Characterization of the Relations between Information-Theoretic Non-malleability, Secrecy, and Authenticity

Akinori Kawachi*[1], Christopher Portmann†[2,3], and Keisuke Tanaka‡[1]

[1]Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan.
[2]Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland.
[3]Group of Applied Physics, University of Geneva, 1211 Geneva, Switzerland.

August 15, 2012

Abstract

Roughly speaking, an encryption scheme is said to be non-malleable, if no adversary can modify a ciphertext so that the resulting message is meaningfully related to the original message. We compare this notion of security to secrecy and authenticity, and provide a complete characterization of their relative strengths. In particular, we show that information-theoretic perfect non-malleability is equivalent to perfect secrecy of two different messages. This implies that for n-bit messages a shared secret key of length roughly $2n$ is necessary to achieve non-malleability, which meets the previously known upper bound. We define approximate non-malleability by relaxing the security conditions and only requiring non-malleability to hold with high probability (over the choice of secret key), and show that any authentication scheme implies approximate non-malleability. Since authentication is possible with a shared secret key of length roughly $\log n$, the same applies to approximate non-malleability.

1 Introduction

There exist many different cryptographic goals to protect information. The most basic is *secrecy*, namely, that the desired information remain unknown to an adversary. Information-theoretic perfect secrecy was already fully characterized by Shannon in the 40's [1]. *Authentication* is another important task, which consists in guaranteeing that the information has not been tampered with, that it really comes from who it claims. Wegman and Carter's seminal work [2] is considered the corner stone in information-theoretic authentication, since it is

*kawachi@is.titech.ac.jp
†chportmann@itp.phys.ethz.ch
‡keisuke@is.titech.ac.jp

the first paper to show that the secret key needed can be much shorter than the message. *Non-malleability* is yet another goal. This notion of security was introduced by Dolev, Dwork and Naor [3] for computational security, and has received quite a lot of attention since.

Roughly speaking, non-malleability is the requirement that an adversary, when given a ciphertext, should not be able to produce a new ciphertext such that the two corresponding messages are "meaningfully related." Or, in other words, the adversary cannot perform a "controlled modification" of the underlying message. For example, if a document such as a contract is encrypted, a dishonest party might try to modify the ciphertext in such a way that he only modifies the amount of money due in the contract. With encryption schemes such as the one-time pad this is perfectly possible, because flipping a bit of the ciphertext flips a bit of the underlying message, even though perfect secrecy is guaranteed.

Shared secret keys are considered a very expensive resource, and thus bounding the length of the key needed and finding schemes which meet this bound are amongst the most important tasks when studying information-theoretic security. In his much celebrated work, Shannon [1] showed that to provide (perfect) secrecy for one message, an encryption scheme requires a shared key at least as long as that message.

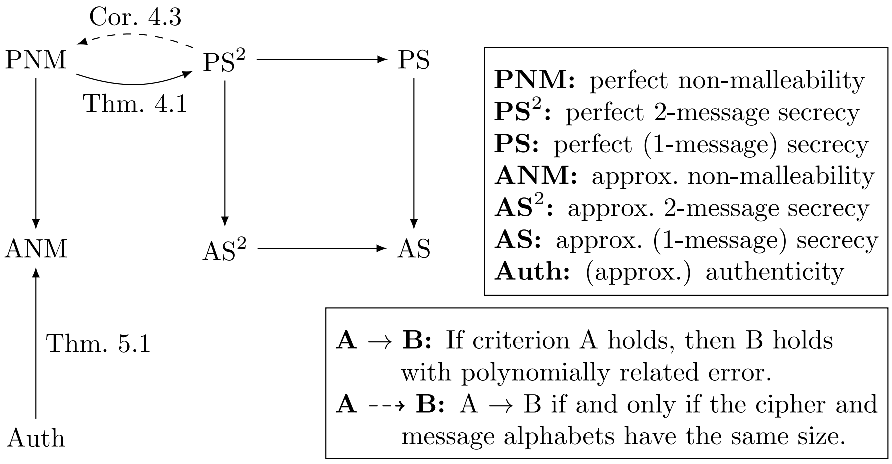
Perfect security can be an expensive or sometimes even an impossible goal to achieve. Relaxing the security conditions and only requiring the security criteria to be met with high probability over the choice of keys often results in great improvements. For example, perfect authentication is impossible: there is always a small chance that a forged message and authentication code (MAC) match.[1] Therefore we can at best guarantee with probability $1 - 1/|\mathcal{Z}|$ that a correctly authenticated message has not been tampered with, where $\mathcal{Z}$ is the alphabet of the MAC appended to the message. To achieve an error of exactly $1/|\mathcal{Z}|$, a shared secret key of length at least n bits is needed [4], where $n = \log|\mathcal{X}|$ is the length of the message. By simply increasing the error from $1/|\mathcal{Z}|$ to $2/|\mathcal{Z}|$, Wegman and Carter [2] showed that the shared secret key needed can be reduced from n to roughly $\log n$ bits.

Previous work on non-malleability. In the case of computational security, several non-malleable schemes have been proposed with semantical "simulation based" security definitions [3] and indistinguishability or "comparison based" security definitions [5, 6]. Many papers focus on comparing and classifying the relative strengths of these different notions of security, both in the public-key setting [7, 8] and computational symmetric-key setting [9].

In the case of information-theoretic security, Hanaoka, Shikata, Hanaoka, and Imai [10, 11] were the first to formalize non-malleable security.[2] To do so, they use the language of entropies and measure the information the adversary

[1] To authenticate a message m, a pair $(m, h_k(m))$ is generated and sent, where k corresponds to the shared secret key, and $h_k(m)$ is the message authentication code (MAC). An adversary wishing to modify the message has to guess the correct $h_k(m')$ corresponding to the new message m' for it to be accepted.

[2] The standard information-theoretic definition of non-malleability (Definition 3.3) is not an immediate adaptation of one of the computational definitions, but differs somewhat in the details. We refer to Section 7 for further comments on this.



PNM: perfect non-malleability
PS^2: perfect 2-message secrecy
PS: perfect (1-message) secrecy
ANM: approx. non-malleability
AS^2: approx. 2-message secrecy
AS: approx. (1-message) secrecy
Auth: (approx.) authenticity

A → B: If criterion A holds, then B holds with polynomially related error.
A ⇢ B: A → B if and only if the cipher and message alphabets have the same size.

Figure 1 – *Complete characterization of the relations between different notions of information-theoretic non-malleability, secrecy, and authenticity.* A directed path between two notions of security means that any scheme providing the first also provides the second. If there is no directed path from one security criterion to another, then there exists an example of a scheme that satisfies the first security definition, but not the second. The dashed arrow means that this relation only holds if the message and cipher alphabets have the same cardinality.

has about the new message given the original one.[3] McAven, Safavi-Naini, and Yung [12] generalized their definition to the case of ciphertexts longer than the message and approximate security. Schemes exist which are known to provide non-malleability and secrecy [10, 11] or non-malleability and authenticity [13]. However, prior to this work, there existed no security reduction between these different notions. These previous works on information-theoretic non-malleability [10–12] did not consider the optimality of the secret key length, and no lower bound on this key length was known.

New results. In this work we provide a complete characterization of the relations between perfect and approximate information-theoretic non-malleability, secrecy, and authenticity, which we illustrate in Figure 1.[4] Only the trivial relations (depicted in Figure 1 by arrows without any reference to a theorem) were previously known.

We first study perfect non-malleability and show that it is equivalent to requiring that the encryption function uniformly maps any two different messages to all possible pairs of two different ciphertexts. This is equivalent to perfect secrecy of two different messages (PS^2 in Figure 1) when the message and ciphertext alphabets have the same size, and strictly stronger if the size of the ciphertext alphabet is larger than that of the message.

An immediate consequence of this is a lower bound on the key needed for perfect non-malleability, namely $\log[|\mathcal{X}|(|\mathcal{X}|-1)]$ bits, where $\mathcal{X}$ is the message

[3] We refer to Section 3 for a precise definition of information-theoretic non-malleability.

[4] We show in Section 6 that this figure is indeed complete: if there is no directed path from one security criterion to another, then there exists an example of a scheme that satisfies the first security definition, but not the second.

alphabet, since this is the key length needed for perfect secrecy of two different messages. This also proves that a scheme by Hanaoka et al. [10, 11] is optimal in the key size.

The converse yields a very easy way to design perfect non-malleable schemes, since we do not need to consider adversary strategies or invalid ciphertexts.

We then relax the security definition of non-malleability to only hold with high probability over the choice of secret key, and define approximate non-malleability (ANM in Figure 1).[5] We prove that any authentication scheme with error ε (Auth in Figure 1) is a non-malleable scheme with error $\varepsilon' \leq 2\sqrt{\varepsilon}$, even though the formal definition of non-malleability does not consider the adversary to have failed if his choice of forged ciphertext is invalid. This answers an open question by Hanaoka [11].

This also means that authentication techniques such as approximate strong 2-universal hashing provide approximate non-malleability with a shared secret key of length roughly $2 \log\log |\mathcal{X}| + 3 \log \frac{1}{\varepsilon}$ [14], where $\mathcal{X}$ is the message alphabet and ε the error probability.[6]

Intuition and comparison to computational security. Intuitively, there are two ways to achieve information-theoretic non-malleable security:

1. If any two different message are uniformly mapped to all pairs of different ciphertexts, then the adversary's falsified message gets mapped at random to any message when decrypted, and thus not be correlated to the original message.
2. If the legitimate players nearly always detect tampering with the ciphertext, the falsified ciphertext gets decrypted to "invalid" and thus not be correlated to the original message.

However, the second method is not sufficient to provide perfect security, since there is always a small chance that the adversary can fool the legitimate players into accepting a falsified message, in which case the underlying messages can be correlated. So requiring perfect security forces any non-malleable scheme to uniformly map different messages to all pairs of different ciphertexts, which is stronger than 2-message security.

