\bar{l}-1} \in \{0,1\}^{128}$ freely, i.e., it must not be part of an A-path for some input. The probability of $\mathcal{A}$ finding (S_{l-1}, X_{l-1}, K),

[40] While not visible in this representation, by definition of A-paths, Y_1 and S_1 (respectively Y_{l-1} and S_{l-1}) differ not only in their first 64 bits but also from bit 65 to 192, where the key is XORed.

$(\overline{S}_{\bar{l}-1}, \overline{X}_{\bar{l}-1}, \overline{K})$[41] such that $Y_{l-1} \neq \overline{Y}_{\bar{l}-1}$ and Eq. (2) holds with q queries, equals the probability of finding a collision in a list of q uniformly distributed elements. Using Theorem 17 for $q \leq 2^{127}$, the latter can be bounded from above by

$$1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right).$$

Next, we turn our attention to deriving an upper bound for $\Pr[\mathsf{PP}_{\text{A}}]$. We will see that finding problematic paths is hard, even for a plain sponge without Ascon's blinding mechanisms, which is why we reduce to this setting. We consider a sponge-based hash function $\mathcal{H}$ obtained from the permutation ρ with rate 256 for the first round of absorption and rate 196 for all remaining ones. Further, its initial state is given by $0^{256} \parallel IV$ and the output produced by $\mathcal{H}$ has length 128. Analogously to A-paths, we also model evaluations of $\mathcal{H}$ by paths in the directed graph G. For $s \geq 1$, a *PS-path* of length s is a sequence of $2s$ nodes

$$Y_0, S_1, Y_1, S_2, \ldots, S_{s-1}, Y_{s-1}, S_s$$

with

1. $\lfloor Y_0 \rfloor_{64} = IV$,
2. $\lfloor Y_i \rfloor_{128} = \lfloor S_i \rfloor_{128}$ for all $i \in \{1, \ldots, s-1\}$, and
3. G contains edges from Y_{i-1} to S_i for all $i \in \{1, \ldots, s\}$.

We define the *input* of a PS-path as $I := (Z_0, \ldots, Z_{s-1}) \in \{0,1\}^{256} \times \{0,1\}^{192} \times \cdots \times \{0,1\}^{192}$ for $Z_0 = \lceil Y_0 \rceil_{256}$ and $Z_i = \lceil S_i \rceil_{192} \oplus \lceil Y_i \rceil_{192}$ for all $i \in \{1, \ldots, s-1\}$. As a notation for PS-paths that incorporates the input, we write

$$Z_0 | Y_0 \to \cdots \to S_{s-2} | Z_{s-2} | Y_{s-2} \to S_{s-1} | Z_{s-1} | Y_{s-1} \to S_s \,.$$

The notion of problematic A-paths can be directly transferred to PS-paths. The event that $\mathcal{A}$ finds a pair of problematic PS-paths is denoted by PP_{PS}.

We next observe that a pair of problematic A-paths, can also be considered as a pair of problematic PS-paths, i.e., in particular the event PP_{A} implies the event PP_{PS}. Let $(P_{\text{A}}, \overline{P}_{\text{A}})$ be a pair of problematic A-paths, i.e.,

$$\begin{aligned} P_{\text{A}} &= (K \parallel N) | Y_0 \to \cdots \to S_{l-2} | X_{l-2} | Y_{l-2} \to S_{l-1} | X_{l-1} | Y_{l-1} \to S_l \\ \overline{P}_{\text{A}} &= (\overline{K} \parallel \overline{N}) | \overline{Y}_0 \to \cdots \to \overline{S}_{\bar{l}-2} | \overline{X}_{\bar{l}-2} | \overline{Y}_{\bar{l}-2} \to \overline{S}_{\bar{l}-1} | \overline{X}_{\bar{l}-1} | \overline{Y}_{\bar{l}-1} \to \overline{S}_{\bar{l}} \,. \end{aligned}$$

