How to wrap it up – A formally verified proposal for the use of authenticated wrapping in PKCS#11

Full version (June 26, 2019)

Alexander Dax, Robert Künnemann, Sven Tangermann and Michael Backes

Abstract

Being the most widely used and comprehensive standard for hardware security modules, cryptographic tokens and smart cards, PKCS#11 has been the subject of academic study for years. PKCS#11 provides a key store that is separate from the application, so that, ideally, an application never sees a key in the clear. Again and again, researchers have pointed out the need for an import/export mechanism that ensures the integrity of the permissions associated to a key. With version 2.40, for the first time, the standard included authenticated deterministic encryption schemes. The interface to this operation is insecure, however, so that an application can get the key in the clear, subverting the purpose of using a hardware security module.

This work proposes a formal model for the secure use of authenticated deterministic encryption in PKCS#11, including concrete API changes to allow for secure policies to be implemented. Owing to the authenticated encryption mechanism, the policy we propose provides more functionality than any policy proposed so far and can be implemented without access to a random number generator. Our results cover modes of operation that rely on unique initialisation vectors (IVs), like GCM or CCM, but also modes that generate synthetic IVs. We furthermore provide a proof for the deduction soundness of our modelling of deterministic encryption in Böhl et. al.'s composable deduction soundness framework.

Contents

• *All authors are with the CISPA Helmholtz Center for Information Security at the Saarland Informatics Campus.*

Manuscript received May 6, 2019; minor revisions on May 10 and Jun 26.

1 INTRODUCTION

PKCS#11 is one of the *Public-Key Cryptography Standards* and was defined by RSA Security in 1994. By now, it is the most prevalent standard for operating hardware security modules (HSM), but also smart cards and cryptographic libraries. It defines an API intended to separate usage and storage of cryptographic secrets so that application code can only access these secrets indirectly, via handles. The hope is that HSMs provide a higher level of security than the multi-purpose machines running the respective application. This is reasonable: HSMs are designed for security and have less functionality and therefore a smaller attack surface, making them easier to secure. Consequently, PKCS#11 is used throughout the public-key infrastructure and the banking network.

In contrast to this stated goal, raising the level of security, many versions and configurations of PKCS#11 allow for attacks on the logical level [9, 13, 18, 10]. Here, a perfectly valid chain of commands leads to the exposure of sensitive key material to the application, defeating the purpose of separating the (possibly vulnerable) application from the (supposedly secure) hardware implementation — and thus defeating their purpose. Formal methods have been used to identify configurations that are secure [18, 10, 29]. In this context, a configuration or policy refers to a specification of the device's behaviour that implements a subset of the standard, e.g., PKCS#11 with the restriction that all keys generated must have a certain attribute set. In order to be secure, the two most functional secure policies [10, 29] either have to limit the ability to transfer keys between devices [29] or have some keys degrade in functionality after transfer, i.e., after transfer, they cannot be used for operations that were permitted prior to transfer [10]. Recent versions of PKCS#11 have adopted various security extensions (e.g., wrapping/unwrapping templates, 'wrap-with-trusted'), but none of these improve upon this lack of functionality. Fundamentally, the problem is that the export mechanism for keys (*key wrapping*, i.e., encrypting a key with another key) does not provide a way to authenticate the attributes or the role that a key should be imported with.

Authenticated encryption with associated data (AEAD) provides a solution to this problem [37]. AEAD was not available in 1994, when PKCS#11 was invented. Academic development started around 2000 [25], standardisation followed suit in 2004 [19]. With version 2.40, support for two AEAD schemes was finally added to the set of supported algorithms in PKCS#11, but as Steel pointed out [45], the interface that v2.40 provides allows for a two-time pad attack. The application is able to set the initialisation vector (IV). If it chooses to use the same IV twice, wrapping can be used to decrypt and obtain keys in the clear. Figure 1 depicts why this attack works. Both GCM and CCM are based on CTR-mode. If we leave out the computation of the message authentication tag, it is easy to see that any cyphertext can be decrypted by XORing it with the keystream that is deterministically generated from the IV. Requesting an encryption with the same IV is essentially a decryption without the authenticity check.

This attack demonstrates that the mere support of AEAD schemes is not enough, a suitable interface needs to be provided, too. Unfortunately, this is not a trivial task. As keys can be present on several devices at the same time, each

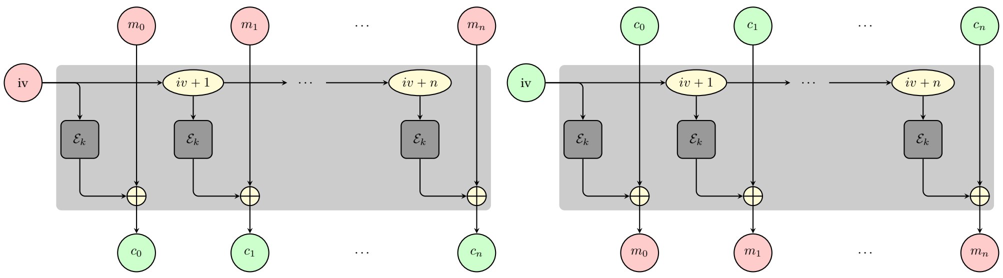


Fig. 1: Example on key extraction using CTR-mode. By supplying the same IV twice, the attacker can wrap a key and then encrypt the resulting wrapping, again using the same IV. This leads to the leakage of the key.

device individually needs to ensure that, globally, an IV is not used twice. Hence in this paper, we tackle the following questions:

I How can we guarantee global uniqueness even on devices that lack a random number generator (RNG)?
II Using authenticated encryption, is it possible to create a secure PKCS#11 configuration that is strictly more powerful than those proposed so far?

Contributions

The contributions of this paper can be summarized as follows:

1) We answer (I) and affirm (II) by proposing a secure PKCS#11 configuration that uses authenticated encryption.
2) We formally verify this proposal in the symbolic model and provide custom heuristics that allow for automated proof generation. These results apply to the previously proposed modes of operation GCM and CCM.
3) We put forward a *deduction soundness* [8] result, which is a necessary condition for computational soundness. It justifies the symbolic abstraction of AEAD and is of independent interest for protocol verification. Besides AEAD, it also supports hash functions, public-key cryptography, digital signatures and MAC.
4) The PKCS#11 technical committee considered SIV mode [41] as an alternative to GCM/CCM as it does not require an initialisation vector [44]. We derive a construction to obtain an AEAD scheme out of SIV mode (in fact, any deterministic authenticated encryption scheme). This construction cancels out if we use it in a particular way. With only slight syntactical modification to our model we can thus derive a similar policy for SIV mode while reusing the deduction soundness result, model and heuristics.

2 PKCS#11

PKCS#11 provides applications an interface to cryptographic implementations ranging from cryptographic libraries to smart cards and HSMs. Once an application establishes a *session* to a device (*slot* in PKCS#11 parlance), it identifies as a Security Officer (SO), or a normal user. The SO may initialise a slot and set a PIN for the normal user. Only if this PIN is set, the normal user can login. As we consider the case where the application or the host computer are malicious, we will abstract away from this and assume the attacker has complete control over a session.

PKCS#11 exposes so-called *objects*, e.g., keys and certificates, to the user or attacker. They are referred to indirectly, via *handles*. Handles do not reveal any information about the object they refer to. Objects have *attributes*, some of which are specific to their type (e.g., public keys of type `CKK_RSA` have a public exponent). Some however, are general for all keys, and control how they can be used. E.g.:

- `CKA_SENSITIVE` marks keys that ought not to be read out in the clear.
- `CKA_DECRYPT` marks keys that can be used to decrypt cyphertexts.
- `CKA_WRAP` marks wrapping keys: If `C_WrapKey` is given two handles, and the first has `CKA_WRAP`, it uses the key referred to by the first to encrypt the second. Wrapping is used to export keys. Additional constraints apply to the attributes associated to the second key, but we omit them for simplicity. To import, the function `C_UnwrapKey` takes a handle and a wrapping (the cyphertext resulting from `C_WrapKey`), decrypts the latter with the key referred to by the handle, stores the results and returns a handle to the newly generated object.

Typically, a given implementation supports only some of the functionality specified by PKCS#11, first, because the standard is extensive and contains many legacy algorithms, but also because the full standard is insecure. Clulow's attack provides a nice and concise example [13]:

1) A key is generated and marked `CKA_SENSITIVE`, `CKA_DECRYPT` and `CKA_WRAP`.
2) The key is used to wrap itself, obtaining an encryption of itself.
3) The key is used to decrypt the wrapping from the previous step, obtaining the key in the clear.

This attack and others have prompted vendors to limit the functionality offered by their respective implementations, which are often dubbed *configurations* or *policies*. Some vendors introduced proprietary functionality, e.g., marking `CKA_DECRYPT` and `CKA_WRAP` as conflicting, but even those were prone to attacks [10]. With version 2.20, wrapping and

function	description	rule	comment
Object management functions			
`C_CreateObject`	creates an object	(4)	only by SO during setup
`C_GetObjectSize`, `C_GetAttributeValue`	gather information about object	—	not useful to adversary
`C_FindObjects*`	find objects	—	not useful to adversary
`C_CopyObject`	creates a copy of an object	—	not useful to adversary (in this configuration)
`C_DestroyObject`	destroys an object	—	not useful to adversary
`C_SetAttributeValue`	modifies object's attribute	—	forbidden by policy
Key-management functions			
`C_GenerateKey`	generates a secret key	(3)	generated with level l and universally unique handle h
`C_DeriveKey`	derives a key from a base key	(9)	base keys and derived keys must have level 2, key-derivation needs random salt, universally unique handle h
`C_GenerateKeyPair`	generates a public / private key pair	—	always level 2; asymm. wrapping keys permits 'Trojan wrapped key attack', thus not modelled (only key-usage)
`C_WrapKey`	wraps (encrypts) a key	(7)	wrapping key must have larger level than argument key; internal IV generation (e.g., like `C_EncryptMessage`); authenticate level and handle as additional data
`C_UnwrapKey`	unwraps (decrypts) a key and stores it	(8)	level and handle of new key have to match additional authenticated data
Key-usage functions			
`C_Encrypt`	encrypts single-part data	(5)	require $l = 2$, internal IV generation
`C_Decrypt`	decrypts single-part data	(6)	require $l = 2$
message digest, signature, MAC, RNG etc.		—	require $l = 2$, not modelled (only key-usage)

TABLE 1: PKCS#11 operations for object and key management, and corresponding rules in our modelling (cf. Section 5).

unwrapping templates were introduced to control what keys can be wrapped, and what attributes objects created via unwrapping can obtain. While this mechanism provides more flexibility, it does not improve the expressiveness of policies – these templates can essentially be hard-coded in the device. In effect, two secure configurations were discovered that are incomparable to each other, but more or equally functional to the others found so far [10, 29].

The fundamental problem is that a wrapping does not contain authenticated information about the role of the key prior to the wrapping, i.e., its intended use. Hence, if it is possible to wrap two keys that have different roles, it is not clear which attributes a (re-)imported key should obtain upon unwrapping — it could originate from a key with either role. For instance, the first of the two most functional policies [10] allows for two different roles to be wrapped, but to be secure, the attributes obtained upon unwrapping provide less capabilities than either role — the keys 'degrade'. The second of them [29] allows only keys of a single role specific role to be wrapped.

Before OASIS took over standardisation from RSA Security in 2012, RSA drafted, but never published, version 2.30. Based on this draft, OASIS published version 2.40 in 2015, introducing support for AEAD schemes. AEAD schemes can be used for wrapping, which finally provides a way of authenticating a key's attributes upon wrapping. Unfortunately, the API requires the user to set the initialization vector, which allows for a simple attack where some vector is used twice [45]. The security of the schemes supported (AES-CCM and AES-GCM) relies on the uniqueness of the initialisation vector, hence the upcoming standard 3.00 is planned to support device-internal nonce generation for encryption/decryption.

The present work was motivated by the drafting of the standard. We announced our results on the OASIS PKCS#11 mailing list and stressed the need to support device-internal nonce generation for wrapping and encryption [17]. Assuming that support for this is present in v3.00, our policy provides a template for the secure use of PKCS#11. The current version at time of writing is version 2.40 with errata 01 [33]. The most recent proposal for PKCS#11 v3.00 is working draft 5 [34].

3 Policy

Our policy implements three central ideas: a key-hierarchy, globally unique counters and authentication of handles.

Key-hierarchy: keys are created with a given level, i.e., a natural number, and may only be used to wrap and unwrap keys of a lower level. If we extend this to payload data, we can assign level 1 to payload and level 2 to encryption keys that cannot wrap. Ergo, wrapping keys must have level 3 or higher. When wrapping a key, we authenticate the level of the enclosed key with the encryption. Upon unwrapping, this level is restored. To be consistent with that, decryption only succeeds if the cyphertext is authentic w.r.t. level 1. This already prevents Clulow's attack, as wrappings will never be decrypted, since whatever level the wrapping key was created with must be larger or equal to three.

Globally unique counters: The deduction soundness result that we will present in Section 7 holds only for protocols that guarantee that AEADs are created with a unique initialisation vector. This is necessary, as otherwise, for counter-mode based schemes like GCM and CCM, key-wrapping can immediately be used to decrypt. The simplest way to ensure this is to choose the IV randomly, however, many low-cost devices do not have a random number generator. We thus describe a secure low-cost alternative that is slightly more involved. We require each device to have a unique *device identifier* at initialisation time, e.g., a serial number with a unique vendor id. For every encryption, a running counter is increased, so that the combination of this unique public value and the running counter is unique in the network. Hence, even if a key is shared between two devices, the initialisation vector remains unique. Practically, this combination can be a simple concatenation: if the serial number and the counter have 64 bit, they match the blocksize of AES. For an HSM that can run 10M encryptions per second, it would take about 60'000 years to repeat a counter. In terms of the soundness of our deduction rules, any other way of combining those is sound, as long as it provides an injective mapping into the set of initialisation vectors (or is indistinguishable from one).

Authentication of handles: The third novelty to our policy is the authentication of handles. Usually, handles are assigned through a running counter or are simply the memory address where the key is physically stored. If a key is exported to another device, it most likely receives a new handle. Instead, we chose a unique handle at key-generation time, and ensure that, no matter on which device, this handle always resolves to the same key. We call this property *handle integrity*.

We discus the relevant parts of PKCS#11 in the follow-up. Table 1 gives an overview, see [34, Table 30] for the full list.

3.1 Object-management

The main security goal is to keep keys secret from the possibly malicious host. Hence, for the operation of the device, we disallow direct key imports via `C_CreateObject`. Nevertheless, in order to import keys via `C_UnwrapKey`, at least one key must be shared initially. A common practice is to have the security officer (SO) set up shared keys using `C_CreateObject`. Thus this function may only be used by the SO, which we assume happens only during setup or in an otherwise safe environment and only with trusted PKCS#11 tokens, i.e., tokens implementing our policy by vendors that guarantee the uniqueness of their unique device identifiers.

As the key-hierarchy is static, so are the attributes. We thus disable the function `C_SetAttributeValue` altogether.

We allow the user to inspect the device using functions like `C_GetObjectSize` and `C_FindObjects` and its siblings. As the adversary has full control, this information is redundant to him and w.l.o.g., we omit them from our model. Similar for `C_DestroyObject`. As our model assumed no limit on space for storing keys, any attack using it can be transformed into an attack that does not delete objects.

3.2 Key-management

In our policy, normal users can create new objects via `C_GenerateKey`, `C_GenerateKeyPair`, `C_DeriveKey` or `C_UnwrapKey`. `C_GenerateKey` and `C_GenerateKeyPair` create a new symmetric or asymmetric key, `C_DeriveKey` derives a new key from an existing one, and `C_UnwrapKey` decrypts a wrapping, i.e., an AEAD that was output by `C_WrapKey`, and imports its content. We thus consider these four functions plus `C_WrapKey` the key-management core of PKCS#11.

`C_GenerateKey` *and* `C_GenerateKeyPair`*:* Keys are generated and then stored with a level and a universally unique handle. The level is provided by the user by setting the attribute parameter `CK_ATTRIBUTE_PTR`. The handle can be chosen randomly from a sufficiently large space or using any another technique/mechanism for creating universally unique identifier [36]. This ensures handle integrity without central coordination. The details of the precise encoding from levels to `CK_ATTRIBUTE_PTR` are not important, but the token has to enforce that the level is correctly encoded. In general, the level can be represented using a vendor-specific PKCS#11 attribute that encodes this number in an integer. If there is a suitable upper bound, these levels can also be encoded in standard PKCS#11 attributes, e.g., if the bound is 4, the values of `CKA_WRAP` and `CKA_ENCRYPT` can be used to encode a binary representation of each level between 1 and 4. As wrapping with asymmetric keys is fundamentally flawed (asymmetric wrapping keys can be used to inject keys whose values are known to the attacker [13]), asymmetric key generation (`C_GenerateKeyPair`) is restricted to keys of level 2. We hence consider asymmetric encryption keys only for key-usage.

`C_DeriveKey` creates a new key object from a base key. As there is no AEAD scheme in the PKCS#11v2.40 cryptographic mechanism specification that can be used for both wrap/unwrap and key derivation [35, Section 2.11], any key that may be used for key-derivation has level 2 and may only be used to derive keys of level 2. Similar to `C_GenerateKey`, a universally unique handle is created.

`C_WrapKey` creates an authenticated encryption of a key and includes its level and handle as additional authenticated data. This makes sure that keys are reimported with precisely the same attributes. This is not possible with PKCS#11 prior to v2.40, due to the lack of support for AEAD. Note, however, that, even for v2.40, this requires a modification to its specification or a new interface: PKCS#11 v2.40 specifies the initialisation vector to be set from the outside, leading to the aforementioned two-time pad attack. While an implementation may very well ignore the supplied IV and choose it internally, by specification, the function output contains only the cyphertext, not the IV. This is problematic, as it means that the interface cannot be easily changed to communicate the internally generated IV without breaking backwards compatibility. For encryption, the current PKCS#11 v3.00 draft solves this by introducing a new interface for encryption, `C_EncryptMessage`, specifically to support internal nonce-generation for AEAD schemes. `C_EncryptMessage` has an additional parameter that controls whether the pointer to the IV is used as an input or an output. In the current draft, there is no equivalent for key-wrapping, although this would not even require a new interface, but instead just a provision that the (mechanism-specific) IV parameter may be requested to be used for output instead of a new interface. We encourage the inclusion of such a provision in the base specification and making internal IV generation the default for authenticated wrapping. Considering the adaptation of `C_EncryptMessage`, we deem this a realistic proposal. Moreover, internal nonce-generation follows from the FIPS requirement on GCM: 'The probability that the authenticated encryption function ever [across all instances] will be invoked with the same IV and the same key on two (or more) distinct sets of input data shall be no greater than 2^{-32}.' [19, 4]

`C_UnwrapKey` decrypts a wrapping, verifies its authenticity and stores the decryption as a new key. It takes the following parameters: the handle of the wrapping key, an attribute template specifying the attributes that the newly generated object should obtain, and a handle for this newly generated object. The initialisation vector is supplied as the mechanism parameter. Our policy is to reject any handle and any template that do not match the authenticated handle and level.[1] In contrast to previous policies, it is thus not possible to reimport a key on the same device under different handles — there is no need to, as all instances of a key are guaranteed to have the same attributes. Thread-safe implementations should thus check if the requested handle is present on the device before unwrapping, relying on locks only to synchronize concurrent unwrap, key-generation and key-derivation actions.

3.3 Key-usage

PKCS#11 supports a variety of functions for creating message digests, signatures, MACs or random numbers. All of these operate on payload data, hence, we impose that the keys must have level 2. We impose no further restrictions beyond PKCS#11's standard requirements, e.g., MACs can only be computed with MACing keys, etc.

For AEAD encryption and decryption specifically, we require that the authenticated header contains the level $l = 1$ (for payload data). This prohibits encryptions to be confused with wrappings and thus 'trojan key' attacks [13], where unwrapping injects dishonest key material into the store. The same policy applies to encryption for multi-part data (`C_EncryptInit`, `C_EncryptUpdate` and `C_EncryptFinal`), however, our model only covers encryption and decryption for single-part data.

Similar to prior work [18, 22, 10, 20], we will only model key-usage functions that could possibly interfere with key-management, i.e., symmetric encryption and decryption, as indicated by Clulow's attack. Keys that do not support encryption can, by the standard, not be used to create or import wrappings, and hence do not interact with the key-management. By our policy, asymmetric encryption falls into the same category. Extending the model to cover non-key-management operations is straight-forward, but unlikely to lead to new insights with respect to the security of policies.

3.4 Limitations

The policy we propose is based on a *static* key-hierarchy: This reduces the flexibility when setting up keys. Similarly, a popular best practice for HSMs is to disallow the modification of attributes for all users but the SO.

To benefit from handle authentication, existing applications have to be modified to make use of this feature by validating the authenticity of the handle provided. In current applications, objects are identified by a user-specified attribute `CKA_LABEL`. `C_FindObjects` is used to obtain all handles associated to objects that have a specified label and these handles are used without further validation. Instead, the handle should be specified (in place of the label) to identify keys. Practically, however, this is not always possible, as handles are implementation-dependent and cannot always be chosen freely. Furthermore, this requires a modification of the application. In the following, we discuss a workaround for both issues. The handle (in the sense of our policy) could be stored within the attribute `CKA_LABEL`. Handle authenticity then pertains to this attribute, which can now be used to identify keys. The advantage is that applications using the previously described method for identifying keys would not require changes. The downside is that this label can neither be set nor modified by the user or SO, but is instead chosen according to the policy upon object creation.

4 Preliminaries

Our analysis takes place in an abstract model of cryptography with an active, Dolev-Yao adversary. The idea is that all implementations are considered participants in a protocol. As the adversary is active and has access to all of them, he can send arbitrary commands to them and combine their outputs. This represents a network where all hosts are under adversarial control. We analyzed this model with Tamarin [43], a protocol verifier with support for (stateful) security protocols.

Terms and equational theories

Cryptographic messages are represented by a term algebra over public names PN, fresh names FN and variables $\mathcal{V}$. Let Σ be a signature, i.e., a set of function symbols, each with an arity. We write f/n when function symbol f is of arity n, e.g., $\mathsf{pair}/2$ is a function symbol for pairs. Let Terms be the set of terms built over Σ, PN, FN and $\mathcal{V}$, e.g., $\mathsf{pair}(t, t') \in$ Terms, which we will abbreviate $\langle t, t' \rangle$.

Equality is defined by means of an equational theory E, i.e., a finite set of equations between terms. E induces a binary relation $=_E$ that is closed under application of function symbols, bijective renaming of names and substitution of variables by terms.

1. In addition, we recommend checking that the authenticated level is smaller than the wrapping key's level to provide resilience against key compromise. Our model, however, does not consider key compromise.