In the computational symmetric-key setting, Katz and Yung [9] show that if the adversary has access to the encryption oracle, non-malleability is stronger than secrecy, and otherwise, authentication is sufficient to achieve non-malleability. In their model, if the adversary has access to the encryption oracle, he can use it to generate valid signatures, making authentication impossible. This forces any non-malleable scheme to use the first method outlined above, which is stronger than secrecy. But if the adversary does not have access to the encryption oracle, the second method works and authentication is sufficient.

In the public-key setting, non-malleability has a somewhat different meaning. Since all players use the same public-key to encrypt messages, there is no need to (correctly) establish who encrypted a message in order to decrypt it. Thus, if an adversary intercepts and modifies a ciphertext, he can also replace the name of the sender with his own. There is no need for the adversary to impersonify the

[5] As discussed in Appendix A.2, McAven et al.'s definition of approximate non-malleability [12] does not capture the notion of "security with high probability." We therefore redefine approximate non-malleability to reflect this concept.

[6] Since authentication does not imply secrecy, approximate non-malleability does not imply secrecy either. We refer to Section 5 for more details on this.

legitimate player as in the symmetric-key setting. For example, if a municipality makes a public offer for a construction contract, and company A encrypts his bid with the given public-key, company B can try to generate a new ciphertext corresponding to a smaller bid, and send that as his own bid. To prevent this, the scheme has to provide non-malleable security [3]. But since no impersonification or falsification takes place, authentication does not help, and any public-key non-malleable scheme is always stronger than secrecy [8, 15].

Structure of this paper. We start in Section 2 by introducing the notation and defining the symmetric-key encryption model used for information-theoretic security. In Section 3 we then define the different notions of perfect and approximate security needed in this work, namely secrecy, non-malleability, and authenticity. In Section 4 we prove the first main result about the relation between perfect non-malleability and perfect secrecy of two messages. In Section 5 we consider approximate security, and prove the second main result, that approximate non-malleability can be achieved by any authentication scheme. In Section 6 we then show that the relations depicted in Figure 1 are complete by providing a proof for all necessary separations. And finally in Section 7 we conclude with several remarks on the consequences of these results and a discussion of alternative information-theoretic non-malleable security definitions.

2 Preliminaries

2.1 Notation

In this paper we use calligraphic letters for alphabets (e.g., $\mathcal{X}$), lowercase letters for elements of these sets (e.g., $x \in \mathcal{X}$) and uppercase letters for random variables (e.g., X). We write $P_X(x)$ for the probability that X takes the value x. For two random variables X and Y with joint probability distribution $P_{XY}(\cdot,\cdot)$, we write $X|_{Y=y}$ to denote the random variable X given $Y = y$, and $P_{X|Y}(\cdot|y) := \frac{P_{XY}(\cdot,y)}{P_Y(y)}$ for the corresponding distribution. We also denote by $X \cdot Y$ the random variable with distribution $P_{X \cdot Y}(x, y) := P_X(x)P_Y(y)$. Note that unless X and Y are independent, $X \cdot Y \neq XY$.

To measure the distance between two random variables over a common alphabet we use the variational distance (sometimes also called statistical distance) and write

$$d(X,Y) = \frac{1}{2}\sum_{x\in\mathcal{X}} |P_X(x) - P_Y(x)|.$$

We denote the expected variational distance between X and Y over a third random variable Z by

$$d(X,Y|Z) := \frac{1}{2}\sum_{x,z} P_Z(z)\big|P_{X|Z}(x|z) - P_{Y|Z}(x|z)\big|.$$

The variational distance will be used in particular to measure how close two random variables (over possibly different alphabets) are to being independent from each other, i.e., we are interested in $d(XY, X \cdot Y)$. In this case, conditioning on a third random variable Z results in

$$d(XY, X \cdot Y|Z) = \frac{1}{2} \sum_{x,y,z} P_Z(z) \left| P_{XY|Z}(x,y|z) - P_{X|Z}(x|z) P_{Y|Z}(y|z) \right|.$$

For an alphabet $\mathcal{X}$ and a random variable X distributed over $\mathcal{X}$, we call *domain of* X and write $\mathscr{D}(X)$ the subset of $\mathcal{X}$ with non-zero probability, that is $\mathscr{D}(X) = \{x \in \mathcal{X} : P_X(x) > 0\}$. We will often be interested in several random variables (usually two) $X_1 \cdots X_\ell$, each one defined over the same alphabet $\mathcal{X}$, but such that $\mathscr{D}(X_1 \cdots X_\ell)$ consists only of tuples of all different elements, i.e., for any $i, j \in [\ell]$, $i \neq j$, $\Pr[X_i = X_j] = 0$. So we introduce the notation

$$\mathcal{X}^{\times \ell}_{\text{diff}} := \{(x_1, \ldots, x_\ell) \in \mathcal{X}^{\times \ell} : \forall i, j \in [\ell], i \neq j \Rightarrow x_i \neq x_j\}$$

for the subset over which these random variables are defined, and say that they are *different*.

We write $H(X)$ for the (Shannon) entropy of X and $I(X;Y) := H(X) + H(Y) - H(XY)$ for the mutual information between X and Y. This notation extends in the usual way for conditional entropies, e.g., $H(X|Y)$, $I(X;Y|Z)$.

2.2 Symmetric-key model

To achieve information-theoretic security, we consider the symmetric-key model, in which the two honest parties wishing to communicate share a secret key $k \in \mathcal{K}$. No matter what notion of security is desired — whether it be secrecy, non-malleability, or authenticity — the protocol follows the same steps. To transmit a message m, the sender applies a function f_k to the message, obtaining $c = f_k(m)$, which we refer to as the ciphertext. This is transmitted on an insecure channel to the receiver, who applies the inverse function, $m = f_k^{-1}(c)$. Since decryption must always be possible (if the ciphertext was not tampered with during transmission), the functions $\{f_k\}_{k \in \mathcal{K}}$ must be injective. If c has been modified, then there might not be any corresponding message m, in which case the decryption results in $\perp$.

In the following we loosely refer to any such scheme as an encryption scheme, and to the corresponding operations as encryption and decryption, even when secrecy is not required.

Definition 2.1. A symmetric-key encryption scheme is defined by a set of keys $k \in \mathcal{K}$, a probability distributions $P_K(\cdot)$ over these keys and injective *encryption functions* $f_k : \mathcal{X} \to \mathcal{Y}$ associated with each key. The *decryption functions* are defined as

$$\begin{aligned} g_k : \mathcal{Y} &\to \mathcal{X} \cup \{\perp\} \\ c &\mapsto \begin{cases} f_k^{-1}(c) & \text{if this is well defined.} \\ \perp & \text{otherwise.} \end{cases} \end{aligned}$$

The two legitimate players wishing to securely communicate a message m must share the key $k \in \mathcal{K}$ with probability $P_K(k)$ at the beginning of the protocol. The sender creates the ciphertext $c = f_k(m)$ and transmits it on an insecure channel to the receiver, who applies the decryption function $\tilde{m} = g_k(\tilde{c})$ to whatever (possibly modified) ciphertext $\tilde{c}$ he receives.

In the following we usually describe the messages, ciphertexts and keys by random variables M, C and K respectively, with $C = f_K(M)$.

3 Information-theoretic security notions

In this section we define the three notions of security, secrecy, non-malleability, and authenticity, in Sections 3.1, 3.2, and 3.3 respectively. All these definitions already appear in the literature, except the definition of approximate non-malleability (Definition 3.4), which is slightly different from previous ones [12]. Definition 3.4 is however a straightforward generalization of perfect non-malleability (Definition 3.3, [10, 11]).

3.1 Secrecy

Since in the symmetric-key model described in Section 2.2 the ciphertext is sent on an insecure channel, an adversary can intercept it, and try to gain information about the message from it. So for a given message random variable M, an encryption scheme is considered to provide perfect secrecy if the adversary cannot learn anything about the message given the ciphertext, no matter how much time and computation power he has, that is, if

$$H(M|C) = H(M) \text{ or } I(M;C) = 0, \tag{1}$$

as already defined by Shannon [1] in the 40's.

When we design an encryption scheme, we do not want it to be secure for some random variable M_1 with distribution $P_{M_1}(\cdot)$, but insecure for some other random variable M_2 with distribution $P_{M_2}(\cdot)$. Ideally, the scheme should still be secure, no matter how the messages are distributed over the message space, as long as they are independent from the key. We therefore require that Eq. (1) be fulfilled for all distributions $P_M(\cdot)$ on $\mathcal{X}$ independent from the key, i.e., for all M such that $I(M;K) = 0$.

Eq. (1) is called *perfect* secrecy, since the adversary's information is zero. However, in most practical situation, it is sufficient to have *approximate* secrecy, in which the adversary's probability (over the choice of keys) of noticing a difference between the real situation and the ideal one in which the ciphertext is independent from the message, is bounded by some very small ε. We therefore do not require any more that the message and ciphertext be perfectly independent, but that they be ε-close to independent according to the variational distance.[7]

Definition 3.1. An encryption scheme is said to provide *ε-approximate secrecy* (AS) if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — we have

$$d(MC, M \cdot C) \leq \varepsilon, \tag{2}$$

where C is the resulting ciphertext random variable.

If $\varepsilon = 0$, Eq. (2) is equivalent to Eq. (1), and we say that the scheme provides *perfect secrecy* (PS).

This secrecy criterion is defined for encrypting one message. If the key is much larger than the message, the same encryption function and key could be used several times to encrypt different messages and still preserve secrecy. Since

[7] There exist several alternative ways to formulate approximate secrecy. We give a brief overview of these in Appendix A.1, and show that they are equivalent.

we only need a security definition for the secrecy of two messages in this work, we restrict the following definition to two messages. Generalizing it to any number of messages is however straightforward.