By defining

$$\begin{aligned} Z_0 &= K \parallel N \in \{0,1\}^{256} & \overline{Z}_0 &= \overline{K} \parallel \overline{N} \in \{0,1\}^{256} \\ Z_1 &= X_1 \parallel K \in \{0,1\}^{192} & \overline{Z}_1 &= \overline{X}_1 \parallel \overline{K} \in \{0,1\}^{192} \\ Z_i &= X_i \parallel 0^{128} \in \{0,1\}^{192} & \overline{Z}_i &= \overline{X}_i \parallel 0^{128} \in \{0,1\}^{192} \\ Z_{l-1} &= X_{l-1} \parallel K \in \{0,1\}^{192} & \overline{Z}_{\bar{l}-1} &= \overline{X}_{\bar{l}-1} \parallel \overline{K} \in \{0,1\}^{192} \end{aligned}$$

[41] Note that, $\mathcal{A}$ must choose $(S_{l-1}, X_{l-1}, K) \neq (\overline{S}_{\bar{l}-1}, \overline{X}_{\bar{l}-1}, \overline{K})$ to ensure $Y_{l-1} \neq \overline{Y}_{\bar{l}-1}$.

we obtain the following presentation of $(P_{\text{A}}, \overline{P}_{\text{A}})$ as PS-paths:

$$\begin{aligned} P_{\text{PS}} =& Z_0|Y_0 \to \cdots \to S_{l-2}|Z_{l-2}|Y_{l-2} \to S_{l-1}|Z_{l-1}|Y_{l-1} \to S_l \\ \overline{P}_{\text{PS}} =& \overline{Z}_0|\overline{Y}_0 \to \cdots \to \overline{S}_{\bar{l}-2}|\overline{Z}_{\bar{l}-2}|\overline{Y}_{\bar{l}-2} \to \overline{S}_{\bar{l}-1}|\overline{Z}_{\bar{l}-1}|\overline{Y}_{\bar{l}-1} \to \overline{S}_{\bar{l}}\,. \end{aligned}$$

Visualization for this is provided in Fig. 31. As we neither change $(Y_{l-1}, \overline{Y}_{\bar{l}-1})$ nor any of the edges, the paths $(P_{\text{PS}}, \overline{P}_{\text{PS}})$ form a pair of problematic PS-paths. Thus, we have shown that PP_{A} implies PP_{PS}, hence $\Pr[\mathsf{PP}_{\text{A}}] \leq \Pr[\mathsf{PP}_{\text{PS}}]$. This allows us to focus on the plain sponge setting for the rest of the proof. More precisely, we show that it is hard to find a pair of problematic PS-paths, i.e., we derive an upper bound for $\Pr[\mathsf{PP}_{\text{PS}}]$.

For this, we define the following auxiliary events:

1. Event E_t (target hitting query):
   $\mathcal{A}$ makes a query Y to ρ such that $\lfloor \rho(Y) \rfloor_{64} = IV$ or $\mathcal{A}$ makes a query S to ρ^{-1} such that $\lfloor \rho^{-1}(S) \rfloor_{64} = IV$.
2. Event E_c (colliding queries):
   $\mathcal{A}$ makes queries $Y \neq \overline{Y}$ to ρ such that $\lfloor \rho(Y) \rfloor_{128} = \lfloor \rho(\overline{Y}) \rfloor_{128}$ or $\mathcal{A}$ makes queries Y to ρ and $\overline{S}$ to ρ^{-1} such that $\lfloor \rho(Y) \rfloor_{128} = \lfloor \rho^{-1}(\overline{S}) \rfloor_{128}$.

Next, we show that if $\mathcal{A}$ triggers PP_{PS}, then it triggers one of the events defined above.

For the proof assume that PP_{PS} holds and denote the problematic paths $\mathcal{A}$ finds by $(P_{\text{PS}}, \overline{P}_{\text{PS}})$. We first consider the case that there is at least one backward edge in $(P_{\text{PS}}, \overline{P}_{\text{PS}})$. We assume w.l.o.g. that P_{PS} contains at least one backward edge. Note that for each PS-path that contains at least one backward edge, we can define a corresponding minimal PS-path containing exactly one backward edge. To do this—starting from the end of the path—all edges are removed, until the last edge of the path is a backward edge and all other edges (if any remain) are forward edges. For sake of simplicity, we write P_{PS} also for the minimal path corresponding to P_{PS} in the following and write it as

$$P_{\text{PS}} = Z_0|Y_0 \to \cdots \to S_{s-2}|Z_{s-2}|Y_{s-2} \to S_{s-1}|Z_{s-1}|Y_{s-1} \to S_s\,.$$