Example 1. *Our model employs the following equational theory. Unary function symbols* fst *and* snd *model projection on pairs:*

$$\mathsf{fst}(\langle x, y\rangle) = x \qquad \mathsf{snd}(\langle x, y\rangle) = y$$

Hence $\mathsf{fst}(\mathsf{snd}(\langle x, \langle y, z\rangle\rangle)) =_E y$*. We use* $\mathsf{true}/0$ *to model a constant truth value. We model AEAD using* $\mathsf{senc}/4$*, which expects a key, an initialisation vector, some authentication data and a message. The following equations apply:*

$$\begin{aligned} \mathsf{sdec}(k, iv, h, \mathsf{senc}(k, iv, h, m)) &= m \\ \mathsf{sdecSuc}(k, iv, h, \mathsf{senc}(k, iv, h, m)) &= \mathsf{true}() \\ \mathsf{getHeader}(\mathsf{senc}(k, iv, h, m)) &= h \\ \mathsf{getIV}(\mathsf{senc}(k, iv, h, m)) &= iv \end{aligned}$$

We use the two-ary function symbol $\cup^\#$ *to model multiset union. Written in infix notation, the following equations for associativity and commutativity apply:*

$$x \cup^\# (y \cup^\# z) = (x \cup^\# y) \cup^\# z \qquad x \cup^\# y = (y \cup^\# x)$$

This function symbol is built into Tamarin. We will use it to model natural numbers. We also include a symbol $\mathsf{kdf}/2$ *for key-derivation, without any equations.*

Multiset Rewriting

In the Tamarin protocol prover, the protocol itself, its state and its behavior are modeled using a multiset of facts and rewriting rules operating on this set. The state of the system is a multiset of ground facts $\mathcal{G}$, where a fact $F(t_1, ..., t_k)$ of arity k is ground if all k terms $t_1, ..., t_k$ are ground. Further, there are predefined fact symbols for special purposes. The state of the adversary's knowledge is encoded using the fact symbol !K. Freshness information is denoted with the fact symbol Fr and messages on the network are represented by In and Out. Multiset rewriting rules are denoted by $l \dashv[\ a\ ]\!\!\to r$, where l is the premise, r is the conclusion and a labels so-called *actions*. *Linear* facts used in the premise are consumed by the transition rule. An exclamation mark in front of a fact symbol indicates that it is *persistent* and can be consumed arbitrarily often. For example, freshness Fr is a linear fact, whereas adversarial knowledge !K is a permanent fact.

Example 2. *To express, e.g., a key hierarchy or a counter, we need to identify natural numbers. We can model them using Tamarin's built-in support for multisets: a multiset with n elements* $1 \in PN$ *represents n. The following two rules ensure that terms t for which a fact* $!\mathsf{Nat}(t)$ *or action* $\mathsf{IsNat}(t)$ *exists are always multisets consisting only of* $1 \in PN$*.*

$$\dashv[\ \mathsf{IsNat}(1)\ ]\!\!\to !\mathsf{Nat}(1) \tag{1}$$

$$!\mathsf{Nat}(n) \dashv[\ \mathsf{IsNat}(n \cup^\# 1)\ ]\!\!\to !\mathsf{Nat}(n \cup^\# 1) \tag{2}$$

Intuitively, we say that a rewriting step is possible if all facts in l are in the current state S. In the resulting state, all linear facts from l are removed and all facts in r are added. We will formulate this intuition in the following, but need some preliminaries first. We use lfacts and pfacts to denote the linear, respectively, the permanent facts in a set, set to turn a multiset into a set and mset to turn a set into a multiset. We mark the multiset equivalents of the subset relation, set difference and set union with a $\#$ superscript, i.e. $\subset^\#$, $\setminus^\#$ and $\cup^\#$.

We define a labeled transition relation $\to_{\mathcal{M}} \subset G^\# \times P(G) \times G^\#$, where $G^\#$ denotes a multiset of ground facts and $\mathcal{M}$ denotes a set of ground instantiations of multiset rules, as follows:

$$\frac{l \dashv[\ a\ ]\!\!\to r \in \mathcal{M} \quad \mathit{lfacts}(l) \subset^\# S \quad \mathit{pfacts}(l) \subset \mathit{set}(S)}{S \xrightarrow{\mathit{set}(a)}_{\mathcal{M}} (S \setminus^\# \mathit{lfacts}(l)) \cup^\# \{\ r\ \}^\#}$$

Consider, e.g., the following application of (2):

$$\{\ !\mathsf{Nat}(1)\ \}^\# \xrightarrow{\mathsf{IsNat}(1 \cup^\# 1)} \{\ !\mathsf{Nat}(1), !\mathsf{Nat}(1 \cup^\# 1)\ \}^\#.$$

Using the labelled transition relation, we can define executions of some model $\mathcal{M}$ as a set of traces:

$$\begin{aligned} \{(A_1, \ldots, A_n) \quad | \; & \exists S_1, \ldots, S_n \in \mathcal{G}^\#.\ \varnothing \xrightarrow{A_1}_{\mathcal{M}} \ldots \xrightarrow{A_n}_{\mathcal{M}} S_n \\ & \wedge \forall i \neq j.\ \forall x.\ \ S_{i+1} \setminus^\# S_i = \{\mathsf{Fr}(x)\} \\ & \implies S_{j+1} \setminus^\# S_j \neq \{\mathsf{Fr}(x)\}\} \end{aligned}$$

Combining the previous transition with an application (1), we obtain the trace $(\mathsf{IsNat}(1), \mathsf{IsNat}(1 \cup^\# 1))$. The side condition ensures that fresh variables are instantiated with unique fresh names.

Tamarin combines a user-defined set of rules describing the protocol itself with the builtin rules for message deduction MD depicted in Figure 2. They represent a standard Dolev-Yao attacker who obtains knowledge (!K) by eavesdropping on the network (Out), creating fresh names, or by using public values. This knowledge can be combined by applying function symbols f/k. Known terms can be sent to the network.

$$\begin{array}{rcl}
\mathsf{Out}(x) & -[\ ]\!\!\rightarrow & !\mathsf{K}(x) \\
\mathsf{Fr}(x : \mathit{fresh}) & -[\ ]\!\!\rightarrow & !\mathsf{K}(x : \mathit{fresh}) \\
 & -[\ ]\!\!\rightarrow & !\mathsf{K}(x : \mathit{pub}) \\
!\mathsf{K}(x_1), \ldots, !\mathsf{K}(x_k) & -[\ ]\!\!\rightarrow & !\mathsf{K}(f(x_1, \ldots, x_k)) \\
!\mathsf{K}(x) & -[\ K(x)\ ]\!\!\rightarrow & \mathsf{In}(x)
\end{array}$$

Fig. 2: The set of rules MD.

5 Formal Modelling

We present the multiset rewrite rules used to formalise the policy described in Section 3 and Table 1.

Devices

At any time, a new device can be introduced to the network. This device has a fresh identifier dev, and its device counter is initialised to $1 \in PN$, representing the natural number 1. Previous work [10, 18, 29] abstracted all PKCS#11 devices in the network with a single store. As we want to tackle the problem of locally generating network-wide unique IVs, we need to capture the absence of a secure channel between these devices, and thus model them individually.

$$\begin{aligned}\mathsf{Fr}(dev), !\mathsf{Nat}(1) \ -[\ \mathsf{DCtrIs}(dev, '1')\ ]\!\!\rightarrow \\ !\mathsf{D}(dev), \mathsf{DCtr}(dev, 1)\end{aligned}$$

Each device ($!\mathsf{D}(dev)$), obtains a fresh identifier ($\mathsf{Fr}(dev)$), which links it to the initial counter value ($\mathsf{DCtr}(dev, 1)$). The action DCtrIs is used in the lemma `counter_mono` (cf. Section 6) to refer to this counter and show each counter is monotonically increasing.

Key-generation

When a new key is created, it is stored along with its level, a freshly chosen handle and a natural number l on the store of dev, represented by the fact $!\mathsf{Store}(dev, h, k, l)$. The rules from Example 2 are part of our model and ensure that l represents a natural number. The handle and the level of the key are handed out to the adversary ($\mathsf{Out}(\langle h, l\rangle)$).

$$\begin{aligned}&!\mathsf{D}(dev), !\mathsf{Nat}(l), \mathsf{Fr}(k), \mathsf{Fr}(h) \\ &-[\ \mathsf{CreateK}(h, k, l), \mathsf{StoreK}(dev, h, k, l)\ ]\!\!\rightarrow \\ &!\mathsf{Store}(dev, h, k, l), \mathsf{Out}(\langle h, l\rangle)\end{aligned} \tag{3}$$

The action CreateK marks the creation of a key along with its level and attribute. It is referenced by lemma `key_int_conf` to say that keys imported via unwrapping were honestly generated at an earlier point (i.e., no trojan keys can exist). StoreK, by contrast, marks that a key is added to the store, which includes import via unwrap and key-derivation.

A second rule additionally contains $!\mathsf{D}(dev')$ in the premise and $!\mathsf{Store}(dev', h, k, l)$ in the conclusion and is used to model a trusted set-up phase where a common key is established on two devices.

$$\begin{aligned}\cdots, !\mathsf{D}(dev') \ -[\ \cdots, \mathsf{StoreK}(dev', h, k, l)\ ]\!\!\rightarrow \\ \cdots, !\mathsf{Store}(dev', h, k, l)\end{aligned} \tag{4}$$

Note that devices only need to produce fresh names during key-generation. Hence, w.l.o.g., a device without RNG is represented by an adversary that chooses to never employ an instance of the key-generation rule where dev is instantiated to this device. Devices without RNG exist and are useful: lightweight authentication tokens can, e.g., obtain a master key via a trusted set-up, and subsequently import keys via unwrapping.

Encryption and decryption of payload data

Encryption (`C_Encrypt`) expects some payload m and encrypts it with the authenticated header affirming the level as 1 (payload data) and, for uniformity, an empty handle value $\epsilon \in PN$. For simplicity, the handle h is not required as an explicit input – the adversary chooses the appropriate instantiation of this handle anyway. We set the initialisation vector to $\langle dev, ctr\rangle$, which, as we will show, ensures the network-wide uniqueness of the IV.

$$\begin{aligned}&!\mathsf{Nat}(ctr \cup^{\#} 1), !\mathsf{D}(dev), !\mathsf{Store}(dev, h, k, l), \\ &\mathsf{DCtr}(dev, ctr), \mathsf{In}(m) \ -[\ \mathsf{UseK}(dev, h, k, l), \\ &\mathsf{DCtrIs}(dev, ctr \cup^{\#} 1), \mathsf{IV}(\langle dev, ctr\rangle)\ ]\!\!\rightarrow \\ &\mathsf{DCtr}(dev, ctr \cup^{\#} 1), \mathsf{Out}(\mathsf{senc}(k, \langle dev, ctr\rangle, \langle 1, \epsilon\rangle, m)\end{aligned} \tag{5}$$

As before, DCtrIs records the new counter value ($\mathsf{DCtr}(dev, ctr \cup^{\#} 1)$) to ensure monotonicity. IV marks the use of the IV. The lemma `uniqueness_IV` will ensure that no two instances of this action have the same value, which is a cryptographic requirement for AEAD schemes. Finally, UseK marks the use of a key with the handle and level that were assumed. Lemma `key_usage` will ensure that any key used was created or imported with exactly this handle and level.

Decryption (`C_Decrypt`) verifies that the authenticated tag is $\langle 1, \epsilon\rangle$. Let $iv = \mathsf{getIV}(c)$, $t = \mathsf{getHeader}(c)$ and $m = \mathsf{sdec}(k, iv, t, c)$ in

$$\begin{aligned}&!\mathsf{D}(dev), !\mathsf{Store}(dev, h, k, l), \mathsf{In}(c)\\ &-\!\![\ \mathsf{UseK}(dev, h, k, l), \mathsf{Decrypt}(m),\\ &\quad \mathsf{IsTrue}(\mathsf{sdecSuc}(k, iv, t, c)), \mathsf{Eq}(t, \langle 1, \epsilon\rangle)\ ]\!\!\rightarrow\\ &\mathsf{Out}(m)\end{aligned} \tag{6}$$

Again, UseK tracks the use of the key. $\mathsf{Decrypt}(m)$ will be used in the lemma `origin` to state that any knowledge obtained by the output message m was known by the adversary before invoking decryption.

We use the action IsTrue to check whether the decryption was successful: every lemma φ presented in the next section is verified w.r.t. the subset of traces for which the condition

$$\alpha := (\forall\, a, i. \mathsf{IsTrue}(a)@i \implies a =_E \mathsf{true}())^2$$

holds true. This is achieved by showing $\alpha \implies \varphi$ on the entire set of traces. For every trace where the term $\mathsf{sdecSuc}(k, iv, t, c)$ is unequal to $\mathsf{true}()$ (modulo E), the property is trivially true and thus the property is valid iff α holds for all traces that adhere to the restriction. Tamarin conveniently allows specifying several so-called *restrictions* α, which apply to all lemmas in this way.

Key-wrapping

Wrapping proceeds in the same vein. A key on the device ($!\mathsf{Store}(dev, h_w, k_w, l_w)$) can be used to encrypt another key ($!\mathsf{Store}(dev, h_e, k_e, l_e)$). Again, let $iv = \langle dev, ctr\rangle$.

$$\begin{aligned}&!\mathsf{Nat}(ctr \cup^{\#} 1), !\mathsf{D}(dev), !\mathsf{Store}(dev, h_w, k_w, l_w),\\ &!\mathsf{Store}(dev, h_e, k_e, l_e), \mathsf{DCtr}(dev, ctr)\\ &-\!\![\ \mathsf{UseK}(dev, h_w, k_w, l_w), \mathsf{DCtrIs}(dev, ctr \cup^{\#} 1),\\ &\qquad \mathsf{IV}(iv), \mathsf{Lt}(el, wl)\ ]\!\!\rightarrow\\ &\mathsf{DCtr}(dev, ctr \cup^{\#} 1), \mathsf{Out}(\mathsf{senc}(k_w, iv, \langle l_e, h_e\rangle, k_e))\end{aligned} \tag{7}$$

The output $\mathsf{senc}(k_w, iv, \langle l_e, h_e\rangle, k_e)$ constitutes the wrapping of k_e under k_w with additional authenticated data $\langle l_e, h_e\rangle$ for the previous handle and level of k_e on device dev. Again, UseK, DCtrIs and IV track the state of keys, counters and the IV $iv = \langle dev, ctr\rangle$. Similar to IsTrue, the action Lt ensures the wrapped key has a lower level than the wrapping key by imposing another restriction on traces: for every action $\mathsf{Lt}(a, b)$, there is a non-empty a' such that $a \cup^{\#} a' = b$, i.e., a represents a (strictly) smaller number than b. This avoids key-cycles.

Unwrapping

To unwrap (`C_UnwrapKey`), a device is called with a handle to a wrapping key (i.e., a key of level $\geqslant 3$) and an authenticated encryption c. It decrypts c, and stores the resulting key along with the authenticated handle and level for future use ($!\mathsf{Store}(dev, h_e, k_e, l_e)$). Let $iv = \mathsf{getIV}(c)$, $t = \langle l_e, h_e\rangle = \mathsf{getHeader}(c)$ and $k_e = \mathsf{sdec}(k, iv, t, c)$ in

$$\begin{aligned}&!\mathsf{Nat}(l_e), !\mathsf{D}(dev), !\mathsf{Store}(dev, h, k, l), \mathsf{In}(c)\\ &-\!\![\ \mathsf{UseK}(dev, h, k, l), \mathsf{ImportK}(dev, h_e, k_e, l_e), \mathsf{Neq}(l_e, 1),\\ &\mathsf{StoreK}(dev, h_e, k_e, l_e), \mathsf{IsTrue}(\mathsf{sdecSuc}(k, iv, t, c))\ ]\!\!\rightarrow\\ &!\mathsf{Store}(dev, h_e, k_e, l_e)\end{aligned} \tag{8}$$

As before, UseK marks the use of the wrapping key and StoreK their addition to the store. IsTrue ensures that $\mathsf{sdecSuc}(k, iv, t, c) =_E \mathsf{true}()$. $\mathsf{ImportK}$ marks that the key contained in the wrapping has been imported, and not created. It will be referred to by lemma `key_int_conf` (cf. Section 6) to say that any key imported by wrapping was once created on some device.

By our deduction soundness result, the cyphertext c in our model contains the authenticated header and IV in the clear. Hence it represents the 'raw' cyphertext, as well as the other parameters supplied to `C_Unwrap`.

2. $F@i$ denotes that action F appears at position i in the trace.

Key-derivation

Key-derivation (`C_DeriveKey`) is restricted to key-usage keys, i.e., keys of level 2. Recall that we omitted pure key-usage like MACs from the model, except for AEAD encryption and decryption. We therefore model key-derivation with AEAD base keys to represent derivation from other keys of level 2. The Fr-facts in the premise model the generation of a globally unique handle, as well as a random salt r, which is used to derive the new key as $\mathsf{kdf}(k, r)$. Let $two = 1 \cup^{\#} 1$ in

$$\begin{aligned}&!\mathsf{D}(dev), !\mathsf{Store}(dev, h, k, two), \mathsf{Fr}(r), \mathsf{Fr}(h')\\ &-\!\![\ \mathsf{UseK}(dev, h, k, two), \mathsf{StoreK}(dev, h', \mathsf{kdf}(k, r), two),\\ &\quad \mathsf{CreateK}(h', \mathsf{kdf}(k, r), two)\ ]\!\!\rightarrow\\ &\quad !\mathsf{Store}(dev, h', \mathsf{kdf}(k, r), two)\end{aligned} \tag{9}$$

As before, UseK marks the use of the key k. Similar to key-generation, this rule is marked with StoreK (as the derived key is added to the store of dev), as well as CreateK (as the key $\mathsf{kdf}(k, r)$ is created).

6 Results for AES-GCM/CCM

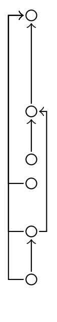
dep.	lemma	description	steps	seconds
	origin	Any messages obtained by decryption were encrypted before and all keys imported via unwrapping were either created on the device or known to the adversary at some point. $(!\mathsf{D}(m)@i \implies \exists j.!\mathsf{K}(m)@j \wedge j < i) \wedge (\mathsf{ImportK}(dev, h, k, l)@i \implies (\exists j.\mathsf{CreateK}(h, k, l)@i \wedge j < i) \vee \exists j'.!\mathsf{K}(k)@j' \wedge j' < i)$.	1597	72
	counter_mono	The device counter is monotonically increasing. $\mathsf{DCtrls}(d, c)@i \wedge \mathsf{DCtrls}(d, c')@j \wedge i < j \implies \exists z.c' =_E z + c$.	1880	77
	uniqueness_IV	No IV is used twice, no matter on which device. $\mathsf{IV}(t)@i \wedge \mathsf{IV}(t)@j \implies i = j$.	8	16
	key_usage	All keys that are used were created by unwrapping, key-derivation or key-generation. $\mathsf{UseK}(d, h, k, l)@i \implies \exists j.\mathsf{StoreK}(d, h, k, l)@j \wedge j < i$.	78	17
	key_int_conf	All keys are created on some device $(\mathsf{ImportK}(d, h, k, l)@i \implies \exists j.\mathsf{CreateK}(h, k, l) \wedge j < i)$ and are never known $(\neg(\mathsf{CreateK}(h, k, l)@i \wedge K(k)@j))$.	428	45
	key_level_handle	Keys always retain the level and handle they were created with. $\mathsf{StoreK}(d, h, k, l)@i \wedge \mathsf{StoreK}(d', h', k, l')@j \implies l =_E l' \wedge h =_E h'$.	170	21

TABLE 2: Proof lemmas and their dependencies. We use $F@i$ to denote that an action F appears at position i in a trace. For brevity, unbound variables are to be read as universally quantified.

The stated purpose of PKCS#11 is to separate secret data from untrusted code accessing the interface. Hence our main goal is to ensure that no key generated on the device can leak to the adversary. Nevertheless, there are two additional integrity properties that we consider important, but that have been largely overlooked by prior work. First, the integrity of the keys themselves: each key on the device was created on some honest device; it is not possible to import trojan keys. Second, the integrity of the mapping from handles to keys: each key, on whichever device it may be placed, will always have the same level and the same handle. The latter property is a new feature of our policy that is meant to ensure that no attacker can confuse an honest application into using an insecure or deprecated key by altering the assignment from handles to keys.

We verify these properties using two helping lemmas (see Table 2). These lemmas were stated manually, but proven automatically. The first one (`origin`), establishes that any knowledge obtained through decryption was available beforehand, and that all keys imported via wrapping were either originally created on some device, or was otherwise known by the adversary before. The first conjunct of `origin` prunes cases where decryption is used to derive a term of arbitrary form from an encryption. Intuitively, when Tamarin's backward search algorithm is trying to prove that a certain term cannot be deduced, e.g., a key stored on the device, it considers all rules that have a matching Out-fact. The rule for decryption (6) by itself could output any term t, as long as $c = \mathsf{senc}(k, iv, \langle 1, \epsilon \rangle, t)$ is input, and thus known to the adversary. This c itself could come from rule (6), which, without `origin`, creates a loop. This conjunct establishes that the content of the cyphertext must have been known prior to using the decryption rule. As knowledge facts are permanent, the application of rule (6) is superfluous if $!\mathsf{K}(t)$ is already present, and thus this step can be pruned. The second conjunct can be used to either establish the freshness of keys (both rules containing $\mathsf{CreateK}(k)$ have the premise $\mathsf{Fr}(k)$), or to pinpoint an earlier leak of a key, which helps in the inductive steps of many of the follow-up lemmas.

The second helping lemma (`counter_monotonicity`) establishes that on each device, the counter is monotonically increasing. Proving it is just a matter of considering all pairs of rules where the action DCtrls occurs, but when applied, it readily entails the relationship between any two counter-values once their temporal relation can be established.

With these lemmas in place, we show: First (`key_usage`), that all keys that are used by an honest token were put in the store either by unwrapping (8), by key-derivation (9) or by key-generation (3); and that the attribute and handle remain unchanged. Second (`key_int_conf`, first conjunct), if they were created by unwrapping, they were previously generated by key-generation or key-derivation with the same attribute and handle, but possibly on a different device. Together, this means that all keys that are used were honestly generated, and that throughout their use, they are associated with the same attributes and handle. Third (`key_int_conf`, second conjunct), all keys are confidential: it is not possible for any key that was created on the device to be deduced by the adversary. In Tamarin, this is expressed by referring to the action !K in the message deduction rule for adversarial output (see Figure 2): the adversary cannot output a key created on some device. Fourth, whenever a key is added to the store on any device, it is associated with the same level and handle.

Finally, the deduction soundness result in the next section comes with a proof obligation for the protocol: whenever a term $\mathsf{senc}(k, iv, h, m)$ is output, the tuple (k, iv, h) needs to be unique. Lemma `uniqueness_IV` establishes the stronger property that iv itself is distinct within all such terms.

All these lemmas can be shown automatically using a custom heuristic that prioritizes goals relevant to IV generation. We report the number of proof steps and the verification time per lemma in Table 2. Both were measured on a 3.1 GHz Intel Core i7 with 16GB RAM. A full proof took about four minutes. As we present a new policy of PKCS#11 with new features, we cannot compare the verification time with previous efforts. The closest work to ours also used Tamarin and reported a runtime of half an hour on a dedicated computation server [29]. The structure of the proof, in particular the choice of the helping lemmas and their order, follows the structure in this paper, albeit adapted to our model. We thus feel confident that our helping lemmas and heuristics can be reused for other policies that guarantee key and attribute integrity.

7 Justifying the symbolic abstraction

Symbolic models in the literature that include symmetric encryption usually imply authenticity of the cyphertext. In the cryptographic setting, this is called non-malleability. They do, however, not account for the choice of the IV. This is reasonable, as in most cases, this choice is part of the encryption scheme itself, and not a protocol task. For the configuration we discussed in the last section, however, IV generation is part of the protocol itself and hence cannot be abstracted away.