Definition 3.2. An encryption scheme is said to provide 2-*message ε-approximate secrecy* (AS^2) if for all pairs of *different* message random variables $M_1 M_2$ on $\mathcal{X}^{\times 2}_{\text{diff}}$ independent from the key — i.e., $I(M_1 M_2; K) = 0$ and $\Pr[M_1 = M_2] = 0$ — we have

$$d(M_1 M_2 C_1 C_2, M_1 M_2 \cdot C_1 C_2) \leq \varepsilon, \tag{3}$$

where C_1 and C_2 are the resulting ciphertext random variables, i.e., $C_i = f_K(M_i)$ for $i = 1, 2$.

If $\varepsilon = 0$, Eq. (3) is equivalent to

$$I(M_1 M_2; C_1 C_2) = 0,$$

and we say that the encryption scheme provides 2-*message perfect secrecy* (PS^2).

When the same key is used to encrypt two messages, and these messages are identical (respectively different), their ciphertexts are necessarily identical (respectively different) too, since the encryption scheme is deterministic and uses the same key each time. It is therefore impossible for $I(M_1 M_2; C_1 C_2) = 0$ for all random variables $M_1 M_2$ defined over $\mathcal{X}^{\times 2}$, since the adversary can always learn which messages are identical or different, hence the restriction to *different* messages defined on $\mathcal{X}^{\times 2}_{\text{diff}}$.

3.2 Non-malleability

As briefly explained in Section 1, an encryption scheme is said to be malleable if an adversary can perform a controlled modification of an encrypted message, that is, modify a ciphertext in such a way that the new message resulting from decrypting the modified ciphertext is meaningfully related to the original message. An encryption scheme is then non-malleable, if the adversary cannot perform such a controlled modification of the message.

Let the original message be given by a random variable M, and let C be the corresponding ciphertext when encrypted with the key K. An adversary trying to perform a controlled modification of the message replaces the ciphertext with another ciphertext $\tilde{C}$, which, after decryption, becomes the message $\tilde{M}$. For simplicity we assume for the moment that the message and ciphertext alphabets have the same size, since otherwise the ciphertext $\tilde{C}$ generated by the adversary might be invalid.

If the encryption scheme is malleable, the adversary can thus create an $\tilde{M}$ which is meaningfully related to M, that is, which satisfies some specific relation $\mathcal{R}(M, \tilde{M})$ with high probability. Thus, if we give M to this adversary (who already holds C and $\tilde{C}$) he will have some information about $\tilde{M}$ — he knows that it satisfies this relation $\mathcal{R}$ — more information than if he had not created $\tilde{C}$ and only held M and C. If on the other hand the scheme is non-malleable, then he cannot create an $\tilde{M}$ to satisfy any relation $\mathcal{R}$. So given M, C and $\tilde{C}$, he does not know any more about $\tilde{M}$ than if he has only M and C.

Let us illustrate this with the one-time pad. The one-time pad is a malleable encryption scheme, because if an adversary flips some bits of the ciphertext, he also flips the same bits of the message, and can thus decide how to modify the message even without knowing what this message is. So if after flipping some bits of the ciphertext C to create $\tilde{C}$, the adversary is then given the message M, he can reconstruct $\tilde{M}$ by flipping the same bits of M. An observer who does not know how the adversary created $\tilde{C}$ would only learn from M that $\tilde{M}$ is different, but no more. So an adversary who holds $MC\tilde{C}$ would know more about $\tilde{M}$ than an observer who only holds MC, but does not know how $\tilde{C}$ was created, i.e.,

$$H(\tilde{M}|MC\tilde{C}) < H(\tilde{M}|MC).$$

On the other hand, if the encryption scheme is non-malleable, then as described above, the adversary does not know more about $\tilde{M}$ than had he not created $\tilde{C}$, so

$$H(\tilde{M}|MC\tilde{C}) = H(\tilde{M}|MC). \tag{4}$$

Note that this is equivalent to $I(\tilde{M};\tilde{C}|MC) = 0$. Criterion (4) was first proposed by Hanaoka et al. [10] (see also [11]) to define information-theoretic non-malleability. Following [12], we generalize their definition to the case where the ciphertext alphabet can be larger than the message alphabet, by extending the message alphabet to $\bar{\mathcal{X}} := \mathcal{X} \cup \{\perp\}$, as described in Section 2.2.

Definition 3.3. An encryption scheme is said to provide *perfect non-malleability* (PNM), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C};K|MC) = 0$ — we have

$$I(\tilde{M};\tilde{C}|MC) = 0,$$

where $\tilde{M}$ is defined on $\mathcal{X} \cup \{\perp\}$ and takes the value $\tilde{M} = \perp$ whenever $\tilde{C}$ is invalid.

There are several important remarks to make about this definition. The first concerns the domains of M and $\tilde{C}$. M is chosen by the legitimate players, so we can require that they choose it independently from the key. $\tilde{C}$ is however chosen by the adversary, who might make it depend on whatever information he holds about the secret key, e.g., information leaked to him from the ciphertext C. What is more, the legitimate players can decide to make (part of) the message public, or the adversary might know it by some other means. The pair MC leaks (much) more information about the key, and hence we need to allow the adversary to make his choice of $\tilde{C}$ depend on this. Thus, in general we have $H(K|\tilde{C}) < H(K)$, or equivalently $I(\tilde{C};K) > 0$. But the adversary should not get any information about K from any other source than MC, i.e., $\tilde{C}$ should not depend on any other part of K than that leaked by MC. Expressed with entropies, this means that we must have $H(K|MC\tilde{C}) = H(K|MC)$, or equivalently $I(\tilde{C};K|MC) = H(K|MC) - H(K|MC\tilde{C}) = 0$, which is one of the conditions of Definition 3.3.

The second remark concerns the condition $\Pr[C = \tilde{C}] = 0$. The adversary can always choose whether to modify the ciphertext or not, and hence can always decide whether $\tilde{M}$ is equal to or different from M. Criterion (4) can thus never be satisfied for a general ciphertext $\tilde{C}$, since knowledge of the pair $C\tilde{C}$ will always tell us whether $M = \tilde{M}$ or $M \neq \tilde{M}$. But since this cannot be avoided, it is of no concern either. As the informal definition of non-malleability states, we are only interested in modifications of the original message, and hence restrict our attention to this case.

Thirdly, we consider it important to extend the message alphabet to include "$\perp$" and not simply declare the adversary to be unsuccessful if he produces an invalid ciphertext. This is because we do not want the adversary to have the ability to generate an invalid ciphertext given that the message has certain properties, but not for other messages. We refer to Section 7 for a more detailed discussion of this.

As in Section 3.1, we are interested in generalizing the security notion to hold only with high probability over the choice of keys. Instead of requiring $\tilde{M}$ and $\tilde{C}$ to be perfectly independent given M and C, we require them to be ε-close to independent.

Definition 3.4. An encryption scheme is said to provide *ε-approximate non-malleability* (ANM), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C};K|MC) = 0$ — we have

$$d\Big(\tilde{M}\tilde{C}, \tilde{M} \cdot \tilde{C}\Big|MC\Big) \leq \varepsilon,$$

where $\tilde{M}$ is defined on $\mathcal{X} \cup \{\perp\}$ and takes the value $\tilde{m} = \perp$ whenever $\tilde{C}$ is invalid.

It is immediate from this definition that ANM with $\varepsilon = 0$ is equivalent to PNM.

Definition 3.4 can be seen as average case security over the values of C. Let $M = m$ be fixed, and suppose the scheme is insecure when $C = c$. If this occurs only with negligible probability (i.e., the keys mapping m to c are only chosen with negligible probability), then the scheme can still be secure according to Definition 3.4. We can strengthen this definition by considering the worst case over C.[8]

Definition 3.5. An encryption scheme is said to provide *ε-approximate strong non-malleability* (ASNM), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C};K|MC) = 0$ — we have for all $(m, c, \tilde{c}) \in \mathscr{D}(MC\tilde{C})$,

$$d\Big(\tilde{M}|_{MC\tilde{C}=mc\tilde{c}}, \tilde{M}|_{MC=mc}\Big) \leq \varepsilon.$$

Since Definition 3.4 can be rewritten as

$$\frac{1}{2}\sum_{m,c,\tilde{m},\tilde{c}} P_{MC\tilde{C}}(m,c,\tilde{c})\Big|P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m,c,\tilde{c}) - P_{\tilde{M}|MC}(\tilde{m}|m,c)\Big| \leq \varepsilon$$

and Definition 3.5 is equivalent to

$$\frac{1}{2}\sum_{\tilde{m}\in\tilde{\mathcal{X}}} \Big|P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m,c,\tilde{c}) - P_{\tilde{M}|MC}(\tilde{m}|m,c)\Big| \leq \varepsilon,$$

[8] Definition 3.4 already requires the scheme to be secure for all choices of M and $\tilde{C}$, which is equivalent to taking the worst case over M and $\tilde{C}$ (see Appendix A.1 for a proof of this in the case of secrecy).

we clearly have that any ASNM scheme is also ANM.

We refer to Appendix A.2 for a discussion of an alternative stronger approximate non-malleability definition.

3.3 Authentication

In an authentication protocol, the goal is not to provide any form of secrecy, but to be sure that the message has not been tampered with, i.e., that it really comes from the legitimate party. Since no secrecy is needed, authentication schemes usually append some MAC to the message, which is sent in clear, i.e., $f_k(m) = (m, h_k(m))$, where h_k is some hash function. Upon reception of $\tilde{c} = (\tilde{m}, \tilde{s})$, the party sharing the secret key k and wishing to authenticate the message simply checks if $\tilde{s} = h_k(\tilde{m})$, thus

$$g_k(\tilde{c}) = \begin{cases} \tilde{m} & \text{if } \tilde{s} = h_k(\tilde{m}) \\ \perp & \text{otherwise.} \end{cases}$$

In terms of random variables, the adversary who intercepts the ciphertext to replace it with his own obtains C. But even if C does not contain a clear copy of M, just like for non-malleability we have to assume that (part of) the message might be public, or that the adversary knows it by some other means. Hence when he creates the ciphertext $\tilde{C}$ he can make it depend on the part of the key leaked by MC, but not on any other part of K, i.e., $H(K|MC\tilde{C}) = H(K|MC)$, or equivalently $I(\tilde{C}; K|MC) = H(K|MC) - H(K|MC\tilde{C}) = 0$. The authentication scheme is successful if $\tilde{M} = \perp$ whenever the adversary modifies C.

