McOE: A Foolproof On-Line Authenticated Encryption Scheme

Ewan Fleischmann, Christian Forler, and Stefan Lucks

Professur für Mediensicherheit, Bauhaus-Universität Weimar, Germany

Abstract. On-Line Authenticated Encryption (OAE) combines confidentiality with data integrity and is on-line computable. Most block cipher-based schemes for Authenticated Encryption can be run on-line and are provably secure against *nonce-respecting* adversaries. But they fail badly for more general adversaries. This is not a theoretical observation only – in practice, the reuse of nonces is a frequent issue[1].
In recent years, cryptographers developed *misuse-resistant* schemes for Authenticated Encryption. These guarantee excellent security even against general adversaries which are allowed to reuse nonces. Their disadvantage is that encryption can be performed in an off-line way, only.
This paper considers OAE schemes dealing both with nonce-respecting *and* with general adversaries. It introduces McOE, an efficient design for OAE schemes. The construction is based on a 'simple' block cipher and is on-line computable. It provably guarantees reasonable security against general adversaries as well as standard security against nonce-respecting adversaries.

Keywords: authenticated encryption, online encryption, provable security, misuse resistant

1 Introduction

On-Line Authenticated Encryption (OAE). Application software often requires a network channel that guarantees the privacy and authenticity of data being communicated between two parties. Cryptographic schemes able to meet both of these goals are commonly referred to as Authenticated Encryption (AE) schemes. The ISO/IEC 19772:2009 standard for AE [16] defines generic composition (Encrypt-then-MAC [3]) and five dedicated AE schemes: OCB2 [33], SIV [36] (denoted as "Key Wrap" in [16]), CCM [10], EAX [5], and GCM [29]. To integrate an AE-secure channel most seamlessly into a typical software architecture, application developers expect it to encrypt in an *on-line* manner meaning that the i-th ciphertext block can be written before the $(i+1)$-th plaintext block has to be read. A restriction to off-line encryption, where usually the entire plaintext must be known in advance (or read more than once) is an encumbrance to software architects.

Nonces and their reuse. Goldwasser and Micali [14] formalized encryption schemes as stateful or probabilistic, because otherwise important security properties are lost. Rogaway [32, 34, 35] proposed a unified point of view, by always defining a cryptographic scheme as a deterministic algorithm that takes a user supplied nonce (a *number used once*). So the application programmer – and not the encryption scheme – is responsible for flipping coins or maintaining state. This reflects cryptographic practice since the algorithm itself is often implemented by a multi-purpose cryptographic library which is more or less application-agnostic.

In theory, the concept of a nonce is simple. In practice, it is challenging to ensure that a nonce is *never* reused. Flawed implementations of nonces are ubiquitous [7, 15, 23, 39, 40]. Apart

[1] A prominent example is the PlayStation 3 'jailbreak' [15], where application developers used a constant that was actually supposed to be a nonce for a digital signature scheme.

secure ...	**against nonce-respecting adversaries**	**ag. nonce-reusing adversaries**
on-line	CCFB[28] CHM[17] CIP[18] CWC[24] EAX[5] GCM[29] IACBC[21] IAPM[21] **McOE** OCB1-3[35, 33, 25] RPC[8] TAE[26] XCBC[13]	**McOE** (this paper)
off-line	BTM[19] CCM[10] HBS[20] SIV[36] SSH-CTR[31]	BTM[19] HBS[20] SIV[36]

Table 1. Classification of provably secure block cipher-based AE Schemes. CCM and SSH-CTR are considered off-line because encryption requires prior knowledge of the message length. Note that McOE, because of being on-line, satisfies a slightly weaker security definition against nonce-reusing adversaries than SIV, HBS, and BTM.

	privacy attack workload	authenticity attack workload
CCFB [28]	$O(1)$	$O(1)$
CCM [10]	$O(1)$	$\ll 2^{(n/2)}$ [11]
CHM [17]	$O(1)$	$O(1)$
CIP [18]	$O(1)$	$O(1)$
CWC [24]	$O(1)$	$O(1)$
EAX [5]	$O(1)$	$O(1)$
GCM [29]	$O(1)$	$O(1)$
IACBC [21]	$O(1)$	$O(1)$

	privacy attack workload	authenticity attack workload
IAPM [21]	$O(1)$	$O(1)$
OCB1 [35]	$O(1)$	$O(1)$
OCB2 [33]	$O(1)$	$O(1)$
OCB3 [25]	$O(1)$	$O(1)$
RPC [8]	$O(1)$	$O(1)$
TAE [26]	$O(1)$	$O(1)$
XCBC [13]	$O(2^{n/4})$	?

Table 2. Overview of our **nonce-reuse** attacks on published AE schemes, excluding SIV, HBS and BTM, which have been explicitly designed to resist nonce-reuse. Almost all attacks achieve an advantage close to 1. An "attack workload" of X means that the adversary is restricted to at most X units of time and at most X chosen texts. Details are given in Appendix A.

from implementation failures, there are fundamental reasons why software developers can't always prevent nonce reuse. A persistently stored counter, which is increased and written back each time a new nonce is needed, may be reseted by a backup – usually after some previous data loss. Similarly, the internal and persistent state of an application may be duplicated when a virtual machine is cloned, etc.

Related Work and Our Contribution. This paper is about achieving *both simultaneously*: security against nonce-reusing adversaries (sometimes also called nonce-misusing adversaries) *and* support for on-line-encryption in terms of an AE scheme. None of the ISO/IEC 19772:2009 schemes – in fact, no previously published AE scheme at all – achieves both of these goals, cf. Table 1. In this table, we classify a vast variety of provably secure block cipher-based AE scheme with respect to their on-line-ability and against which adversaries (nonce-respecting versus -reusing) they are proven secure. To close the apparent gap in the upper-right of Table 1, we considered the schemes in the upper-left, whether they would fit. As it turned out, we actually found nonce-reuse attacks for *all* of those schemes, cf. Table 2 and Appendix A. In this paper we present a new construction for efficient AE schemes, called McOE, that is actually able to fill the apparent gap in the upper-right. We argue that closing this gap is both practically relevant and theoretically interesting.

Initial Value (IV) based AE schemes maximally forgiving of repeated IV's have been adressed in [36], coining the notion of "misuse resistance" and proposing SIV as a solution. SIV and related schemes (HBS [20] and BTM [19]) actually provide excellent security against nonce-reusing adversaries, though there are other potential misuse cases, cf. Appendix A.2. Their main disadvantage

is that they are inherently off-line: For encryption, one must either keep the entire plaintext in memory, or read the plaintext twice.

Ideally, an adversary seeing the encryptions of two (equal-length) plaintexts P_1 and P_2 can't even decide if $P_1 = P_2$ or not. When using a nonce more than once, deciding about $P_1 = P_2$ is easy. SIV and its relatives ensure that nothing else is feasible for nonce-reusing adversaries. In the case of on-line encryption, where the first few bits of the encryption of a lengthy message must not depend on the last few bits of that message, there is unavoidably something beyond $P_1 = P_2$. The adversary can compare any two ciphertexts for their longest common prefix, and then conclude about common prefixes of the secret plaintexts. Our notion of *misuse resistance* means that this is all the adversary can gain. Even in the case of a nonce-reuse, the adversary

1. can't do anything beyond determining the length of common plaintext prefixes and
2. the scheme still provides the usual level authenticity for AE (INT-CTXT).

The first property is common for on-line ciphers/permutations (OPRP) [1]. Recently, [38] studied the design of on-line ciphers from tweakable block ciphers bearing some similarities to our approach. In constrast to McOE, the constructions from [38] provide no authentication. Also, McOE can employ either a tweakable block cipher, or a related-key secure untweaked block cipher with key-lenght $\geq$ block size.

Design Principles for AE Schemes. The question how to provide authenticated encryption (without stating that name) when given a secure on-line cipher is studied in [2], the revised and full version of [1]. The first idea in [2] only provides security if all messages are of the same length. The second idea repairs that by prepending the message's length to the message, at the cost of being off-line, since the message length must be known at the beginning of the encryption process. The third idea is to prepend and append a random W to a message M and then to perform the on-line encryption of $(W||M||W)$. This looks promising, but the same W is used for two different purposes, putting different constraints on the generation of W. For privacy, it suffices that W behaves like a nonce, not requiring secrecy or unpredictability. Even if W is not a nonce, but the same W is used for the encryption of several messages, all the adversary can determine are the lengths of common plaintexts prefixes, as we required for nonce-reuse. On the other hand, authenticity actually assumes a *secret or unpredictable* W, rather than a nonce. If the adversary can guess W before choosing a message, she asks for the authenticated encryption of $(M||W)$. Then she can predict the authenticated encryption of M without actually asking for it.

The McOE construction replaces the "random" W by a proper nonce and a *key-dependent* value τ, performing a nonce-dependent on-line encryption of $(M||\tau)$. The encryption can also depend on some associated data, which turns McOE into a scheme for OAEAD *(On-Line Authenticated Encryption with Associated Data).*

Roadmap. In Section 2 we describe a block cipher based OAE scheme – a concrete instance of McOE – and provide perfomance data when the scheme is instantiated with either the AES or Threefish as the underlying block cipher. Section 3 deals with general notions and definitions, and Section 4 defines the security of OAE. The main result of the paper, the full McOE construction and its analysis, is presented in Section 5. The discussion in Section 6 concludes the paper. The appendix deals with misuse attacks against published AE schemes, and provides some proof supplements.