We thus provide some justification for the equational theory we use to model AEAD, which was introduced in Example 1, by showing a necessary, but not sufficient, condition for the soundness of the symbolic attacker. As we will see, we have to impose a condition on the protocol. Luckily, this condition can be proven to hold using Tamarin.

Formal models rely on an abstract representation of cryptography for efficient tool support. The relationship between results in this formal model and the complexity-theoretic model of cryptography was first established by Abadi and Rogaway [1] under the name of *computational soundness*. Computational soundness says that each attack that occurs with non-negligible probability in the computational model is represented in the symbolic model. It thus ensures that the symbolic model and the semantics of the protocol calculus are adequate models of the cryptographic primitives and the behaviour of the protocol parties.

Rather than extending the existing body of work with an additional computational soundness result for a small set of primitives, we opted to extend the deduction soundness framework [16] by Cortier and Warinschi. The distinguishing feature of this framework is that it allows for the composition of deduction soundness results for different primitives. As PKCS#11 covers many different cryptographic primitives this is a very useful feature. The downside is that deduction soundness does not guarantee computational soundness. The research question of defining a composable framework for computational soundness is still open, thus we opted for extending Böhl et. al.'s deduction soundness result [8] at the expense of a weaker guarantee. Their result includes public key encryption, secret key encryption, signatures, MACs, hashes[3] and also public data structures. All these primitives are supported by PKCS#11, and thus it is very attractive to use this model and be able to reason about higher-level protocols building on our PKCS#11 configuration.

We extend Böhl et. al.'s result with deterministic authenticated encryption, so we can reason about schemes like AES-GCM and AES-CCM as supported by PKCS#11. We can only sketch the result here, and refer to Appendix D to Appendix G for the details. We keep the notation minimal in this section and use Böhl et. al.'s notation in the appendices.

Cryptographic requirements

We introduce a cryptographic security notion, DAE-N security, which is a version of DAE security [40, Definition 1], modified to give the adversary access to the IV. DAE [40] security is logically equivalent to AEAD security [38] and formalises the confidentiality and authenticity for AEAD. Our modification, DAE-N security, differs from DAE security [40] in that oracles can be called with arbitrary IVs, as long as they do not repeat.[4]

Definition 1 (Deterministic Authenticated Encryption with IVs)**.** *Let $\Pi = (Gen, Enc, Dec)$ be an IV-based authenticated encryption scheme that can handle an associated header. That means: Given IV space S, associated data [5] space $\mathcal{H}_{AD}$ and message $\mathcal{M}$, the encryption algorithm Enc takes as input a key $k \xleftarrow{\$} Gen(1^\eta)$, an IV $n \in S$, a string of associated data H, with $H \in \mathcal{H}_{AD}$ and a message m with $m \in \mathcal{M}$. It returns a cyphertext $c = Enc(k, n, H, m)$ with $c \in \mathcal{M}$. Decryption takes a key $k \xleftarrow{\$} Gen(1^\eta)$, an IV $n \in S$, a string of associated data H, with $H \in \mathcal{H}_{AD}$ and a cyphertext c with $c \in \mathcal{M}$ as input and returns m with $m \in \mathcal{M} \cup \{\perp\}$.*

3. PKCS#11 supports a SHA-1-based key-derivation mechanism.
4. DAE-N security can also be seen as a weaker version of Rogaway's notion of misuse-resistant AE (MRAE) security [40, Definition 5]. GCM and CCM mode provide AEAD security and thus DAE-N security, but not MRAE security. If used appropriately, SIV mode provides both MRAE and DAE-N security.
5. In the context of our work *header*, *additional data* and *associated data* are interchangeable terms.

The DAE-N-advantage of an attacker A with access to two oracles (the first called left-hand, the second called right-hand) in Π is defined

$$\mathrm{Adv}_{\Pi}^{dae-n}(A) = |\Pr[A^{O_k^{Enc}(\cdot,\cdot,\cdot),O_k^{Dec}(\cdot,\cdot,\cdot)} = 1] \\ - \Pr[A^{\$(\cdot,\cdot,\cdot),\perp(\cdot,\cdot,\cdot)} = 1]|$$

where $k \xleftarrow{\$} Gen(1^\eta)$ and $O_k^{Enc}(\cdot,\cdot,\cdot)$ and $O_k^{Dec}(\cdot,\cdot,\cdot)$ denote an encryption oracle and a decryption oracle, respectively. Further, let $\$(\cdot,\cdot,\cdot)$ be an algorithm returning a random bitstring c with $c \in \mathcal{M}$ and $\perp(\cdot,\cdot,\cdot)$ an algorithm always returning $\perp$.

The adversary may not repeat an IV in a left-query and may not ask a right-query (H, IV, Y) if some previous left-query (H, IV, X) returned Y. (MRAE security defines Adv^{mrae} *just the same, but restricts the adversary to not repeat a left-query and may not ask a right-query (H, IV, Y) if some previous left-query (H, IV, X) returned Y.)*

A scheme Π is DAE-N secure iff, for all ppt algorithms A,

$$Adv_{\Pi}^{dae-n}(A) \leqslant \mathtt{negl}(\eta)$$

for a negligible function $\mathtt{negl}()$ *and a security parameter η.*

AEAD security [39] has been proven for CCM by Jonnson [24] and for GCM by McGrew and Viega [32]. In Appendix A, we show that this implies DAE-N security.

Symbolic model and deduction relation

We represent the equations in Example 1 in the deduction soundness framework as a typed symbolic model and deduction relation $\vdash$ between a set of terms the adversary knows, and a term the adversary can deduce from this set. A term is deducible if it can be constructed from other deducible terms or obtained by applying decryption and similar operations. In our case, the symbolic model consists of a two randomized function k_c^l and k_h^l, representing AEAD key generation, an encryption function E with four arguments, and a function con_S that transforms terms into IVs. The superscript l marks k_c and k_h as randomized, as opposed to E and con_S which are deterministic. Both k_c and k_h are implemented in an identical way, but different symbols are used to mark keys that may be corrupted initially, and keys that shall not be revealed. This is ensured by the protocol conditions below. We use k_x to make statements that hold for both k_h and k_c.[6]

For each l, k_x^l represents a different, randomly chosen key. The types make sure that the first argument to encryption is always a key and that the second is an IV. The other two arguments, the authenticated information and the message, can be arbitrary terms. The deduction relation is defined by the following four rules:

$$\frac{k_x^l()\ con_S(n)\ H\ m}{E(k_x^l(), con_S(n), H, m)} \qquad \frac{E(k_x^l(), con_S(n), H, m)}{con_S(n)}$$

$$\frac{E(k_x^l(), con_S(n), H, m)}{H} \qquad \frac{E(k_c^l(), con_S(n), H, m)}{m}$$

From top left to bottom right, they allow (a) the attacker to construct encryptions if he knows all inputs, (b) to extract the IV, (c) to extract the authentication information and (d) to deduce the message if the key may be corrupted initially. The protocol conditions in the following paragraph ensure that the adversary only learns keys that are initially corrupted, and hence (d) correctly represents the first equation in Example 1, as w.l.o.g., the symbolic adversary corrupts all keys k_c^l from the start.

Implementation

An implementation consist of a Turing machine that computes each function symbol, a length function that for each term predicts the length its corresponding bitstring has, an interpretation function that defines how bitstrings are interpreted as terms and a `valid` predicate that restricts the operations an attacker can perform. The latter is used to define protocol conditions. These are necessary for soundness results that have only standard assumptions on the cryptographic primitives, as the following example illustrates. It is well known that IND-CCA security does not guarantee anything in the presence of key-cycles [3]. Hence soundness can only hold if the deduction soundness attacker (and thus the protocol) is restricted to not produce them. Alternatively, stronger notions of security such as key-dependent message security can be used. There is a trade-off between protocol conditions and requirements on the cryptographic algorithms.

In our case, the Turing machine implementation is constructed using a DAE-N secure encryption scheme and an injective function that maps the bitstring representation of any terms in S into the IV space. For GCM and CCM, e.g., this can be a simple concatenation if we can guarantee that all terms in S can be represented in 32 bit. Our result is parametric in this S. We define the bitstring representation of an encryption to contain the authenticated information and the IV in the clear. The length function and the interpretation function are straight-forward. (See Appendix E for details.)

The validity predicate enforces the following protocol conditions (paraphrased for simplicity):

1) AEAD keys k^l can only occur in the first position of an E-term or in an initial corruption query.
2) No n in $E(k^l, con_S(n), H, m)$ occurs twice for the same l.
3) Whenever $con_S(t)$ appears in some term, $t \in S$.

6. There is a similar distinction for E that we gloss over here, but is explained in detail in Appendix B.

Proof overview

Due to its size (about 15 pages), we need to refer to the full version [17] for the proof, but will outline its structure here. We show deduction soundness in a stepwise proof over four games, starting from the deduction soundness game. This game is used to state that an adversary can never generate a bitstring that can be parsed to a term that he should not be able to deduce according to $\vdash$. In this game, the adversary interacts with an oracle that gives him access to the bitstring representation of terms of his choice. In the first step, it is shown that terms only collide with negligible probability. In the second, cyphertexts under honest keys are replaced with random bitstrings. In the third, the winning condition is made stricter by adding a rule to the deduction system that allows the adversary to generate honest cyphertexts — any adversary that can distinguish between the deduction soundness game with or without this rule is able to break the authentication property of the scheme. In the fourth and final step, it is shown that the modified deduction system is compatible with Cortier and Warinschi's notion of composability.

At this point, the model is not yet suited for key-wrapping, as keys can only appear at key positions and thus not be encrypted. Böhl et. al.'s framework handles this in an additional step. Function symbols carry an annotation to mark some of their input positions as *forgetful*; in our case, the fourth position of E. We show that a forgetful implementation, i.e., an implementation that substitutes each input at a forgetful function with a random bitstring of the same length, is also deduction sound. This allows us to relax the first condition of the validity predicate:

1) AEAD keys k^l can only occur in an initial corruption query, in the first position of an E-term, or as a subterm of a forgetful position of a function symbol that we compose with (but not E itself).

The last disjunct implicitly excludes key-cycles: by composing our AEAD model and implementation M_{AEAD} with (a renamed version) of itself, M'_{AEAD}, keys of M_{AEAD} can encrypt keys of M'_{AEAD}, but not vice versa.

Relation to our model

Our model has to make sure that for all possible traces, all three conditions of the validity predicate hold. The first condition can be checked syntactically: keys are indeed only output within encryptions terms, where they occur at position one or four. The only use at the fourth position is in the rule for key-wrapping.[7] There, key-cycles are avoided by means of the restriction $\mathsf{Lt}(el, wl)$. The lemma `key_level_handle` ensures that the level associated to each key is always the same. We can hence iteratively apply the compositionality result for all keys of level 1, $1+1$, etc.; the restriction associated to Lt makes sure that keys in the fourth position are always of lower level than the key at position one.

As a side-effect, however, the dynamic corruption of encryption keys is not guaranteed to be deduction sound. This is unfortunate, because the policy we propose implements a key-hierarchy to limit the potential damage due to wrapping keys that leak, e.g. due to side-channel attacks or brute-forcing.

Consequently, we refrained from formalising this property, as the soundness of a model that includes dynamic corruption cannot be guaranteed. There is an existing proposal that permits computational soundness without such protocol restrictions [5] that applies to a hybrid encryption scheme based on CBC-mode and an arbitrary MAC [46]. We leave extending these results and investigating the resistance against key-leakage to future work.

The second condition requires the protocol to make sure each IV is only used once per key, for all protocol traces. This is guaranteed by the lemma `uniqueness_IV`, which can be verified using Tamarin.

The third restriction can be checked syntactically, if we fix an implementation of con_S. For instance, we can set S to the set of terms $\langle t_1, t_2\rangle$ such that t_1 has a suitable type for device ids, e.g., $\{0,1\}^{32}$ and t_2 represents $\{1,\dots,2^{32}\}$. We then define the implementation of con_S to decode their bitstring representation and concatenate them.

Limitations of deduction soundness

We stress that deduction soundness is only a necessary criterion for computational soundness, as it only argues about the term representation and the deduction relation, but not the process representation. Our symbolic results do not necessarily carry over to the computational model. However, it was helpful in determining the validity conditions. Cortier and Warinschi point out that, in addition to deduction soundness, a so-called *commutation property* is necessary to establish computational soundness [16]. It is not known how to do this in a modular manner.

Roughly speaking, deduction soundness by itself talks about secrecy, not integrity. We opted for deduction soundness because of the composability it offers. How to obtain composability and computational soundness at the same time remains an interesting open question but we consider this question out of the scope of this paper.

8 Results for SIV

As we have pointed out before, user-provided IVs constitute a considerable attack vector. An alternative to generating IVs internally is to get rid of them altogether. Rogaway proposed a construction where the initialisation vector is synthesised from the authenticated information and the message using a hash function [41] (see Figure 3).

We can readily apply the deduction soundness result to SIV mode, if we apply the construction sketched in Figure 4.

As Rogaway showed, this construction can turn SIV mode into a MRAE secure scheme [41, Section 7], which implies DAE-N security. Interestingly, the construction effectively vanishes if either iv or h are always set to the empty string

7. The payload in the rule for encryption (5) is guaranteed to not be a key by lemma `key_int_conf`.

dep.	lemma	steps	seconds
○	origin	2087	103
○	counter_mono	1880	79
○	uniqueness_IV	8	16
○	key_usage	86	18
○	key_int_conf	443	46
○	key_level_handle	170	22

TABLE 3: Results for SIV mode.

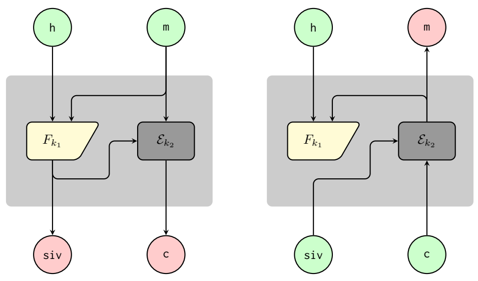


Fig. 3: SIV encryption (left) and decryption (right).

ε. We can therefore argue about SIV mode by slightly modifying our model so that the third position of $\mathsf{senc}/4$, i.e., the authenticated information, is always set to ϵ. As concatenation cancels out, SIV by itself is a valid cryptographic implementation and the existing deduction soundness result applies. We must still ensure uniqueness of iv, so we include the device identifier and counter in the header. Finally, we thus verify all lemmas from Section 6 in 284 seconds overall (see Table 3 for details). SIV mode was considered for inclusion in PKCS#11 v3.0 [44], but as of now, it is *not* supported [35]. If SIV was supported, no modification to `C_WrapKey` would be necessary (as opposed to internal-nonce generation for GCM/CCM).

9 Related work

The search for logical attacks on security APIs goes back to Longley and Rigby [31] and Bond and Anderson [9]. There is a huge body of work specifically on PKCS#11 [10, 14, 18], but there have also been academic proposals for new APIs [28, 15, 27]. While attacks were often a driving factor, a lot of effort was directed towards finding configurations that are secure, i.e., that preserve secrecy of keys.

There are three major approaches to the analysis of PKCS#11 configurations. The first is using program verification techniques, but this was not automated and therefore has largely been discarded [20, 21]. The second approach is using security type-checking on the implementation, e.g., C-code [12] or a domain-specific language [2]. This technique was used to show secrecy of keys against a Dolev-Yao attacker, but the type-system needs to be modified to reflect new cryptographic primitives like AEAD encryption. With the third approach, adoption of new primitives is easier. Here, protocol verification techniques are used. Essentially, the security token is the only participant in a protocol, and the API-level adversary is represented by the network attacker. Early results were based on model-checking [18] and thus limited to a fixed number of keys, but under certain assumptions, the soundness for an unbounded number of keys can be established [22]. The high degree of automation even allows for automated attack reconstruction [10]. More flexibility can be achieved by using protocol verification tools in the unbounded model, as existing results for the soundness of a bounded model do not apply if the API itself is modified, e.g., by introduction of stronger cryptographic primitives [29]. To our knowledge, the two

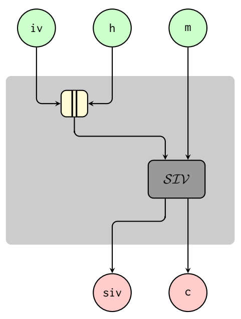


Fig. 4: DAE-N/MRAE secure scheme from SIV mode.

most functional yet secure configurations that were discovered either have keys that lose functionality on wrapping and reimporting [10] or do not allow to export wrapping keys [10, 29].

In contrast to finding configuration which are secure against logical attacks, cryptographic security proofs for Security APIs [27, 11] achieve stronger guarantees, but have not been automated so far. Even though some results retain compatibility with PKCS#11 [42], their focus is on secure design, not identification of secure configurations. Furthermore, following cryptographic necessity, the proposed design forbids that keys may be used for more than one purpose, e.g., the keys used for wrapping and encryption need to be separated by design, in contrast to the policy identified here. While this is cryptographic good practice, PKCS#11 policies often provide this functionality to allow for more flexibility in HSM-based protocols.

The idea of relating symbolic abstractions to cryptographic security notions goes back to Abadi and Rogaway's introduction of computational soundness [1]. Various results established the soundness of symmetric encryption [6], signatures [7], and hash function [23], just to name a few. Most results exclude key-cycles [6], however, it is possible to overcome this limitation by strengthening the cryptographic requirements [3] or the Dolev-Yao attacker [30]. A priori, these results do not compose, hence Cortier and Warinschi proposed *deduction soundness* [16] as a framework that allows for some amount of composability. Subsequent work in this framework covered most cryptographic primitives present in PKCS#11, including MACs, hashes, signatures, symmetric and public key encryption [8]. To be sure that we handle device-internal nonce generation correctly, we introduce deterministic authenticated encryption with associated data to this framework.

10 Conclusion

We summarize our suggestions for PKCS#11 version 3.0 and other Security APIs and point out challenges in the protocol verification approach.

The addition of AEAD schemes to PKCS#11 has shown great potential for functional and secure key-management policies. It is vital that HSMs can guarantee network-wide unique IVs, thus this should be mandated for key-wrapping. The current interface does not provide this IV in the function output, which is making a device-internal generation impossible or at least unnecessarily complicated. The attributes attached to a key should be authenticated with the wrapping, and AES keys should either be usable for wrapping and unwrapping, or for encryption and decryption. In contrast to previous policies, the authenticity of a key's attribute is guaranteed and thus both encryption and wrapping keys can be wrapped. While we proposed this policy for PKCS#11, it is also compatible with the Key Management Interoperability Protocol (KMIP) [26], an independent standard for key-management that is also governed by OASIS. KMIP allows for (but does not default to) authenticating attributes when exporting and importing keys. It provides support for the GCM and CCM modes of operation as well as internal IV generation.

Our approach was based on protocol verification, which was flexible enough to handle the introduction of new primitives, however, finding the correct equations and protocol conditions is not easy. Despite the huge body of work in computational soundness, there was no result that gave an answer right away. No computational soundness results covers the range of cryptographic primitives supported by PKCS#11. While Böhl's deduction soundness result does, thanks to its composability, it provides weaker guarantees. We thus encourage future research to consolidate existing knowledge on computational soundness and to facilitate the adoption of new primitives by investigating the composability of computationally sound cryptographic primitives.

Acknowledgements

This work has been partially funded by the German Research Foundation (DFG) via the collaborative research center "Methods and Tools for Understanding and Controlling Privacy" (SFB 1223), project B3.

References

[1] Martin Abadi and Phillip Rogaway. "Reconciling two views of cryptography (the computational soundness of formal encryption)". In: *Journal of cryptology* 15.2 (2002), pp. 103–127.

[2] Pedro Adão, Riccardo Focardi, and Flaminia L. Luccio. "Type-Based Analysis of Generic Key Management APIs". In: *CSF 2013*. 2013, pp. 97–111.

[3] Pedro Adão et al. "Soundness of Formal Encryption in the Presence of Key-Cycles". In: *Computer Security – ESORICS 2005*. Springer Berlin Heidelberg, 2005, pp. 374–396.

[4] *Annex A: Approved Security Functions for FIPS PUB 140-2,Security Requirements for Cryptographic Modules*. Tech. rep. NIST, 2018.

[5] Michael Backes, Ankit Malik, and Dominique Unruh. "Computational Soundness Without Protocol Restrictions". In: *Proceedings of the 2012 ACM Conference on Computer and Communications Security*. CCS '12. ACM, 2012, pp. 699–711.

[6] Michael Backes and Birgit Pfitzmann. "Symmetric encryption in a simulatable Dolev-Yao style cryptographic library". In: *Computer Security Foundations Workshop, 2004. Proceedings. 17th IEEE*. IEEE. 2004, pp. 204–218.

[7] Michael Backes, Birgit Pfitzmann, and Michael Waidner. "A composable cryptographic library with nested operations". In: *Proceedings of the 10th ACM conference on Computer and communications security*. ACM. 2003, pp. 220–230.

[8] Florian Böhl, Véronique Cortier, and Bogdan Warinschi. "Deduction soundness: prove one, get five for free". In: *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. ACM. 2013, pp. 1261–1272.

[9] M. Bond and R. Anderson. "API level attacks on embedded systems". In: *IEEE Computer Magazine* (2001), pp. 67–75.

[10] Matteo Bortolozzo et al. "Attacking and Fixing PKCS#11 Security Tokens". In: *17th ACM Conference on Computer and Communications Security (CCS'10)*. ACM, 2010, pp. 260–269.

[11] C. Cachin and N. Chandran. "A Secure Cryptographic Token Interface". In: *Proc. 22th IEEE Computer Security Foundation Symposium (CSF'09)*. IEEE Comp. Soc. Press, 2009, pp. 141–153.

[12] Matteo Centenaro, Riccardo Focardi, and Flaminia L. Luccio. "Type-based analysis of key management in PKCS#11 cryptographic devices". In: *Journal of Computer Security* 21.6 (2013).

[13] Jolyon Clulow. "On the Security of PKCS #11". In: *Cryptographic Hardware and Embedded Systems - CHES 2003*. Springer-Verlag, 2003, pp. 411–425.

[14] V. Cortier, G. Keighren, and G. Steel. "Automatic Analysis of the Security of XOR-based Key Management Schemes". In: *TACAS 2007*. LNCS. Springer, 2007.

[15] Véronique Cortier, Graham Steel, and Cyrille Wiedling. "Revoke and let live: a secure key revocation API for cryptographic devices". In: *CCS 2012*. ACM, 2012.

[16] Veronique Cortier and Bogdan Warinschi. "A composable computational soundness notion". In: *Proceedings of the 18th ACM conference on Computer and communications security*. ACM. 2011, pp. 63–74.

[17] Alexander Dax et al. *How to wrap it up – A formally verified proposal for the use of authenticated wrapping in PKCS#11*. Tech. rep. https://eprint.iacr.org/2019/462. 2019.

[18] Stéphanie Delaune, Steve Kremer, and Graham Steel. "Formal Analysis of PKCS#11 and Proprietary Extensions". In: *Journal of Computer Security* 18.6 (2010), pp. 1211–1245.

[19] Morris J. Dworkin. *SP 800-38C. Recommendation for Block Cipher Modes of Operation: The CCM Mode for Authentication and Confidentiality*. Tech. rep. 2004.