Definition 3.6. An encryption scheme is said to provide *ε-authenticity* (Auth), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M; K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C}; K|MC) = 0$ — we have

$$P_{\tilde{M}}(\perp) \geq 1 - \varepsilon.$$

Unlike secrecy, authenticity can only be defined with high probability over the choice of keys, since it is always possible that an adversary might be lucky and choose a valid ciphertext. Definition 3.6 can however still be strengthened a little, since it corresponds to average case security over C. Similar to what we did for approximate non-malleability in Section 3.2, we can define strong authenticity.[9]

Definition 3.7. An encryption scheme is said to provide *ε-strong authenticity* (SAuth), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M; K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C}; K|MC) = 0$ — we have for all $(m, c, \tilde{c}) \in \mathscr{D}(MC\tilde{C})$,

$$P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) \geq 1 - \varepsilon.$$

[9] Just like for non-malleability, Definition 3.6 is already worst case over M and $\tilde{C}$, because the definition must hold for all M and $\tilde{C}$ (see Appendix A.1 for a proof of this in the case of secrecy).

Any SAuth scheme also provides Auth, since for any $\tilde{m}$,

$$\sum_{m,c,\tilde{c}} P_{MC\tilde{C}}(m,c,\tilde{c}) P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m,c,\tilde{c}) = P_{\tilde{M}}(\tilde{m}).$$

We note that these two definitions could equivalently have been written with the variational distance notation. Abusing slightly notation, we write $\perp$ for the random variable on $\mathcal{X} \cup \{\perp\}$ which takes value $\perp$ with probability 1. Then Definitions 3.6 and 3.7 are equivalent to $d(\tilde{M}, \perp) \leq \varepsilon$ and $d(\tilde{M}|_{MC\tilde{C}=mc\tilde{c}}, \perp) \leq \varepsilon$ respectively.

The notation used in these two definitions — in particular the use of random variables — is not quite standard. We use it for compatibility with the definitions of non-malleability. These authenticity definitions are however identical to what is found in textbooks, e.g., [16]. We additionally give a proof in Appendix B.2, Lemma B.2, that ε-approximate 2-strong universal hashing forms an ε-authentication scheme according to both these definitions.

4 Non-malleability and 2-message secrecy

The main result of this section, stated here under as Theorem 4.1 is that information-theoretic perfect non-malleability (PNM) is equivalent to uniformly mapping any pair of different messages to all possible pairs of different ciphertexts. As noted in Corollary 4.3, this means PNM is equivalent to 2-message perfect secrecy (PS^2) if the message and ciphertext alphabets have the same size, and strictly stronger than PS^2 if the ciphertext alphabet is larger. This immediately gives a lower bound on the necessary key size for PNM, and an easy way to design and prove the secrecy of these schemes.

Theorem 4.1. *Let $\{f_k : \mathcal{X} \to \mathcal{Y}\}_{k\in\mathcal{K}}$ be a set of encryption functions with key given by K and $|\mathcal{Y}| > 2$. The corresponding encryption scheme provides perfect non-malleability (PNM) if and only if for any two different random variables M_1 and M_2 with domain $\mathscr{D}(M_1M_2) \subseteq \mathcal{X}^{\times 2}_{\text{diff}}$ and independent from the key — i.e., $\Pr[M_1 = M_2] = 0$ and $I(M_1M_2;K) = 0$ — and any values $(m_1, m_2) \in \mathscr{D}(M_1M_2)$ and $(c_1, c_2) \in \mathcal{Y}^{\times 2}_{\text{diff}}$,*

$$P_{C_1C_2|M_1M_2}(c_1, c_2|m_1, m_2) = \frac{1}{|\mathcal{Y}^{\times 2}_{\text{diff}}|}. \tag{5}$$

Note that this theorem immediately implies the equivalence between PNM and PS^2 if $|\mathcal{X}| = |\mathcal{Y}|$.

Eq. (5) and PNM both define some probability distribution on a key and two pairs of messages and ciphertexts. They however both apply to different subsets of all possible tuples. For example, Eq. (5) makes a statement about the distribution of two ciphertext random variables C_1C_2, given that the messages M_1M_2 are independent from the key K. It does not describe what the probability distribution of C_1C_2 should be if M_1M_2 and K are correlated. PNM on the other hand does not require $M\tilde{M}$ and K to be independent, and it is not hard to find examples where the two are indeed correlated.

To prove Theorem 4.1, we first consider a simpler setting in which we remove M_1C_1 and MC from their respective equations. This results in the following lemma.

Lemma 4.2. *Let $\{f_k : \mathcal{X} \to \mathcal{Y}\}_{k\in\mathcal{K}}$ be the encryption functions from a symmetric-key encryption scheme. For any random variable M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and any $m \in \mathcal{X}$ and $c \in \mathcal{Y}$, we have*

$$P_{C|M}(c|m) = \frac{1}{|\mathcal{Y}|} \tag{6}$$

if and only if for any random variable $\tilde{C}$ on the ciphertext alphabet $\mathcal{Y}$ independent from the key — i.e., $I(\tilde{C};K) = 0$ — we have

$$I(\tilde{M};\tilde{C}) = 0, \tag{7}$$

where $\tilde{M} = g_K(\tilde{C})$ and g_k is the decryption function corresponding to f_k with range $\mathcal{X} \cup \{\perp\}$.

We provide a proof of Lemma 4.2 in Appendix C as Lemma C.1.

Removing M_1C_1 and MC from their respective equations can be interpreted as these random variables taking constant values. Lemma 4.2 can thus be seen as showing that Theorem 4.1 holds for fixed values of MC and M_1C_1. To finish the proof, we need to show that it holds for arbitrary MC and M_1C_1 under the given restrictions.

Proof of Theorem 4.1. We start with the "if direction" (Eq. (5) $\implies$ PNM). Note that for any random variables X, Y and Z, $I(X;Y|Z) = 0$ if and only if for all $z \in \mathscr{D}(Z)$, $I(X;Y|Z = z) = 0$. So to prove PNM, it is sufficient to fix m_1 and c_1 arbitrarily, and show that $I(\tilde{M};\tilde{C}|MC = m_1c_1) = 0$.

We define new random variables M', C' and K' on alphabets $\mathcal{X}' := \mathcal{X} \setminus \{m_1\}$, $\mathcal{Y}' := \mathcal{Y} \setminus \{c_1\}$ and $\mathcal{K}' := \{k \in \mathcal{K} : f_k(m_1) = c_1\}$ with joint distribution

$$P_{M'C'K'}(m,c,k) := P_{M_2C_2K|M_1C_1}(m,c,k|m_1,c_1). \tag{8}$$

It follows from Eq. (5) that $P_{C_2|M_1M_2C_1}(c_2|m_1,m_2,c_1) = \frac{1}{|\mathcal{Y}|-1}$. Hence from Eq. (8),

$$P_{C'|M'}(c|m) = \frac{P_{M_2C_2|M_1C_1}(m,c|m_1,c_1)}{P_{M_2|M_1C_1}(m|m_1,c_1)} = P_{C_2|M_1M_2C_1}(c|m_1,m,c_1) = \frac{1}{|\mathcal{Y}'|}.$$

All the conditions are gathered to apply Lemma 4.2, which tells us that for any $\tilde{C}'$ defined on $\mathcal{Y}'$ with $I(\tilde{C}';K') = 0$, $I(\tilde{M}';\tilde{C}') = 0$. By repeating this for different values of M_1 and C_1, we can extend $\tilde{C}'$ to any random variable $\tilde{C}$ such that $I(\tilde{C};K|M_1C_1) = 0$, but otherwise arbitrarily correlated to M_1 and C_1, and with $I(\tilde{M};\tilde{C}|M_1C_1) = 0$.

Now for the "only if direction" (PNM $\implies$ Eq. (5)). Let M_1 be any random variable with $\mathscr{D}(M_1) = \mathcal{X}$ and pick any values $(m_1,m_2) \in \mathcal{X}^{\times 2}_{\text{diff}}$ and $(c_1,c_2) \in \mathcal{Y}^{\times 2}_{\text{diff}}$. Since the scheme provides PNM we have $I(\tilde{M};\tilde{C}|M_1C_1)$ for any $\tilde{C}$ with $I(\tilde{C};K|M_1C_1) = 0$ and $\Pr[\tilde{C} = C_1] = 0$. Similar to what we did above, we define random variables $\tilde{M}', \tilde{C}', K'$ and $\tilde{M}'', \tilde{C}'', K''$ as

$$P_{\tilde{M}'\tilde{C}'K'}(m,c,k) := P_{\tilde{M}\tilde{C}K|M_1C_1}(m,c,k|m_1,c_1),$$
$$P_{\tilde{M}''\tilde{C}''K''}(m,c,k) := P_{\tilde{M}\tilde{C}K|M_1C_1}(m,c,k|m_2,c_2).$$

So $I(\tilde{M}';\tilde{C}') = I(\tilde{M}'';\tilde{C}'') = 0$. We can now apply Lemma 4.2 to the encryption functions of K' and K'' respectively, and get that for any M' and M'' on $\mathcal{X} \setminus \{m_1\}$ and $\mathcal{X} \setminus \{m_2\}$ and independent from K' and K'' respectively,

$$\begin{aligned} P_{C'|M'}(c'|m') &= \frac{1}{|\mathcal{Y}|-1}, \\ P_{C''|M''}(c''|m'') &= \frac{1}{|\mathcal{Y}|-1}. \end{aligned} \tag{9}$$