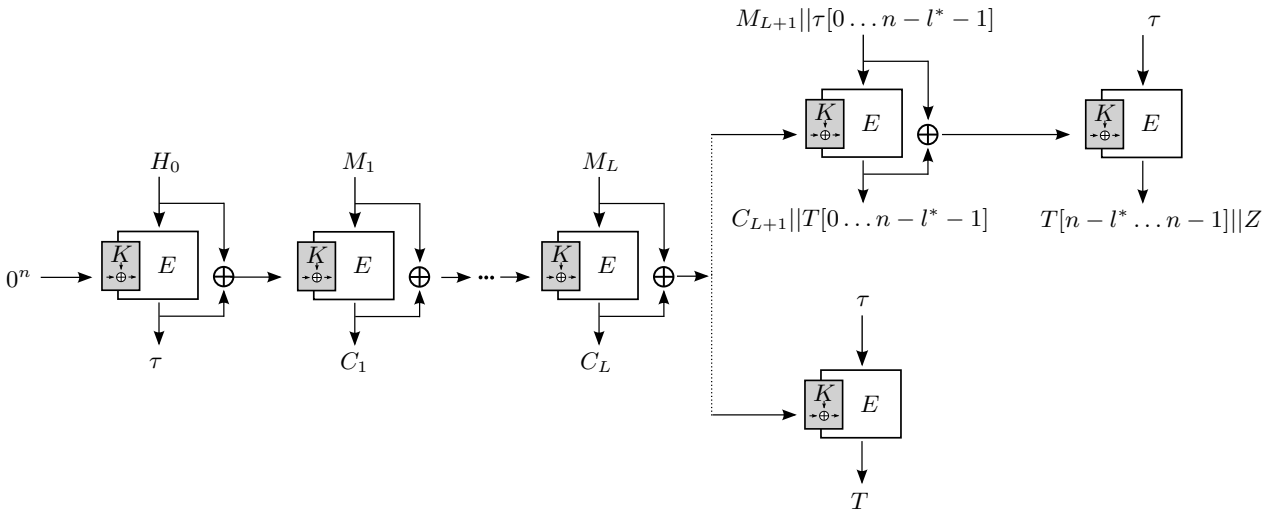

Fig. 1. The McOE-AES/McOE-Threefish encryption process. If, after the last complete message block has been encrypted, there is some incomplete block left, McOE performs tag-splitting (upper variant). Else, the tag can be computed without splitting (lower variant). The key used for the block cipher E is computed by the injective function $K \oplus V$ which is given the secret key K an the chaining value input V. The tag returned is the n-bit value T. The $n-l$-bit value Z is discarded. The ciphertext in the tag-splitting case is $(T, C_1, \ldots, C_L, C_{L+1})$ and otherwise $(T, C_1, \ldots, C_L)$. The decryption process works in a similar way from 'left to right', only the block cipher component E_φ is replaced by its counterpart E_φ^{-1} apart from one exception: the first call computing T_M.

2 Practical On-Line Authenticated Encryption using AES and Threefish

We start with the fruits of our analysis by giving two concrete instances of OAE schemes including performance data and reference source code[2]. One instance, McOE-AES uses AES-128 as the core component while McOE-Threefish uses the block cipher Threefish-512, a cipher with 512-bit block size and key size, which is the core working component inside the SHA-3 finalist Skein[30].

We also introduce the *tag-splitting* (TS) method for processing messages whose length is not a multiple of the block length. Without TS, we would have to pad such messages and then encrypt the padded messages – resulting in an expanded ciphertext. TS is similar to a well-known length preserving method called ciphertext stealing (CTS), *e.g.* [9]. Note that CTS requires to process the last block before the last-but-one, which is not possible for McOE.

Let E_K be a block cipher taking a k-bit key K and a plaintext/ciphertext of size n-bit. Note that for our chosen instances, AES-128 and Threefish-512, we have $n = k$. The pseudo code for these two McOE instances is given in Table 3 – on the top without TS, on the bottom including TS.

The algorithms without TS, **EncryptAuthenticate** and **DecryptAuthenticate**, are simplified algorithms for messages that are aligned on n-bit boundaries, *i.e.* $M = (M_1, \ldots, M_L) \in (\{0,1\}^n)^L$ for some integer L. The TS-variants **EncryptAuthenticateSplitTag** and **DecryptAuthenticateSplitTag**, can handle arbitrarily sized messages, *i.e.*, $M = (M_1, \ldots, M_L, M_{L+1}) \in (\{0,1\}^n)^L||\{0,1\}^{l^*}$ where L and l^* are integers with $0 \leq l^* < n$ and $'||'$ denotes the string concatenation operator. See Figure 1 and Table 3.

[2] The reference source code is available on request; it will be published as open source.

EncryptAuthenticate(H_0, M)
1. $\tau \leftarrow E_K(H_0)$
2. $U \leftarrow H_0 \oplus \tau \oplus K$
3. for i from 1 to L loop
   $C_i \leftarrow E_U(M_i)$
   $U \leftarrow M_i \oplus C_i \oplus K$
   end loop
4. $T \leftarrow E_W(\tau)$
5. return $(C_1, \ldots, C_L, T)$

DecryptAuthenticate(H_0, C, T)
1. $\tau \leftarrow E_K(H_0)$
2. $U \leftarrow H_0 \oplus \tau \oplus K$
3. for i from 1 to L loop
   $M_i \leftarrow E_U^{-1}(C_i)$
   $U \leftarrow M_i \oplus C_i \oplus K$
   end loop
4. $T' \leftarrow E_U(\tau')$
5. if $T = T'$ then return $(M_1, \ldots, M_L)$
   else return $\perp$

EncryptAuthenticateSplitTag(H_0, M)
1. $\tau \leftarrow E_K(H_0)$
2. $U \leftarrow H_0 \oplus \tau \oplus K$
3. for i from 1 to L loop
   $C_i \leftarrow E_U(M_i)$
   $U \leftarrow M_i \oplus C_i \oplus K$
   end loop
4. $M^* \leftarrow M_{L+1} || \tau[0 \ldots n - l^* - 1]$
5. $C^* \leftarrow E_U(M^*)$
6. Parse $C_{L+1} || T[0 \ldots n - l^* - 1] \leftarrow C^*$
7. $U \leftarrow M^* \oplus C^* \oplus K$
8. $C^{**} \leftarrow E_U(\tau)$
9. $T[n - l^* \ldots n - 1] \leftarrow C^{**}[0 \ldots l^* - 1]$
10. return $(C_1, \ldots, C_L, C^*_{L+1}, T)$

DecryptAuthenticateSplitTag(H_0, C, T)
1. $\tau \leftarrow E_K(H_0)$
2. $U \leftarrow H_0 \oplus \tau \oplus K$
3. for i from 1 to L loop
   $M_i \leftarrow E_U^{-1}(C_i)$
   $U \leftarrow M_i \oplus C_i \oplus K$
   end loop
4. $C^* \leftarrow C_{L+1} || T[0 \ldots n - l^* - 1]$
5. $M^* \leftarrow E_U^{-1}(M^*)$
6. Parse $M_{L+1} || \tau'[0 \ldots n - l^* - 1] \leftarrow M^*$
7. $U \leftarrow M^* \oplus C^* \oplus K$
8. $C^{**} \leftarrow E_U(\tau)$
9. $T'[n - l^* \ldots n - 1] \leftarrow C^{**}[0 \ldots l^* - 1]$
10. if $\tau'[0 \ldots n - l^* - 1] = \tau[0 \ldots n - l^* - 1]$
    and $T'[n - l^* \ldots n - 1] = T[n - l^* \ldots n - 1]$
    then return $(C_1, \ldots, C_L, C_{L+1}, T)$
    else return $\perp$

Table 3. Instances of McOE: Left side is for messages whose size is evenly divisible by the block size n; Right side is for arbitrarily sized messages (TS-variant); see text for details

Note that the algorithms with TS are compatible to the simpler variants without. I.e., **EncryptAuthenticate**$(H_0, (M_1, \ldots, M_L))$ will give the same ciphertext as **EncryptAuthenticateSplitTag**$(H_0, (M_1, \ldots, M_L, M_{L+1}))$, where the additional M_{L+1}-block is of length zero.

McOE-AES and McOE-Threefish are easily extended to smoothly handle associated data, *i.e.* data that is not encrypted but only authenticated. This is discussed in more detail in Section 5.

Remarks. Note that we actually do need related key resistance for the block cipher E since the adversary can 'partially control' some relations among keys used in the computation.

3 On-Line Authenticated Encryption and Related Notions

3.1 Definitions

Length of Longest Common Prefix (LLCP$_n$). The length of a string $x \in \{0,1\}^n$ is denoted by $|x| := n$. For integers $n, \ell, d \geq 1$, set $D_n^d = (\{0,1\}^n)^d$, and $D_n^* := \bigcup_{d \geq 0} D_n^d$, and $D_{\ell,n} = \bigcup_{0 \leq d \leq \ell} D_n^d$. Note that D_n^0 only contains the empty string. For $M \in D_n^d$; we write $M = (M_1, \ldots, M_d)$ with $M_1, \ldots, M_d \in D_n$. For $P, R \in D_n^*$, say, $P \in D_n^p$ and $R \in D_n^r$, we define the *length of the longest*

Block cipher	impl. type	Message length in Bytes									
		16	40	128	256	576	1000	1500	4096	10000	32768
AES-128	software	62.9	42.6	26.8	24.3	22.9	22.6	22.3	33	21.8	21.8
AES-128	AES-NI	72.2	19.4	12.7	11.7	11.2	11.1	10.9	10.8	10.7	10.7
Threefish-512	software	82.3	33.9	13.3	10.1	8.3	7.9	7.6	7	6.9	6.8

Table 4. Performance values (cycles-per-byte), measured on an Core i5 540M for AES-128 and Threefish-512. The AES implementation is based on Gladman [12]. The Threefish implementation is based on the NIST/SHA-3 reference source as provided by the Skein authors [30].

common n-prefix of P and R as

$$\mathrm{LLCP}_n(P,R) = \max\{i \le \min\{p,r\} \mid P_1 = R_1, \ldots, P_i = R_i\}.$$

For a non-empty set $\mathcal{Q}$ of strings in D_n^* we define $\mathrm{LLCP}_n(\mathcal{Q},P)$ as $\max_{q\in\mathcal{Q}}\{\mathrm{LLCP}_n(q,P)\}$. For example, if $P \in \mathcal{Q}$, then $\mathrm{LLCP}_n(\mathcal{Q},P) = |P|/n$.

For convenience, we introduce a notation for a *restriction on a set*. If $\mathcal{Q} = \{0,1\}^a \times \{0,1\}^b \times \{0,1\}^c$, we write $\mathcal{Q}_{|b,c} = \{(B,C) \mid \exists A : (A,B,C) \in \mathcal{Q}\}$. This generalizes in the obvious way.

3.2 Block Ciphers and On-Line Permutations

Block Ciphers. An (k,n) block cipher is a keyed family of permutations consisting of two paired algorithms $E : \{0,1\}^k \times D_n \to D_n$ and $E^{-1} : \{0,1\}^k \times D_n \to D_n$, accepting a k-bit key and an input from D_n for some $k, n > 0$. For $n > 0$, $Block(k,n)$ is the set of all (k,n) block ciphers. For any $E \in Block(k,n)$ and a fixed key $K \in \{0,1\}^k$, the decryption $E_K^{-1}(Y) := E^{-1}(K,Y)$ is the inverse function of encryption $E_K(X) := E(K,X)$, so that $E_K^{-1}(E_K(X)) = X$ holds for any $X \in D_n$.

We follow the usual convention to write oracles, that are provided to an algorithm, as superscripts. We define the related key PRP-security of a block cipher E by the success probability of an adversary trying to differentiate between the block cipher and a random permutation.