We further distinguish the following two sub-cases:

Case 1: $s = 1$
The path is simply $Z_0|Y_0 \to S_1$ and $\mathcal{A}$ queried S_1 to ρ^{-1}. By construction, we have $\lfloor Y_0 \rfloor_{64} = IV$ and $\rho^{-1}(S_1) = Y_0$, hence in particular $\lfloor \rho^{-1}(S_1) \rfloor_{64} = \lfloor Y_0 \rfloor_{64} = IV$. Thus $\mathcal{A}$'s query triggered event E_t.

Case 2: $s \geq 2$
The path is $Z_0|Y_0 \to \cdots \to S_{s-2}|Z_{s-2}|Y_{s-2} \to S_{s-1}|Z_{s-1}|Y_{s-1} \to S_s$. Except for the last edge, all edges are forward edges. By construction, we have $\lfloor S_{s-1} \rfloor_{128} = \lfloor Y_{s-1} \rfloor_{128}$. Furthermore, S_{s-1} is the result of querying Y_{s-2} to ρ (forward edge) and Y_{s-1} is the result of querying S_s to ρ^{-1} (backward edge). This yields that these two queries trigger event E_c.

We now consider the case that the penultimate states Y_{s-1} and $\overline{Y}_{\overline{s}-1}$ are equal, but P_{PS} and $\overline{P}_{\text{PS}}$ contain no backward edges. For such a pair of paths $(P_{\text{PS}}, \overline{P}_{\text{PS}})$, we define the corresponding minimal pair of PS-paths by choosing $s+\overline{s}$ minimal such that $(Z_0, \ldots, Z_{s-1}) \neq (\overline{Z}_0, \ldots, \overline{Z}_{\overline{s}-1})$ and $Y_{s-1} = \overline{Y}_{\overline{s}-1}$ still hold. We consider the minimal pair of paths corresponding to $(P_{\text{PS}}, \overline{P}_{\text{PS}})$ and—for sake of simplicity—also denote them by $(P_{\text{PS}}, \overline{P}_{\text{PS}})$. As before we use the following representation

$$\begin{aligned} P_{\text{PS}} =& Z_0|Y_0 \to \cdots \to S_{s-2}|Z_{s-2}|Y_{s-2} \to S_{s-1}|Z_{s-1}|Y_{s-1} \to S_s \\ \overline{P}_{\text{PS}} =& \overline{Z}_0|\overline{Y}_0 \to \cdots \to \overline{S}_{\overline{s}-2}|\overline{Z}_{\overline{s}-2}|\overline{Y}_{\overline{s}-2} \to \overline{S}_{\overline{s}-1}|\overline{Z}_{\overline{s}-1}|\overline{Y}_{\overline{s}-1} \to \overline{S}_{\overline{s}}\,. \end{aligned}$$

Without loss of generality, we further assume $s \leq \overline{s}$ and distinguish between the following three sub-cases:

Case 1: $s = 1 \wedge \overline{s} = 1$
The paths are simply $Z_0|Y_0 \to S_1$ and $\overline{Z}_0|\overline{Y}_0 \to \overline{S}_1$. By construction, we have $\lceil Y_0 \rceil_r = Z_0$ and $\lceil \overline{Y}_0 \rceil_r = \overline{Z}_0$ which leads to a contradiction as $Z_0 \neq \overline{Z}_0$ (recall that problematic paths have different inputs) but $Y_0 = \overline{Y}_0$. Thus, this case cannot occur.