Let M_2 be any random variable such that $\mathscr{D}(M_1M_2) \subseteq \mathcal{X}^{\times 2}_{\text{diff}}$, $m_1, m_2 \in \mathscr{D}(M_2)$ and $I(M_1M_2;K) = 0$, and choose the M' and M'' from Eq. (9) such that

$$\begin{aligned} P_{M'C'}(m,c) &= P_{M_2C_2|M_1C_1}(m,c|m_1,c_1) \\ \text{and } P_{M''C''}(m,c) &= P_{M_2C_2|M_1C_1}(m,c|m_2,c_2). \end{aligned}$$

We then have

$$\begin{aligned} P_{C_1C_2|M_1M_2}(c_1,c_2|m_1,m_2) &= P_{C_1|M_1M_2}(c_1|m_1,m_2)P_{C_2|M_1M_2C_1}(c_2|m_1,m_2,c_1) \\ &= P_{C_1|M_1}(c_1|m_1)\frac{1}{|\mathcal{Y}|-1}, \\ P_{C_1C_2|M_1M_2}(c_2,c_1|m_2,m_1) &= P_{C_1|M_1}(c_2|m_2)\frac{1}{|\mathcal{Y}|-1}. \end{aligned} \tag{10}$$

Since the same encryption function with the same key is applied to m_1 and m_2, we must have $P_{C_1C_2|M_1M_2}(c_1,c_2|m_1,m_2) = P_{C_1C_2|M_1M_2}(c_2,c_1|m_2,m_1)$, and hence for all $(m_1,m_2) \in \mathcal{X}^{\times 2}_{\text{diff}}$ and $(c_1,c_2) \in \mathcal{Y}^{\times 2}_{\text{diff}}$,

$$P_{C_1|M_1}(c_1|m_1) = P_{C_1|M_1}(c_2|m_2).$$

Since $|\mathcal{Y}| > 2$, this implies that for any $(m_1,m_2) \in \mathcal{X}^{\times 2}_{\text{diff}}$ and $(c_1,c_2,c_3) \in \mathcal{Y}^{\times 3}_{\text{diff}}$,

$$P_{C_1|M_1}(c_1|m_1) = P_{C_1|M_1}(c_3|m_2) = P_{C_1|M_1}(c_2|m_1).$$

Since $\sum_c P_{C_1|M_1}(c|m) = 1$, we get $P_{C_1|M_1}(c|m) = \frac{1}{|\mathcal{Y}|}$. Putting this in Eq. (10) proves the theorem. □

Theorem 4.1 equates PNM with a uniform mapping from pairs of different messages to different ciphertexts (Eq. (5)). This latter condition is slightly different from PS^2. Corollary 4.3 makes the correspondence between PNM and PS^2 explicit.

Corollary 4.3. *For any symmetric-key encryption scheme with ciphertext alphabet size $|\mathcal{Y}| > 2$,*[10]

$$\begin{aligned} PNM_{|\mathcal{X}|=|\mathcal{Y}|} &\Leftrightarrow PS^2_{|\mathcal{X}|=|\mathcal{Y}|}, \\ PNM_{|\mathcal{X}|<|\mathcal{Y}|} &\Rightarrow PS^2_{|\mathcal{X}|<|\mathcal{Y}|}, \\ PNM_{|\mathcal{X}|<|\mathcal{Y}|} &\not\Leftarrow PS^2_{|\mathcal{X}|<|\mathcal{Y}|}. \end{aligned}$$

[10] By PNM$_{|\mathcal{X}|=|\mathcal{Y}|}$, PNM$_{|\mathcal{X}|<|\mathcal{Y}|}$, etc., we simply mean encryption functions with message and ciphertext alphabet sizes corresponding to the subscript and meeting the corresponding security definitions. We did not formally introduce this notation, because it is quite intuitive and is not used anywhere else. All other results about PNM, PS^2, etc., apply to all message and ciphertext alphabet sizes if not clearly stated otherwise.

Proof. Eq. (5) is clearly a sufficient condition to imply PS^2, no matter what the ciphertext length is. So from Theorem 4.1 we immediately have

$$\mathrm{PNM}_{|\mathcal{X}|=|\mathcal{Y}|} \Rightarrow \mathrm{PS}^2_{|\mathcal{X}|=|\mathcal{Y}|},$$
$$\mathrm{PNM}_{|\mathcal{X}|<|\mathcal{Y}|} \Rightarrow \mathrm{PS}^2_{|\mathcal{X}|<|\mathcal{Y}|}.$$

If $|\mathcal{X}| = |\mathcal{Y}|$, then for any scheme providing PS^2, and random variables M_1M_2 uniformly distributed on $\mathcal{X}^{\times 2}_{\mathrm{diff}}$,

$$H(C_1C_2|M_1M_2) = H(C_1C_2) \geq \log |\mathcal{X}^{\times 2}_{\mathrm{diff}}| = \log |\mathcal{Y}^{\times 2}_{\mathrm{diff}}|.$$

The inequality above holds because the entropy of the ciphertexts must be at least as large as the entropy of the messages. Thus Eq. (5) holds as well, which means that

$$\mathrm{PNM}_{|\mathcal{X}|=|\mathcal{Y}|} \Leftarrow \mathrm{PS}^2_{|\mathcal{X}|=|\mathcal{Y}|}.$$

Finally, to show that

$$\mathrm{PNM}_{|\mathcal{X}|<|\mathcal{Y}|} \not\Leftarrow \mathrm{PS}^2_{|\mathcal{X}|<|\mathcal{Y}|},$$

we give an example in Lemma 6.1 of an encryption scheme with $|\mathcal{X}| < |\mathcal{Y}|$ and providing PS^2, but not satisfying Eq. (5). □

We note that the requirement that $|\mathcal{Y}| > 2$ is essential, since otherwise PNM does not even imply PS. This can easily be seen by considering the following example. Let $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and the encryption function be the identity function. For such a small alphabet $H(\tilde{M}|MC) = H(\tilde{M}|MC\tilde{C}) = 0$, because as $\Pr[\tilde{M} = M] = 0$, once $M = m$ is fixed, $\tilde{M}$ can only take the other value, and hence has zero entropy. This scheme thus provides PNM, because the ciphertext $\tilde{C}$ chosen by the adversary provides no information about $\tilde{M}$.

An important consequence of Theorem 4.1 is that we get an immediate lower bound on the size of the secret key needed for PNM for any ciphertext size.

Corollary 4.4. *If an encryption scheme with key K provides PNM, then*

$$H(K) \geq \log |\mathcal{Y}^{\times 2}_{\mathrm{diff}}| = \log |\mathcal{Y}|(|\mathcal{Y}| - 1).$$

This immediately implies that the perfect non-malleable scheme proposed by Hanaoka et al. [10, 11] is optimal in the key size.[11] We describe this scheme in Appendix B.1 for completeness.

5 Non-malleability and authentication

We first show in Theorem 5.1 that any authentication scheme provides approximate non-malleability. Then in Theorem 5.2 we show that the same holds when we replace the notions of authenticity and non-malleability with strong authenticity (Definition 3.7) and strong approximate non-malleability (Definition 3.5).

Theorem 5.1. *Any scheme which provides ε-authenticity also provides $(\sqrt{\varepsilon}+\varepsilon)$-non-malleability.*

[11] This scheme is also optimal in the ciphertext size, since $|\mathcal{X}| = |\mathcal{Y}|$.

Proof. For all $(m, c) \in \mathscr{D}(MC)$, let $\varepsilon_{m,c} := 1 - P_{\tilde{M}|MC}(\perp|m, c)$. So we have $\sum_{m,c} P_{MC}(m, c)\varepsilon_{m,c} \leq \varepsilon$. Note that

$$\begin{aligned} P_{\tilde{M}|MC}(\perp|m, c) &= \frac{\sum_{\tilde{c}} P_{\tilde{M}\tilde{C}MC}(\perp, \tilde{c}, m, c)}{P_{MC}(m, c)} \\ &= \sum_{\tilde{c}} P_{\tilde{C}|MC}(\tilde{c}|m, c) P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}). \end{aligned}$$

From Lemma C.2 we then have that

$$\frac{1}{2} \sum_{\tilde{c}} P_{\tilde{C}|MC}(\tilde{c}|m, c) \Big| P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) - P_{\tilde{M}|MC}(\perp|m, c) \Big| \leq \sqrt{\varepsilon_{m,c}}.$$

Using Jensen's inequality we get

$$\frac{1}{2} \sum_{m,c,\tilde{c}} P_{MC\tilde{C}}(m, c, \tilde{c}) \Big| P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) - P_{\tilde{M}|MC}(\perp|m, c) \Big| \leq \sqrt{\varepsilon}.$$

Putting this in the definition of non-malleability we finally obtain

$$\begin{aligned} &\frac{1}{2} \sum_{m,c,\tilde{m},\tilde{c}} P_{MC\tilde{C}}(m, c, \tilde{c}) \Big| P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m, c, \tilde{c}) - P_{\tilde{M}|MC}(\tilde{m}|m, c) \Big| \\ &\quad \leq \frac{1}{2} \sum_{m,c,\tilde{c}} P_{MC\tilde{C}}(m, c, \tilde{c}) \Big| P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) - P_{\tilde{M}|MC}(\perp|m, c) \Big| \\ &\qquad + \sum_{\tilde{m} \in \mathcal{X}} P_{\tilde{M}}(\tilde{m}) \\ &\quad \leq \sqrt{\varepsilon} + \varepsilon. \end{aligned}$$

□

Roughly the same statement holds if we replace non-malleability and authenticity with strong non-malleability and strong authenticity.