[20] Sibylle B. Fröschle and Nils Sommer. "Reasoning with Past to Prove PKCS#11 Keys Secure". In: *7th International Workshop on Formal Aspects in Security and Trust (FAST'10)*. LNCS. 2010, pp. 96–110.

[21] Sibylle Fröschle and Nils Sommer. "Concepts and Proofs for Configuring PKCS#11". In: *FAST 2011*. LNCS. Springer, 2012.

[22] Sibylle Fröschle and Graham Steel. "Analysing PKCS#11 Key Management APIs with Unbounded Fresh Data". In: *Joint Workshop on Automated Reasoning for Security Protocol Analysis and Issues in the Theory of Security (ARSPA-WITS'09)*. LNCS. Springer, 2009.

[23] Romain Janvier, Yassine Lakhnech, and Laurent Mazaré. "Completing the picture: Soundness of formal encryption in the presence of active adversaries". In: *European Symposium on Programming*. Springer. 2005, pp. 172–185.

[24] Jakob Jonsson. "On the Security of CTR + CBC-MAC". In: *Revised Papers from the 9th Annual International Workshop on Selected Areas in Cryptography*. SAC '02. Springer-Verlag, 2003, pp. 76–93.

[25] Charanjit S. Jutla. "Encryption Modes with Almost Free Message Integrity". In: *Advances in Cryptology — EUROCRYPT 2001*. Springer Berlin Heidelberg, 2001, pp. 529–544.

[26] *Key Management Interoperability Protocol Specification Version 1.4*. Tech. rep. OASIS, 2017.

[27] Steve Kremer, Robert Künnemann, and Graham Steel. "Universally Composable Key-Management". In: *ESORICS 2013*. LNCS. Springer, 2013.

[28] Steve Kremer, Graham Steel, and Bogdan Warinschi. "Security for Key Management Interfaces". In: *CSF 2011*. IEEE Computer Society, 2011, pp. 66–82.

[29] Robert Künnemann. "Automated backward analysis of PKCS#11 v2.20". In: *4th Conference on Principles of Security and Trust (POST'15)*. LNCS. Springer, 2015, pp. 219–238.

[30] Peeter Laud. "Encryption cycles and two views of cryptography". In: *Proceedings of the 7th Nordic Workshop on Secure IT Systems (NORDSEC)*. 31. Citeseer. 2002, pp. 85–100.

[31] Dennis Longley and Simon Rigby. "An Automatic Search for Security Flaws in Key Management Schemes". In: *Computers and Security* 11.1 (1992), pp. 75–89.

[32] David A. McGrew and John Viega. *The Security and Performance of the Galois/Counter Mode of Operation (Full Version)*. Cryptology ePrint Archive, Report 2004/193. https://eprint.iacr.org/2004/193. 2004.

[33] *PKCS #11 Cryptographic Token Interface Base Specification Version 2.40 Plus Errata 01*. Tech. rep. May 2016.

[34] *PKCS #11 Cryptographic Token Interface Base Specification Version 3.0 Working Draft 05*. Tech. rep. July 2018.

[35] *PKCS #11 Cryptographic Token Interface Current Mechanisms Specification Version 2.40 Plus Errata 01*. Tech. rep. May 2016.

[36] P. Leach, M. Mealling, and R. Salz. *A Universally Unique IDentifier (UUID) URN Namespace*. RFC 4122 (Proposed Standard). RFC. RFC Editor, July 2005. URL: https://www.rfc-editor.org/rfc/rfc4122.txt.

[37] Phillip Rogaway. "Authenticated-encryption with Associated-data". In: *Proceedings of the 9th ACM Conference on Computer and Communications Security*. CCS '02. ACM, 2002, pp. 98–107.

[38] Phillip Rogaway. "Authenticated-encryption with associated-data". In: *Proceedings of the 9th ACM conference on Computer and communications security*. ACM. 2002, pp. 98–107.

[39] Phillip Rogaway. "Evaluation of some blockcipher modes of operation". In: *Cryptography Research and Evaluation Committees (CRYPTREC) for the Government of Japan* (2011).

[40] Phillip Rogaway and Thomas Shrimpton. "A provable-security treatment of the key-wrap problem". In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer. 2006, pp. 373–390.

[41] Phillip Rogaway and Thomas Shrimpton. "The SIV mode of operation for deterministic authenticated-encryption (key wrap) and misuse-resistant nonce-based authenticated-encryption". In: *Aug* 20 (2007), p. 3.

[42] Guillaume Scerri and Ryan Stanley-Oakes. "Analysis of Key Wrapping APIs: Generic Policies, Computational Security". In: *IEEE 29th Computer Security Foundations Symposium, CSF 2016, Lisbon, Portugal, June 27 - July 1, 2016*. IEEE Computer Society, 2016, pp. 281–295.

[43] Benedikt Schmidt et al. "The TAMARIN Prover for the Symbolic Analysis of Security Protocols". In: *25th International Conference on Computer Aided Verification (CAV'13)*. LNCS. Springer, 2013, pp. 696–701.

[44] Graham Steel. mail on OASIS PKCS11 mailing list. June 2016.

[45] Graham Steel. *Attacks on Key-Wrapping in PKCS#11 v2.40*. June 2016. URL: https://cryptosense.com/blog/attacks-on-key-wrapping-in-pkcs11-v2-40/.

[46] Dominique Unruh. *Programmable encryption and key-dependent messages*. Cryptology ePrint Archive, Report 2012/423. https://eprint.iacr.org/2012/423. 2012.

APPENDIX A
AEAD SECURITY IMPLIES DAE-N SECURITY

We now want to show that for "AE with AD" schemes that are secure considering privacy and authenticity as defined above, it holds that those schemes are also DAE-N secure.

Lemma 1 (AEAD security [40, Proposition 8]). *Let $\Pi = (Gen, Enc, Dec)$ be a authenticated encryption with associated data with AD space $\mathcal{H}_{AD}$, IV space $\mathcal{N}$ and message space $\mathcal{M}$[38]. Let $\mathcal{A}$ be an adversary with access to two oracles. Suppose $\mathcal{A}$ runs in time $\sqcup$ and asks q_L queries to its left oracle, these totaling u_L bits, and asks q_R queries to its right oracle, these totaling u_R bits. Then there exist adversaries D and F such that*

$$Adv_{\Pi}^{dae-n}(\mathcal{A}) \leqslant Adv_{\Pi}^{priv}(D) + q_R\, Adv_{\Pi}^{auth}(F)$$

where D runs in time $\sqcup O(u_L + u_R)$ and asks q_L queries totaling u_L bits, and F runs in time $\sqcup + O(u_L + u_R)$, asking at most q_L left-queries and one right-query, these totaling at most $u_l + u_R$ bits.

Proof. This proof is exactly the proof of [40, Proposition 8], however, instead of the modified syntax for deterministic authenticity/privacy, the original syntax [38] needs to be employed, i.e., the oracles take a third input for the IV, hence $O(\cdot,\cdot)$ is replaced by $O(\cdot,\cdot,\cdot)$ for every oracle. Both definitions restrict the adversary to not query the same IV twice. □

APPENDIX B
DEDUCTION SOUNDNESS: BACKGROUND

Cortier and Warinschi [16] defined the notion of *deduction soundness* which was a first step towards a framework, helping to prove the computational soundness of cryptographic protocols more easily. Böhl et al. [8] extended their work by getting rid of some prior limitations and adding more security primitives (like MACs or hashes) to their framework.

Our contribution, is to extend the work of Böhl et al. by adding AEAD schemes to the composition theorems. To achieve this in a more readable way, we will try to use a very similar notation and writing style.

B.1 Symbolic model

Symbolic models in our sense are used to represent an abstract and formal environment to verify the security of cryptographic protocols for specific security notions in the "symbolic world". Extending the work of Böhl et al. we will adopt their symbolic model making the extension more easy and increase the readability of the work.

We define our symbolic model $\mathcal{M}$ to be a tuple of the form $\mathcal{M} = (\mathcal{T}, \leqslant, \Sigma, \mathcal{D})$, where $\mathcal{T}$ defines the set of data types, $\leqslant$ the corresponding sub type relation, Σ the signature and $\mathcal{D}$ defines the deduction system. We will define them in detail step by step in the following paragraphs.

Data types $\mathcal{T}$ and sub type relation $\leqslant$

We define the symbolic model to have data types $\mathcal{T}$, which we require

- to have a base type $\top$, and
- to have a sub type relation $\leqslant$.

We require $\leqslant$ to be a preorder and that for every $\tau \in \mathcal{T}$, it holds that $\tau \leqslant \top$ (every type $\tau \in \mathcal{T} \backslash \top$ is a sub type of the base type $\top$).

Signature Σ and variables

Let the signature Σ be a set of function symbols and corresponding arities, such that each function symbol f has an arity of the form

$$ar(f) = \tau_1 \times ... \times \tau_n \to \tau$$

with n being a natural number (n $\geqslant$ 0) and $\tau, \tau_i \in \mathcal{T}$ for $0 \leqslant i \leqslant n$.
For function symbols c with 0 arguments (meaning n = 0), we call c a *constant*.
We also refer to τ as the type of the function f (type(f) = τ), with $\tau \neq \top$ for all function symbols. The only exception to this rule are garbage terms of the base type $\top$, called $g_\top$.
Garbage of type τ: We require from our signature to contain constants g_τ, such that for all $\tau \in \mathcal{T}$, g_τ represents the garbage of type τ. Typically bit-strings produced by an adversary, which cannot be parsed to a meaningful term, are associated with garbage terms. Those constants are randomized. Randomization is denoted with *labels* (see *Labels and randomness*) applied to function symbols.
In addition to function symbols, we also define an infinite set of *variables* `variables` where for each variable $x \in$ `variables`, x is of some type $\tau \in \mathcal{T}$.

Labels and randomness

For each symbolic model $\mathcal{M}$ we fix an infinite set of *labels* `labels = labelsH` $\cup$ `labelsA` where both sets are disjoint and `labelsH` denotes an infinite set of *honest labels*, while `labelsA` denotes an infinite set of *adversarial labels*. Labels denote randomness, therefore it is essential, that we distinguish between honest and adversarial labels.

Example 3. $t := enc^{l_h}(k, m)$ *with* $l_h \in$ `labelsH` *and ar(enc) =* $\tau_{keys} \times \top \rightarrow \tau_{cipher}$

Example 4. $t := n^{l_a}()$ *with* $l_a \in$ `labelsA` *and ar(g) =* τ_{nonce}

Example 5. $t := f(t_1, t_2)$ *with ar(f) =* $\tau_1 \times \tau_2 \rightarrow \tau$

We see one application of an honest, randomized function in Example 3, whereas we see an adversarial, randomized function application in Example 4. In Example 5 we see an application of an deterministic function.

Set of terms

We define `Terms` $(\Sigma, \mathcal{T}, \leqslant) := \bigcup_{\tau \in \mathcal{T}}$ *set of terms of type* τ, where the latter is defined inductively by

$$\begin{aligned} t :=& \\ &\mid x \\ &\mid f(t_1, .., t_n) \\ &\mid f^l(t_1, .., t_n) \quad l \in \texttt{labels} \end{aligned}$$

where t is a term of type τ, x is a variable of type τ, $f(t_1, .., t_n)$ is an application of a deterministic $f \in \Sigma$ and $f^l(t_1, .., t_n)$ is an application of a randomized $f \in \Sigma$ For $f(t_1, .., t_n)$ and $f^l(t_1, .., t_n)$, we further require that each t_i is of type τ_i' with $\tau_i' \leqslant \tau_i$ and that the arity of f is $ar(f) = \tau_1 \times ... \times \tau_n \rightarrow \tau$.
We also write `Terms` instead of `Terms` $(\Sigma, \mathcal{T}, \leqslant)$ if $\mathcal{M}$ is clear from the context and also write $t = f^l(t_1, .., t_n)$ for general terms (also for deterministic function symbols).

Substitutions

For $x_i, t_i \in \mathcal{T}$ and $i \in \{1, .., n\}$ we write $\delta = \{x_1 = t_1, .., x_n = t_n\}$ to refer to a substitution with domain $dom(\delta) = \{x_1 .., x_n\}$. A substitution $\delta = \{x_1 = t_1, .., x_n = t_n\}$ is *well-typed* if for each $i \in \{1, .., n\} : t_i \leqslant x_i$. In our symbolic model we only consider well-typed substitutions. We write $\delta(t) = t\delta$ for an application of a substitution δ of a term t.

Deduction systems

Each symbolic model contains a deduction relation $\vdash$: $2^{\texttt{Terms}} \times$ `Terms` , which is used to model the capabilities of a symbolic adversary to retrieve information from terms.
With $T \vdash t$ we describe the capability of a symbolic adversary to build t from the set of terms T and say t is *deducible* from T. Usually deduction relations are defined through *deduction systems* which we formally specify as follows.

Definition 2. *(DEDUCTION SYSTEM). A deduction system D is a set of rules $\frac{t_1, .., t_n}{t}$ such that $t_1, .., t_n, t \in$ `Terms` $(\Sigma, \mathcal{T}, \leqslant)$. The deduction relation $\vdash_D \subseteq 2^{\texttt{Terms}} \times$ `Terms` associated to D is the smallest relation satisfying:*

- *$T \vdash_D t$ for any $t \in T \subseteq$ `Terms` $(\Sigma, \mathcal{T}, \leqslant)$*
- *if $T \vdash_D t_1\delta, .., T \vdash_D t_n\delta$ for some substitution δ and $\frac{t_1, .., t_n}{t} \in D$ then $T \vdash_D t\delta$.*

Review of Definition 1 by Böhl, Cortier, and Warinschi [8]

Further we require for all deduction systems the rule $\frac{}{g_\top^l}$ for all $g_\top^l \in \Sigma$ and $l \in$ `labelsA` .

Note that we omit D from $\vdash_D$ if D is clear from the context.

Additional Notes:

- *For δ being a substitution we say that $\frac{t_1\delta, .., t_n\delta}{t\delta}$ is an instantiation of $\frac{t_1, .., t_n}{t} \in D$*
- *We require deduction relations to be efficiently decidable. Knowing this, for $T \vdash t$ we can always efficiently find a sequence $\pi = T \xrightarrow{a_1} T_1 \xrightarrow{a_2} .. \xrightarrow{a_n} T_n$ such that for all $i \in \{1, .., n\}$ it holds that:*

1) $a_i = \frac{t_1, .., t_n}{t}$
2) $t_1, ..t_n \in T_{i-1}$
3) $t' \notin T_{i-1}$
4) $t' \in T_i$
5) $t \in T_n$

We refer to π as an deduction proof for $T \vdash t$.

B.2 Implementation

As well as the symbolic model, we also redefine the corresponding implementation of Böhl, Cortier, and Warinschi [8].
We define $\mathcal{I}$ as an implementation of a symbolic model $\mathcal{M}$ with $\mathcal{I} = (M_\eta, [\![\cdot]\!]_\eta, len_\eta, \texttt{open}_\eta, \texttt{valid}_\eta)_\eta$ being a family of tuples. In the following, we will omit the security parameter η for readability.

The components of $\mathcal{I}$ are defined as follows:

- ***Turing machine** M*
  Provides concrete algorithms for the corresponding function symbols in the signature of the symbolic mode. Those algorithms operate on bitstrings.
- ***Mapping function** $[\![\cdot]\!]$*
  The function $[\![\cdot]\!] : \mathcal{T} \to 2^{\{0,1\}^*}$ maps each type to a set of bitstrings.
- ***Length function** len*
  The function $len : \texttt{Terms} \to \mathbb{N}$ returns for each term the length of its corresponding bitstring.
- ***Interpretation algorithm** `open`*
  The algorithm `open` interprets bitstrings as terms.
- ***The `valid` predicate***
  To get computational soundness results we usually have to restrict the implementation. In our very case, the `valid` function takes care of this. The `valid` function takes a trace $\mathbb{T}$ as an input (we will formally introduce traces later on in Definition 7) and returns a boolean value representing the validity of this trace.

Additionally we have the following requirements on the implementation $\mathcal{I}$ (to the symbolic model $\mathcal{M}$ corresponding).

Requirements:

1) *For each $\tau \in \mathcal{T}$ we require $[\![\tau]\!] \subseteq \{0,1\}^\eta$ to be nonempty. For the base type $\top$ we require $[\![\top]\!] = \{0,1\}^*$ and that for all $\tau, \tau' \in \mathcal{T}$ it holds that, if $\tau \leqslant \tau'$ then $[\![\tau]\!] \subset [\![\tau']\!]$ and otherwise $[\![\tau]\!] \cap [\![\tau']\!] = \emptyset$.*
   Further we define
   - *$[\![\mathcal{T}]\!]$ as $\bigcup_{\tau \in \mathcal{T}/\{\top\}} [\![\tau]\!]$*
   - *$\langle c_1, .., c_n, \tau \rangle$ as a bijective function that encodes $c_1, .., c_n$ to c' with $c' \in [\![\tau]\!]$.*
2) *The Turing Machine M is required to be deterministic, but has a random tape $\mathcal{R}$ provided at each run. Precisely, we require for each $f \in \Sigma$ which is not a garbage symbol and has arity $ar(f) := \tau_1 \times .. \times \tau_n \to \tau$, that M calculates a function (Mf) on input f. The domain of (Mf) should be $[\![\tau_1]\!] \times .. \times [\![\tau_n]\!] \times \{0,1\}^*$ and the range $[\![\tau]\!]$.*
   To generate a bitstring for a term $t = f^l(t_1, .., t_n)$ we apply (Mf) to the bitstrings of the subterms t_i of t and use some randomness (except for deterministic function symbols). The recursively resulting bitstring $[\![t]\!]$ is called concrete interpretation of t.

Interpretations

In cryptographical applications, it occurs that the same random values occur multiple times within the same term, for instance, if a random nonce appears multiple times in an encryption. To deal with such occurrences within our interpretation, we will use a partially defined mapping $L : \{0,1\}^* \to \texttt{HTerms}$ from arbitrary bitstrings to so-called Hybrid Terms, which acts as a library.

Hybrid Terms** are either garbage terms or $t = f^l(c_1, .., c_n)$ where $f \in \Sigma$ with $ar(f) := \tau_1 \times .. \times \tau_n \to \tau$ and $c_i \in [\![\tau_i]\!]$ for $i \in \{1, .., n\}$. The **domain of L** $dom(L) \subseteq 2^{\{0,1\}^*}$ is the set of bitstrings for which L is already defined. The **interpretation of bitstrings** $c \in dom(L)$ **with respect to L** is defined as $L[\![c]\!] := f^l(\ L[\![c_1]\!], .., L[\![c_n]\!]\ )$ if $L(c) = f^l(c_1, .., c_n)$ and $L[\![c]\!] := L(c)$ for $L(c)$ being a garbage term. A **complete mapping** is given, if $\forall (c, f^l(c_1, .., c_n)) \in L : c_1, .., c_n \in dom(L)$. **Generating function
In the following we will define a generate function using a mapping L introduced in section B.2. This generate function takes a term $t = f^l(t_1, .., t_n)$ as input and produces a bitstring c corresponding to t.

To produce c, the Turing machine M has to use the random tape $\mathcal{R}$ as a source of randomness, using an algorithm $\mathcal{R}(t) : \texttt{Terms} \to \{0,1\}^\eta$ which produces a different random bitstring r depending on a term $t \in \texttt{Terms}$.

Listing 1: `generate`

```
generate$_{M,\mathcal{R}}(t, L)$:
    if for some $c \in dom(L)$ we have $L[\![c]\!] == t$ then
        return $(c, L)$
    else
        for $i \in \{1, ..., n\}$ do
            let $(c_i, L) :=$ generate$_{M,\mathcal{R}}(t_i, L)$
        let $r := \mathcal{R}(t)$
        let $c := (Mf)(c_1, ..., c_n; r)$
        let $L(c) := f^l(a_1, ..., a_n))(l \in$ labesH$)$
        return $(c, L)$
```

Additional notes:

- *`generate` also updates L.*
- *`generate` depends on M and $\mathcal{R}$, but we will omit them if they are clear from the context.*
- *If L is complete, it holds that for all terms $t \in \texttt{Terms}$ with $t = f^l(t_1, .., t_n)$ and $(c, L') :=$ `generate` (t, L), L' is also complete.*
- *For all subterms $t' = f^l(t_1, .., t_n)$ of t where $l \in \texttt{labelsA}$, there is a $c \in dom(L)$ with $L[\![c]\!] = t'$*

- *There is no t'' being a subterm of t, s.t. $t'' = g^l()$ and $l \in$* `labelsH` *.*

The last point ensures that all bitstrings generated by the `generate` *function only contain non-garbage function applications where the function symbols carry an honest label $l \in$* `labelsH` *.*

Parsing function

In addition to the `generate` *function which produces bitstrings from terms, we require from the implementation the definition of a function which produces terms from bitstrings. The function takes a bitstring c and a mapping L as input and returns a term $t \in$* `Terms` *and an updated mapping L'.*

Since the function depends on the `open` *: $\{0,1\}^* \times$* `libs` *$\rightarrow \{0,1\}^* \times$* `HTerms` *function provided by an concrete implementation, we have to omit an exact definition of parse but require the following structure:*

Listing 2: `parse`

```
parse$(c, L)$:
    if $c \in dom(L)$ then
        return $(L[\![c]\!], L)$
    else
        let $L_h := \{(\bar{c}, f^l(...)) \in L | l \in \texttt{labelsH}\}$
        let $L := (\bigcup_{(\bar{c},\cdot) \in L} \texttt{open}(\bar{c}, L_h))$
        let $G := \{(\bar{c}, g_\top^{l(\bar{c})})\} \quad (l \in \texttt{labelsH})$
        while $G \neq \emptyset$ do
            let $L := (L \backslash G) \cup (\bigcup_{(\bar{c},\cdot) \in G} \texttt{open}(\bar{c}, L_h))$
            let $G := \{(\bar{c}, g_\top^{l(\bar{c})}) | (\bar{c}, f^l(.., \bar{c}, ..)) \in L \wedge \bar{c} \in dom(L)\}$
        return $(L[\![c]\!], L)$
```

Additional notes:

- *has to provide a concrete context such that different* `open` *functions can be composed*
- `open` *is only allowed to use honestly generated bitstrings*
- *Foreign bitstrings, meaning bitstrings of a data type which is not part of the implementation, are ignored by* `open`
- `open` *functions are commutative due to the properties stated above, which is important for the composition theorems in F.*

Good implementation

Starting in this section we will describe several requirements and restrictions towards the behavior of an implementation. The term good implementation refers to those implementations which satisfy all properties defined in this section.

Length regularity

We require that $len(f^l(t_1, ..., t_n)) := |(Mf)(c_1, ..., c_n, r)|$ only depends on the length of c_i. Such a length function len is equivalent to a set of length functions $len_f : \mathbb{N}^n \rightarrow \mathbb{N}$ for each function symbol $f \in \Sigma$ with $ar(f) := \tau_1 \times ... \times \tau_n \rightarrow \tau$. This equivalence is needed to be able to compose length functions of different implementations as described in Section B.4.

Collision freeness

Problems occur if values in the mapping L are overwritten. Assume a bitstring $c \in dom(L)$ is then parsed to a term t' instead of the intended term t. This could prevent an adversary from winning the deduction soundness game, so a good implementation should be collision free. To define what collision freeness means we also need to introduce the notion of a supplementary transparent function to model the capability of an adversary to choose arbitrary bitstrings for arguments of type $\top$. This is needed since transparent implementations also need to be collision free, because we want them to be composable, too.