Definition 1. *Let $E \in Block(k,n)$ and denote by E^{-1} the corresponding inverse. Let $\varphi : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^k$. A fixed related key adversary A has access to an E oracle with two parameters such that she can query either $E_{\varphi(K,\cdot)}(\cdot)$ or its inverse. Let* $\textsc{Perm}(n,n)$ *be the set of n-bit permutations such that the first parameter models the permutation and the second parameter the value that is to be permuted,* i.e. *for $\pi \in \textsc{Perm}(n,n)$ it holds that $\pi(Z,\cdot)$ is a random permutation for any given value of Z. The related-key (RK) advantage [27] of A in breaking E is then defined as*

$$\mathbf{Adv}_E^{\text{RK-CPA-PRP}}(A) = |\Pr[K \xleftarrow{\$} \{0,1\}^k : A^{E_{\varphi(K,\cdot)}(\cdot)} \Rightarrow 1] - \Pr[\pi \xleftarrow{\$} \mathit{Perm(n,n)} : A^{\pi(\cdot,\cdot)} \Rightarrow 1]|$$

$$\mathbf{Adv}_{E,E^{-1}}^{\text{RK-CCA-PRP}}(A) = |\Pr[K \xleftarrow{\$} \{0,1\}^k : A^{E_{\varphi(K,\cdot)}(\cdot),E^{-1}_{\varphi(K,\cdot)}(\cdot)} \Rightarrow 1] - \Pr[\pi \xleftarrow{\$} \mathit{Perm(n,n)} : A^{\pi(\cdot,\cdot),\pi^{-1}(\cdot,\cdot)} \Rightarrow 1]|.$$

On-Line Permutations. We aim for larger permutations that not only permute single blocks but can handle multiple/variable block messages. Such a permutation, from D_n^* to D_n^*, is $(n-)$on-line if the i-th block of the output is determined completely by the first i blocks of the input.

Definition 2. *Let $n, k, v \geq 0$, $K \in \{0,1\}^k, H_0 \in D_n$. A function $\Pi : \{0,1\}^k \times D_n^* \to D_n^*$ is an (n)-on-line permutation if for any fixed K, H_0 the function $\Pi(K, H_0, \cdot)$ is a permutation and there exists for any message $M = (M_1, M_2, \ldots, M_m)$ a family of functions $\widetilde{\pi}^i : \{0,1\}^k \times \{0,1\}^v \times D_n^i \to D_n$, $i = 1, \ldots, m$ such that*

$$\begin{aligned} \Pi(K, H_0, M) = \; & \widetilde{\pi}_K^1(H_0, M_1) || \widetilde{\pi}_K^2(H_0, M[1..2]) \\ & || \ldots || \\ & \widetilde{\pi}_K^{m-1}(H_0, M[1..m-1]) || \widetilde{\pi}_K^m(H_0, M[1..m]), \end{aligned}$$

where $M[a \ldots b] := M_a || M_{a+1} || \ldots || M_b$ with "$||$" being the concatenation of strings, holds.

An encryption scheme is (n)-on-line if the encryption function is (n)-on-line. A thorough discussion of on-line encryption and its properties can be found in [1].

3.3 Authenticated Encryption (With Associated Data)

An authenticated encryption scheme is a tuple $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. Its aim is to provide privacy and data integrity. The key generation function $\mathcal{K}$ takes no input and returns a randomly chosen key K from the key space, *e.g.* from $\{0,1\}^k$. The encryption algorithm $\mathcal{E}$ and the decryption algorithm $\mathcal{D}$ are deterministic algorithms that map values from $\{0,1\}^k \times \mathcal{H} \times D_n^*$ to a string or – if the input is invalid – the value $\perp$. The header $\mathcal{H}$ consists either only of the initial value/nonce $H_0 \in D_n$ (if no data is to be authenticated/checked in the encryption/decryption process) or is a combination of H_0 and a value from D_n^*. So $\mathcal{H} \subset D_n^+$ in either case. For sake of convenience, we usually write $\mathcal{E}_K^H(M)$ for $\mathcal{E}(K, H, M)$ and $\mathcal{D}_K^H(M)$ for $\mathcal{D}(K, H, M)$, where the message M is chosen from D_n^*, $H \in \mathcal{H}$ and a key from the key space. We require $\mathcal{D}_K^{\mathcal{H}}(\mathcal{E}_K^{\mathcal{H}}(M)) = M$ for any possible K, M, H, and define the tag size for a message $M \in D_n^*$ and header $H \in \mathcal{H}$ as $\textsc{tag}(H, M) := |\mathcal{E}_K^H(M)| - |M|$. We denote an authenticated encryption scheme with the requirement that the initial vector H_0 is only used once a *nonce based* scheme. Otherwise, we call such a scheme *deterministic*. Similarly, we call an adversary *nonce-respecting* (NR) if no nonce is used twice for any query. Otherwise, the adversary is called *nonce-ignoring* (NI).

4 Security Notions for On-Line Authenticated Encryption

Authenticated (On-Line) Encryption tries to achieve privacy and authenticity at the same time. Therefore we need security notions that to handle this twofold goal. For AE, there have been notions and their relations been introduced for deterministic [37] and nonce based [3, 4, 22, 32, 35] AE schemes. In order to have one convenient toolset of notions, we adopt the notion of CCA3 security suggested in [37] as a *natural strengthening* of CCA2 security.

We parameterize our definition in order to define different – but closely related – notions by explicitly stating whether we mean an on-line or off-line scheme, $\omega \in \{\textsc{ae}, \textsc{aoe}\}$ and stating the adversary behavior as either nonce-respecting or nonce-ignoring, $\nu \in \{\textsc{nr}, \textsc{ni}\}$.

Definition 3 (CCA3(ω, ν)). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an authenticated encryption scheme with header space $\mathcal{H}$ and message space D_n^*, and fix an adversary A. The advantage of A breaking Π is defined as*

$$\mathbf{Adv}_{\Pi}^{\mathrm{CCA3}(\omega,\nu)}(A) = |\Pr\left[K \xleftarrow{\$} \mathcal{K} : A^{\mathcal{E}_K(\cdot,\cdot),\mathcal{D}_K(\cdot,\cdot)} \Rightarrow 1\right] - \Pr\left[A^{\$^{\omega}(\cdot,\cdot),\perp(\cdot,\cdot)} \Rightarrow 1\right]|.$$

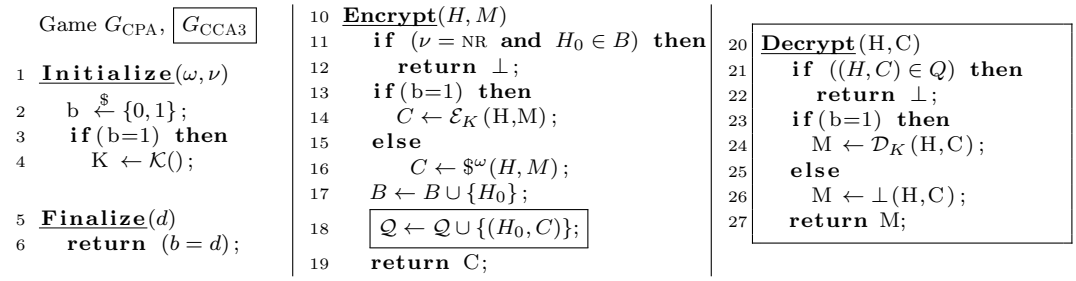

Fig. 2. $G_{\text{CPA}}(\omega, \nu)$ is the $\text{CPA}_{\Pi}^{(\omega,\nu)}$-Game and $G_{\text{CCA3}}(\omega, \nu)$ the $\text{CCA3}_{\Pi}^{(\omega,\nu)}$-Game where $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. Game G_{CCA3} contains the code in the box while G_{CPA} does not. The oracle $\$^{\text{AE}}(H, M)$ returns a string of length $|M| + \text{TAG}(H, M)$, this string is on-line compatible if $\omega =$ AOE. H_0 denotes the first block of the header representing the nonce/initial value.

The adversary's random-bits oracle, $\$^{\text{AE}}(\cdot,\cdot)$ or $\$^{\text{AOE}}(\cdot,\cdot)$, returns on a query with header $H \in \mathcal{H}$ and plaintext $X \in D_n^*$ a random string of length $|\mathcal{E}_K(M)|$ which is either on-line or not, depending on the variable ω. The $\perp(\cdot,\cdot)$ oracle returns $\perp$ on every input. We assume *wlog.* that the adversary A never ask a query which answer is already known. It is easy to see that we can rewrite the term given in Definition 3 as

$$\mathbf{Adv}_{\Pi}^{\text{CCA3}(\omega,\nu)}(A) = |\Pr\left[K \xleftarrow{\$} \mathcal{K} : A^{\mathcal{E}_K(\cdot,\cdot),\mathcal{D}_K(\cdot,\cdot)} \Rightarrow 1\right] - \Pr\left[K \xleftarrow{\$} \mathcal{K} : A^{\mathcal{E}_K(\cdot,\cdot),\perp(\cdot,\cdot)} \Rightarrow 1\right] \quad (1)$$

$$+ \Pr\left[K \xleftarrow{\$} \mathcal{K} : A^{\mathcal{E}_K(\cdot,\cdot),\perp(\cdot,\cdot)} \Rightarrow 1\right] - \Pr\left[A^{\$^{\omega}(\cdot,\cdot),\perp(\cdot,\cdot)} \Rightarrow 1\right]|. \quad (2)$$

One can interpret (1) as the advantage that an adversary has on the integrity of the ciphertext and (2) as the advantage that an CPA adversary has on the privacy. Using this decomposition as a motivational starting point, we now define ciphertext integrity and what we mean by a CPA adversary an authenticated encryption schemes. From now on, our definitions are based on the game playing methodology. For example, we can restate Definition 3 using the game G_{CCA3} given in Figure 2 as

$$\mathbf{Adv}_{\Pi}^{\text{CCA3}(\omega,\nu)}(A) = 2|\Pr[A^{G_{\text{CCA3}}(\omega,\nu)} \Rightarrow 1] - 0.5|.$$

We denote $\mathbf{Adv}_{\Pi}^{\text{CCA3}(\omega,\nu)}(q, t, \ell)$ as the maximum advantage over all $\text{CCA3}(\omega, \nu)$ adversaries run in time at most t, ask a total maximum of q queries to $\mathcal{E}$ and $\mathcal{D}$, and whose total query length is not more than ℓ blocks.

4.1 Privacy and Integrity Notions for Authenticated Encryption Schemes.

Similarly, we define the privacy and integrity of an authenticated (on-line) encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ with header space D_n^+, message space D_n^* and tag-size function $\text{TAG}(H, M)$ as follows.