Theorem 5.2. *Any scheme which provides ε-strong authenticity also provides $3\varepsilon/2$-strong non-malleability.*

Proof. A scheme provides SAuth with error ε if for all $(m, c, \tilde{c}) \in \mathscr{D}(MC\tilde{C})$,

$$P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) \geq 1 - \varepsilon.$$

$$\begin{aligned} \text{So } P_{\tilde{M}|MC}(\perp|m, c) &= \frac{\sum_{\tilde{c}} P_{\tilde{M}\tilde{C}MC}(\perp, \tilde{c}, m, c)}{P_{MC}(m, c)} \\ &= \sum_{\tilde{c}} P_{\tilde{C}|MC}(\tilde{c}|m, c) P_{\tilde{M}|MC\tilde{C}}(\perp|m, c, \tilde{c}) \\ &\geq 1 - \varepsilon. \end{aligned}$$

This also implies that

$$\begin{aligned} \sum_{\tilde{m} \in \mathcal{X}} P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m, c, \tilde{c}) &\leq \varepsilon \\ \text{and } \sum_{\tilde{m} \in \mathcal{X}} P_{\tilde{M}|MC}(\tilde{m}|m, c) &\leq \varepsilon. \end{aligned}$$

Putting this together we get

$$\begin{aligned}
\frac{1}{2}\sum_{\tilde{m}\in\bar{\mathcal{X}}}&\Big|P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m,c,\tilde{c})-P_{\tilde{M}|MC}(\tilde{m}|m,c)\Big| \\
&\leq \frac{1}{2}\Big|P_{\tilde{M}|MC\tilde{C}}(\perp|m,c,\tilde{c})-P_{\tilde{M}|MC}(\perp|m,c)\Big| \\
&\qquad\qquad + \frac{1}{2}\sum_{\tilde{m}\in\mathcal{X}}P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m,c,\tilde{c})+\frac{1}{2}\sum_{\tilde{m}\in\mathcal{X}}P_{\tilde{M}|MC}(\tilde{m}|m,c) \\
&\leq \frac{3\varepsilon}{2} \qquad\qquad \square
\end{aligned}$$

ε-(strong) authentication can be achieved with a shared key of length

$$\log|\mathcal{K}| \leq 2\log\log|\mathcal{X}| + 3\log\frac{1}{\varepsilon} \tag{11}$$

by using approximate strong 2-universal hashing. For completeness we show this in Appendix B.2. The parameters of Eq. (11) are from a specific family of approximate strong 2-universal hash functions by Bierbrauer et al. [14]. We refer to an expository paper on 2-universal hashing by Stinson [17] for an overview of constructions.

We note that since (approximate) secrecy is only possible if the key is as long as the message, this means that ANM does not imply secrecy. This might seem surprising at first, because in the public-key setting non-malleability does imply secrecy [7]. This difference between non-malleability and secrecy in the symmetric-key setting has however already been noted by Katz and Yung [9].[12]

6 Completing the picture

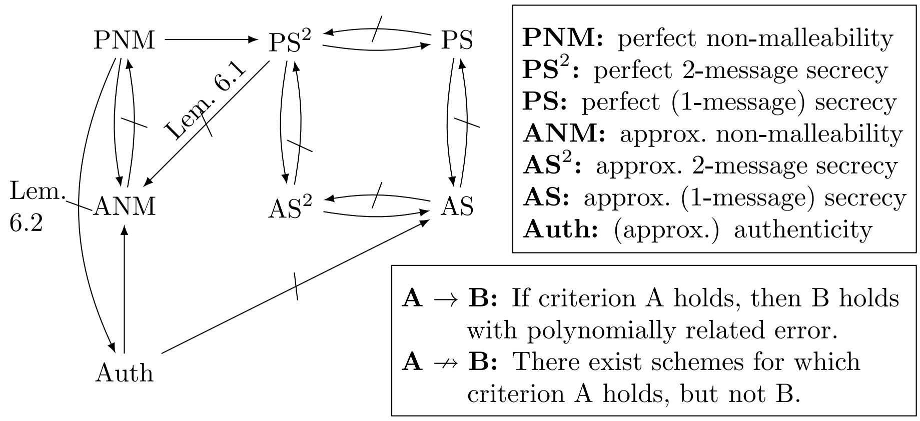
As described in Section 1, the relations depicted in Figure 1 are complete: if there is no directed path between two security definitions, then there exists an encryption scheme satisfying the first definition but not the second. Unlike the more detailed treatment given to non-malleability in the previous sections, we do not make a distinction here between the cases in which the message and ciphertexts have the same cardinality or not. We refer to Remark 6.3 for a further comment on this.

The complete list of separations amongst these notions is quite long. We fortunately only need to prove this property for a subset of all pairs of security definitions for which it holds, and the others follow immediately. We depict this subset of separations in Figure 2 by crossed arrows.

Since the separations amongst authentication and the different notions of secrecy are all previously known, we simply discuss and provide proofs for the separations between non-malleability and secrecy, and non-malleability and authenticity here. We use the following shorthand:

$$\{\mathrm{A}_i\}_{i=1}^{m} \not\Rightarrow \{\mathrm{B}_j\}_{j=1}^{n},$$

[12] In [9], the adversary is declared unsuccessful if the message produced is invalid, in which case it is trivial that authenticity is sufficient to achieve approximate non-malleability. We refer to Section 7 for a further discussion of how to handle invalid messages.



PNM: perfect non-malleability
PS^2: perfect 2-message secrecy
PS: perfect (1-message) secrecy
ANM: approx. non-malleability
AS^2: approx. 2-message secrecy
AS: approx. (1-message) secrecy
Auth: (approx.) authenticity

A $\rightarrow$ B: If criterion A holds, then B holds with polynomially related error.
A $\nrightarrow$ B: There exist schemes for which criterion A holds, but not B.

Figure 2 – *Complete characterization of the relations between different notions of information-theoretic non-malleability, secrecy, and authenticity with explicit separations.* This is an extension of Figure 1 in which some separations — relations for which there exist examples of schemes satisfying the first security definition but not the second — are explicitly depicted by crossed arrows. All other separations — all pairs of security definitions not connected by a directed path — can immediately be deduced from the ones that are depicted.

means that for any pair $(\mathrm{A}_i, \mathrm{B}_j)$, $\mathrm{A}_i \nRightarrow \mathrm{B}_j$, i.e., there are examples of schemes satisfying A_i but not B_j.

The complete list of separations thus reads

$$\mathrm{ANM}, \mathrm{PNM} \nRightarrow \mathrm{Auth}, \tag{12}$$

$$\mathrm{ANM} \nRightarrow \mathrm{AS}, \mathrm{AS}^2, \mathrm{PS}, \mathrm{PS}^2, \tag{13}$$

$$\mathrm{Auth} \nRightarrow \mathrm{PNM}, \tag{14}$$

$$\mathrm{AS}, \mathrm{AS}^2, \mathrm{PS}, \mathrm{PS}^2 \nRightarrow \mathrm{ANM}, \mathrm{PNM}, \tag{15}$$

Since PNM $\Rightarrow$ ANM, a scheme which satisfies PNM but not Auth, also satisfies ANM but not Auth. Thus to prove Eq. (12), it is sufficient to show that PNM $\nRightarrow$ Auth, which we do in Lemma 6.2.

Similarly, to prove Eq. (13) it is sufficient to find a scheme which satisfies ANM but not AS. Since AS^2, PS, and PS^2 are all stronger than AS, this example will not satisfy them either. Auth is stronger than ANM as shown in Theorem 5.1. So a scheme which provides authenticity but not AS is also sufficient to prove Eq. (13). This result is widely known since Wegman and Carter [2] showed that authentication is possible with a key of length roughly $\log \log |\mathcal{X}|$. Eq. (13) thus follows immediately from this and Theorem 5.1, and we do not need to provide a proof here.

Likewise, Eq. (14) does not require a proof either, since from Theorem 4.1 PNM is stronger than AS, the same example which shows the separation between Auth and AS also shows the separation between Auth and PNM.

And finally to prove Eq. (15), we need to show that PS^2 $\nRightarrow$ ANM, from which the other relations follow. We do this in Lemma 6.1.

Lemma 6.1. *There exists an encryption scheme with functions $\{f_k : \mathcal{X} \rightarrow \mathcal{Y}\}_{k \in \mathcal{K}}$ which provides 2-message perfect security (PS^2), but not ε-approximate non-malleability (ANM) for any $\varepsilon \in o(1)$.*

Proof. Without loss of generality, let $\mathcal{X} = \{0,1\}^n = \mathrm{GF}(2^n)$ and $\mathcal{Y} = \{0,\dots,2^n\}$. We choose $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \{0,1\}$, where $\mathcal{K}_1 = \mathrm{GF}(2^n) \setminus \{0\}$ and $\mathcal{K}_2 = \mathrm{GF}(2^n)$, and with $(k_1,k_2) \in \mathcal{K}_1 \times \mathcal{K}_2$ chosen with uniform probability. The last bit of the key is 0 with some constant probability $\alpha < \frac{1}{2}$. We define functions $h_0, h_1 : \mathcal{X} \to \mathcal{Y}$ which both map any element from $\{0,1\}^n \setminus \{0^n\}$ to $\{1,\dots,2^n-1\}$ in the obvious way, but $h_0(0^n) = 0$ and $h_1(0^n) = 2^n$. The encryption function is then

$$f_{(k_1,k_2,b)}(m) = h_b(k_1 m + k_2).$$

This function clearly provides PS^2, but does not uniformly map $\mathcal{X}^{\times 2}_{\text{diff}}$ to $\mathcal{Y}^{\times 2}_{\text{diff}}$, so by Theorem 4.1 it does not provide PNM. To see that it does not provide ANM either, note that an adversary can always choose to replace the legitimate cipher by some $\tilde{c} \in \{1,\dots,2^n-1\}$ if he wants the resulting message to be $\perp$ with probability 0 and $\tilde{c} = 0$ for the message to be $\perp$ with probability $1-\alpha$. If the adversary chooses between these two options with probability $\frac{1}{2}$ if $c \neq 0$ and chooses $\tilde{c} = 2^n$ if $c = 0$, then $d\Big(\tilde{M}\tilde{C}, \tilde{M} \cdot \tilde{C} \Big| MC\Big) \geq \big(1 - \frac{\alpha}{2^n}\big)\frac{1-\alpha}{2}$. □

We note that Lemma 6.1 does not hold if we consider an alternative definition of non-malleability, in which the adversary is not successful if he can control whether the received ciphertext is invalid or not. We refer to Section 7 for a further discussion of this.