In Definition 3 we introduce the mentioned supplementary transparent functions and in Definition 4 we finally introduce the notion of collision freeness.

Definition 3. `Supplementary transparent functions`

For a set of bitstrings $\mathcal{B} \subset \{0,1\}^*$ we define the transparent model $\mathcal{M}_{supp}^{tran}(\mathcal{B})$ as follows:

- *$\mathcal{T}_{supp}^{tran} := \{\top, \tau_{supp}^{tran}\}$. τ_{supp}^{tran} is a subtype of $\top$.*
- *$\Sigma_{supp}^{tran} := \{f_c \;:\; c \in \mathcal{B}\}$ (all function symbols are deterministic)*
- *$D_{supp}^{tran} := \{\frac{}{f_c()} \;:\; c \in \mathcal{B}\}$*

and an implementation $\mathcal{I}_{supp}^{tran}(\mathcal{B})$ as follows:

- *$[\![\tau_{supp}^{tran}]\!] := \mathcal{B}$*
- *$(M_{supp}^{tran} f_c)()$ returns c*
- *$(M_{supp}^{tran} func)(c)$ returns f_c if $c \in \mathcal{B}$, $\perp$ otherwise*

(Review of Definition 2 by Böhl, Cortier, and Warinschi [8])

Definition 4. `Collision-free implementation`

Let $DS'_{\mathcal{M},\mathcal{I},A}(\eta)$ be the deduction soundness game from Listing 5 where we replace the `generate` *function by the function* `generate'` *from Listing 1. We say an implementation $\mathcal{I}$ is collision free if for all p.p.t. adversaries A*

$$Pr[DS_{\mathcal{M}\cup\mathcal{M}_{supp}^{tran}([\![\mathcal{T}]\!]),\mathcal{I}\cup\mathcal{I}_{supp}^{tran}([\![\mathcal{T}]\!]),A}(\eta) = 1]$$
$$-Pr[DS'_{\mathcal{M}\cup\mathcal{M}_{supp}^{tran}([\![\mathcal{T}]\!]),\mathcal{I}\cup\mathcal{I}_{supp}^{tran}([\![\mathcal{T}]\!]),A}(\eta) = 1]$$

is negligible.
(Review of Definition 3 by Böhl, Cortier, and Warinschi [8])

Type safety
We define type safety in the following sense

Definition 5. `Type safe implementation`
We say that an implementation $\mathcal{I}$ of a symbolic model $\mathcal{M}$ is type safe if
(i) $\mathsf{open}(c, L) = (c, g^l_\top)$ *for* $l \in$ `labelsA` *if* $c \notin [\![\mathcal{T}]\!]$.
(ii) $\mathsf{open}(c, L) = \mathsf{open}(c, L|[\![\mathcal{T}]\!])$ *where* $L|\,[\![\mathcal{T}]\!] := \{(c, h) \in L : \exists \tau \in \mathcal{T}\backslash\{\top\} : c \in [\![\tau]\!]\}$.
(Review of Definition 4 by Böhl, Cortier, and Warinschi [8])

Intuitively, (i) *states that* open *should not work on foreign bitstrings and only return garbage of type* $\top$*, while* (ii) *states that the behavior of the* open *function is not altered by any foreign bitstring in L. This guarantees that composed* open *functions do not interfere with each other, meaning they do not work on each others bitstrings.*
Since we demand to run in polynomial time in size of the mapping L*, we therefore demand a reasonable run time for* open *(since is based on* open *).* **`valid` *requirements***
Before we can start to define our requirements on the `valid` *function, we will now introduce the notions of queries and traces.*

Definition 6. *Query*
Queries in our case are requests sent by the adversary to the deduction soundness game (or any variation of it). A query q is either "`generate` t*", "*`sgenerate` t*", "c" or "*`init` T, H*".*

Definition 7. *Trace*
We define a trace $\mathbb{T}$ *as a time ordered list of queries* $\mathbb{T} := q_1 + q_2 + \ldots + q_n$

Knowing this, we now state the requirements on `valid` *which we need to be able to compose* `valid` *functions later on.*

(i) *If* `valid`$(\mathbb{T} + q)$ = `true` *then for any variation* $\hat{q}$ *of* q*, which we define explicitly below, it holds that* `valid`$(\mathbb{T} + \hat{q})$ = `true`.
If q = "generate t" and $\hat{q}$ *= "generate* $\hat{t}$*":*
Here, a variation $\hat{t}$ *of* t *is defined as follows: Any subterm* $f^l(t_1, ..., t_n)$ *of* t *where* f *is a foreign function symbol, i.e.* $f \notin \Sigma$*, can be replaced by* $\hat{f}^l(\hat{t}_1, ..., \hat{t}_n)$ *where* $\hat{f}$ *is another foreign function symbol and for all* $i \in \{1, ..., n\}$ $\hat{t}_i$ *is either* t_j *for some* $j \in \{1, ..., n\}$ *(where each* t_j *only can be used once) or* $\hat{t}_i$ *does not contain any function symbols from* Σ*. As a special case, if for example* $\hat{f}$ *is "empty", we may replace* $f^l(t_1, ..., t_n)$ *with a term* $\hat{t}_1$*.*
If q = "init T, H" and $\hat{q}$ *= "init* $\hat{T}, \hat{H}$*":*
$T = (t_1, ..., t_n)$ *and* $\hat{T} = (\hat{t}_1, ..., \hat{t}_n)$ *where for each* $i \in \{1, ..., n\}$ $\hat{t}_i$ *is a variation of* t_i*; and* $H = (h_1, ..., h_m)$ *and* $\hat{H} = (\hat{h}_1, ..., \hat{h}_m)$ *where for each* $i \in \{1, ..., m\}$ $\hat{h}_i$ *is a variation of* h_i*.*
(ii) *We demand that, if* `valid`$(\mathbb{T} + q)$ = `true` *and term* $t \in q$*, then for any* $t \in st(t)$ *it holds that* `valid`$(\mathbb{T}+$ "`sgenerate` t'"$)$ = `true`.
(iii) *The run time of* `valid`$(\mathbb{T})$ *should be polynomial and should only depend on* $\mathbb{T}$.

To conclude: If an implementation $\mathcal{I}$ *is length regular, collision free, type safe and if the requirements on* `valid` *hold, we call* $\mathcal{I}$ *a good implementation.*

B.3 Transparent functions

To conclude the idea from Definition 3, we will now start to formally define the notion of transparent functions (and also transparent symbolic models and transparent implementations). We need those notions, because we want to achieve our soundness results in the presence of any public data structure, for instance: tuples, arrays, XML files etc. We call those public data structures transparent functions and will define transparent symbolic models and transparent implementations using such functions in the following.
Transparent symbolic models:
A transparent symbolic model $\mathcal{M}_{tran} = (\mathcal{T}_{tran}, \leqslant_{tran}, \Sigma_{tran}, \mathcal{D}_{tran})$ *is defined similar to the symbolic model from B.1 but having a specific definition of the deduction system* $\mathcal{D}_{tran}$ *stated below:*

$$\left\{ \begin{array}{cc} \dfrac{t_1 \cdots t_n}{f^l(t_1, \cdots, t_n)} & l \in labelsA, f \in \Sigma_{tran} \\ \\ \dfrac{f^l(t_1, \cdots, t_n)}{t_i} & 1 \leqslant i \leqslant n, l \in labelsA, f \in \Sigma_{tran} \end{array} \right\}$$

Transparent implementations:
A transparent implementation $\mathcal{I}_{tran} = (M_{tran}, [\![\cdot]\!]_{tran}, len_{tran}, open_{tran}, valid_{tran})$ *of a symbolic model is also an implementation, meaning it meets the same requirements as in B.2. For the Turing machine* M_{tran} *we require two additional modes of operations:* $func$

and $proj$. We define them right below:
For all function symbols $f \in \Sigma$ with arity $ar(f) = \tau_1 \times \ldots \times \tau_n \to \tau$ we define

$$(M_{tran}\ func) : \{0,1\}^* \to \Sigma \cup \{\bot\}$$
$$(M_{tran}\ prod\ f\ i) : \{0,1\}^* \to \{0,1\}^*$$

such that for all $c_i \in [\![\tau_i]\!]$ with $1 \leqslant i \leqslant n$ and $r \in \{0,1\}^\eta$

$$(M_{tran}\ func)((M_{tran}\ f)(c_1,\ldots,c_n;r)) = f$$
$$(M_{tran}\ prod\ f\ i)((M_{tran}\ f)(c_1,\ldots,c_n;r)) = c_i$$

and require that $(M_{tran}\ func) = \bot$ for all $c \notin [\![\mathcal{T}]\!]$. We also require that $(M_{tran}\ func)$ and $(M_{tran}\ prod\ f\ i)$ both run in polynomial time (depending on η).
The bitstring mapping $[\![\cdot]\!]_{tran}$ as well as the length function len_{tran} are not defined any different as in B.2. The open function $open_{tran}$ on the other hand is defined explicitly below:

Listing 3: open$_{tran}$

```
$open_{tran}$(c, L)
    if c $\in [\![\mathcal{T}]\!] \cap$ dom(L) then
        return (c, L(c))
    else if $(M_{tran}\ func)(c) == \bot$ then
        find unique $\tau \in \mathcal{T}$ s.t. $c \in [\![\tau]\!]$ and $c \notin [\![\tau']\!]$
            for all $\tau' \in \mathcal{T}$ with $[\![\tau']\!] \subsetneq [\![\tau]\!]$
        return $(c, g_\tau^{l(c)})$ $(l(c) \in labelsA)$
    else
        let $f := (M_{tran}\ func)(c)$ with
            ar($f$) = $\tau_1 \times \ldots \times \tau_n \to \tau$
        if  $(M_{tran}\ prod\ f\ i) == \bot$
                for any $i \in [1,n]$ then
            return $(c, g_\tau^{l(c)})$ $(l(c) \in labelsA)$
        else
            for $i = 1$ to $n$ do
                let $c_i := (M_{tran}\ prod\ f\ i)(c)$
            return $(c, f^{l(c)}(c_1,\ldots,c_n))(l(c) \in labelsA)$
```

With this function, we can easily show that the implementation is type safe by using the two properties required by Definition 5. The first property is already required above (see $func$) and the second property holds because L is not used by the $open_{tran}$ function.
Also transparent functions are not restricted in any way so we define $valid_{tran}(\mathbb{T})$ = true for all possible traces $\mathbb{T}$.

B.4 Composition

In this we will explain how to compose two symbolic models $\mathcal{M}_1 = (\mathcal{T}_1, \leqslant_1, \Sigma_1, D_2)$ and $\mathcal{M}_2 = (\mathcal{T}_2, \leqslant_2, \Sigma_2, D_2)$ and there corresponding implementations $\mathcal{I}_1$ and $\mathcal{I}_2$ in a general way.

We define the composition of $\mathcal{M}_1$ and $\mathcal{M}_2$ as

$$\mathcal{M}' = (\mathcal{T}_1 \cup \mathcal{T}_2, \leqslant_1 \cup \leqslant_2, \Sigma_1 \cup \Sigma_2, D_1 \cup D_2)$$

if
1) $\Sigma_1 \cap \Sigma_2 = \{g_\top\}$
2) $\mathcal{T}_1 \cap \mathcal{T}_2 = \{\top\}$

To compose two implementations $\mathcal{I}_1$ and $\mathcal{I}_2$ we require that
i) *for all $\tau_1 \in \mathcal{T}_1 \backslash \{\top\}, \tau_2 \in \mathcal{T}_2 \backslash \{\top\}$ we have $[\![\tau_1]\!] \cap [\![\tau_2]\!] = \varnothing$*
ii) *and further that $\mathcal{I}' = \mathcal{I}_1 \cup \mathcal{I}_2$ is a good implementation of $\mathcal{M}'$.*

To fulfill requirement ii) we now start to define $\mathcal{I}' := \mathcal{I}_1 \cup \mathcal{I}_2$. The Turing machine $M' = M_1 \cup M_2$ returns $(M' f) := (M_1 f)$ if $f \in \Sigma_1$ or $(M' f) := (M_2 f)$ otherwise.
For all $\tau \in \mathcal{T}_1$ we set $[\![\tau]\!]' := [\![\tau]\!]_1$ and for all $\tau \in \mathcal{T}_2$ we set $[\![\tau]\!]' := [\![\tau]\!]_2$. Further we require all $[\![\top]\!]' = [\![\top]\!]_1 = [\![\top]\!]_2 = \{0,1\}^*$
As we saw in section B.2, the length functions are easy composable. We have $len_1 := len_{f_1} : \mathbb{N}^N \to \mathbb{N}$ with $f_1 \in \Sigma_1$ and $len_2 := len_{f_2} : \mathbb{N}^N \to \mathbb{N}$ with $f_2 \in \Sigma_2$. We simply define $len' := len_1 \cup len_2$.
We define the composition of the open *function in the following way:*

Listing 4: open$'$

```
$(\mathsf{open}_1 \circ \mathsf{open}_2)(c, L)$:
    let $(c,t) := \mathsf{open}_1(c, L)$
    if $t = g_\top^l$ for some $l \in$ labelsA  then
        return $\mathsf{open}_2(c, L)$
    else
        return $(c,t)$
```

Therefore we have $\mathsf{open}' := \mathsf{open}_1 \circ \mathsf{open}_2$.

From the valid predicate `valid'`*, we demand both* $\mathtt{valid}_1$ *and* $\mathtt{valid}_2$ *to remain valid, so we write* $\mathtt{valid}'(\mathbb{T}) := \mathtt{valid}_1(\mathbb{T}) \wedge \mathtt{valid}_2(\mathbb{T})$.

After defining the composition $\mathcal{I}' = (M', [\![\cdot]\!]', len', \mathsf{open}', \mathtt{valid}')$ *we need to show that* $\mathcal{I}'$ *is a good implementation B.2 of* $\mathcal{M}'$.

Type safety *holds since* $\mathcal{T}_1 \cap \mathcal{T}_2 = \{\top\}$ *is given by requirement (2.) of composing symbolic models. This implies that* $[\![\mathcal{T}_1]\!]' \cap [\![\mathcal{T}_2]\!]' = [\![\top]\!]'$ *and since we have* $\mathcal{I}_1$ *and* $\mathcal{I}_2$ *being type safe and* $\mathsf{open}_1 \circ \mathsf{open}_2 = \mathsf{open}_2 \circ \mathsf{open}_1$ *we can easily conclude that* $\mathcal{I}'$ *is also type safe (see B.2 for concrete requirements).*

Length regularity *is automatically given for* len' *since* $\Sigma_1 \cap \Sigma_2 = \{g'_\top\}$ *leads to* $len_1 \cap len_2 = \varnothing$*. Hence, for any input* $f \in \Sigma_1$ $len'(f^l(t_1, ..., t_n))$ *returns* $len_1(f^l(t_1, ..., t_n))$*, what is length regular by assumption, and returns* $len_2(f^l(t_1, ..., t_n))$ *for* $f \in \Sigma_2$ *otherwise, what is naturally also length regular by assumption.*

The **`valid` *requirements*** *are fulfilled by* `valid'` *by construction since both* $\mathtt{valid}_1$ *and* $\mathtt{valid}_2$ *must hold for* `valid'` *to hold.*

Collision freeness *is the second requirement for composing implementations as well as one of the requirements of a good implementation. As a strategy to show that* $\mathcal{I}'$ *is collision free we show that additional requirements for* $\mathtt{valid}_1$ *and* $\mathtt{valid}_2$ *hold. Those requirements are concretized in Lemma 2.*

Lemma 2. *Let* $\mathcal{M}_1, \mathcal{M}_2$ *be symbolic models with implementations* $\mathcal{I}_1$ *and* $\mathcal{I}_2$*, respectively. If in addition to requirements (1.),(2.) [of the symbolic model composition] and i) [of the implementation composition] the following requirements for* $\mathtt{valid}'(\mathbb{T}) := \mathtt{valid}_1(\mathbb{T}) \wedge \mathtt{valid}_2(\mathbb{T})$ *hold:*

1) *Let* $\hat{\mathbb{T}}$ *be* $\mathbb{T}$ *with all silent* `generate` *"*`sgenerate` t*" replaced with normal* `generate` *queries "*`generate` t*". Then* $\mathtt{valid}'(\mathbb{T}) \Rightarrow \mathtt{valid}'(\hat{\mathbb{T}})$.
2) *Let* $x \in \{1, 2\}$*. If* $\mathtt{valid}_x$*("*`init` T, H*"), then for each* $t \in T \cup H$ *all function symbols in* t *are from* Σ_x *or no function symbol in* t *is from* Σ_x.
3) *Let* $x \in \{1, 2\}$*. Let* $\hat{\mathbb{T}}$ *be an expansion of* $\pi = q_1 + ... + q_n$ *in the following sense: A* q_i *= "*`generate` t*" for* $i \in \{1, ..., n\}$ *is replaced with* $q_i^1, ..., q_i^m$ *where* q_i^j *= "*`sgenerate` t_j*,* $t_j \in st(t)$ *and* t_j *does not contain function symbols from* Σ_x *for* $j \in \{1, ..., m\}$*. Then* $\mathtt{valid}'_x(\mathbb{T}) \Rightarrow \mathtt{valid}'_x(\hat{\mathbb{T}})$.

then $(\mathcal{M}_1, \mathcal{I}_1)$ *and* $(\mathcal{M}_2, \mathcal{I}_2)$ *are compatible.*
Review of Lemma 1 by Böhl, Cortier, and Warinschi [8].

Proof. One can prove Lemma 2 by a sequence of games. First one would show that there is no adversary A who can distinguish whether he plays $DS'_{\mathcal{M}' \cup \mathcal{M}^{tran}_{supp}([\![\mathcal{T}']\!]), \mathcal{I}' \cup \mathcal{I}^{tran}_{supp}([\![\mathcal{T}']\!]), A}(\eta)$ or the same game in which the `generate'` function is modified with

```
if $c \in dom(L) \cap [\![\mathcal{T}_1]\!]$ then
    exit game with return value 1 (collision)
```

instead of

```
if $c \in dom(L)$ then
    exit game with return value 1 (collision)
```

After showing that the games described above are indistinguishable, one shows that the latter game is indistinguishable from the original DS game.
Using Definition 4, the indistinguishability of the games described above suffices to proof that $\mathcal{I}'$ is collision free. (For the full proof see Böhl, Cortier, and Warinschi [8]) □

B.5 Deduction soundness

Finally, in this section we will recall the notion of deduction soundness [8]. We define deduction soundness with the help of a game, the DS game (see Listing 5).
To explain deduction soundness, we first will introduce the notion of a parameterized transparent symbolic model and its corresponding parameterized transparent implementation and then continue with explaining the idea behind the DS game.

Definition 8. *(Parameterization)*
A mapping $\mathcal{M}_{tran}(\nu)$ *from a bitstring* ν *to a transparent symbolic model* $\mathcal{M}_{tran}$ *is called a parameterized transparent symbolic model. Respectively, we define a parameterized transparent implementation* $\mathcal{I}_{tran}(\nu)$ *as mapping from the bitstring* ν *to a transparent implementation* $\mathcal{I}_{tran}$*. Further,* $\mathcal{I}_{tran}(\nu)$ *is an implementation of* $\mathcal{M}_{tran}$ *and the length(*ν*) ¡ length(p(*η*)) where p is a polynomial function.*
We also define ν *to be good, if* $\mathcal{I}_{tran}(\nu)$ *is a good implementation (see section B.2) of* $\mathcal{M}_{tran}(\nu)$.

The DS game is defined 1) *between a challenger and an adversary* A 2) *to show that* $\mathcal{I}$ *is a deduction sound implementation of* $\mathcal{M}$*, where* $\mathcal{M}(\nu) := \mathcal{M} \cup \mathcal{M}_{tran}$ *and* $\mathcal{I}(\nu) := \mathcal{I} \cup \mathcal{I}_{tran}$.

1) *The challenger keeps a set of requested terms* S*, a mapping* L *and a trace of queries* $\mathbb{T}$*, while the adversary provides a parameter* ν*. At first, with an "*`init` T, H*" query, bitstrings corresponding to a set of terms* T *and a set of hidden terms* H *are generated. Then*

the adversary has the possibility to parse bitstrings to terms and to generate bitstrings from terms, while the valid *predicate must always hold on $\mathbb{T}$ during the game or the adversary losses (return 0). On the other hand, the adversary wins the DS game, if he is able to create a bitstring which can be parsed, but is not deducible from S (which would be a non-Dolev-Yao term).*

2) *Deduction soundness of an implementation $\mathcal{I}$ with respect to a symbolic model $\mathcal{M}$ is illustrated in the following Definition.*

Listing 5: Deduction Soundness game

```
$DS_{\mathcal{M}(\nu),\mathcal{I}(\nu),A}(\eta)$:
    Let $S$ := $\varnothing$
    Let $L$ := $\varnothing$
    Lat $\mathbb{T}$ := $\varnothing$
    $\mathcal{R} \leftarrow \{0,1\}^*$

    Receive parameter $\nu$ from $A$

    on request "init $T, H$" do
        add "init $T$" to $\mathbb{T}$
        if valid($\mathbb{T}$) then
            let $S$ := $S \cup T$
            let $C$ := $\varnothing$
            for each $t \in T$ do
                let $(c, L)$ := generate$(t, L)$
                let $C$ := $C \cup \{c\}$
            for each $t \in H$ do
                let $(c, L)$ := generate$(t, L)$
            send $C$ to $A$
        else
            return 0 ($A$ is invalid)

    on request "sgenerate $t$" do
        if valid($\mathbb{T}$ + "sgenerate $t$") then
            let $(c, L)$ := generate$(t, L)$

    on request "generate $t$" do
        add "generate $t$" to $\mathbb{T}$
        if valid($\mathbb{T}$) then
            let $S$ := $S \cup \{t\}$
            let $(c, L)$ := generate$(t, L)$
            send $c$ to $A$
        else
            return 0 ($A$ is invalid)

    on request "parse $c$" do
        let $(t, L)$ := parse$(c, L)$
        if $S \vdash_D t$ then
            send $t$ to $A$
        else
            return 1 ($A$ produced a non-DY term)
```

Definition 9. *(Deduction Soundness)*
Let $\mathcal{M}$ be a symbolic model and $\mathcal{I}$ be an implementation of $\mathcal{M}$. We call $\mathcal{I}$ a deduction sound implementation of $\mathcal{M}$, if for all parameterized transparent symbolic models $\mathcal{M}_{tran}(\nu)$ and for all parameterized transparent implementations $\mathcal{I}_{tran}(\nu)$ of $\mathcal{M}_{tran}$ that are composable with $\mathcal{M}$ and $\mathcal{I}$ (see requirements from section B.4) we have

$$Pr[DS_{\mathcal{M}\cup\mathcal{M}_{tran}(\nu),\mathcal{I}\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]$$

to be negligible for all probabilistic polynomial time adversaries A sending only good parameters ν where DS is the deduction soundness game from Listing 5. Note that $\mathcal{M} \cup \mathcal{M}_{tran}(\nu)$ can be generically composed to a parameterized symbolic model $\mathcal{M}'(\nu)$ and parameterized implementation $\mathcal{I}'(\nu)$, respectively.
(Review of Definition 5 Böhl, Cortier, and Warinschi [8])

Intuitively, we call an implementation $\mathcal{I}$ deduction sound with respect to its corresponding symbolic model $\mathcal{M}$, if the deduction relation $\vdash_D$ reflects all capabilities of the computational adversary. Concretely, this means that no adversary can produce, given a set of terms S, a bitstrings c corresponding to a term t with $S \nvdash t$.