Case 2: $s = 1 \wedge \overline{s} \geq 2$
The paths are

$$\begin{aligned} & Z_0|Y_0 \to S_1 \\ & \overline{Z}_0|\overline{Y}_0 \to \cdots \to \overline{S}_{\overline{s}-2}|\overline{Z}_{\overline{s}-2}|\overline{Y}_{\overline{s}-2} \to \overline{S}_{\overline{s}-1}|\overline{Z}_{\overline{s}-1}|\overline{Y}_{\overline{s}-1} \to \overline{S}_{\overline{s}}\,. \end{aligned}$$

We have $\lfloor Y_0 \rfloor_{64} = IV$ by construction and $Y_0 = \overline{Y}_{\overline{s}-1}$ by assumption. This allows to deduce $\left\lfloor \overline{S}_{\overline{s}-1} \right\rfloor_{64} = IV$. Since $\overline{S}_{\overline{s}-1}$ is the result of querying $\overline{Y}_{\overline{s}-2}$ to ρ, this query triggered event E_t.

Case 3: $s \geq 2 \wedge \overline{s} \geq 2$
The paths are

$$\begin{aligned} & Z_0|Y_0 \to \cdots \to S_{s-2}|Z_{s-2}|Y_{s-2} \to S_{s-1}|Z_{s-1}|Y_{s-1} \to S_s \\ & \overline{Z}_0|\overline{Y}_0 \to \cdots \to \overline{S}_{\overline{s}-2}|\overline{Z}_{\overline{s}-2}|\overline{Y}_{\overline{s}-2} \to \overline{S}_{\overline{s}-1}|\overline{Z}_{\overline{s}-1}|\overline{Y}_{\overline{s}-1} \to \overline{S}_{\overline{s}}\,. \end{aligned}$$

Consider the penultimate edges of both paths, i.e., the edges from Y_{s-2} to S_{s-1} and from $\overline{Y}_{\overline{s}-2}$ to $\overline{S}_{\overline{s}-1}$, which are both forward edges. By assumption, $Y_{s-1} = \overline{Y}_{\overline{s}-1}$ holds. We have to distinguish two more cases based on the inputs Z_{s-1} and $\overline{Z}_{\overline{s}-1}$:

Case 3.1: $Z_{s-1} = \overline{Z}_{\overline{s}-1}$
From $Y_{s-1} = \overline{Y}_{\overline{s}-1}$ and $Z_{s-1} = \overline{Z}_{\overline{s}-1}$, we can conclude that $S_{s-1} = \overline{S}_{\overline{s}-1}$. Then, however, one can obtain a pair of shorter paths by dropping the last edges of both P and $\overline{P}$ while maintaining the desired property. Thus this case is impossible as it contradicts the minimality of the paths.

Case 3.2: $Z_{s-1} \neq \overline{Z}_{\overline{s}-1}$
As $Y_{s-1} = \overline{Y}_{\overline{s}-1}$ and $Z_{s-1} \neq \overline{Z}_{\overline{s}-1}$, we can deduce $S_{s-1} \neq \overline{S}_{\overline{s}-1}$ and $\lfloor S_{s-1} \rfloor_{128} = \left\lfloor \overline{S}_{\overline{s}-1} \right\rfloor_{128}$. Then $Y_{s-2} \neq \overline{Y}_{\overline{s}-2}$ and the queries on Y_{s-2} and $\overline{Y}_{\overline{s}-2}$ triggered event E_c.

Collecting the above, yields

$$\Pr[\mathsf{PP}_{\mathrm{A}}] \leq \Pr[\mathsf{E}_t \vee \mathsf{E}_c] \leq \Pr[\mathsf{E}_t] + \Pr[\mathsf{E}_c]\,.$$

We apply Lemma 19 for $c = 2^{64}$ and obtain $\Pr[\mathsf{E}_t] \leq \frac{q}{2^{63}}$ and once again for $c = 2^{128}$, which gives $\Pr[\mathsf{E}_c] \leq \frac{q(q-1)}{2^{128}}$. In total, we have shown