This last separation is trivial.

Lemma 6.2. *There exists an encryption scheme with functions $\{f_k : \mathcal{X} \to \mathcal{Y}\}_{k \in \mathcal{K}}$ such that $|\mathcal{X}| = |\mathcal{Y}|$, and which provides perfect non-malleability (PNM), but not ε-authenticity for any $\varepsilon \in o(1)$.*

Proof. The PNM scheme by Hanaoka et al. [10, 11] (see also Appendix B.1) is such an example. □

Remark 6.3. If we were to make a distinction between the cases in which the message and ciphertexts have the same cardinality or not, the landscape of separations changes a bit.

The separation $\mathrm{PS}^2 \nRightarrow \mathrm{ANM}, \mathrm{PNM}$ only holds for $|\mathcal{Y}| > |\mathcal{X}|$. When the message and ciphertexts have the same cardinality, any scheme providing PS^2 also provides ANM, PNM. However $\mathrm{AS}, \mathrm{AS}^2, \mathrm{PS} \nRightarrow \mathrm{ANM}, \mathrm{PNM}$ is true even when $|\mathcal{X}| = |\mathcal{Y}|$.

The proof that $\mathrm{ANM} \nRightarrow \mathrm{AS}, \mathrm{AS}^2$ also only holds for $|\mathcal{Y}| > |\mathcal{X}|$. We do not know if it is still valid for $|\mathcal{Y}| = |\mathcal{X}|$.

7 Concluding remarks

In this work we studied information-theoretic non-malleability, extending a line of research initiated by Hanaoka et al. [10]. The formal definitions used to capture the intuitive notion of non-malleability follow these previous works [10–12]. There exist however alternative ways to characterize the same notion. We discuss them briefly in this section.

Unifying the definitions. Although the works on computational and information-theoretic non-malleability in the symmetric-key setting use the same informal definition, the tools used to formalize this definition are different: the former computes the probability that the falsified message is related to the original message in the real and ideal case [9], the latter measures the indistinguishability of message and ciphertext distributions between the real and ideal case. In the case of quantum messages, yet another definition has been proposed, in which a scheme provides non-malleability if the decrypted message is always an arbitrary convex combination of the original message and the result of decrypting a garbage state that the adversary sent [18]. It remains open to prove formally that these definitions are indeed equivalent when the messages are classical, and the distinguisher in the computational security definition is unlimited and does not access oracles.

Invalid ciphertexts. In the formal definition of non-malleability, we chose that the adversary is allowed to pick invalid ciphertexts and still be successful. We could have considered an alternative weaker definition, in which the adversary automatically fails when this happens. In the public-key setting, both ways of treating invalid ciphertexts can be found, and there is no clear consensus as to how to deal with this case. Pass et al. [8] investigate the differences between the two notions in detail. They point out how the stronger notion in which the adversary can produce an invalid ciphertext makes a critical difference in certain situations, in particular for composability.

In the case of information-theoretic security, if we had defined the adversary to be unsuccessful when he picks an invalid ciphertext, then perfect non-malleability would have been exactly equivalent to 2-message perfect secrecy, and not strictly strong for a ciphertext longer than the message. And authenticity would trivially imply approximate non-malleability, instead of requiring some work.

Accessing oracles. When considering computational security, the adversary usually has access at various stages to a decryption oracle.[13] In information-theoretic security, when the adversary is computationally unbounded, unlimited access to an oracle is not possible. McAven et al. [12] and Portmann and Tanaka [19] propose security definitions in which the adversary can make ℓ queries to an oracle. The definitions of non-malleability used in this work can be seen as allowing the adversary 1 query to an encryption oracle, after which he has to choose his forged ciphertext $\widetilde{C}$. By generalizing this to ℓ-queries to either encryption or decryption oracles, we can define various notions of ℓ-non-malleability.

We conjecture that the results from this work on the relations between 1-non-malleability, 2-message security, and the 2-universal hashing used for authentication, directly generalize to ℓ-non-malleability, $(\ell + 1)$-message security and $(\ell + 1)$-universal hashing.

[13] In the case of computational symmetric-key cryptography, he may also access an encryption oracle.

Acknowledgments

AK was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B), 18300002. CP was supported by the Swiss National Science Foundation through the National Centre of Competence in Research 'Quantum Science and Technology', the European Research Council – ERC (grant no. 258932). KT was supported in part by I-System Co. Ltd.

Appendices

A More on information-theoretic security notions

A.1 Approximate secrecy

Requiring a security criterion to hold for all random variables X is equivalent to worst case security over X, that is, equivalent to requiring the same condition to hold for every value $x \in \mathscr{D}(X)$. We illustrate this in Lemma A.1 by showing that the definition of approximate secrecy given in Definition 3.1 is equivalent to requiring that the ciphertext distributions for any two messages m and n be ε-close.

The same can be shown for any security criterion, in particular for ANM (Definition 3.4).

Lemma A.1. *Let $\{f_k : \mathcal{X} \to \mathcal{Y}\}_{k \in \mathcal{K}}$ be a set of injective functions indexed by keys distributed according to the random variable K, let M be a message random variable independent from the key — i.e., $I(M;K) = 0$ — and $C = f_K(M)$. Then the three following secrecy criteria are equivalent up to a multiplicative constant.*

$$\forall m, n \in \mathscr{D}(M),\ d(C|_{M=m}, C|_{M=n}) \leq \varepsilon_1, \tag{16}$$

$$\forall m \in \mathscr{D}(M),\ d(C|_{M=m}, C) \leq \varepsilon_2, \tag{17}$$

$$\forall M \text{ such that } I(M;K) = 0,\ d(MC, M \cdot C) \leq \varepsilon_3, \tag{18}$$

Proof. (16) $\implies$ (17) for $\varepsilon_2 = \varepsilon_1$ because

$$\begin{aligned} d(C|_{M=m}, C) &= \frac{1}{2} \sum_{c \in \mathcal{Y}} \left| P_{C|M}(c|m) - P_C(c) \right| \\ &= \frac{1}{2} \sum_{c \in \mathcal{Y}} \left| P_{C|M}(c|m) - \sum_{n \in \mathcal{X}} P_M(n) P_{C|M}(c|n) \right| \\ &\leq \frac{1}{2} \sum_{(n,c) \in \mathcal{X} \times \mathcal{Y}} P_M(n) \left| P_{C|M}(c|m) - P_{C|M}(c|n) \right| \\ &\leq \sum_{n \in \mathcal{X}} P_M(n) \varepsilon_1 = \varepsilon_1. \end{aligned}$$

Similarly, (17) $\implies$ (18) for $\varepsilon_3 = \varepsilon_2$ because

$$\begin{aligned} d(MC, M \cdot C) &= \frac{1}{2} \sum_{(m,c) \in \mathcal{X} \times \mathcal{Y}} |P_{MC}(m,c) - P_M(m) P_C(c)| \\ &= \frac{1}{2} \sum_{(m,c) \in \mathcal{X} \times \mathcal{Y}} P_M(m) \big| P_{C|M}(c|m) - P_C(c) \big| \\ &\leq \sum_{m \in \mathcal{X}} P_M(m) \varepsilon_2 = \varepsilon_2. \end{aligned}$$

To prove that (18) $\implies$ (16) for $\varepsilon_1 = 2\varepsilon_3$, we consider the message random variable M which takes the values m and n, each with probability $1/2$. We then have

$$\begin{aligned} 2\varepsilon_3 &\geq \sum_{(x,c) \in \mathcal{X} \times \mathcal{Y}} P_M(x) \big| P_{C|M}(c|x) - P_C(c) \big| \\ &= \frac{1}{2} \sum_{c \in \mathcal{Y}} \big| P_{C|M}(c|m) - P_C(c) \big| + \frac{1}{2} \sum_{c \in \mathcal{Y}} \big| P_C(c) - P_{C|M}(c|n) \big| \\ &\geq \frac{1}{2} \sum_{c \in \mathcal{Y}} \big| P_{C|M}(c|m) - P_{C|M}(c|n) \big|. \end{aligned} \tag{19}$$

Note that Eq. (19) is independent from the distribution of M. And since Eq. (18) must hold for all M, the same reasoning can be made for all pairs $m, n \in \mathcal{X}$, thus Eq. (16) holds for any M and all $m, n \in \mathscr{D}(M)$. □

Expressing AS in these different ways makes it easier to understand how this security criterion captures "security with high probability." We can construct a scheme where some bad key value k_0 maps any message m to $m||0$, where $||$ stands for the concatenation of two strings, but the other keys uniformly map m to all possible $c||1$. If this bad value k_0 is chosen with negligible probability (less than ε), then the scheme is secure even if some ciphertexts (all those of form $m||0$) are completely insecure, because these ciphertexts also occur with negligible probability.

Remark A.2. Similar to what we do for approximate non-malleability in the next section, the definition of approximate secrecy used so far can be strengthened. Instead of requiring that for all m, $\sum_c |P_{C|M}(c|m) - P_C(c)| \leq \varepsilon$, we require that for all m and all c,

$$|P_{C|M}(c|m) - P_C(c)| \leq \varepsilon P_C(c).$$

We note that the encryption scheme described a paragraph higher is not secure anymore according to this new criterion.