Definition 4. *Let $G_{CPA}(\omega, \nu)$ be the $CPA_{\Pi}^{\omega,\nu}$ game given in Figure 2. Fix an adversary A. The advantage of A breaking Π is defined as*

$$\mathbf{Adv}_{\Pi}^{CPA(\omega,\nu)}(A) = 2|\Pr[A^{G_{CPA}(\omega,\nu)} \Rightarrow 1] - 0.5|.$$

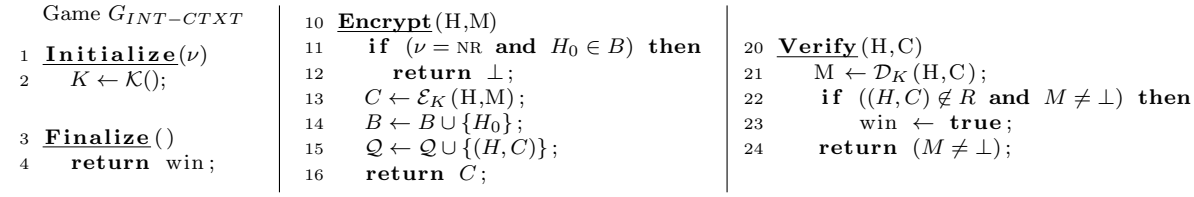

Fig. 3. Game $G_{INT-CTXT}(\nu)$ is the INT-CTXT$^{\omega,\nu}_{\Pi}$ game where $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$. H_0 denotes the first block of the header representing the nonce/initial value.

.

Definition 5. *Let $G_{\text{INT-CTXT}}(\nu)$ be the* INT-CTXT$^{\nu}_{\Pi}$ *game given in Figure 3. Fix an adversary A. The advantage of A breaking Π is defined as*

$$\mathbf{Adv}^{\text{INT-CTXT}(\nu)}_{\Pi}(A) = \Pr[A^{G_{\text{INT-CTXT}}(\nu)} \Rightarrow 1].$$

We denote $\mathbf{Adv}^{CPA(\omega,\nu)}_{\Pi}(q,t,\ell)$ and $\mathbf{Adv}^{\text{INT-CTXT}(\nu)}_{\Pi}(q,t,\ell)$ as the maximum advantage over all $CPA(\omega,\nu)$ resp. INT-CTXT(ν) adversaries run in time at most t, ask a total maximum of q queries to $\mathcal{E}$ and $\mathcal{D}$, and whose total query length is not more than ℓ blocks.

4.2 CCA3 is equal to INT-CTXT plus CPA.

We now give a generalization of Theorem 3.2 from Bellare and Namprempre [3]. It simply states the equivalence of a scheme being CCA3 secure and both INT-CTXT and CPA secure. These statements hold in the on-line and non-on-line case.

Theorem 1. *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an authenticated encryption scheme. Fix $\omega \in \{\text{AE}, \text{AOE}\}$ and $\nu \in \{\text{NR}, \text{NI}\}$. Let A be an* CCA3$(\omega,\nu)_{\Pi}$*-adversary running in time t, making q queries with a total length of at most ℓ blocks. Then there are an* CPA(ω,ν)*-adversary A_p and an* INT-CTXT(ω,ν)*-adversary A_c such that*

$$\mathbf{Adv}^{\text{CCA3}(\omega,\nu)}_{\Pi}(A) \leq \mathbf{Adv}^{\text{CPA}(\omega,\nu)}_{\Pi}(A_p) + \mathbf{Adv}^{\text{INT-CTXT}(\omega,\nu)}_{\Pi}(A_c).$$

Furthermore, A_c and A_p run in time $O(t)$ and both make at most q queries in each case.

The proof is given in Appendix B.

5 The On-Line Authenticated Encryption Scheme McOE

In this section, we present McOE, a construction for an OAE scheme. We prove that McOE achieves our two-fold goal of being secure against a nonce-respecting adversary and also guaranteeing a certain minimum security against a nonce-ignoring adversary.

Since we already have presented two McOE instances in Section 2, we proceed by formally defining McOE and giving its pseudocode. Indeed this is very similar to the results presented in Section 2, but here our definitions are slightly more general. Instead of fixing the key computation function to $K \oplus V$, where V is the chaining value and K the secret key, we here use a key derivation function $\varphi(K,V)$. By this we make sure that our proof also works for tweakable block ciphers - with K as key and V as tweak - leading to more efficient design.

Definition 6 (McOE). *Let $k, n \in \mathbb{N}$, $E \in \texttt{Block}(k, n)$ and $\varphi : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^k$ such that $\varphi(K, \cdot)$ is injective. The encryption function takes a header $H \in D_n^{L_H}$, a message M and returns a ciphertext C and a tag $T \in D_n$. The decryption function takes a header $H \in D_n^{L_H}$, a ciphertext C and a tag $T \in D_n$ and returns either a plaintext M or the fail symbol $\perp$.*

(i) 'Non-TS'. *Let $M, C \in D_N^L$ for some integer L, then* McOE *is defined by the algorithms* ***EncryptAuthenticate*** *and* ***DecryptAuthenticate*** *given in Table 5.*

(ii) 'TS'. *Let $M, C \in D_N^L || \{0,1\}^{l^*}$ for some integers L and l^*, $0 < l^* < n$, then* McOE*/TS is defined by the algorithms* ***EncryptAuthenticateSplitTag*** *and* ***DecryptAuthenticateSplitTag*** *given in Table 5.*

We now proceed to show the security of McOE. For this we use the results of Theorem 1 and show the INT-CTXT and RK-CPA-PRP security separately.

Theorem 2.

(i) *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a* McOE *scheme as in Definition 6 (i),* i.e. *$\mathcal{K}$ is the key derivation function, $\mathcal{E} =$* **EncryptAuthenticate** *and $\mathcal{D} =$* **DecryptAuthenticate***. We further assume that the block cipher E is secure against related key attacks. Then*

$$\mathbf{Adv}_{\Pi}^{\text{CCA3(AOE,NI)}}(q, \ell, t) \leq \frac{2(q+\ell)(q+\ell+1) + 3q + 2\ell}{2^n - (q+\ell)} + 2\mathbf{Adv}_{E}^{\text{RK-CCA-PRP}}(q+\ell).$$

(ii) *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a* McOE *scheme as in Definition 6 (ii),* i.e. *$\mathcal{K}$ is the key derivation function, $\mathcal{E} =$* **EncryptAuthenticateSplitTag** *and $\mathcal{D} =$* **EncryptAuthenticateSplitTag***. We further assume that the block cipher E is secure against related key attacks. Then*

$$\mathbf{Adv}_{\Pi}^{\text{CCA3(AOE,NI)}}(q, \ell, t) \leq \frac{2(q+\ell)(q+\ell+1) + 3q + 2\ell}{2^n - (2q+\ell)} + 2\mathbf{Adv}_{E}^{\text{RK-CCA-PRP}}(2q+\ell) + \frac{q}{2^{n/2} - (2q+\ell)}.$$

Proof. The proof of (i) follows from Theorem 1 together with Lemmas 1 and 2. Similarly, (ii) also follows from Theorem 1 together with Lemmas 3 and 4. Lemmas 3 and 4 are stated – due to the lack of space without complete proofs – in Appendix C.

Lemma 1. *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a* McOE *scheme as in Definition 6 (i). Let q be the number of total queries an adversary A is allowed to ask and ℓ be an integer representing the total length in blocks of the queries to $\mathcal{E}$ and $\mathcal{D}$. Then,*

$$\mathbf{Adv}_{\Pi}^{\text{INT-CTXT(NI)}}(q, \ell, t) \leq \frac{(q+\ell)(q+\ell+1)}{2^n - (q+\ell)} + \frac{2q+\ell}{2^n - (q+\ell)} + \mathbf{Adv}_{E}^{\text{RK-CCA-PRP}}(q+\ell).$$

Proof (Lemma 1). Our bound is derived by game playing arguments. Consider games G_1-G_3 of Figure 4 and a fixed adversary A asking at most q queries with a total length of at most ℓ blocks. The functions **Initialize** and **Finalize** are identical for all games in this proof. Lets denote G_0 as the Game INT-CTXT(NI) as defined in Figure 3. Definition 5 states that

$$\mathbf{Adv}_{\Pi}^{\text{INT-CTXT(NI)}}(A) \leq \Pr[A^{G_0} \Rightarrow 1].$$

```
Initialize()
  (K, L) ←$ K();
  B ← {0^v};

Encrypt(H, M) Game G_1
  A ← A ∪ H_0;
  τ ← E_{φ(K,0^n)}(H_0);
  U ← φ(K, H_0 ⊕ τ);
  for i from 1 to L_H do
    U ← φ(K, H_i ⊕ E_U(H_i));
  for i from 1 to L do
    C_i ← E_U(M_i);
    U ← φ(K, C_i ⊕ M_i);
  T ← E_U(τ);
  Q ← (H, M, C, T);
  return (C_1, ..., C_L, T);

Encrypt(H, M) Game G_2, [G_3]
  A ← A ∪ H_0;
  p ← LLCP_n(Q_{|H,M}, (H, M));
  τ ← E_{φ(K,0^n)}(H_0);
  U ← φ(K, H_0 ⊕ τ);
  if(p = 0 and U ∈ B) then
    bad ← true; [U ←$ {0,1}^n \ B;]
  B ← B ∪ U;
  for i = 1, ..., L_H do
    U ← φ(K, H_i ⊕ E_U(H_i));
    if (U ∈ B and i ≥ p) then
      bad ← true; [U ←$ {0,1}^n \ B;]
    B ← B ∪ U;
  for i = 1, ..., L do
    C_i ← E_U(M_i);
    U ← φ(K, C_i ⊕ M_i);
    if (U ∈ B and i ≥ p) then
      bad ← true; [U ←$ {0,1}^n \ B;]
    B ← B ∪ U;
  T ← E_U(τ);
  Q ← (H, M, C, T);
  return (C_1, ..., C_L, T);

Finalize()
  return win;

Verify(H, C, T) Game G_1
  A ← A ∪ H_0;
  τ ← E_{φ(K,0^n)}(H_0);
  U ← φ(K, H_0 ⊕ τ);
  for i from 1 to L_H do
    U ← φ(K, H_i ⊕ E_U(H_i));
  for i from 1 to L do
    M_i ← E_U^{-1}(C_i);
    U ← φ(K, C_i ⊕ M_i);
  T' ← E_U(τ);
  if(T = T' and (H, C) ∉ Q_{|H,C}) then
    win ← true;
  Q ← (H, ⊥, C, ⊥);
  A ← A ∪ H_0;
  return (T = T');

Verify(H, C, T) Game G_2, [G_3]
  p ← LLCP_n(Q_{|H,M}, (H, M));
  τ ← E_{φ(K,0^n)}(H_0);
  U ← φ(K, H_0 ⊕ τ);
  if(p = 0 and U ∈ B) then
    bad ← true; [U ←$ {0,1}^n \ B;]
  B ← B ∪ U;
  for i = 1, ..., L_H do
    U ← φ(K, H_i ⊕ E_U(H_i));
    if (U ∈ B and i ≥ p) then
      bad ← true; [U ←$ {0,1}^n \ B;]
    B ← B ∪ U;
  for i = 1, ..., L − 1 do
    M_i ← E_U^{-1}(C_i);
    U ← φ(K, C_i ⊕ M_i);
    if (U ∈ B and i ≥ p) then
      bad ← true; [U ←$ {0,1}^n \ B;]
    B ← B ∪ U;
  M_L ← E_U^{-1}(C_L);
  U ← φ(K, C_L ⊕ M_L);
  if (U ∈ B and H_0 ∉ A) then
    bad ← true; [U ←$ {0,1}^n \ B;]
  T' ← E_U(τ);
  if (T = T' and (H, C, T) ∉ Q_{|H,C,T}) then
    win ← true;
  Q ← (H, ⊥, C, ⊥);
  B ← B ∪ U;
  return (T = T');
```

Fig. 4. Games G_1-G_3 for the proof of Lemma 1. Game G_3 contains the code in the box while G_2 does not.