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{CMT}}_{\textsc{Ascon}}(\mathcal{A}) &\leq \mathbf{Adv}^{\mathsf{TagColl}}_{\textsc{Ascon}}(\mathcal{A}) \\
&\leq \Pr[\mathsf{CP} \wedge \neg\mathsf{PP}_{\mathrm{A}}] + \Pr[\mathsf{PP}_{\mathrm{A}}] \\
&\leq \Pr[\mathsf{CP} \wedge \neg\mathsf{PP}_{\mathrm{A}}] + \Pr[\mathsf{E}_t] + \Pr[\mathsf{E}_c] \\
&\leq 1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right) + \frac{q}{2^{63}} + \frac{q(q-1)}{2^{128}}\,,
\end{aligned}$$

which finishes the proof. □

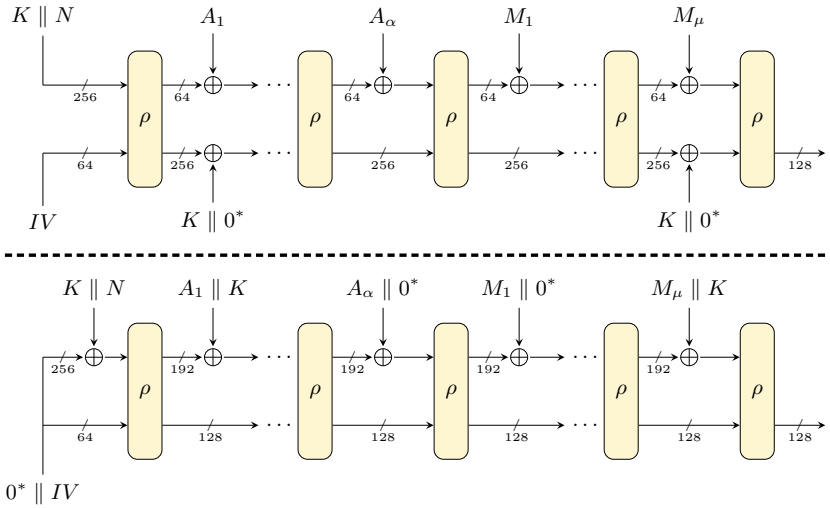

Fig. 31: Illustration of a proof step for Ascon (Theorem 10).

$\textsc{Ascon.Enc}(K, N, A, M)$

$S \leftarrow \textsc{Enc}_{\mathcal{E}}(K, N, A)$
$(C, T) \leftarrow \textsc{Enc}_{\mathcal{M}}(K, S, M)$
return (C, T)

$\textsc{Enc}_{\mathcal{E}}(K, N, A)$

$A_1, \ldots, A_\alpha \xleftarrow{r} \mathtt{pad}_{10^*}(A, r)$
$Y \leftarrow IV \parallel K \parallel N$
$S \leftarrow \rho^a(Y)$
$S \leftarrow S \oplus 0^{n-\kappa} \parallel K$
for $i = 1, \ldots, \alpha$
    $Y \leftarrow S \oplus (A_i \parallel 0^c)$
    $S \leftarrow \rho^b(Y)$
$S \leftarrow S \oplus 0^{n-1} \parallel 1$
return S

$\textsc{Enc}_{\mathcal{M}}(K, S, M)$

$M_1, \ldots, M_\mu \xleftarrow{r} \mathtt{pad}_{10^*}(M, r)$
$Y \leftarrow S \oplus (M_1 \parallel 0^c)$
$C_1 \leftarrow \lceil Y \rceil_r$
for $i = 2, \ldots, \mu$
    $S \leftarrow \rho^b(Y)$
    $Y \leftarrow S \oplus (M_i \parallel 0^c)$
    $C_i \leftarrow \lceil Y \rceil_r$
$C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
$Y \leftarrow Y \oplus (0^r \parallel K \parallel 0^{c-\kappa})$
$S \leftarrow \rho^a(Y)$
$T \leftarrow \lfloor S \rfloor_\tau \oplus K$
return (C, T)

Fig. 32: Pseudocode of Ascon [31] in terms of $\textsc{Enc}_{\mathcal{E}}$ and $\textsc{Enc}_{\mathcal{M}}$.