Additionally, since the DS game does not prevent collisions, we naturally want to show that no adversary can produce those collisions with an overwhelming probability. Since only generate *function calls can produce collisions by producing a bitstring which is already in L (the structure of the function on the other hand prevents collisions), it suffices to show that the following lemma 3 holds, which is already proven by Böhl, Cortier, and Warinschi [8].*

Lemma 3. *Let $DS'_{\mathcal{M}(\nu),\mathcal{I}(\nu),A}(\eta)$ be the deduction soundness game where we replace the* `generate` *function by the collision aware generate function (see Figure 3 [8]). Then no p.p.t. adversary A can distinguish $DS_{\mathcal{M}(\nu),\mathcal{I}(\nu),A}(\eta)$ from $DS'_{\mathcal{M}(\nu),\mathcal{I}(\nu),A}(\eta)$ with non-negligible probability. (Note that the transparent functions are already included $\mathcal{M}(\nu)$ and $\mathcal{I}(\nu)$ in here.)*
(Review of Lemma 2 Böhl, Cortier, and Warinschi [8])

Appendix C
Deduction soundness of AEAD schemes

The advantage of deduction soundness is that it is relatively easy to extend. Böhl, Cortier, and Warinschi [8] already added public datastructures, public key encryption, signatures, secret key encryption, MACs and hashes to their framework. Our contribution to this will be, to extend the framework with *authenticated encryption schemes with associated data*.
We will define a symbolic model $\mathcal{M}_{AEAD}$ and a corresponding implementation $\mathcal{I}_{AEAD}$. Then we will show that for any symbolic model $\mathcal{M}$ which is composable with $\mathcal{M}_{AEAD}$ (see Section B.4) and implementation $\mathcal{I}$ where $\mathcal{I}$ is a deduction sound implementation of $\mathcal{M}$, it holds that the composition $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a deduction sound implementation of $\mathcal{M} \cup \mathcal{M}_{AEAD}$ if $\mathcal{I}$ is composable with $\mathcal{I}_{AEAD}$.
To achieve this, we will use the notion of DAE-N security (Definition 1) and rewrite the definition in a game-like way to fit into the syntax of our computational model.

Computational preliminaries

Listing 6: DAE-N game

```
DAE-N$_A^{(Gen, Enc, Dec)}(\eta)$:
    b $\xleftarrow{\$}$ $\{0,1\}$
    oracles := $\emptyset$

    on request "new oracle" do
        let r $\xleftarrow{\$}$ $\{0,1\}^\eta$
        let k := Gen(1$^\eta$, r)
        oracles.add(k)
        let ciphers$_k$ := $\emptyset$
        send k to $A$

    on request "$O_k^{Enc}(n,H,m)$" do
        if k $\notin$ oracles then
            send $\perp$ to $A$
        else
            if b == 0 then
                let c':= Enc$_k$(n, H, m)
                let c $\xleftarrow{\$}$ $\{0,1\}^{|c'|}$
                ciphers$_k$.add((c, m))
                send c to $A$
            else
                send Enc$_k$(n, H, m) to $A$

    on request "$O_k^{Dec}(n,H,c)$" do
        if k $\notin$ oracles then
            send $\perp$ to $A$
        else
            if b == 0 then
                if (c, m) $\in$ ciphers$_k$
                           for some m
                then
                    send m to $A$
                else
                    send $\perp$ to $A$
            else
                send Dec$_k$(n, H, m) to $A$

    on request "guess b'" do
        if b == b' then
            return 1
        else
            return 0
```

Intuitively, the adversary A which now plays DAE-N game still tries to distinguish whether he interacts with real oracles or with some fake oracles. Concretely, a bit b is chosen at random in the beginning of the game, which decides whether the adversary gets a response from a real oracle (if b = 1) or from a fake oracle (if b = 0). If the adversary is able to send a request "guess b'" (and $b' == b$) with a probability significantly higher than $\frac{1}{2}$, he can break DAE-N security of the encryption scheme. Note that we additionally added a random input parameter r to the key generation algorithm to clarify that all oracles use a different source of randomness.

APPENDIX D
SYMBOLIC MODEL

At first we define the symbolic model $\mathcal{M}_{AEAD} = (\mathcal{T}_{AEAD}, \leqslant_{AEAD}, \Sigma_{AEAD}, \mathcal{D}_{AEAD})$:

Signature Σ_{AEAD}:

$$\begin{aligned} k_x &: \tau^{k_x}_{AEAD} \\ con_S &: \top \to \tau^{n}_{AEAD} \\ E_x &: \tau^{k_x}_{AEAD} \times \tau^{n}_{AEAD} \times \top \times \top \to \tau^{cipher}_{AEAD} \end{aligned}$$

are the featured function symbols, with x $\in \{h, c\}$ and S being a set of possible nonces.
The randomized function k_h returns honest keys while k_c returns corrupted keys.
The deterministic function con_S maps an arbitrary input value to a nonce from the set S.
The deterministic function E_h returns an honest cyphertext using an honest key, a nonce, and two additional arbitrary values as input.
The only difference of E_c to E_h is that E_c uses some corrupted key as input and returns a corrupted cyphertext.

Set of types $\mathcal{T}_{AEAD}$

$$\mathcal{T}_{AEAD} = \{\top, \tau^{k_x}_{AEAD}, \tau^{n}_{AEAD}, \tau^{cipher}_{AEAD}\}$$

Sub type relation $\leqslant_{AEAD}$

All types introduced above are direct sub types of the base type $\top$.

Deduction System $\mathcal{D}_{AEAD}$:

$$\frac{k^l_x()\ con_S(n)\ H\ m}{E_x(\ k^l_x(), con_S(n), H, m)} \qquad \frac{E_x(\ k^l_x(), con_S(n), H, m)}{con_S(n)}$$

$$\frac{E_x(\ k^l_x(), con_S(n), H, m)}{H} \qquad \frac{E_c(\ k^l_c(), con_S(n), H, m)}{m}$$

APPENDIX E
IMPLEMENTATION

We now define a concrete implementation $\mathcal{I}_{AEAD} = (M_{AEAD}, [\![\cdot]\!]_{AEAD}, len_{AEAD}, open_{AEAD}, valid_{AEAD})$ for *authenticated encryption schemes with associated data*. The implementation uses some DAE-N secure authenticated secret key encryption scheme $\Pi_{DAEN} = (DAEN.Gen, DAEN.Enc, DAEN.Dec)$. Π_{DAEN} additionally is collision free by construction since a collision would break the authenticity of the encryption scheme. We fix a set of bitstrings $S' \subset \{0,1\}^*$, which we will later require to correspond to a specified set of terms used to derive IVs. We also fix an arbitrary injective and efficiently computable function $\iota : S' \to [\![\tau^{N}_{AEAD}]\!]$. We can now define the model and implementation as follows:

Turing machine M_{AEAD}

Now we will give the computational interpretations of the function symbols defined in D.

- $(M_{AEAD}\ k_x)(r)$: Let $k := DAEN.Gen(1^\eta, r)$.
  Return $\langle k, \tau^{k_x}_{AEAD}\rangle$
- $(M_{AEAD}\ con)(n)$: Return $\langle \iota(n), \tau^{n}_{AEAD}\rangle$
- $(M_{AEAD} E_x)(\hat{k}, \hat{n}', H, m)$: Parse $\hat{k}$ as $\langle k_x \tau^{k_x}_{AEAD}\rangle$.
  Parse $\hat{n}$ as $\langle n', \tau^{n}_{AEAD}\rangle$.
  Let $c := DAEN.Enc_k(n', H, m)$
  Let $\hat{H} := \langle H, \top\rangle$ Return $\langle c, \hat{n}', \hat{H}, k, \tau^{cipher}_{AEAD}\rangle$

are the computational interpretations of k_x, con and E_x, respectively.

The mapping function $[\![\cdot]\!]_{AEAD}$

The mapping function $[\![\cdot]\!]_{AEAD} : \mathcal{T} \to 2^{\{0,1\}^*}$ is a function which maps each type $\tau \in \mathcal{T}$ to a set of bitstrings. $[\![\cdot]\!]_{AEAD}$ should fulfill all conditions stated in B.2. Namely, we demand from our concrete implementation that $[\![\top]\!] = \{0,1\}^*$ and for each $x \in \{\tau^{k_x}_{AEAD}, \tau^{n}_{AEAD}, \tau^{cipher}_{AEAD}\}$ that $[\![x]\!] \subset \{0,1\}^\eta$. Further, we demand that for all $x, y \in \{\tau^{k_x}_{AEAD}, \tau^{n}_{AEAD}, \tau^{cipher}_{AEAD}\}$ $[\![x]\!] \cap [\![y]\!] = \varnothing$.

The length function len_{AEAD}

The function len_{AEAD} : **Terms** $\to \mathbb{N}$ computes the length of a term if interpreted as a bitstring, in other words $len_{AEAD}(f^l(t_1, t_2, ..., t_n)) := |(M_{AEAD} f)(c_1, c_2, ..., c_n; r)|$. The length regularity of len_{AEAD} (and therefore of our implementation) follows directly from the construction of AEAD schemes.

The interpretation function open$_{AEAD}$

The open function for AEAD is defined as follows:

Listing 7: open$_{AEAD}$

```
open_AEAD(c, L)
    if c ∈ $[\![\mathcal{T}_{AEAD}]\!]_{AEAD} \cap dom$(L) then
        return (c, L(c))
    else if c = $\langle k, \tau^{k_x}_{AEAD}\rangle$ then
        return (c, $g^{l(c)}_{AEAD^{k_x}}$)
    else if c = $\langle n', \tau^{n}_{AEAD}\rangle$ then
        return $(c, \tau^{n}_{AEAD})$
    else if c = $\langle c', \tau^{cipher}_{AEAD}\rangle$ then
        extract $\hat{H}$ = (H, ⊤) from c'
        extract $\hat{n}$ = $(n', \tau^{n}_{AEAD})$ from c'
        for each $(\hat{k}, k^h_x())$ ∈ L do
            parse $\hat{k}$ as $\langle k, \tau^{k_x}_{AEAD}\rangle$
            let m := DAEN.Dec(k, $n'$, H, c')
            if m ≠ ⊥ then
                return (c, $E^{l(c)}_x(\hat{k}, \hat{n}, \hat{H}$, m))
        return $(c, g^{l(c)}_{AEAD^{cipher}})$
    else
        return $(c, g^{l(c)}_{\top})$
```

open$_{AEAD}$ returns $(c, g^{l(c)}_{\top})$ for any foreign bitstring and its behavior is in all cases independent on any foreign bitstring in the library. Since open$_{AEAD}$ therefore fulfills the conditions of Definition 5, our implementation is *type safe*.

The `valid`$_{AEAD}$ predicate

The `valid`$_{AEAD}$ predicate is dependent on a set of terms S that specifies which terms can be turned into IVs by ι. As the IV space is typically finite (e.g. for GCM mode), and ι is injective, S needs to be restricted, too. Our result is parametric in S, S' and ι. We may define S as a subset of the set of terms that is defined by composition, e.g., to derive S from a transparent model. We therefore fix some model $\mathcal{M} = (\mathcal{T}, \leqslant, \Sigma, D)$ and its deduction sound implementation $\mathcal{I}$ to compose with, such that $\mathcal{M}$ and $\mathcal{I}$ are composable with $\mathcal{M}_{AEAD}$ and $\mathcal{I}_{AEAD}$ regarding the requirements from Section B.4. We then choose $S \subset \text{Terms}(\Sigma \cup \Sigma_{AEAD}, \mathcal{T} \cup \mathcal{T}_{AEAD}, \leqslant \cup \leqslant_{AEAD})$ such that any bitstring representation for any term $t \in S$ is in S'.

Formally, for any A and any a parameterized transparent symbolic model $\mathcal{M}_{tran}(\nu)$ with a corresponding parameterized implementation $\mathcal{I}_{tran}(\nu)$ such that $\mathcal{M}_{tran}(\nu)$ and $\mathcal{I}_{tran}(\nu)$ are composable with $\mathcal{M} \cup \mathcal{M}_{AEAD}$ and $\mathcal{I} \cup \mathcal{I}_{AEAD}$ regarding the requirements from Section B.4 for ν being send by the adversary A, we require that for the library L at any moment in any instance of the deduction soundness game

$$\text{DS}_{(\mathcal{M} \cup \mathcal{M}_{AEAD}) \cup \mathcal{M}_{tran}(\nu), (\mathcal{I} \cup \mathcal{I}_{AEAD}) \cup \mathcal{I}_{tran}(\nu)}(\eta)$$

it holds that $\forall s'. L[\![s']\!] \in S \iff s' \in S'$.

For an appropriately chosen S, we can now define `valid`$_{AEAD}$ as follows:

1) We demand that the trace $\mathbb{T}$ starts with exactly one `init` query "`init` T, H" where at least one of them could be an empty list.
2) The adversary is not allowed to use E_x in the the `init` query.
3) i) For the query "`init` T, H" it should hold that:
    * the function symbol k_c should only occur in a term $k^l_c() \in T$.
    * the function symbol k_h should only occur in a term $k^l_h() \in H$.

ii) For each label l of k_x^l, l should be unique in $T \cup H$.
iii) Whenever $k_x^l()$ occurs in a generate query, $k_x^l()$ must have occurred in the init query before.
iv) Except generation, $k_x^l()$ should only occur in E_x as its first argument.
This rules guarantee that all keys are generated in the init query.

4) No tuple of $(cons_S(n), H, m)$ occurs twice in some trace $\mathbb{T}$. In other words, we demand that for every term $E_x(k_x^l(), cons_S(n), H, m)$ $(cons_S(n), H, m)$ is different in each "init T, H", "generate t" or "sgenerate t" queries. (For all terms $E_x(k_x^l(), cons_S(n), H, m)$, $E_x(k_x^l(), cons_S(n'), H', m') \in \mathbb{T}$ it has to hold that

$$(cons_S(n), H, m) \neq (cons_S(n'), H', m').$$

5) For each term $cons_S(n)$, $n \in S$.

These rules guarantee that all keys which may be used by the adversary are generated in the init query. They also guarantee that the adversary has to decide which keys are corrupted and which keys are honest during initialization, because we only allow static corruption of keys. Furthermore, to prevent key cycles, keys are only allowed to be used for encryption and decryption. At last, the rules guarantee the freshness of the used nonces.

For all parse and generate requests of the adversary on the trace $\mathbb{T}$ all valid_{AEAD} conditions must be fulfilled.

Appendix F
AEAD composability

In this section we will finally prove AEAD schemes to be composable with the deduction soundness framework. Like Böhl, Cortier, and Warinschi [8], we will proof the composability of AEAD schemes similar to the already existing proofs of e.g. *public key encryption schemes*, by using different games, which we show to be indistinguishable from each other. In other words we will show an implementation of an AEAD scheme to be deduction sound in respect to its symbolic model. This is concretized in the following theorem.

Theorem 1. *Let $\mathcal{M}$ be a symbolic model and $\mathcal{I}$ its corresponding, deduction sound implementation. If $\mathcal{M}$ and $\mathcal{I}$ are composable with $\mathcal{M}_{AEAD}$ and $\mathcal{I}_{AEAD}$ regarding the requirements from section B.4 and if the encryption scheme is DAE-N secure, then $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a deduction sound implementation of $\mathcal{M} \cup \mathcal{M}_{AEAD}$.*

Proof. Let A be a p.p.t. adversary and let $\mathcal{M}_{tran}(\nu)$ be a parameterized transparent symbolic model with a corresponding parameterized implementation $\mathcal{I}_{tran}(\nu)$. Furthermore, let $\mathcal{M}_{tran}(\nu)$ and $\mathcal{I}_{tran}(\nu)$ be composable with $\mathcal{M} \cup \mathcal{M}_{AEAD}$ and $\mathcal{I} \cup \mathcal{I}_{AEAD}$ regarding the requirements from B.4 for ν being send by the adversary A.
We have to show that the success probability of A winning the deduction soundness game
$\text{DS}_{(\mathcal{M} \cup \mathcal{M}_{AEAD}) \cup \mathcal{M}_{tran}(\nu), (\mathcal{I} \cup \mathcal{I}_{AEAD}) \cup \mathcal{I}_{tran}(\nu)}(\eta)$ is negligible.
Intuition:
A wins the deduction soundness game if A can provide a bitstring which corresponds to a term t for which it holds that A cannot deduce t from the previously generated terms, i.e. A does not "know" t symbolically. Only one opportunity to achieve this is added by adding AEAD.
Proof strategy intuition:
First we show that an adversary cannot find any collision to break the game. In the next step we replace honest encryptions by random bitstrings. Using the fact that the scheme is DAE-N secure, an adversary cannot learn anything about the original message except its length. In the third step we add another deduction rule to give the adversary the opportunity to create honest encryption of chosen messages. If the adversary would note a difference between this two steps it would break the authentication of cyphertexts, but this could only happen with negligible probability. In the last step we show that the encryption can be simulated with parameterized transparent functions. So if an adversary would come up with a non-DY term this would lead to a non-DY request in the DS game of $\mathcal{M}$ and $\mathcal{I}$. But this would contradict our assumption of $\mathcal{I}$ being a deduction sound implementation of $\mathcal{M}$.

F.1 *Game 0*

In this game, adversary A plays the original deduction soundness game from B.5.

$$\text{DS}_{(\mathcal{M} \cup \mathcal{M}_{AEAD}) \cup \mathcal{M}_{tran}(\nu), (\mathcal{I} \cup \mathcal{I}_{AEAD}) \cup \mathcal{I}_{tran}(\nu)}(\eta)$$

F.2 *Game 1*

In this game we replace the generate function in the deduction soundness game by the collision aware generate function 8. If A could distinguish between Game 0 and Game 1, this could be used to break the encryption scheme.

Listing 8: generate'

```
generate'$_{M,\mathcal{R}}(t, L)$:
    if for some $c \in dom(L)$ we have $L[\![c]\!] == t$ then
        return $(c, L)$
```

```
else
    for $i \in \{1, ..., n\}$ do
        let $(c_i, L) := \texttt{generate}_{M,\mathcal{R}}(t_i, L)$
    let $r := \mathcal{R}(t)$
    let $c := (Mf)(c_1, ..., c_n; r)$
    let $L(c) := f^l(a_1, ..., a_n))(l \in \texttt{labesH})$
    return $(c, L)$
```

F.3 Claim: Game 0 and Game 1 are indistinguishable.

Since $\mathcal{I} \cup \mathcal{I}_{tran}(\nu)$ is a collision-free implementation the only possibility for the adversary to find collisions is in $\mathcal{I}_{AEAD}$. This on the other hand would only happen if the probability of $DAEN.Enc(k, n, H, m) = DAEN.Enc(k, n', H', m')$ with $(n, H, m) \neq (n', H', m')$ is non-negligible. This would contradict the DAE-N security of AEAD, more concretely, this would break the authentication of the scheme. Knowing that $(\mathcal{I} \cup \mathcal{I}_{AEAD}) \cup \mathcal{I}_{tran}(\nu)$ is a collision-free implementation Game 0 and Game 1 are indistinguishable by 3.

F.4 *Game 2*

In this game we make a few changes with respect to Game 1 concerning the possibilities of the adversary. We take his option of learning something from cyphertexts or about honest keys, by replacing the cyphertexts created under honest keys by random bitstrings. We abstract from just using random bitstrings by using encryptions of random bitstrings under the same keys (because the adversary still needs the possibility to retrieve the nonce and the header) and also replace all honest keys in the library by random bitstrings with the same length as the keys. Because of the properties of the DAE-N secure scheme, A should not notice any difference between Game 1 and Game 2.
Concretely, we make the following changes:

1) Replace cyphertexts created under honest keys by encryptions of random bitstrings
   In the `generate` (and `generate'`) function we insert

   ```
   if $t == E_h(k_h^l(), cons_S(n), H, m)$ then
       let $r' := \mathcal{R}(c_4)$
       let $c := (M\ E_h)(c_1, c_2, c_3, r')$
   else
       let $c := (M\ f)(c_1, ..., c_n; r)$
   end if
   ```

   instead of the original line

   ```
   let $c := (M\ f)(c_1, ..., c_n; r)$
   ```

   c_1 resembles the bitstring of the key, c_2 resembles the bitstring of the nonce, c_3 resembles the bitstring of the header and c_4 resembles the bitstring of the message. With $\mathcal{R}(c_4)$ we create a random value with the same length as c_4. If the term t represents an encryption under an honest key the corresponding bitstring is generated as usual (by generating all the subterms first) replace c_4 with a random bitstring of length $|c_4|$, otherwise the generation of bitstrings works like in Game 1.
2) Replace honest keys in the library by random bitstrings
   If a honest key should be generated, instead of calling (M$_{AEAD}$ k_h)(r), we pick a random $r \xleftarrow{\$} \{0,1\}^\eta$ of length η. Then we normally generate the key $k \xleftarrow{\$}$ DAEN.Gen($1^\eta, r$). Then we generate a second random value $r' \xleftarrow{\$} \{0,1\}^{|k|}$ with the same length as the key. We then use $\hat{k} := \langle r', \tau_{AEAD}^{k_x} \rangle$ as the corresponding bitstring for $k_h()$ and remember k as the true key. The parse function is the rewritten such that the remembered key for is used instead of the random value in the library.

F.5 Claim: Game 1 and Game 2 are indistinguishable.

Intuition:

To show that Game 1 and Game 2 are indistinguishable, we use reduction, i.e. we show that if an adversary A can distinguish Game 1 from Game 2, we could use A to construct an adversary B which could win the oracle-based DAE-N game with non-negligible probability.

So we first show how we would construct B from A, i.e we show how B generates and parses bitstrings. Then we analyze our construction and show indistinguishability between the two games:

1) **How to generate bitstrings.**
   First, for each key generation request for honest keys (for all "generate k_h^l" queries), B does not call the `generate` function, but instead request a new oracle from the DAE-N game and hence receives a new encryption and a corresponding decryption oracle O_k^{Enc} and O_k^{Dec} (which encrypt and decrypt the given message if $b == 1$ or encrypt an random bitstring and return always $\perp$ for unknown queries otherwise). Furthermore, B picks a random bitstring

k' in a way such that $\hat{k} := \langle k', \tau^{k_h}_{AEAD} \rangle \in [\![\tau^{k_h}_{AEAD}]\!]_{AEAD}$ and then adds $(\hat{k}, k^l_h())$ to L (A is not able to learn the value of k what is required in the implementation E).
All other "generate t" queries are handled as before (by calling the generate function).
To use the encryption oracle O^{Enc}_k to get honest encryptions, we need to change the generate function in the following way:
In the generate function

```
if $t == E_h(k^l_h(), cons_S(n), H, m)$ then
    let $c := \langle O^{Enc}_{c_1}(c2, c3, c4), c'_1, \tau^{cipher}_{AEAD} \rangle$
else
    let $c := (M\ f)(c_1, ..., c_n; r)$
end if
```

is inserted instead of the original line

```
let $c := (M\ f)(c_1, ..., c_n; r)$
```

c_1 resembles the bitstring of the key used by the oracle, c_2 resembles the bitstring of the nonce, c_3 resembles the bitstring of the header and c_4 resembles the bitstring of the message. c'_1 resembles the bitstring of the the random value which was picked by B instead of the original key (which is unknown to B)
This change produces encryptions like in Game 1 if $b == 1$ (if the oracle produces encryption of the message) and produces encryptions like in Game 2 otherwise (The oracle produces encryptions of a random string).