In G_1, the encryption and verify placeholders are replaced by their specific McOE counterparts as of Definition 6. Clearly, $\Pr[A^{G_0} \Rightarrow 1] = \Pr[A^{G_1} \Rightarrow 1]$. We now discuss the differences between G_1 and G_2. The set B is initialized to $\{0^v\}$ and then collects new key-input values U which are computed during the encryption or verification process (in lines 206, 211, 217, 227, 232, 238 and 243). We note that, since φ is injective, a collision for the chaining values follows if there is a collision in the U values.

In lines 202 and 223, the function LLCP_n is invoked. Finally, the variable `bad` is set to `true` if one of the if-Conditions in lines 205, 211, 216, 226, 231, 237 or 242 is `true`. *None* of these modifications affect the values returned to the adversary and therefore

$$\Pr[A^{G_1} \Rightarrow 1] = \Pr[A^{G_2} \Rightarrow 1].$$

For our further discussion we require another game G_4 which is explained in more detail later in this proof. [3] It follows that

$$\begin{aligned}
\Pr[A^{G_2} \Rightarrow 1] &= \Pr[A^{G_3} \Rightarrow 1] + |\Pr[A^{G_2} \Rightarrow 1] - \Pr[A^{G_3} \Rightarrow 1| \\
&\leq \Pr[A^{G_3} \Rightarrow 1] + \Pr[A^{G_3} \mathit{sets}\ \texttt{bad}] \\
&\leq \Pr[A^{G_4} \Rightarrow 1] + |\Pr[A^{G_3} \Rightarrow 1] - \Pr[A^{G_4} \Rightarrow 1]| + \Pr[A^{G_3} \mathit{sets}\ \texttt{bad}] \qquad (3)
\end{aligned}$$

We now proceed to upper bound any of the three terms contained in (3) – in right to left order. The success probability of game G_3 does not differ from the success probability of G_2 unless a chaining value U occurs twice. In this case, the adversary must (i) either have 'found' a collision for $E_{\varphi(K,A)}(B) \oplus B$, *i.e.* she stumble over (A, B) and (A', B') such that $E_{\varphi(K,A)}(B) \oplus B = E_{\varphi(K,A')}(B') \oplus B'$ or, (ii), must have found a preimage of 0^n, which is always the starting point of our chain. Note that that value 0^n is initially stored in the set B. In both cases, the variable `bad` would have been set to `true`, and it follows [6] that

$$\Pr[A^{G_3} \mathit{sets}\ \texttt{bad}] \leq \frac{(q+\ell)(q+\ell+1)}{2^n - (q+\ell)} + \frac{q+\ell}{2^n - (q+\ell)}$$

We now describe the new game G_4. It is equal to G_3 *except* that the block cipher E (and E^{-1}) is replaced by a randomly chosen functions **EncryptBlock** (and **DecryptBlock**), which is modeled as a pseudo random permutation (PRP). We assume that they are implemented via lazy sampling. More precisely, the call $E_K(A)$ is replaced by an invocation of $\mathbf{EncryptBlock}_K(A)$ and the call $E_K^{-1}(A)$ is replaced by an invocation of $\mathbf{DecryptBlock}_K(A)$. We now upper bound the difference between G_3 and G_4. So, by definition of G_4, we have

$$|\Pr[A^{G_3} \Rightarrow 1] - \Pr[A^{G_4} \Rightarrow 1]| \leq \mathbf{Adv}_E^{\text{RK-CCA-PRP}}(q+\ell).$$

Finally, we have to upper bound the advantage for the adversary A to win the game G_4. A can only win this game if the condition in line 245 (resp. 445 for game G_4) is `true`. As usual, we assume *wlog.* that A doesn't ask a question if the answer is already known which implies that $(H, C, T) \notin \mathcal{Q}_{|H,C,T}$. For our analysis we distinguish between three cases. So we formally adjust line 244 (*i.e.* choose as the tag computation operation either E or E^{-1}) such that we always have enough randomness left for our result.

[3] Since the difference is very minor, we do not provide an extra figure.

Case 1: H_0 *has already been used in a Encrypt or Verify query before and* $U \in B$. Since we already have computed T' in the past, the chance of success is upper bounded by $\Pr[E_U^{-1}(T) = T']$ which can be upper bounded by $1/(2^n - (q + \ell))$.

Case 2: H_0 *has never been used before, also* U *has never been used as a chaining value.* Then the tagging operation uses a 'new key' – essentially due since φ is injective – and therefore the output of $E_U(\tau)$ is uniformly distributed and the success probability is $\leq 1/2^n$.

Case 3: $H_0 \in A$ *but* U *has never been used as a chaining value.* the chance of success is upper bounded by $\Pr[E_U^{-1}(T) = T']$ which can be upper bounded by $1/2^n$.

Note that the 'missing' fourth case has been explicitly excluded by line 242 (resp. 442). Since these three cases are mutually exclusive, we can upper bound the success probability for q queries as

$$\Pr[A^{G_4} \Rightarrow 1]| \leq \frac{q}{2^n - (q + \ell)}.$$

Our claim follows by adding up the individual bounds. □

Lemma 2. *Let* $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ *be a* McOE *scheme as in Definition 6 (i). Let* q *be the number of total queries an adversary* A *is allowed to ask and* ℓ *be an integer representing the total length of the queries to* $\mathcal{E}$ *and* $\mathcal{D}$*. Then,*

$$\mathbf{Adv}_{\Pi}^{\text{CPA(AOE,NI)}}(q, \ell, t) \leq \frac{(q + \ell)(q + \ell + 1)}{2^n - (q + \ell)} + \frac{q + \ell}{2^n - (q + \ell)} + \mathbf{Adv}_{E}^{\text{RK-CPA-PRP}}(q + \ell)$$

Proof (of Lemma 2). Our bound is derived by game playing arguments. Consider games G_1 and G_2 of Figure 5. The functions **Initialize** and **Finalize** are identical for any of those games.

At first we investigate the differences between the $CPA(\text{AOE}, \text{NI})$ game from Figure 2 and G_1 from Figure 5. In G_1 we have replaced $\mathcal{E}$ by its definition of McOE, and $\$^w$ by an on-line encryption oracle that just models a 'perfect' OPRP, *i.e.* for two plaintexts with an equal prefix it returns two ciphertexts that also share a prefix of the same length. We again assume this oracle to be implemented by lazy sampling. Then, set B does only collect new chaining values (lines 106, 111, 122) in order to intercept the occurrence of two equal chaining values which do lead – due to the injectivity of φ – to two equal keys for the encryption of a block.

In line 101, the oracle LLCP_n^* is invoked returning (j, p) where p denotes the length of the prefix determined via LLCP_n and j denotes the index in $\mathcal{Q}_{|H,M}$. In line 113, we denote by C_i^j the i-th n-bit block of the j-th entry in $\mathcal{Q}_{|C}$. If a message block M_i is part of the longest common prefix, we look up the ciphertext in the query history. Finally, the variable `bad` is set to `true` if (one of) the conditions of lines 104/204, 109/209 or 111/211 is `true`. These changes do not affect the success probability of an adversary, because the output of the oracle remains unchanged. More precisely, the distribution of the output does not change. This means that – using a new game G_3 described shortly –

$$\mathbf{Adv}_{\Pi}^{\text{CPA(AOE, NI)}}(A) = 2 \cdot |\Pr[A^{G_1} \Rightarrow 1] - 0.5|.$$

Therefore, by common game playing arguments,

$$\begin{aligned}
\Pr[A^{G_1} \Rightarrow 1] &= \Pr[A^{G_2} \Rightarrow 1] + |\Pr[A^{G_1} \Rightarrow 1] - \Pr[A^{G_2} \Rightarrow 1| \\
&\leq \Pr[A^{G_2} \Rightarrow 1] + \Pr[A^{G_2} \textit{sets } \texttt{bad}] \\
&\leq \Pr[A^{G_3} \Rightarrow 1] + |\Pr[A^{G_2} \Rightarrow 1] - \Pr[A^{G_3} \Rightarrow 1]| + \Pr[A^{G_2} \textit{sets } \texttt{bad}].
\end{aligned}$$

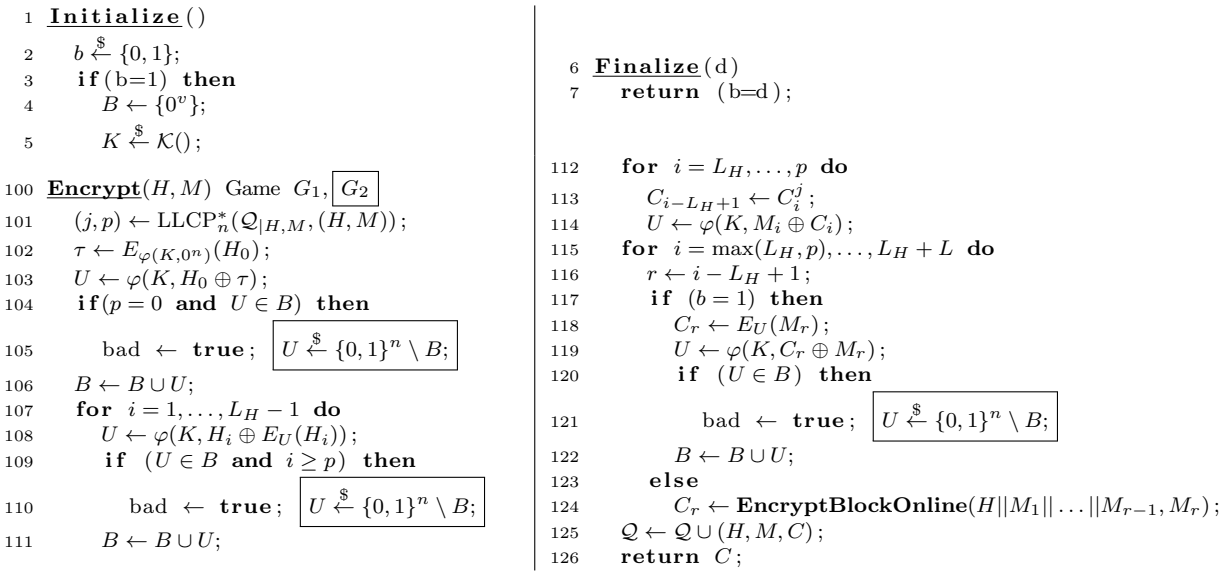

Fig. 5. Games G_1 and G_2 for the proof of Lemma 2. Game G_2 contains the code in the box while G_1 does not.