B.9 Proof of Theorem 11 (SCHWAEMM)

Proof. Firstly, we observe that finding different inputs to $\text{SCHWAEMM}_{IV}.\text{ENC}$ that result in the same ciphertext is easy. However, breaking CMT security also includes finding colliding tags, which is what we focus on in the following. An adversary $\mathcal{A}$ that wins the game CMT against SCHWAEMM_{IV} its output denoted by $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$, in particular finds a tag collision, i.e., wins the game $\mathsf{TagColl}$ (see Fig. 16). Hence we can deduce

$$\mathbf{Adv}^{\mathsf{CMT}}_{\text{SCHWAEMM}_{IV}}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{TagColl}}_{\text{SCHWAEMM}_{IV}}(\mathcal{A})\,.$$

As a next step, we pass over to a plain sponge construction. For this, we construct a $\mathsf{ShiftedColl}_{128}$ adversary $\mathcal{B}$ against a sponge hash function $\mathcal{H}$ obtained from the permutation ρ with rate 256, capacity 128 and output length 128. Further, its initial state is given by $0^{256} \parallel IV$. First, $\mathcal{B}$ runs $\mathcal{A}$, which outputs $(K, N, A, M), (\overline{K}, \overline{N}, \overline{A}, \overline{M})$. For every query that $\mathcal{A}$ makes to ρ, the adversary $\mathcal{B}$ makes the same query to its own permutation and sends the response back to $\mathcal{A}$. Then $\mathcal{B}$ computes the state S_i (and respectively $\overline{S}_i$) after the i-th application of the permutation in SCHWAEMM_{IV} evaluated in (K, N, A, M) (and respectively $(\overline{K}, \overline{N}, \overline{A}, \overline{M})$). Denote by $S_{i,r}$ and $S_{i,c}$ (and respectively $\overline{S}_{i,r}$ and $\overline{S}_{i,c}$) the rate and the capacity part of the state S_i (and respectively $\overline{S}_i$). The adversary $\mathcal{B}$ then outputs

$$\begin{aligned}
X = {} & K \parallel N \parallel \\
& (S_{1,r} \oplus \xi_2(S_{1,r}, A_1) \oplus \omega_{c,r}(S_{1,c})) \parallel \ldots \parallel \\
& (S_{\alpha,r} \oplus \xi_2(S_{\alpha,r}, A_\alpha) \oplus \omega_{c,r}(S_{\alpha,c})) \parallel \\
& (S_{\alpha+1,r} \oplus \xi_2(S_{\alpha+1,r}, M_1) \oplus \omega_{c,r}(S_{\alpha+1,c})) \parallel \ldots \parallel \\
& (S_{\alpha+\mu,r} \oplus \xi_2(S_{\alpha+\mu,r}, M_\mu) \oplus \omega_{c,r}(S_{\alpha+\mu,c}) \\
\overline{X} = {} & \overline{K} \parallel \overline{N} \parallel \\
& (\overline{S}_{1,r} \oplus \xi_2(\overline{S}_{1,r}, \overline{A}_1) \oplus \omega_{c,r}(\overline{S}_{1,c})) \parallel \ldots \parallel \\
& (\overline{S}_{\overline{\alpha},r} \oplus \xi_2(\overline{S}_{\overline{\alpha},r}, \overline{A}_{\overline{\alpha}}) \oplus \omega_{c,r}(\overline{S}_{\overline{\alpha},c})) \parallel \\
& (\overline{S}_{\overline{\alpha}+1,r} \oplus \xi_2(\overline{S}_{\overline{\alpha}+1,r}, \overline{M}_1) \oplus \omega_{c,r}(\overline{S}_{\overline{\alpha}+1,c})) \parallel \ldots \parallel \\
& (\overline{S}_{\overline{\alpha}+\overline{\mu},r} \oplus \xi_2(\overline{S}_{\overline{\alpha}+\overline{\mu},r}, \overline{M}_{\overline{\mu}}) \oplus \omega_{c,r}(\overline{S}_{\overline{\alpha}+\overline{\mu},c})\,,
\end{aligned}$$

which guarantees that $\mathcal{H}(X)$ (and respectively $\mathcal{H}(\overline{X})$) models SCHWAEMM_{IV} evaluated on (K, N, A, M) (and $(\overline{K}, \overline{N}, \overline{A}, \overline{M})$, respectively). More precisely, instead of the state-update-function and rate-whitening applied in SCHWAEMM_{IV}, for $\mathcal{H}$ we XOR a suitable value which imitates these operations. A visualization for this is provided in Fig. 33.