2) **How to parse bitstrings.**
For B to be able to deal with adversarial encryption under honest keys, the open$_{AEAD}$ function has to be modified. Concretely, the open$_{AEAD}$ function

```
if $(\langle k, \tau^{k_h}_{AEAD} \rangle, k^l_h()) \in L$ then
    let $m := O^{Dec}_k(cons_S(n), H, c')$
else
    parse $\hat{k}$ as $\langle k, \tau^{k_x}_{AEAD} \rangle$
    let $m := DAEN.Dec(k, cons_S(n), H, c')$
end if
```

is inserted instead of the original line

```
parse $\hat{k}$ as $\langle k, \tau^{k_x}_{AEAD} \rangle$
let $m := DAEN.Dec(k, cons_S(n), H, c')$
```

So B simply uses the decryption oracle in the simulation instead of the key to decrypt (For DAEN secure schemes this should be same concerning decryption).
Further observation:

- Using the open$_{AEAD}$ function, any adversary A that learns the bitstring representation of an honest key (with non-negligible probability) is able to win the DAE-N game with non-negligible probability:
If open$_{AEAD}$ is called with input $c \in [\![\tau^{k_h}_{AEAD}]\!]_{AEAD}$ ($L[\![c]\!] = \langle k, \tau^{k_h}_{AEAD} \rangle$), B will parse c as $\langle k, \tau^{k_h}_{AEAD} \rangle$, pick a random message $m \xleftarrow{\$} \{0,1\}^\eta$ and compute $x := DAEN.Dec(k, cons_S(n), H, O^{enc}_k(cons_S(n), H, m))$. If x is the message ($x == m$), B knows that real encryption/decryption is used and sends "guess 1" to the DAE-N game. This leads B to win the DAE-N game with overwhelming probability.
- At this point, we require 4) of the valid$_{AEAD}$ predicate to hold. If this would not be the case, an adversary A could trivially win the oracle based DAE-N game by querying the same pair of IV, header and message twice. Since encryption is deterministic, the same oracle would produce the same cyphertext in both queries if its a real oracle and different cyphertexts in the other case.

3) **Analysis and conclusion**
We can split the analysis in two cases: The oracle produces real encryptions i) or the oracle encrypts random bitstrings ii):

i) The oracle produces real encryptions:
$\Rightarrow$ B simulates Game 1 for A despite having random values instead of the honest keys in the library L.
A can only detect a difference if

   * A guesses one of the random bitstrings, or
   * parses a bitstring belonging to a key of Game 1.

The first case can only happen with negligible probability depending on η, while the second case would lead to B winning the DAE-N game with overwhelming probability as described in the *parsing* paragraph above.
$\Rightarrow$ If the oracle produces real encryptions, A cannot distinguish the simulation.

ii) The oracle produces "fake" encryptions:
B simulates Game 2 for A perfectly, meaning every correct guess of A in distinguishing the two games leads to B winning the DAE-N game.

So we see that every correct guess of A in distinguishing Game 1 from Game 2 leads to B guessing correct in the DAE-N game. But because of the scheme being DAE-N secure this can only happen with negligible probability
$\Rightarrow A$ cannot distinguish Game 1 from Game 2.

F.6 *Game 3*

In this game we also add another deduction rule to our deduction system to allow the adversary to get encryptions of chosen messages under honest keys. We add

$$\frac{con_S(n)\ H\ m}{E_h(k_h^l(), con_S(n), h, m)}$$

to the deduction system. Additionally, we define $\vdash_2$ as the deduction relation of Game 2, where the relation is based on the deduction system $D_2 = D \cup D_{AEAD} \cup D_{tran}$. Further we define $\vdash_3$ as the deduction relation of Game 3, where the relation is based on deduction system D_3 (Which is the same as D_2, despite the rule that was added above).
The idea behind this game is to use the authenticity of the encryption scheme. We show that an adversary which can distinguish between Game 2 and Game 3 could be used to break authentication (A could forge a cyphertext).

F.7 Claim: Game 2 and Game 3 are indistinguishable

Intuition:
The idea behind this proof is to show that an adversary A that can craft a bitstring c, which can be parsed to a term t for which it holds that $S \nvdash_2 t$ (t is a non-DY term in Game 2) and $S \vdash_3 t$ (t is a DY term in Game 3), can be used to construct a second adversary B which plays the DAE-N game. We first (1) show how to construct B and how B simulates A and then (2) show how B can win the DAE-N game using A

(1) Construction of B:
B simulates Game 2 for A in a completely analogous way as in the proof above. Concretely, the way the `init` requests are handled is perfectly analogous as well as the way the `generate` and $\texttt{open}_{AEAD}$ functions are altered. The simulation from the proof above perfectly simulates Game 2, which means that adversary A should not notice any difference.

(2) In the case in which A sends a "`parse` c" query to B such that c can be parsed to a term t ($t := L[\![c]\!]$) and t is non-DY in Game 2 ($S \nvdash_2 t$) and DY in Game 3 ($S \vdash_3 t$) , we claim that t must contain a forged encryption under an honest key, i.e., a term $t_f = E_h(k_h^l(), con_S(n), H, m)$ such that $t \in st(t)$ but $t \notin st(S)$.
For contradiction, we assume t does not contain a forgery under an honest key and further let π be a proof for $S \vdash_3 t$. By this assumption we can, first, remove all instantiations of the deduction rules

$$\frac{k_h^l()\ con_S(n)\ H\ m}{E_h(k_h^l(), con_S(n), h, m)},$$

from our proof π, as A should not know any honest key due to the static corruption requirement. Second, all instances of the deduction rule

$$\frac{con_S(n)\ H\ m}{E_h(k_h^l(), con_S(n), h, m)}$$

can be removed, as they directly produce a forgery in the above sense. These removals would give us a new proof π' which should be the same by our assumption (π' is a proof for $S \vdash_3 t$). But π' is also a proof for $S \vdash_2 t$, because the only differences in the two deduction relations are the instantiations of the deduction rule

$$\frac{con_S(n)\ H\ m}{E_h(k_h^l(), con_S(n), h, m)}$$

which are not part of π'. This on the other hand contradicts our initial assumption.
Hence,t must contain a forged encryption
Since c can be parsed, the library L contains t and all of the corresponding bitstrings to its subterms. If at some point in time A generates an honest encryption, A must have guessed the random bitstring k' (corresponding to the key k) which was never used to generate/compute a bitstring sent to A before. Hence, this can only happen with negligible probability.
This means A can create a forgery of encryption without knowing the bitstring of the key with overwhelming probability. B then could use this forgery and send it to its decryption oracle. If the oracle returns $\perp$ it means that the oracle must be fake (because a valid encryption should be decrypted to a message $\neq \perp$) and B then sends a "guess 0" query to the DAE-N game instance and "guess 1" otherwise. In this case B wins the DAE-N game with overwhelming probability.

Due to our scheme being DAE-N secure, the probability of crafting a valid forgery should be negligible. Because the only possibility to distinguish Game 2 and Game 3 is to craft a bitstring c which then is parsed to a term t with $S \nvdash_2 t$ and $S \vdash_3 t$ (Because the only difference in Game 2 and Game 3 is the deduction system/relation), it also holds that both games are indistinguishable.

F.8 *Game 4*

Intuition:

In this game a new adversary B plays the deduction soundness game for $\mathcal{M}$ and $\mathcal{I}$ and simulates Game 3 for A by adding transparent functions to simulate authenticated encryption with associated data.

We first define a transparent symbolic model for authenticated encryption with associated data. Second we show how to convert terms used in $\mathcal{M}_{AEAD}$ and map them to their corresponding terms in the transparent symbolic model.

Transparent symbolic model for authenticated encryption with associated data

For $\mathcal{M}^{tran}_{AEAD}(\nu)$ we will use the same typeset and the same subtype relation as for $\mathcal{M}_{AEAD}$. The deduction system of $\mathcal{M}^{tran}_{AEAD}(\nu)$ is the same as in B.3. We further describe the signature Σ^{tran}_{AEAD} in the following:

- deterministic $f_{k^l_x()}$
  ar($f_{k^l_x()}$) = $\tau^{k_x}_{AEAD}$ for all labels $l \in \nu$.
- deterministic $f_{cons_S(n)}$
  ar($f_{cons_S(n)}$) = τ^{n}_{AEAD} for all $n \in S$.
- deterministic $f_{cons_S(\cdot)}$
  ar($f_{cons_S(\cdot)}$) = $\top \rightarrow \tau^{n}_{AEAD}$.
- deterministic $f_{E_h(k^l_h(),cons_S(n),H,r)}$
  ar($f_{E_h(k^l_h(),cons_S(n),H,r)}$) = τ^{cipher}_{AEAD} for all labels $l \in \nu$, $n \in S$ and r representing a random bitstring.
- deterministic $f_{E_h(k^l_h(),cons_S(n),\cdot,\cdot)}$
  ar($f_{E_h(k^l_h(),cons_S(n),\cdot,\cdot)}$) = $\top \times \top \rightarrow \tau^{cipher}_{AEAD}$ for all labels $l \in \nu$ and $n \in S$.
- deterministic $f_{E_c(k^l_c(),cons_S(n),\cdot,\cdot)}$
  ar($f_{E_c(k^l_c(),cons_S(n),\cdot,\cdot)}$) = $\top \times \top \rightarrow \tau^{cipher}_{AEAD}$ for all labels $l \in \nu$ and $n \in S$.

where ν is an encoded triple (l, k, k') with $l \in$ labels, k being a symmetric key and k' being a value representing honest keys in the library.

Now we specify the corresponding transparent implementation $\mathcal{I}^{tran}_{AEAD}(\nu)$. We use the bitstring mapping and the length function of $\mathcal{I}_{AEAD}$. We also take the open function and the valid predicate from B.3. We now define the Turing machine M^{tran}_{AEAD} for $\mathcal{I}^{tran}_{AEAD}(\nu)$:

- (M^{tran}_{AEAD} $f_{k^l_x()}$)(r) returns $\langle k', \tau^{k_x}_{AEAD}\rangle$
- (M^{tran}_{AEAD} $f_{cons_S(n)}$)(r) returns $\langle n', \tau^{n}_{AEAD}\rangle$
- (M^{tran}_{AEAD} $f_{cons_S(\cdot)}$)(r) returns (M_{AEAD} $cons_S$)($\langle n, \top\rangle$)
- (M^{tran}_{AEAD} $f_{E_h(k^l_h(),cons_S(n),H,r)}$)(r) returns ($\mathrm{M}_{AEAD}$ E_h)($\langle k, \tau^{k_x}_{AEAD}\rangle$, $\langle n', \tau^{n}_{AEAD}\rangle$, H, r) where r has the length of the bitstring corresponding to m.
- (M^{tran}_{AEAD} $f_{E_h(k^l_h(),cons_S(n),\cdot,\cdot)}$)(H, m, r) returns ($\mathrm{M}_{AEAD}$ E_h)($\langle k, \tau^{k_x}_{AEAD}\rangle$, $\langle n', \tau^{n}_{AEAD}\rangle$, H, m)
- (M^{tran}_{AEAD} $f_{E_c(k^l_c(),cons_S(n),\cdot,\cdot)}$)(H, m, r) returns ($\mathrm{M}_{AEAD}$ E_c)($\langle k, \tau^{k_x}_{AEAD}\rangle$, $\langle n', \tau^{n}_{AEAD}\rangle$, H, m)

Now we define the mode of operation $func$ as follows:

```
($\mathrm{M}^{tran}_{AEAD}$ $func$)(b):
    if b == $\langle k', \tau^{k_x}_{AEAD}\rangle$ for some $(l,k,k')$ $\in \nu$ then
        return $f_{k^l_x()}$
    if b == $\langle n', \tau^{n}_{AEAD}\rangle$ then
        return $f_{cons_S(\cdot)}$
    if b $\in \tau^{cipher}_{AEAD}$ then
        parse b as $\langle c, \tau^{cipher}_{AEAD}\rangle$
        get nonce n' and header h' from c
        for each $(l,k,k')$ $\in \nu$ do
            let m := AEAD.DEC(k, n', H', c)
            if m $\neq \perp$ then
                if $l$ belongs to an honest key then
                    return $f_{E_h(k^l_h(),cons_S(n),\cdot,\cdot)}$
                else
                    return $f_{E_c(k^l_c(),cons_S(n),\cdot,\cdot)}$
    return $\perp$
```

The mode of operation $proj$ shall be defined as in B.3.

Converting terms:

The adversary A uses function symbols $f \in \Sigma_{AEAD}$, so B has to map those function symbols to the corresponding function symbols of Σ^{tran}_{AEAD}. To achieve this we define a function $convert$ as follows:

- $convert(f^l(t_1, ..., t_n)) = f^l(convert(t_1), ..., convert(t_n))$
  for all $f \notin \Sigma^{tran}_{AEAD}$.
- $convert(k^l_x()) = f_{k^l_x()}$
- $convert(E_h(k^l_h(), cons_S(n), H, m)) = f^{l(m,r)}_{E_h(k^l_h(),cons_S(n),H,r)}$ where r has the length of the bitstring corresponding to m.

- $convert(E_h(k^l_h(), cons_S(n), H, m)) = f_{E_h(k^l_h(), cons_S(n), \cdot, \cdot)}(convert(H))(convert(m))$
- $convert(E_c(k^l_c(), cons_S(n), H, m)) = f_{E_c(k^l_c(), cons_S(n), \cdot, \cdot)}(convert(H))(convert(m))$

Further we define that for a list of terms T $convert(T) = \{convert(t) : t \in T\}$.

Simulation:

As already stated above, B simulates Game 3 for A while playing the deduction soundness game $\text{DS}_{\mathcal{M} \cup (\mathcal{M}^{tran}_{AEAD}(\nu) \cup \mathcal{M}_{tran}(\nu')), \mathcal{I} \cup (\mathcal{I}^{tran}_{AEAD}(\nu) \cup \mathcal{I}_{tran}(\nu'))}(\eta)$. Because of $\mathcal{M}^{tran}_{AEAD}(\nu)$ and $\mathcal{M}_{tran}(\nu')$ being both parameterized transparent symbolic models we could compose them to one parameterized transparent symbolic model by definition $\mathcal{M}'_{tran}(\nu \| \nu')$ since both ν and ν' must be good. This also holds for the implementation analogously. We do not combine them though for readability but we use it in the proof later on.

In the following we describe how the queries sent by A are handled by B:

- `init`
  Initially, $\nu = \varnothing$. At first A sends "T, H" to B and starts to fill ν; for each $k^l_x() \in T$ B generates a key $k \xleftarrow{\$}$ DAEN.Gen$(1^\eta, r)$ with $r \xleftarrow{\$} \{0,1\}^\eta$. After this, B adds (l, k, k') to ν where k' is k if x = c and $k' \xleftarrow{\$} \{0,1\}^{|k|}$ otherwise. B then sends $\nu \| \nu'$ to its game and queries "`init` $convert(T)$, $convert(H)$". After that, B also queries "sgenerate $k^l_h()$" for each $k^l_h() \in T$.
- `generate`
  For each "`generate` t" request by A, B adds this request to $\mathbb{T}$ and also sends "`generate` t" to its own game. The response is then send to A again.
  For each subterm t' of t where $t' = E_h(k^l_h(), cons_S(n), H, m)$ is an honest encryption B also queries "sgenerate $convert(m)$"
  The same strategy is used for "`sgenerate` t" queries.
- `parse`
  For each "`parse` c" request by A, B sends "`parse` c" to its own game and receives a term t'. B then converts t' to t ($t = convert^{-1}(t')$) and sends t back to A.

F.9 Claim: Game 3 and Game 4 are indistinguishable.

Intuition:

We now show that the simulation by B of Game 3 is indistinguishable for A from the original game. We will do this in two steps: We first show that (1) every valid trace produced by A (in Game 3) leads to a valid trace by B (in Game 4) and call them *corresponding* traces. Then we show that for all pairs of corresponding traces (2) the output is the same for Game 3 and Game 4. We achieve this using invariants that hold for a relation between the libraries of both games.

(1) We first show that any trace $\mathbb{T}_A$ produced by A in Game 3 leads to a trace $\mathbb{T}_B$ produced in Game 4 with $valid(\mathbb{T}_A) \Rightarrow valid(\mathbb{T}_B)$. We do this in two steps: The application of the $convert$ function is a variation described in requirement i) of the $valid$ predicates. Further, the additional "`sgenerate` t" queries are a variation described in the $valid$ predicates. Thus, we see that both variation still lead to a valid trace and we see for any term t meeting the requirements of the `generate` function in Game 3, $convert(t)$ meets the requirements in the

$$\text{DS}_{\mathcal{M} \cup (\mathcal{M}^{tran}_{AEAD}(\nu) \cup \mathcal{M}_{tran}(\nu')), \mathcal{I} \cup (\mathcal{I}^{tran}_{AEAD}(\nu) \cup \mathcal{I}_{tran}(\nu'))}(\eta)$$

game.

(2) **Defining the invariants:**
After showing the existence of the corresponding traces, we now have to show that the output of Game 3 and the simulation is the same. We start doing so by defining an invariant.
We assume that the same random coins are used for Game 3 and the simulation. We can do this w.l.o.g. by observing that B and A are using the same amount of randomness to generate keys and that all other uses of random coins coincide. With this we can now define the two invariants:
i) $dom(L_{big}) = dom(L_{small})$
ii) $\forall c \in dom(L_{small}) : convert^{-1}(L_{small}[\![c]\!]) = L_{big}[\![c]\!]$
where L_{big} is the library in Game 2 and L_{small} the library in the simulation.

We now show that the invariants hold for all distinct queries done in both games:

- **Initially**
  Initially, $L_{big} = L_{small} = \varnothing$ so both invariants hold obviously.
- **init T, H**
  We can divide $t \in T \cup H$ into three types according to the `valid`$_{AEAD}$ requirements 2 and 3:
  - $t = k^l_h()$ $(convert(t) = f_{k^l_h()})$
  - $t = cons_S(n)$ $(convert(t) = f_{cons_S()})$
  - t does not contain any function symbol $f \in \Sigma_{AEAD}$ $(convert(t) = t)$

We see that for all nonces $t = cons_S(n)$ that are generated in the big game, B adds $convert(t)$ to its `init` request which leads to both invariants being fulfilled. We also observe that for each term $t = k_h^l()$ in the big game, B adds $convert(t)$ to its `init` request. Doing so leads to

$$dom(L_{big} \backslash \{(c, k_h^l()) : c \in [\![\tau_{AEAD}^{k_h}]\!], l \in labelsH\}) = dom(L_{small})$$

because both generated keys coincide but in the small game the key is not added to the library. The invariants still hold because in the small game , B uses `sgenerate` queries to add the *fake* keys corresponding to the real keys to the library.
Thus, for the `init` query both invariants hold.

- **generate** t
Assuming our invariants hold for both libraries L_{big} and L_{small}, we now show that they still hold after a valid "generate t" query by A.
In the big game we have that
$$(c_{big}, L'_{big}) := generate_{M_{big}, \mathcal{R}}(t, L_{big})$$
while we have
$$(c_{small}, L'_{small}) := generate_{M_{small}, \mathcal{R}}(convert(t), L_{small})$$
in the small game. Note that there might be additional $generate_{M_{small}, \mathcal{R}}$ calls if t contains honest encryptions under honest keys.
Again with a case distinction over t we show that our invariants still hold.
  - M_{big} and M_{small} are never called for k_x and the transparent version of k_x, respectively (according to requirement 2 of the $\texttt{valid}_{AEAD}$ predicates). Therefore, both invariants hold for this case.
  - For a term $t = E_h(k_h^l(), cons_S(n), H, r)$ being an honest encryption, we would call
  $(c_{big}, L'_{big}) := \texttt{generate}_{M_{big}, \mathcal{R}}(t, L_{big})$
  in the big game whereas we would call
  $(c_{small}, L'_{small}) := \texttt{generate}_{M_{small}, \mathcal{R}}(f_{E_h(k_h^l(), cons_S(n), H, r)}(k_h^l)(cons_S(n))(H)(m), L_{small})$
  and $(c_{small}, L'_{small}) := \texttt{generate}_{M_{small}, \mathcal{R}}(m, L_{small})$
  in the small game. Since both resulting bitstrings coincide, both invariants also hold in this case.
  - For all other terms t all keys that are already in library are removed from t by applying the $convert$ function on t but no other changes are done, hence both generated bitstrings coincide. Concluding, both invariants hold.

  This works analogously for "`sgenerate` t" queries.
So we see that our invariants hold for L'_{big} and L'_{small}, implying that the response sent to A is $c = c_{big} = c_{small}$, i.e. the same response in both settings.
- **parse** c
Assuming our invariants hold for both libraries L_{big} and L_{small}, we now show that they still hold after a valid "parse c" query by A. We can show this by observing the cases of c:
If $c \in dom(L_{big})$ there are no changes in L and due to the invariant $ii)$ the response will be the same in both games.
The other case, where $c \notin dom(L_{big})$ (meaning if we have to parse an unknown bitstring) can be reduced to observing the **open** functionality in respect to $[\![\mathcal{T}_{AEAD}]\!]$, because both implementations should be type-safe (so we do not need to care about foreign bitstrings). Thus, because both **open** functions handle those queries the same way (Concretely, they only handle them the same way by applying conversion) our invariants still hold.

F.10 Claim: If A wins, then B wins Game 4.

1) We first showed that playing the original deduction soundness game
$$\mathsf{DS}_{(\mathcal{M} \cup \mathcal{M}_{AEAD}) \cup \mathcal{M}_{tran}(\nu), (\mathcal{I} \cup \mathcal{I}_{AEAD}) \cup \mathcal{I}_{tran}(\nu)}(\eta)$$
from Game 0 is indistinguishable to Game 3.
2) In the last proof we have shown that Game 3 and Game 4 are indistinguishable, meaning due to the invariants from above, we know that for every bitstring sent by A, c is parsed as t in the big game and hence parsed as $convert(t)$ in the small game. This shows, by checking both deduction systems, that if t is non-DY $convert(t)$ is also non-DY. This would be a contradiction to $\mathcal{I}$ being a deduction sound implementation of $\mathcal{M}$
Therefore, A can win Game 3 only with negligible probability.

Taking 1) and 2) into account we see that A can win Game 0 only with negligible probability which concludes our proof. □

Appendix G
Forgetfulness

The drawback of the Theorem 1 and the theorems proven by Böhl, Cortier, and Warinschi [8] lies in the valid predicates, which forbids that keys are used in non-key positions; this means that keys cannot be sent around, i.e. key wrapping is forbidden by $\texttt{valid}_{AEAD}$.

Because deduction soundness does not guarantee that no information about non-DY terms, e.g. a few bits of an used nonce, is leaked by the implementation, Böhl, Cortier, and Warinschi [8] introduced *forgetful symbolic models* and implementations. Intuitively, we mark positions in the forgetful symbolic model as forgetful and the forgetful implementation guarantees that no information about arguments, except their length, at those forgetful positions is leaked to the outside.