The success probability of game G_2 does not differ from the success probability of G_1 unless a chaining value U occurs twice. In this case, the adversary must either have found a collision for $E_{\varphi(K,A)}(B) \oplus B$, *i.e.* she has found (A, B) and (A', B') such that $E_{\varphi(K,A)}(B) \oplus B = E_{\varphi(K,A')}(B') \oplus B'$ or must have found a preimage of 0^n. In both cases, the variable `bad` would have been set to `true`, and it follows again by [6] that

$$\Pr[A^{G_2} sets \texttt{ bad}] \leq \frac{(q+\ell)(q+\ell+1)}{2^n - (q+\ell)} + \frac{q+\ell}{2^n - (q+\ell)}$$

The aforementioned new game G_3 is equal to the game G_2 *except* that the block cipher E (and E^{-1}) is replaced by a randomly chosen functions **EncryptBlock** (and **DecryptBlock**), which is modeled as a pseudo random permutation (PRP). We assume that they are implemented via lazy sampling. More precisely, the call $E_K(A)$ is replaced by an invocation of $\mathbf{EncryptBlock}_K(A)$ and the call $E_K^{-1}(A)$ is replaced by an invocation of $\mathbf{DecryptBlock}_K(A)$. We now upper bound the difference between G_3 and G_4. So, by definition of G_4, we have

$$|\Pr[A^{G_2} \Rightarrow 1] - \Pr[A^{G_3} \Rightarrow 1]| \leq \mathbf{Adv}_E^{\text{RK-CPA-PRP}}(q+\ell).$$

Finally, we have to upper bound the advantage for an adversary A to win the game G_3. Since the U cannot collide and it is not possible to compute a preimage for any query, the algorithm for $b = 0$ *is* an OPRP, and therefore the success probability to win G_3 for any adversary is 0.5, *i.e.* she has not advantage in winning this game. □

6 Discussion and Conclusion

Originally, this research has been inspired by the search for a default authenticated encryption mode of operation for a general-purpose cryptographic library. It should offer, by default, a huge

failure tolerance for practical software developers. Since the well-known schemes as, *e.g.*, OCB and SIV, have not fit our requirements, we searched for other ways to achieve the security and functionality we were looking for. Apart from McOE, generic composition (Encrypt-then-Mac) of a secure on-line cipher for encryption and a secure deterministic MAC for authentication, using two independent keys might be another solution. As it turned out, using McOE, one can save the additional key and the time to generate the MAC by using a slightly tweaked on-line cipher for both encryption and authentication.

Interestingly, on our way to designing McOE, we had to review and question some widespread cryptographic views. The first one being that we, in fact, require some sort of related key security for a *practical* scheme. This draws attention especially to AES where the practical relevance of related-key attacks is still debated. Using a tweakable block cipher with sufficiently long tweak, we can remove the related-key requirement.

Another outcome was that, actually even in the context of hardware accelerated cryptography, the search for high-performance ciphers is not over. It is remarkable that software Threefish-512 performs considerably better as hardware AES-128. This seems to be mainly due to the slow key scheduler hardware implementation of AES-128 and the shorter block size. Naturally, this leads to the question whether a 128 bit block size might be somewhat disadvantageous in nowadays big-memory environments.

References

[1] Mihir Bellare, Alexandra Boldyreva, Lars R. Knudsen, and Chanathip Namprempre. Online Ciphers and the Hash-CBC Construction. In *CRYPTO*, pages 292–309, 2001.

[2] Mihir Bellare, Alexandra Boldyreva, Lars R. Knudsen, and Chanathip Namprempre. Online Ciphers and the Hash-CBC Construction. Cryptology ePrint Archive, Report 2007/197; full version of [1], 2007. `http://eprint.iacr.org/`.

[3] Mihir Bellare and Chanathip Namprempre. Authenticated Encryption: Relations among Notions and Analysis of the Generic Composition Paradigm. *J. Cryptology*, 21(4):469–491, 2008.

[4] Mihir Bellare and Phillip Rogaway. Encode-Then-Encipher Encryption: How to Exploit Nonces or Redundancy in Plaintexts for Efficient Cryptography. In *ASIACRYPT*, pages 317–330, 2000.

[5] Mihir Bellare, Phillip Rogaway, and David Wagner. The EAX Mode of Operation. In *FSE*, pages 389–407, 2004.

[6] John Black, Phillip Rogaway, and Thomas Shrimpton. Black-Box Analysis of the Block-Cipher-Based Hash-Function Constructions from PGV. In *CRYPTO*, pages 320–335, 2002.

[7] Nikita Borisov, Ian Goldberg, and David Wagner. Intercepting Mobile Communications: The Insecurity of 802.11. In *MOBICOM*, pages 180–189, 2001.

[8] Enrico Buonanno, Jonathan Katz, and Moti Yung. Incremental Unforgeable Encryption. In *FSE*, pages 109–124, 2001.

[9] Joan Daemen. *Hash Function and Cipher Design: Strategies Based on Linear and Differential Cryptanalysis.* Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium, March 1995.

[10] Morris Dworkin. *Special Publication 800-38C: Recommendation for block cipher modes of operation: the CCM mode for authentication and confidentiality.* National Institute of Standards and Technology, U.S. Department of Commerce, May 2005.

[11] Pierre-Alain Fouque, Gwenaëlle Martinet, Frédéric Valette, and Sébastien Zimmer. On the Security of the CCM Encryption Mode and of a Slight Variant. In *ACNS*, pages 411–428, 2008.

[12] Brian Gladman. Brian Gladman's AES Implementation, 19th June 2006. `http://gladman.plushost.co.uk/oldsite/AES/index.php`.

[13] Virgil D. Gligor and Pompiliu Donescu. Fast Encryption and Authentication: XCBC Encryption and XECB Authentication Modes. In *FSE*, pages 92–108, 2001.

[14] Shafi Goldwasser and Silvio Micali. Probabilistic Encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[15] George Hotz. Console Hacking 2010 - PS3 Epic Fail. 27th Chaos Communications Congress, 2010. http://events.ccc.de/congress/2010/Fahrplan/attachments/1780_27c3_console_hacking_2010.pdf.

[16] ISO/IEC. *19772:2009, Information technology – Security techniques – Authenticatd Encryption*, 2009.

[17] Tetsu Iwata. New Blockcipher Modes of Operation with Beyond the Birthday Bound Security. In *FSE*, pages 310–327, 2006.

[18] Tetsu Iwata. Authenticated Encryption Mode for Beyond the Birthday Bound Security. In *AFRICACRYPT*, pages 125–142, 2008.

[19] Tetsu Iwata and Kan Yasuda. BTM: A Single-Key, Inverse-Cipher-Free Mode for Deterministic Authenticated Encryption. In *Selected Areas in Cryptography*, pages 313–330, 2009.

[20] Tetsu Iwata and Kan Yasuda. HBS: A Single-Key Mode of Operation for Deterministic Authenticated Encryption. In *FSE*, pages 394–415, 2009.

[21] Charanjit S. Jutla. Encryption Modes with Almost Free Message Integrity. *J. Cryptology*, 21(4):547–578, 2008.

[22] Jonathan Katz and Moti Yung. Unforgeable Encryption and Chosen Ciphertext Secure Modes of Operation. In *FSE*, pages 284–299, 2000.

[23] Tadayoshi Kohno. Attacking and Repairing the WinZip Encryption Scheme. In *ACM Conference on Computer and Communications Security*, pages 72–81, 2004.

[24] Tadayoshi Kohno, John Viega, and Doug Whiting. CWC: A High-Performance Conventional Authenticated Encryption Mode. In *FSE*, pages 408–426, 2004.

[25] Ted Krovetz and Phillip Rogaway. New Blockcipher Modes of Operation with Beyond the Birthday Bound Security. In *FSE*, pages 310–327, 2006.

[26] Moses Liskov, Ronald L. Rivest, and David Wagner. Tweakable Block Ciphers. In *CRYPTO*, pages 31–46, 2002.

[27] Stefan Lucks. Ciphers secure against related-key attacks. In Bimal K. Roy and Willi Meier, editors, *FSE*, volume 3017 of *Lecture Notes in Computer Science*, pages 359–370. Springer, 2004.

[28] Stefan Lucks. Two-Pass Authenticated Encryption Faster Than Generic Composition. In *FSE*, pages 284–298, 2005.

[29] David A. McGrew and John Viega. The Security and Performance of the Galois/Counter Mode (GCM) of Operation. In *In INDOCRYPT, volume 3348 of LNCS*, pages 343–355. Springer, 2004.

[30] Niels Ferguson and Stefan Lucks and Bruce Schneier and Doug Whiting and Mihir Bellare and Tadayoshi Kohno and Jon Callas and Jesse Walker. Skein source code and test vectors. http://www.skein-hash.info/downloads.

[31] Kenneth G. Paterson and Gaven J. Watson. Plaintext-dependent decryption: A formal security treatment of ssh-ctr. In Henri Gilbert, editor, *EUROCRYPT*, volume 6110 of *Lecture Notes in Computer Science*, pages 345–361. Springer, 2010.

[32] Phillip Rogaway. Authenticated-Encryption with Associated-Data. In *ACM Conference on Computer and Communications Security*, pages 98–107, 2002.

[33] Phillip Rogaway. Efficient Instantiations of Tweakable Blockciphers and Refinements to Modes OCB and PMAC. In *ASIACRYPT*, pages 16–31, 2004.

[34] Phillip Rogaway. Nonce-Based Symmetric Encryption. In *FSE*, pages 348–359, 2004.

[35] Phillip Rogaway, Mihir Bellare, John Black, and Ted Krovetz. OCB: a block-cipher mode of operation for efficient authenticated encryption. In *ACM Conference on Computer and Communications Security*, pages 196–205, 2001.

[36] Phillip Rogaway and Thomas Shrimpton. A Provable-Security Treatment of the Key-Wrap Problem. In *EUROCRYPT*, pages 373–390, 2006.

[37] Phillip Rogaway and Thomas Shrimpton. Deterministic Authenticated-Encryption: A Provable-Security Treatment of the Key-Wrap Problem. Cryptology ePrint Archive, Report 2006/221; full version of [36], 2006. http://eprint.iacr.org/.