Next, we show that if $\mathcal{A}$ wins the game $\mathsf{TagColl}$ against SCHWAEMM_{IV}, then the constructed adversary $\mathcal{B}$ wins the game $\mathsf{ShiftedColl}_{128}$ against $\mathcal{H}$. First observe that $\mathcal{A}$ winning implies that $(K, N, A, M) \neq (\overline{K}, \overline{N}, \overline{A}, \overline{M})$ and the corresponding tags $T, \overline{T}$—computed with SCHWAEMM_{IV}—agree. Note that the latter implies that $S_{\alpha+\mu+1,c} \oplus K = \overline{S}_{\overline{\alpha}+\overline{\mu}+1,c} \oplus \overline{K}$, hence by choice of X and $\overline{X}$ it holds that $\mathcal{H}(X) \oplus \lceil X \rceil_{128} = \mathcal{H}(\overline{X}) \oplus \lceil \overline{X} \rceil_{128}$. Further, the fact that

$(K, N, A) \neq (\overline{K}, \overline{N}, \overline{A})$, implies that $X \neq \overline{X}$: for $(K, N) \neq (\overline{K}, \overline{N})$ this is obvious while for $(K, N) = (\overline{K}, \overline{N})$ and $A \neq \overline{A}$, a simple analysis shows that X and $\overline{X}$ differ at the point where the associated data blocks differ for the first time. This implies that $\mathcal{B}$ wins the game $\mathsf{ShiftedColl}_{128}$.

Using the indifferentiability of sponges (cf. Theorem 16), we can replace $\mathcal{H}$ by a random function F, as there exists an efficient simulator for the underlying permutation such that $\mathcal{A}$ cannot distinguish between $\mathcal{H}$ and F. This yields

$$\mathbf{Adv}_{\mathcal{H}}^{\mathsf{ShiftedColl}_{128}}(\mathcal{A}) \leq \mathbf{Adv}_{\mathrm{F}}^{\mathsf{ShiftedColl}_{128}}(\mathcal{A}) + \epsilon\,,$$

for $\epsilon > \frac{(1-2^{-256})q^2+(1+2^{-256})q}{2^{129}}$, which results from the application of Theorem 16.

As a last step, we observe that for a random oracle $\mathrm{F} : \{0,1\}^{\geq 128} \to \{0,1\}^{128}$, it is unlikely that $\mathcal{B}$ wins the game $\mathsf{ShiftedColl}_{128}$. For this, note that an adversary that wins the game $\mathsf{ShiftedColl}_{128}$ against F with q queries, finds a collision in the following list of uniformly distributed elements

$$L = \{\mathrm{F}(X_1) \oplus \lceil X_1 \rceil_{128}\,, \mathrm{F}(X_2) \oplus \lceil X_2 \rceil_{128}\,, \ldots, \mathrm{F}(X_q) \oplus \lceil X_q \rceil_{128}\}\,,$$

for $X_i \in \{0,1\}^{\geq 128}$ being the inputs $\mathcal{B}$ queries to F. By Theorem 17, the probability for this is bounded above by

$$1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right)\,,$$

for $q \leq 2^{127}$. In total, we obtain

$$\begin{aligned}\mathbf{Adv}_{\mathcal{H}}^{\mathsf{ShiftedColl}_{128}}(\mathcal{A}) &\leq \mathbf{Adv}_{\mathrm{F}}^{\mathsf{ShiftedColl}_{128}}(\mathcal{A}) + \epsilon \\ &\leq 1 - \exp\left(\frac{-q(q-1)}{2^{128}}\right) + \epsilon\,,\end{aligned}$$