For this paper to be self-contained, we will first recall the definitions by Böhl et al. Afterwards, we extend our contribution by also showing forgetfulness for (deterministic) AEAD, which puts a large class of protocols (protocols using key wrapping, key exchange protocols) in the scope of our result.

G.1 Review: preliminaries

In this section, we will simply adopt all the needed preliminaries and in the following section the definitions concerning forgetfulness from Böhl, Cortier, and Warinschi [8]. Böhl *et al.*extend the previous setting as follows.

- *Changed hybrid terms for function symbols with forgetful arguments.*
  Let f be a function symbol $f \in \Sigma$ with arity $ar(f) = \tau_1 \times ... \times \tau_n \to \tau$. Let $f(a_1, ..., a_n)$ be the corresponding hybrid term where each a_i is either a bitstring with $a_i \in [\![\tau_i]\!]$ as usual or a term with $type(a_i) = \tau_i$ for the forgetful position i. The definition of completeness of a library as well as the definition of $\mathrm{L}[\![c]\!]$ has to be changed to adapt to the change above.
- *New* valid *requirements*
  We introduce a new signature Σ_{valid} which features function symbols with forgetful position. We change the valid requirements, such that they use Σ_{valid} in the following way:

(i) If $\texttt{valid}(\mathbb{T} + q) = \texttt{true}$ then for any *variation* $\hat{q}$ of q, which we define explicitly below, it holds that $\texttt{valid}(\mathbb{T} + \hat{q}) = \texttt{true}$.
If q = "generate t" and $\hat{q}$ = "generate $\hat{t}$":
Here, a variation $\hat{t}$ of t is defined as follows: Any subterm $f^l(t_1, ..., t_n)$ of t where f is a foreign function symbol, i.e. $f \notin \Sigma \cup \Sigma_{valid}$, can be replaced by $\hat{f}^{\hat{l}}(\hat{t}_1, ..., \hat{t}_n)$ where $\hat{f}$ is another foreign function symbol and for all $i \in \{1, ..., n\}$ $\hat{t}_i$ is either t_j for some $j \in \{1, ..., n\}$ (where each t_j can only be used once) or $\hat{t}_i$ does not contain any function symbols from $\Sigma \cup \Sigma_{valid}$. As a special case, if for example $\hat{f}$ is "empty", we may replace $f^l(t_1, ..., t_n)$ with a term $\hat{t}_1$.
If q = "init T, H" and $\hat{q}$ = "init $\hat{T}, \hat{H}$":
$T = (t_1, ..., t_n)$ and $\hat{T} = (\hat{t}_1, ..., \hat{t}_n)$ where for each $i \in \{1, ..., n\}$ $\hat{t}_i$ is a variation of t_i; and $H = (h_1, ..., h_m)$ and $\hat{H} = (\hat{h}_1, ..., \hat{h}_m)$ where for each $i \in \{1, ..., m\}$ $\hat{h}_i$ is a variation of h_i.

(ii) If $\texttt{valid}(\mathbb{T} + q) = \texttt{true}$ and t is a term which occurs in q, then $\texttt{valid}(\mathbb{T}+$ "sgenerate t'") $= \texttt{true}$ for any t' being a subterm of t and t' not being a subterm of a forgetful position.

(iii) The evaluation of $\texttt{valid}(\mathbb{T})$ can be done in polynomial time (in the length of trace $\mathbb{T}$).

G.2 Review: Forgetful symbolic models and implementations

We call a symbolic model $\mathcal{M}$ a *forgetful symbolic model* if there are function symbols with forgetful positions, i.e. if there are arguments of a function symbol which are marked as forgetful. To define the corresponding forgetful implementation we need to introduce the *oblivious implementation*, which is defined like the implementation seen in Section B.2 but for all forgetful positions, the inputs are natural numbers instead of bitstrings. More precisely, the input will be the length of the actual input of the forgetful position.

Definition 10 (Oblivious implementation)**.** *Let $\mathcal{M}$ be a forgetful symbolic model. $\overline{\mathcal{I}} = (\overline{M}, [\![\cdot]\!], len, \texttt{open}, \texttt{valid})$ is an oblivious implementation of $\mathcal{M}$ if $\mathcal{I}$ is an implementation of $\mathcal{M}$ with a slightly changed signature: For each function symbol $f \in \Sigma$ with arity $ar(f) = \tau_1 \times ... \times \tau_n \to \tau$ the signature of $(\overline{M}f)$ is $\theta(\tau_1) \times ... \times \theta(\tau_n) \times \{0,1\}^\eta \to [\![\tau]\!]$ where $\theta(\tau_i) = \mathbb{N}$ if the ith argument of f is forgetful and $[\![\tau_i]\!]$ otherwise.*

Review of Definition 8 by Böhl, Cortier, and Warinschi [8]

Let $\texttt{generate}^{FIN}$ be a $\texttt{generate}$ function that also deals with forgetful arguments in the following way (*Review of Figure 14 by* Böhl, Cortier, and Warinschi [8]):

Listing 9: $\texttt{generate}^{FIN}$

```
generate$^{FIN}_{M,\mathcal{R}}(t, L)$:
    if for some $c \in dom(L)$ we have $L[\![c]\!] == t$ then
        return $c$
    else
        for $i \in \{1, ..., n\}$ do
        if $i$ is a forgetful argument then
            let $c_i := len(t_i)$
            let $a_i := t_i$
        else
            let $(c_i, L) :=$ generate$_{M,\mathcal{R}}(t_i, L)$
            let $a_i := t_i$
    let $r := \mathcal{R}(t)$
    let $c := (Mf)(c_1, ..., c_n; r)$
    let $L(c) := f^l(a_1, ..., a_n))(l \in$ labesH$)$
    return $(c, L)$
```

Let $\textbf{FIN}^b_{\mathcal{M}(\nu),\mathcal{I}(\nu),\overline{\mathcal{I}}(\nu),A}(\eta)$ be a game in which an adversary tries to distinguish if he interacts with a real implementation $\mathcal{I}$ and an oblivious implementation $\overline{\mathcal{I}}$ (*Review of Figure 15 by* Böhl, Cortier, and Warinschi [8]):

Listing 10: FIN game

```
FIN$^b_{\mathcal{M}(\nu),\mathcal{I}(\nu),\overline{\mathcal{I}}(\nu),A}(\eta)$:
    Let $S$ := $\varnothing$
    Let $L$ := $\varnothing$
    Lat $\mathbb{T}$ := $\varnothing$
    $\mathcal{R}$ ← $\{0,1\}^*$

    if b == 0 then
        let generate := generate$^{FIN}_{M,\mathcal{R}}$
    else
        let generate := generate$_{M,\mathcal{R}}$

    Receive parameter $\nu$ from $A$

    on request "init $T, H$" do
        add "init $T$" to $\mathbb{T}$
        if valid($\mathbb{T}$) then
            let $S$ := $S \cup T$
            let $C$ := $\varnothing$
            for each $t \in T$ do
                let $(c, L)$ := generate$(t, L)$
                let $C$ := $C \cup \{c\}$
            for each $t \in H$ do
                let $(c, L) :=$ generate$(t, L)$
            send $C$ to $A$
        else
            return 0 ($A$ is invalid)

    on request "sgenerate $t$" do
        if valid($\mathbb{T}$ + "sgenerate $t$") then
            let $(c, L) :=$ generate$(t, L)$

    on request "generate $t$" do
        add "generate $t$" to $\mathbb{T}$
        if valid($\mathbb{T}$) then
            let $S$ := $S \cup \{t\}$
            let $(c, L) :=$ generate$(t, L)$
            send $c$ to $A$
        else
            return 0 ($A$ is invalid)

    on request "parse $c$" do
        let $(t, L)$ := parse$(c, L)$
        if $S \vdash_D t$ then
            send $t$ to $A$
        else
            return 1 ($A$ produced a non-DY term)

    on request "guess $b'$" do
        if $b' == b$ then
            return 1 ($A$ wins)
        else
            return 0 ($A$ loses)
```

where S is the set of requested terms, L is the library, $\mathbb{T}$ the trace of queries, $\mathcal{R}$ the random tape and C the list of replies.

Definition 11. *(Forgetful Implementation). We say that an implementation $\mathcal{I} = (M, [\![\cdot]\!], len, \texttt{open}, \texttt{valid})$ is a forgetful implementation of a forgetful symbolic model $\mathcal{M}$ if there is an oblivious implementation $\overline{\mathcal{I}} = (\overline{M}, [\![\cdot]\!], len, \texttt{open}, \texttt{valid})$ such that for all parameterized transparent symbolic models $\mathcal{M}_{tran}(\nu)$ and for all parameterized implementations $\mathcal{I}_{tran}(\nu)$ of $\mathcal{M}_{tran}(\nu)$ that are compatible with $(\mathcal{M}, \mathcal{I})$ we have that*

$$\begin{aligned} &Pr[\textbf{\textit{FIN}}^{0}_{\mathcal{M}\cup\mathcal{M}_{tran},\mathcal{I}\cup\mathcal{I}_{tran},\overline{\mathcal{I}}\cup\mathcal{I}_{tran},A}(\eta) = 1] \\ -&Pr[\textbf{\textit{FIN}}^{1}_{\mathcal{M}\cup\mathcal{M}_{tran},\mathcal{I}\cup\mathcal{I}_{tran},\overline{\mathcal{I}}\cup\mathcal{I}_{tran},A}(\eta) = 1] \end{aligned}$$

is negligible for every p.p.t. adversary A.

(Review of Definition 9 by Böhl, Cortier, and Warinschi [8])

Thus we see that an forgetful implementation $\mathcal{I}$ as defined in Definition 11 is an implementation that is indistinguishable from an oblivious implementation $\overline{\mathcal{I}}$ (see Definition 10) of a forgetful symbolic model $\mathcal{M}$.

Lemma 4. *Let $\mathcal{M}$ be a forgetful symbolic model, $\mathcal{I}$ be an forgetful implementation of $\mathcal{M}$ and $\overline{\mathcal{I}}$ a corresponding oblivious implementation. If $\mathcal{I}$ is deduction sound, then $\overline{\mathcal{I}}$ is deduction sound with respect to the deduction soundness game* **DS**$'$ *that uses* $\texttt{generate}^{FIN}$ *instead of* `generate`*.*

(Review of Lemma 5 by Böhl, Cortier, and Warinschi [8])

G.3 Solving the key wrapping problem

We now will extend our model (and its corresponding implementation) from Section D (and Section E) in a way, s.t. we allow to send keys around. Therefore we have to change $\texttt{valid}_{AEAD}$ to depend on Σ_{valid} (See G.1). Additionally, we replace requirement 3) of the $\texttt{valid}_{AEAD}$ predicate with

3) i) For the query "`init` T, H" it should hold that:
   - ∗ the function symbol k_c should only occur in a term $k^l_c() \in T$.
   - ∗ the function symbol k_h should only occur in a term $k^l_h() \in H$.

   ii) For each label l of k^l_x, l should be unique in $T \cup H$.

   iii) We demand that every time $k^l_x()$ occurs in a `generate` query, $k^l_x()$ must have occurred in the `init` query before.

   iv) Except generation, $k^l_x()$ should only occur in E_x as its first argument or as a sub term of f in a forgetful position with $f \in \Sigma_{valid}$.

Further, we extend $\mathcal{I}_{AEAD}$ to $\mathcal{I}_{AEAD}[\Sigma_{valid}]$, an implementation featuring forgetful positions. Note that Σ_{valid} needs to be fixed at the time of composition.

Further note that if we, i.e. want to send AEAD keys around using authenticated encryption, we need to compose $(\mathcal{M}_{AEAD}, \mathcal{I}_{AEAD})$ with $(\mathcal{M}'_{AEAD}, \mathcal{I}'_{AEAD}[\Sigma_{AEAD}])$ where

1) $(\mathcal{M}_{AEAD}, \mathcal{I}_{AEAD})$ are forgetful regarding the message position of the encryption function symbol E_x,
2) Σ_{AEAD} the signature of $\mathcal{M}_{AEAD}$ and Σ'_{AEAD} of $\mathcal{M}'_{AEAD}$
3) and the `valid` predicate of $\mathcal{I}_{AEAD}$ does not depend on function symbols of Σ'_{AEAD}.

Intuitively, the third point prevents key cycles by stating that dependencies in the `valid` predicates can only be one-directed, allowing E_x to have key symbols of $\mathcal{M}'_{AEAD}$ in its fourth position, but forbids that key symbols of $\mathcal{M}_{AEAD}$ can occur at a non key position of E'_x. We will generalize the composition to allow key wrapping of AEAD keys in Theorem 2 and show that we can use authenticated encryption to wrap keys in Theorem 3.

Böhl, Cortier, and Warinschi [8] already proved forgetfulness of *public key encryption* and *secret key encryption* in respect to the message position as a forgetful position. Now we want to show that we can send AEAD keys (and also other bitstrings produced by $\mathcal{I}_{AEAD}$) using such forgetful implementations. This idea is concretized in Theorem 2.

Theorem 2. *Let $\mathcal{I}$ be a deduction sound forgetful implementation of the forgetful symbolic model $\mathcal{M}$ and let $\mathcal{I}_{AEAD} := \mathcal{I}_{AEAD}[\Sigma]$ where Σ is the signature of $\mathcal{M}$. If $(\mathcal{M}, \mathcal{I})$ is composable with $(\mathcal{M}_{AEAD}, \mathcal{I}_{AEAD})$ and if* `valid` *(of $\mathcal{I}$) does not depend on Σ_{AEAD}, then $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a deduction sound implementation of $\mathcal{M} \cup \mathcal{M}_{AEAD}$.*

Proof. Instead of the five different games used in the previous proof of Theorem 1, we will include another game, so we have to show indistinguishability between six games. In this additional game we will replace the forgetful implementation $\mathcal{I}$ of the forgetful symbolic model $\mathcal{M}$ with its corresponding oblivious implementation $\overline{\mathcal{I}}$. This gives us an additional guarantee: If an adversary A sends the request "`generate` t" with honest keys at a forgetful position of t to the **DS** game then we know that the bitstring interpretations of those honest keys are not used to compute the bitstring corresponding to term t.

Our strategy in proving Theorem 3 will stay the same as in the proof of Theorem 1: We will describe the games and show that they are indistinguishable from each other. For completeness we will also describe the games that do not differ from the proof of Theorem 1.

Game 0 As in proof 1, in this game the adversary A plays the original deduction soundness game from B.5.

$$\mathrm{DS}_{(\mathcal{M}\cup\mathcal{M}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu)}(\eta)$$

Game 1 In Game 1, we replace $\mathcal{I}$ by the corresponding oblivious implementation $\overline{\mathcal{I}}$. By Definition 3 $\overline{\mathcal{I}}$ must exist and $\overline{\mathcal{I}}$ must also be composable with $\mathcal{I}_{AEAD}$. We also replace the function `generate` with `generate`^{FIN} from listing 9. The adversary A then plays

$$\mathrm{DS}_{(\mathcal{M}\cup\mathcal{M}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\overline{\mathcal{I}}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu)}(\eta)$$

Claim: Game 0 and Game 1 are indistinguishable To prove that Game 0 is indistinguishable from Game 1 for A, we first construct an adversary B who plays

$$\mathbf{FIN}^{b}_{\mathcal{M}\cup\mathcal{M}_{tran'}(\nu'),\mathcal{I}\cup\mathcal{I}_{tran'}(\nu'),\overline{\mathcal{I}}\cup\mathcal{I}_{tran}(\nu'),A}(\eta)$$

and simulates Game b for A with $b \in \{0,1\}$. B simulates $\mathcal{I}_{AEAD}$ using transparent functions and checks the validity of A's requests like it would be done in Game 1. This simulation is perfect since B uses transparent functions to simulate $\mathcal{I}_{AEAD}$ and therefore knows all generated terms and does not need to hide any sub terms.

If A could distinguish between Game 0 and Game 1 and would return b' to B (where b' is A's choice in distinguishing the games), then B would simply send a "guess b'" request to its own game and win with non-negligible probability. But, B can only win its game with negligible probability according to Definition 11. Therefore: A can only distinguish between Game 0 and Game 1 with negligible probability.

Game 2 In this game we replace the `generate`^{FIN} function in the deduction soundness game by the collision aware version (like in Game 1 from F.2).

Claim: Game 1 and Game 2 are indistinguishable The proof of indistinguishability is completely analogous to the proof in F.3.

Game 3 In Game 3 we make similar changes to those of Game 2 from F.4. We take the adversary's option of learning something from cyphertexts or about honest keys, by replacing the cyphertexts created under honest keys with random bitstrings. We abstract from just using random bitstrings by using encryptions of random bitstrings under the same keys (because the adversary still needs the possibility to retrieve the nonce and the header) and also replace all honest keys in the library by random bitstrings with the same length as the keys. Because of the properties of the AEAD secure scheme, A should not notice any difference between Game 2 and Game 3.

The concrete changes are completely analogous to those in F.4.

Claim: Game 2 and Game 3 are indistinguishable In this proof, we will use the fact that we have to deal with $\overline{\mathcal{I}}$ instead of $\mathcal{I}$. $\overline{\mathcal{I}}$ guarantees that bitstrings representing honest keys are not used to generate other terms. So if we replace the keys with random bitstrings, the games will be indistinguishable by construction. The rest of the proof is analogous to the proof in Section F.5.

Game 4 In this game, we also give the adversary the possibility to produce honest encryptions of arbitrary messages. The game is constructed like in F.6.

Claim: Game 3 and Game 4 are indistinguishable The proof of indistinguishability is analogous to the proof in F.5. But we need to simulate $\mathcal{I}_{AEAD}$ using transparent functions (for the simulator B) for which we need the fact that `valid` of the forgetful implementation $\mathcal{I}$ does not depend on function symbols from Σ_{AEAD} (what we demand in the `valid` predicate). Game 3 and Game 4 are therefore indistinguishable, too.

Game 5 Game 5 is completely analogous to Game 4 from Section F.8. We create an adversary B who plays the deduction soundness game for $\mathcal{M}$ and $\overline{\mathcal{I}}$ and again simulates Game 4 for A by adding transparent functions to simulate authenticated encryption with associated data.

Claim: Game 4 and Game 5 are indistinguishable We have to show that the simulation by B of Game 4 is indistinguishable for A from the original game. This can be done in two steps analogously to proof F.9. First one shows that every valid trace produced by A (in Game 4) leads to a valid trace by B (in Game 5) and call them *corresponding* traces. Then one can show that for all pairs of corresponding traces the output is the same for Game 4 and Game 5. Additionally we use the indistinguishability of $\mathcal{I}$ from $\overline{\mathcal{I}}$ stated in lemma 4. Analogously to proof F.9 Game 4 and Game 5 are therefore indistinguishable.

Claim: If A wins, then B wins Game 5 Again, B can only win its game with negligible probability with the same arguments as in F.10. Therefore A can only win Game 4 with negligible probability and since Game 4 is indistinguishable from Game 0, the same holds for the winning probability of A playing Game 0.

We can conclude that $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a deduction sound implementation of $\mathcal{M} \cup \mathcal{M}_{AEAD}$. □

Now we want to show that, if we mark the message position of encryption under AEAD as a forgetful position, we are able to send keys around. Let $\overline{\mathcal{M}}_{AEAD}$ be a forgetful symbolic model which is based on $\mathcal{M}_{AEAD}$ but for honest encryption $E_h(k_h^l(), cons_S(n), H, m)$ m is marked as a forgetful position. If we then pick $\mathcal{I}_{AEAD}[\Sigma]$ as the implementation of $\mathcal{M}_{AEAD}$ the the following should hold:

Theorem 3. *Let $\mathcal{I}$ be a deduction sound forgetful implementation of the forgetful symbolic model $\mathcal{M}$ and let $\mathcal{I}_{AEAD} := \mathcal{I}_{AEAD}[\Sigma]$ where Σ is the signature of $\mathcal{M}$. If $(\mathcal{M}, \mathcal{I})$ is composable with $(\overline{\mathcal{M}}_{AEAD}, \mathcal{I}_{AEAD})$, then $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a forgetful implementation of $\mathcal{M} \cup \overline{\mathcal{M}}_{AEAD}$.*

Proof. Intuitively, we want to show that we can use encryption under AEAD to send keys around. Therefore, we have to show that $\mathcal{I} \cup \mathcal{I}_{AEAD}$ is a forgetful implementation of $\mathcal{M} \cup \overline{\mathcal{M}}_{AEAD}$. Using Definition 11 we simply have to show that no adversary A can efficiently distinguish between $\mathcal{I} \cup \mathcal{I}_{AEAD}$ and $\overline{\mathcal{I}} \cup \overline{\mathcal{I}}_{AEAD}$ where $\overline{\mathcal{I}}$ is the corresponding oblivious implementation of $\mathcal{I}$ and $\overline{\mathcal{I}}_{AEAD}$ the corresponding oblivious implementation of $\mathcal{I}_{AEAD}$. For $\overline{\mathcal{I}}_{AEAD}$ we further set the application of the oblivious Turing machine on an honest encryption function symbols $(\overline{M}_{AEAD}\ E_h)(k, n', H, m_{forget}) :=$ $(\overline{M}_{AEAD}\ E_h)(k, n', H, r)$ with r being a random value in the size of m.

Hence, we have to show that

$$\begin{aligned}&Pr[\mathbf{FIN}^1_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\\ -&Pr[\mathbf{FIN}^0_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\end{aligned}$$

is negligible small.

We can prove that this equation holds, using an intermediate step.
Analogously to Game 1 from Theorem 2 we first only replace $\mathcal{I}$ with its corresponding oblivious implementation $\overline{\mathcal{I}}$. Since $\mathcal{I}$ is a forgetful implementation, $\overline{\mathcal{I}}$ must be composable with $\mathcal{I}_{AEAD}$, too. So, analogously to Game 1 we can show that

$$\begin{aligned}&Pr[\mathbf{FIN}^1_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\\ -&Pr[\mathbf{FIN}^0_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\end{aligned}$$

is negligible in respect to the security parameter η. So finally we have to show that

$$\begin{aligned}&Pr[\mathbf{FIN}^0_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\\ -&Pr[\mathbf{FIN}^0_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\end{aligned}$$

is negligible in respect to the security parameter η, too.

In this step we simply replace $\mathcal{I}_{AEAD}$ with its corresponding oblivious implementation $\overline{\mathcal{I}}_{AEAD}$ and replace the keys in the library by random bitstrings.
Since an adversary needs to break authentication of the implementation to be able to distinguish both games, what can only happen with negligible probability since the scheme is DAE-N secure, the statement is valid.

Using both steps, we finally have shown that

$$\begin{aligned}&Pr[\mathbf{FIN}^1_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\\ -\ &Pr[\mathbf{FIN}^0_{(\mathcal{M}\cup\overline{\mathcal{M}}_{AEAD})\cup\mathcal{M}_{tran}(\nu),(\mathcal{I}\cup\mathcal{I}_{AEAD})\cup\mathcal{I}_{tran}(\nu),(\overline{\mathcal{I}}\cup\overline{\mathcal{I}}_{AEAD})\cup\mathcal{I}_{tran}(\nu),A}(\eta) = 1]\\ \leqslant\ &\mathtt{negl}(\eta)\end{aligned}$$

for some negligible function negl() and security parameter η.

□