[38] Phillip Rogaway and Haibin Zhang. Online Ciphers from Tweakable Blockciphers. In *CT-RSA*, pages 237–249, 2011.

[39] Todd Sabin. Vulnerability in Windows NT's SYSKEY encryption. *BindView Security Advisory*, 1999. Availalbe at http://marc.info/?l=ntbugtraq&m=94537191024690&w=4.

[40] Hongjun Wu. The Misuse of RC4 in Microsoft Word and Excel. Cryptology ePrint Archive, Report 2005/007, 2005. http://eprint.iacr.org/.

A Misuse-Attacks: The weak point of current Authenticated Encryption (AE) Schemes.

A.1 Attacking Schemes without Claimed Resistance Against Nonce Reuse

Cipher-block-chaining (CBC) is an unauthenticated encryption mode which is sometimes used as the encryption component of an AE scheme. But one can easily distinguish CBC encryption from a good on-line cipher, if the nonce (or the IV) is constant. The attack from [1] only needs three chosen plaintexts. Counter mode, which has been very popular among the designers of AE schemes, fails terribly in nonce reuse settings, since it generates exactly the same keystream twice when the nonce is reused. It was to be expected that a scheme using counter mode or CBC inherits the nonce reuse issue from that mode. But, as it turned out, common AE schemes also fail at the authenticity frontier, see Table 2 for an overview. This is an unpleasant surprise, since the cryptographic community has known well deterministic MACs for a long time – so why is the authenticity provided by most authenticated encryption schemes so much more fragile than the authenticity provided by well-known MACs?

The following two attack patterns will be used in most of our attacks.

Repeated Keystream. Many AE schemes generate a keystream $S = F_K(V)$ of length $|M|$, depending on the secret key K and the nonce V, and encrypt the message M by computing the ciphertext $C = S \oplus M$, typically by applying a block cipher in counter mode. If the same nonce is used more than once, the following attack straightforwardly breaks the privacy:

- Encrypt a plaintext M under the nonce V to a ciphertext C with tag T.
- Encrypt a plaintext $M' \neq M$ under the same V to a ciphertext C' and a tag T'.
- It turns out that $C' = C \oplus M \oplus M'$ holds.

Linear Tag. Many AE schemes, which generate a keystream $S = F_K(V)$ as above, apply the encrypt-then-authenticate paradigm and allow to rewrite the authentication tag T as

$$T = f(V) \oplus g(C),$$

where V is the nonce, C is the ciphertext, and f and g are some key-dependent functions. This enables the adversary to mount the following attack:

- Encrypt the plaintext M under the nonce V to (C, T) with $T = f(V) \oplus g(C)$.
- Encrypt the plaintext $M' \neq M$ with $|M'| = |M|$ under the nonce $V' \neq V$ to (C', T') with the tag $T' = f(V') \oplus g(C')$.
- Set $M'' := M' \oplus C' \oplus C$. Encrypt M'' under the nonce V' to (C'', T''). Observe $C'' = C$, thus $T'' = f(V') \oplus g(C)$.
- Set $T^* = T \oplus T' \oplus T'' = f(V) \oplus g(C')$, The adversary accepts (C', T^*) under V.

Two-Pass AE(AD) Modes: CWC [24], GCM [29], CCM [10], EAX [5], CHM [17] . All the common two-pass AE(AD) modes, CHM,CWC, GCM, CCM and EAX, use the counter mode as the underlying encryption operation and are thus vulnerable to the repeated keystream attack pattern. Four of them, CHM, CWC, GCM, and EAX, are designed according to the encrypt-then-authenticate paradigm, and are thus vulnerable to the linear tag attack pattern. The designers of CCM followed authenticate-then-encrypt, which seems to defend against the linear tag pattern. Forgery attacks against CCM have been presented in [11], though.

Mixed AE(AD) Modes: RPC [8] and CCFB [28]. RPC combines counter mode and electronic codebook mode. Given an n-bit block cipher E under a key K and a c-bit counter cnt, RPC takes an $(n-c)$-bit plaintext block M_i and computes the ciphertext block $C_i := E_K(M_i||(\text{cnt}+i) \bmod 2^c)$. Authentication is performed locally for each ciphertext block: During decryption, RPC computes $(M_i||X_i) = E_K^{-1}(C_i)$ and accepts M_i as authentic if and only if $X_i = (\text{cnt} + i) \bmod 2^c$. The nonce defines cnt.

Under nonce reuse, the same sequence $(\text{cnt} + i) \bmod 2^c$ of counter values is used for different messages. This makes it easy to attack the privacy – essentially, when encrypting messages of m $(n-c)$-bit blocks, RPC degrades into m independent electronic codebooks. Also, given two authentic ciphertexts, $(C_1^0, \ldots, C_L^0)$ and $(C_1^1, \ldots, C_L^1)$, any ciphertext $(C_1^{\sigma(1)}, \ldots, C_L^{\sigma(L)})$ with $\sigma(i) \in \{0, 1\}$ is valid, since authenticity is verified locally for each $C_i^{\sigma(i)}$.

Similarly to RPC, CCFB is a combination of Counter and CFB mode. Unlike RPC, CCFB generates a single "global" authentication tag. Variants of the repeated keystream pattern applies and the linear tag pattern apply to CCFB.

One-Pass AE(AD) Modes: IAPM [21], OCB1[35], OCB2[33], OCB3[25], TAE [26]. Given a nonce V and a secret key K, IAPM [21] encrypts a plaintext $(M_1, \ldots, M_m)$ to a ciphertext $(C_1, \ldots C_m)$ and an authentication tag T as follows.

Initial step: Generate $m+2$ values $s_0, s_1, \ldots s_{m+1}$ depending on V and K, but not on the plaintext $(M_1, \ldots, M_m)$.
Encryption: $x_0 := V$; For $i \in \{1, \ldots, m\}$: $C_i := E_K(M_i \oplus s_i) \oplus s_i$.
Authentication tag: $T := E_K(s_{m+1} \oplus \sum_{1 \le i \le m} M_i) \oplus s_0$.

Similarly to RPC, IAPM behaves like a set of m independent electronic codebooks and is vulnerable to the same distinguishing attack. A forgery can exploit the fact that two different same-length messages $(M_1, \ldots, M_m)$ and $(M'_1, \ldots, M'_m)$, encrypted under the same nonce, have the same authentication tag $T = E_K(s_{m+1} \oplus \sum_{1 \le i \le m} M_i) \oplus s_0 = E_K(s_{m+1} \oplus \sum_{1 \le i \le m} M'_i) \oplus s_0$ if $\sum_{1 \le i \le m} M_i = \sum_{1 \le i \le m} M'_i$.

As much as our attacks are concerned, OCB1–3 and TAE are quite similar to IAPM, and the attacks are the same.

More One-Pass Modes: IACBC [21] and XCBC [13]. Given a nonce V and a secret key K, IACBC [21] encrypts $(M_1, \ldots, M_m)$ to $(C_1, \ldots C_m)$ and an authentication tag T as follows.

Initial step: Generate $m+1$ values $s_0, s_1, \ldots s_m$ depending on V and K, but not on the plaintext $(M_1, \ldots, M_m)$.
Encryption: $x_0 := V$; For $i \in \{1, \ldots, m\}$: $x_i := E_K(M_i \oplus x_{i-1})$, $C_i := x_i \oplus s_i$.
Authentication tag: $T := E_K(x_m \oplus \sum_{1 \le i \le m} M_i) \oplus s_0$.

The following nonce-reuse attack distinguishes IACBC encryption from an online permutation and also provides an existential forgery. For simplicity, we only consider 1-block messages $V \neq W$, which we also use as nonces: Encrypt W under V to (C_1, T). Encrypt V under W to (C'_1, T'). Encrypt V under V to (C''_1, T''). Set $C'''_1 := C_1 \oplus C'_1 \oplus C''_1$ and $T''' := T \oplus T' \oplus T''$.
(C'''_1, T) is a valid encryption of W under W.

Given a nonce V and secret keys K and K', XCBC encrypts a plaintext $(M_1, \ldots, M_m)$ to a ciphertext $(C_1, \ldots C_m)$ and an authentication tag T as follows.

Initial step: Generate $m+1$ values $s_1, \ldots s_{m+1}$ depending on V and K, but not on the plaintext $(M_1, \ldots, M_m)$.

Encryption:

1. $C_0 := E_K(V)$; $x_0 := E_{K'}(V)$;
2. Generate an additional message word $M_{m+1} := x_0 \oplus M_1 \oplus \cdots \oplus M_m$ for authentication.
3. For $i \in \{1, \ldots, m+1\}$: $x_i := E_K(M_i \oplus x_{i-1})$, $C_i := (x_i + s_i) \bmod 2^n$.

The best attack we have found for XCBC is not quite as damaging as the attacks on the other schemes, as the attack workload is at $O(2^{n/4})$, and the attack only provides a distinguisher, not a forger. For this reuse-nonce chosen-plaintext attack, we ignore the authentication tag: Generate $2^{n/4}$ encryptions of messages M_1^i under a nonce V to C_1^i. Statistically, expect one pair $i \neq j$ such that the least significant $n/2$ bits of C_1^i are identical to the least significant $n/2$ bits of C_1^j. Generate $2^{n/4}$ encryptions of messages (M_1^i, M_2^k) and (M_1^j, M_2^ℓ) under V to (C_1^i, C_2^k) and (C_1^j, C_2^ℓ), where the least significant $n/2$ bits of M_2^k and M_2^ℓ are the same. (Statistically, expect one pair $k \neq \ell$ such that $C_2^k = C_2^\ell$ holds.) Choose an arbitrary M_3. Encrypt (M_1^i, M_2^k, M_3) and (M_1^j, M_2^ℓ, M_3) under V to $(C_1^i, C_2^k, C_3^{i,k})$ and $(C_1^j, C_2^\ell, C_3 j, \ell)$. Observe $C_3^{i,k} = C_3^{j,\ell}$.

A.2 Offline Schemes, Defeating Nonce-Reuse (SIV [36], HBS [20], BTM [19])

Given a nonce N, a message M and associated Data H, these schemes perform two steps:

1. Generate the authentication tag T from H, M, and N.
2. Encrypt M in counter mode, using T as the nonce.

This is inherently offline, because one must finish step 1 before one can start step 2. All of SIV, HBS, and BTM perform counter mode encryption, but employ different MACs schemes to generate the tag T.

This usage of the counter mode is vulnerable in an *online decryption misuse* case, where, during decryption, a would-be plaintext is compromised before the tag has been verified. A chosen-ciphertext adversary can exploit that to determine an unknown keystream and then to decrypt an unknown message.

Another misuse case may apply when nonce-reuse is possible and the sender reads the message twice, once for each of the two steps – if there is any chance that *the message has been modified* between the two read operations.