which finishes the proof of the theorem. □

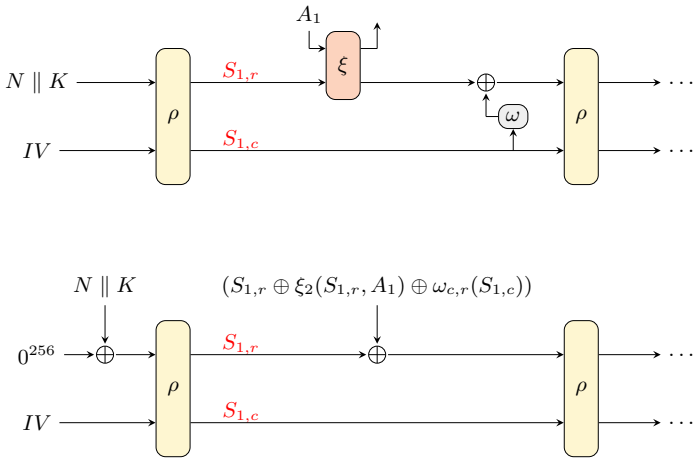

Fig. 33: Illustration of a proof step for SCHWAEMM$_{IV}$ (Theorem 11). SCHWAEMM$_{IV}$ is represented as a plain sponge as is shown for the first XOR in the above figure.

Game $\mathsf{ShiftedColl}_\kappa$

$X, \overline{X} \leftarrow \mathcal{A}()$
if $X = \overline{X}$
  return 0
return $(\mathrm{F}(X) \oplus \mathrm{F}(\overline{X}) = \lceil X \rceil_\kappa \oplus \lceil \overline{X} \rceil_\kappa)$

Fig. 34: Security game $\mathsf{ShiftedColl}_\kappa$ defined for a function $\mathrm{F}\colon \{0,1\}^{\geq\kappa} \to \{0,1\}^\kappa$ and used in the proof for SCHWAEMM$_{IV}$ (Theorem 11).

$\textsc{Schwaemm}.\textsc{Enc}(K, N, A, M)$

$S \leftarrow \textsc{Enc}_{\mathcal{E}}(K, N, A)$
$(C, T) \leftarrow \textsc{Enc}_{\mathcal{M}}(K, S, M)$
return (C, T)

$\textsc{Enc}_{\mathcal{E}}(K, N, A)$

$A_1, \ldots, A_\alpha \xleftarrow{r} \mathsf{pad}_{10^*}(A, r)$
$Y \leftarrow K \parallel N$
$S \leftarrow \rho^a(Y)$
for $i = 1, \ldots, \alpha - 1$
    $(X, \cdot) \leftarrow \xi(\lceil S \rceil_r, A_i)$
    $Y \leftarrow (X \oplus \omega(\lfloor S \rfloor_c)) \parallel \lfloor S \rfloor_c$
    $S \leftarrow \rho^b(Y)$
$(X, \cdot) \leftarrow \xi(\lceil S \rceil_r, A_\alpha)$
$Y \leftarrow X \parallel (\lfloor S \rfloor_c \oplus \iota_A)$
$Y \leftarrow (\lceil Y \rceil_r \oplus \omega(\lfloor Y \rfloor_c)) \parallel \lfloor Y \rfloor_c$
$S \leftarrow \rho^a(Y)$
return S

$\textsc{Enc}_{\mathcal{M}}(K, S, M)$

$M_1, \ldots, M_\mu \xleftarrow{r} \mathsf{pad}_{10^*}(M, r)$
for $i = 1, \ldots, \mu - 1$
    $(X, C_i) \leftarrow \xi(\lceil S \rceil_r, M_i)$
    $Y \leftarrow (X \oplus \omega(\lfloor S \rfloor_c)) \parallel \lfloor S \rfloor_c$
    $S \leftarrow \rho^b(Y)$
$(X, C_\mu) \leftarrow \xi(\lceil S \rceil_r, M_\mu)$
$Y \leftarrow X \parallel (\lfloor S \rfloor_c \oplus \iota_M)$
$Y \leftarrow (\lceil Y \rceil_r \oplus \omega(\lfloor Y \rfloor_c)) \parallel \lfloor Y \rfloor_c$
$C \leftarrow \lceil C_1 \parallel \ldots \parallel C_\mu \rceil_{|M|}$
$S \leftarrow \rho^a(Y)$
$T \leftarrow \lceil S \rceil_\tau \oplus K$
return (C, T)

$\xi(S, I)$

$Y \leftarrow \mathsf{FeistelSwap}(S) \oplus I$
$O \leftarrow S \oplus I$
return (Y, O)

Fig. 35: Pseudocode of Schwaemm [7] in terms of $\textsc{Enc}_{\mathcal{E}}$ and $\textsc{Enc}_{\mathcal{M}}$.