Note that both misuse cases become quite harmless if one replaces the counter mode encryption by the application of an online permutation.

B Proof of Theorem 1

Consider games G_0, G_1, G_2 of Figure 6. For a fixed CCA3(ω, ν) adversary A on the scheme Π it holds that

$$\begin{aligned}\Pr[A_\Pi^{\mathrm{CCA3}(\omega,\nu)} \Rightarrow 1] &= \Pr[A^{G_0} \Rightarrow 1] \\ &= \Pr[A^{G_1} \Rightarrow 1] + (\Pr[A^{G_0} \Rightarrow 1] - \Pr[A^{G_1} \Rightarrow \texttt{true}]) \\ &\leq \Pr[A^{G_1} \Rightarrow 1] + \Pr[A^{G_1} \textit{ sets } \texttt{bad}].\end{aligned}$$

Since the `Decrypt` oraclees of G_1 and G_2 always return $\perp$,

$$\Pr[A^{G_1} \Rightarrow 1] = \Pr[A^{G_2} \Rightarrow 1].$$

Now, we design two adversaries A_c and A_p so that

$$\Pr[A^{G_1} \textit{sets } \texttt{bad}] \leq \Pr[A_c\, {}^{\text{INT-CTXT}(\omega,\nu)}_{\Pi} \Rightarrow 1] \text{ and}$$
$$\Pr[A^{G_2} \Rightarrow 1] \leq \Pr[A_p\, {}^{\text{CPA}(\omega,\nu)}_{\Pi} \Rightarrow 1].$$

A_p: Adversary A_p simply runs A answering A's **Encrypt** queries using its own **Encrypt** oracle, and answers **Decrypt** queries with $\perp$. A_p outputs whatever A outputs.

A_c: Adversary A_c runs A answering A's **Encrypt** queries using its own `Encrypt` oracle. It submits A's **Decrypt** queries to it's **Verify** oracle and, regardless of the response, returns $\perp$. Note that the **Verify** oracle sets `win` to **true** if and only if a fresh **Decrypt** query is valid. Just such a query would set the variable `bad` to **true**. □

```
1  Initialize(ω, ν)
2    b ←$ {0,1};
3    if (b=1) then
4      K ← K();

5  Finalize(d)
6    return (d=b);
```

```
7  Encrypt(H,M)
8    if (ν = NR and H_0 ∈ B) then
9      return ⊥;
10   else
11     B ← B ∪ {H_0};
12   if (b=1) then
13     C ← E_K(H,M);
14   else
15     C ← $^ω(H,M);
16   Q ← Q ∪ {(H,C)};
17   return C;
```

```
100 Decrypt(H,C)        Game G_0, [G_1]
101   M ← ⊥;
102   if ((H,C) ∉ Q and b=1) then
103     M ← D_K(H,C);
104   if (M ≠ ⊥) then
105     bad ← true; [M ← ⊥;]
106   return M;

200 Decrypt(H,C)        Game G_2
201   return ⊥;
```

Fig. 6. Games G_0, G_1 and G_2 for the proof of Theorem 1. Game G_1 contains the code in the box while G_0 does not. H_0 denotes the first block of the header representing the nonce/initial value.

C Proof Sketch for McOE (Tag Splitting)

Lemma 3. *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a* McOE *scheme as in Definition 6 (i). Let q be the number of total queries an adversary A is allowed to ask and ℓ be an integer representing the total length in blocks of the queries to $\mathcal{E}$ and $\mathcal{D}$. Then,*

$$\mathbf{Adv}^{\text{INT-CTXT(NI)}}_{\Pi}(q,\ell,t) \leq \frac{(2q+\ell)(2q+\ell+1)}{2^n-(2q+\ell)} + \frac{2q+\ell}{2^n-(2q+\ell)} + \mathbf{Adv}^{\text{RK-CCA-PRP}}_{E}(2q+\ell) + \frac{q}{2^{n/2}-(2q+\ell)}.$$

Proof. We here give a sketch of the core proof component differing from the proof of Lemma 1. Apart from some technical details, the tag verification arguments are different. For simple reference, we denote the last two chaining values as U_L and U_{L+1}. The success probability of an adversary's **Verify** query, *i.e.*'guessing' the correct tag is now outlined by a case analysis. A new set – in addition to A and B – is introduced which records any encryption processes of τ, *i.e.* the last processing block of an encryption process.

Case 1. In this case, $U_L \in B$ and $U_{L+1} \in B$. Assume that no collision has occurred (since this would have already triggered a **bad** event otherwise). The ciphertext $T[n-l^* \dots n-1]||Z$ is rebuilt using the tag T obtained by the query and using the appropriate entry from the set D. We can then upper bound the the adversary's chance of success by $\Pr[E^{-1}_{U_L+1} = \tau] \leq 1/(2^n - (\ell + 2q))$.

Case 2. In this case, $U_{L+1} \notin B$. It follows that either $U_L \notin B$ or that the inputs of the first tagging computation are different from what has been requested before. So the adversary's chance of success that she has guessed the tag correctly is – for the first tagging-computation – upper bounded by $\Pr(\tau[0..n-l^*-1] = T[0..n-l^*-1]) \leq 1/(2^{n-l^*} - (\ell + 2q)$ and for the second tagging computation by $\Pr(\tau[n-l^*..n-1] = T[n-l^*..1]) \leq 1/(2^{l^*} - (\ell + 2q))$. We note that the output of these two queries is independent and, by definition, that any of those probabilities is always smaller or equal than 1. So the total chance of success can be upper bounded for q queries by

$$\min\left(1, \frac{q}{2^{n-l^*} - (\ell + 2q)}\right) \times \min\left(1, \frac{q}{2^{l^*} - (\ell + 2q)}\right) \leq \frac{q}{2^{n/2} - (\ell + 2q)}.$$

Remarks: This chance of success is worse than compared to the non-tag-splitting variant. But it suffices for our proof since roughly achieving the birthday bound is enough. It essentially comes down to the fact that one of the two tag-encryption blocks is used for at least $n/2$ authentication bits.

Lemma 4. *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a* McOE *scheme as in Definition 6 (i). Let q be the number of total queries an adversary A is allowed to ask and ℓ be an integer representing the total length of the queries to $\mathcal{E}$ and $\mathcal{D}$. Then,*

$$\mathbf{Adv}^{\mathrm{CPA}(\mathrm{AOE,NI})}_{\Pi}(q, \ell, t) \leq \frac{(q+\ell)(q+\ell+1)}{2^n - (2q+\ell)} + \frac{2q+\ell}{2^n - (2q+\ell)} + \mathbf{Adv}^{\text{RK-CPA-PRP}}_{E}(q+\ell).$$

Proof. The proof is very similar to the proof of Lemma 2. Due to the lack of of space it is skipped here and will be available in the full version of the paper.

EncryptAuthenticate(H, M)

1. $\tau \leftarrow E_{\varphi(K,0^n)}(H_0)$
2. $U \leftarrow \varphi(K, H_0 \oplus \tau)$
3. for i from 1 to L_H loop
   $tmp \leftarrow E_U(H_i)$
   $U \leftarrow \varphi(K, H_i \oplus tmp)$
   end loop
4. for i from 1 to L loop
   $C_i \leftarrow E_U(M_i)$
   $U \leftarrow \varphi(K, M_i \oplus C_i)$
   end loop
5. $T \leftarrow E_U(\tau)$
6. return $(C_1, \ldots, C_L, T)$

DecryptAuthenticate(H, C, T)

1. $\tau \leftarrow E_{\varphi(K,0^n)}(H_0)$
2. $U \leftarrow \varphi(K, H_0 \oplus \tau)$
3. for i from 1 to L_H loop
   $tmp \leftarrow E_U(H_i)$
   $U \leftarrow \varphi(K, H_i \oplus tmp)$
   end loop
4. for i from 1 to L loop
   $M_i \leftarrow E_U^{-1}(C_i)$
   $U \leftarrow \varphi(K, M_i \oplus C_i)$
   end loop
5. $T' \leftarrow E_U(\tau')$
6. if $T = T'$ then return $(M_1, \ldots, M_L)$
   else return $\perp$

EncryptAuthenticateSplitTag(H, M)

1. $\tau \leftarrow E_{\varphi(K,0^n)}(H_0)$
2. $U \leftarrow \varphi(K, H_0 \oplus \tau)$
3. for i from 1 to L_H loop
   $tmp \leftarrow E_U(H_i)$
   $U \leftarrow \varphi(K, H_i \oplus tmp)$
   end loop
4. for i from 1 to L loop
   $C_i \leftarrow E_U(M_i)$
   $U \leftarrow \varphi(K, M_i \oplus C_i)$
   end loop
5. $M^* \leftarrow M_{L+1} || \tau[0 \ldots n - l^* - 1]$
6. $C^* \leftarrow E_U(M^*)$
7. Parse $C_{L+1} || T[0 \ldots n - l^* - 1] \leftarrow C^*$
8. $U \leftarrow \varphi(K, M^* \oplus C^*)$
9. $C^{**} \leftarrow E_U(\tau)$
10. $T[n - l^* \ldots n - 1] \leftarrow C^{**}[0 \ldots l^* - 1]$
11. return $(C_1, \ldots, C_L, C^*_{L+1}, T)$

DecryptAuthenticateSplitTag(H, C, T)

1. $\tau \leftarrow E_{\varphi(K,0^n)}(H_0)$
2. $U \leftarrow \varphi(K, H_0 \oplus \tau)$
3. for i from 1 to L_H loop
   $tmp \leftarrow E_U(H_i)$
   $U \leftarrow \varphi(K, H_i \oplus tmp)$
   end loop
4. for i from 1 to L loop
   $M_i \leftarrow E_U^{-1}(C_i)$
   $U \leftarrow \varphi(K, M_i \oplus C_i)$
   end loop
5. $C^* \leftarrow C_{L+1} || T[0 \ldots n - l^* - 1]$
6. $M^* \leftarrow E_U^{-1}(M^*)$
7. Parse $M_{L+1} || \tau'[0 \ldots n - l^* - 1] \leftarrow M^*$
8. $U \leftarrow \varphi(K, M^* \oplus C^*)$
9. $C^{**} \leftarrow E_U(\tau)$
10. $T'[n - l^* \ldots n - 1] \leftarrow C^{**}[0 \ldots l^* - 1]$
11. if $\tau'[0 \ldots n - l^* - 1] = \tau[0 \ldots n - l^* - 1]$
    and $T'[n - l^* \ldots n - 1] = T[n - l^* \ldots n - 1]$
    then return $(C_1, \ldots, C_L, C_{L+1}, T)$
    else return $\perp$

Table 5. Instances of McOE: Left side is for messages whose size is evenly divisible by the block size n; Right side is for arbitrarily sized messages (TS-variant); see text for details