A.2 Approximate non-malleability

In Section 3.2 we introduced two definitions of approximate non-malleability (ANM in Definition 3.4 and ASNM in Definition 3.5), the latter requiring the security to hold for all m, $\tilde{c}$ and c, instead of simply all m and $\tilde{c}$. McAven et al. [12] take this strengthening a step further, and also require strict security no matter what value $\tilde{m}$ is produced.

Definition A.3. An encryption scheme is said to provide *ε-approximate very strong non-malleability* (AVSNM), if for all message random variables M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and all ciphertexts $\tilde{C}$ on $\mathcal{Y}$ different from C and independent from the key given MC — i.e., $\Pr[C = \tilde{C}] = 0$ and $I(\tilde{C};K|MC) = 0$ — we have for all $(m, c, \tilde{c}) \in \mathscr{D}(MC\tilde{C})$ and all $\tilde{m} \in \mathcal{X} \cup \{\perp\}$,

$$\left| P_{\tilde{M}|MC\tilde{C}}(\tilde{m}|m, c, \tilde{c}) - P_{\tilde{M}|MC}(\tilde{m}|m, c) \right| \leq \varepsilon P_{\tilde{M}|MC}(\tilde{m}|m, c).$$

It is again clear that AVSNM implies ASNM. With this definition it is not sufficient if insecure values $\tilde{m}$ occur with negligible probability $P_{\tilde{M}|MC}(\tilde{m}|m, c)$ (over the choice of key), since the error is measured relative to this. No choice of key k should ever result in an insecure $\tilde{m}$ for Definition A.3 to be met. This definition thus provides very strong security, but when only "security with high probability" is desired, Definition 3.4 or 3.5 is sufficient.

We note that the notion of AVSNM cannot be satisfied by an authentication scheme. The idea behind authentication is that with high probability the adversary cannot produce a valid ciphertext ($P_{\tilde{M}}(\perp) \geq 1 - \varepsilon$). If however he is lucky enough to guess right, then the new message created $\tilde{m} \in \mathcal{X}$ is completely determined by the ciphertext $\tilde{c}$. But this only happens with (negligible) probability ε. AVSNM requires that the message $\tilde{m}$ never be determined by the ciphertext $\tilde{c}$, even when this message occurs with negligible probability ($P_{\tilde{M}|MC}(\tilde{m}|mc) \leq \varepsilon$).

The difference between AVSNM and its weaker forms, ANM or ASNM, is the same as between the notion of approximate secrecy defined in Remark A.2 and the various equivalent forms of AS given in Lemma A.1.

B Practical schemes

In this section we give some constructions for perfect and approximate non-malleability. The construction for perfect non-malleability can be found in [11]. The approximate non-malleability scheme is simply authentication using 2-universal hashing. The construction we give is from [14].

B.1 Perfect non-malleability

The results from Section 4 show that the following PNM scheme proposed by Hanaoka [11] is optimal in the key size. Let the message and ciphertext alphabets be some finite field $\mathcal{X} = \mathcal{Y} = \mathrm{GF}(q)$, where q is a prime power. The key is chosen uniformly at random from $k = (k_1, k_2) \in (\mathrm{GF}(q) \setminus \{0\}) \times \mathrm{GF}(q)$, and the encryption functions are defined as

$$c = f_k(m) := k_1 m + k_2. \tag{20}$$

We immediately have from the properties of finite fields that these functions are easily invertible, $m = k_2^{-1}(c - k_2)$, and that any two distinct messages m_1 and m_2 are mapped to all possible distinct ciphertext pairs (c_1, c_2) when choosing different keys. So this scheme provides PS^2 and hence PNM with a key of optimal size $|\mathcal{K}| = q(q - 1)$.

B.2 Approximate non-malleability

As shown in Section 5, to achieve approximate non-malleability, we just need an authentication scheme. Wegman and Carter [2] first noticed how to drastically reduce the shared key size in authentication schemes using approximate strong 2-universal hashing, though Stinson [4] gave the first formal definition of this. We refer to [17] for an overview of 2-universal hashing.

Definition B.1. A set of hash functions $\{h_k : \mathcal{X} \to \mathcal{Z}\}_{k\in\mathcal{K}}$ is said to be ε-approximate strongly 2-universal (ASU$_2$) if for every $x_1, x_2 \in \mathcal{X}$ with $x_1 \neq x_2$ and every $z_1, z_2 \in \mathcal{Z}$,

1) $|\{k \in \mathcal{K} : h_k(x_1) = z_1\}| = \frac{|\mathcal{K}|}{|\mathcal{Z}|}$

2) $|\{k \in \mathcal{K} : h_k(x_1) = z_1, h_k(x_2) = z_2\}| \leq \frac{\varepsilon|\mathcal{K}|}{|\mathcal{Z}|}$.

To construct an authentication scheme using ASU$_2$ hashing, we simply choose each function uniformly at random and define the encryption function to be

$$f_k(m) = (m, h_k(m)). \tag{21}$$

As the following lemma shows, this provides ε-strong authenticity (Definition 3.7).

Lemma B.2. *An ASU$_2$ family of hash functions used as in Eq. (21) provides ε-strong authenticity.*

Proof. Let $m \in \mathcal{X}$ and $c \in \mathcal{Y}$ be the message and ciphertext pair obtained by the adversary, where $\mathcal{Y} = \mathcal{X} \times \mathcal{Z}$ and $c = (m, z) = (m, h_k(m))$. Let the adversary choose the forged ciphertext $\tilde{c} = (\tilde{m}, \tilde{z})$. This ciphertext is valid only if $\tilde{z} = h_k(\tilde{m})$. From Definition B.1 we have that for any choice of $(\tilde{m}, \tilde{z})$ with $\tilde{m} \neq m$,

$$\Pr[\tilde{z} = h_k(\tilde{m}) | z = h_k(m)] = \frac{\Pr[\tilde{z} = h_k(\tilde{m}) \text{ and } z = h_k(m)]}{\Pr[z = h_k(m)]} \leq \varepsilon.$$

It follows that

$$P_{\tilde{M}|MC\tilde{C}}(\perp | m, c, \tilde{c}) \geq 1 - \varepsilon. \qquad \square$$

By Theorem 5.2 this implies that a family of ASU$_2$ hash functions provides $3\varepsilon/2$-approximate strong non-malleability and $3\varepsilon/2$-approximate non-malleability. Bierbrauer et al. [14] show that such a family exists with size

$$\log|\mathcal{K}| \leq 2\log\log|\mathcal{X}| + 3\log\frac{1}{\varepsilon}.$$

C Technical lemmas

In this section we provide a few technical lemmas needed in the main body of this work. The following lemma shows that a ciphertext chosen by an adversary is independent from the corresponding message after decryption if and only if the encryption scheme maps every message to all ciphertext with equal probability.

Lemma C.1. *Let $\{f_k : \mathcal{X} \to \mathcal{Y}\}_{k\in\mathcal{K}}$ be the encryption functions from a symmetric-key encryption scheme. For any random variable M on $\mathcal{X}$ independent from the key — i.e., $I(M;K) = 0$ — and any $m \in \mathcal{X}$ and $c \in \mathcal{Y}$, we have*

$$P_{C|M}(c|m) = \frac{1}{|\mathcal{Y}|}$$

if and only if for any random variable $\tilde{C}$ on the ciphertext alphabet $\mathcal{Y}$ independent from the key — i.e., $I(\tilde{C};K) = 0$ — we have

$$I(\tilde{M};\tilde{C}) = 0,$$

where $\tilde{M} = g_K(\tilde{C})$ and g_k is the decryption function corresponding to f_k with range $\mathcal{X} \cup \{\perp\}$.

Proof. We start with the "if direction" ($I(\tilde{M};\tilde{C}) = 0 \implies P_{C|M}(c|m) = \frac{1}{|\mathcal{Y}|}$). If $I(\tilde{M};\tilde{C}) = 0$ then for any $m \in \bar{\mathcal{X}}$ and $c \in \mathcal{Y}$, $P_{\tilde{M}|\tilde{C}}(m|c) = P_{\tilde{M}}(m)$ and hence for any two $c, c' \in \mathcal{Y}$ and any $m \in \bar{\mathcal{X}}$,

$$P_{\tilde{M}|\tilde{C}}(m|c) = P_{\tilde{M}|\tilde{C}}(m|c'). \tag{22}$$

Since $I(\tilde{C};K) = 0$, for any $m \in \mathcal{X}$ and $c \in \mathcal{Y}$ we have

$$P_{\tilde{M}|\tilde{C}}(m|c) = \sum_{\substack{k\in\mathcal{K}\\ f_k^{-1}(c)=m}} P_K(k) = \sum_{\substack{k\in\mathcal{K}\\ f_k(m)=c}} P_K(k) = P_{C|M}(c|m) \tag{23}$$

for any M with $I(M;K) = 0$. Since the distribution $P_{C|M}$ is well defined, we have for any $m \in \mathcal{X}$ that

$$\sum_{c\in\mathcal{Y}} P_{C|M}(c|m) = 1. \tag{24}$$

Combining Eqs. (22), (23) and (24) we get that for any $m \in \mathcal{X}$ and any $c \in \mathcal{Y}$, $P_{C|M}(c|m) = \frac{1}{|\mathcal{Y}|}$.

The "only if direction" ($P_{C|M}(c|m) = \frac{1}{|\mathcal{Y}|} \implies I(\tilde{M};\tilde{C}) = 0$) works similarly. We have for any $m \in \mathcal{X}$ and $c \in \mathcal{Y}$ that

$$\frac{1}{|\mathcal{Y}|} = P_{C|M}(c|m) = \sum_{\substack{k\in\mathcal{K}\\ f_k(m)=c}} P_K(k) = \sum_{\substack{k\in\mathcal{K}\\ f_k^{-1}(c)=m}} P_K(k) = P_{\tilde{M}|\tilde{C}}(m|c)$$

for any $\tilde{C}$ independent from the key. Furthermore

$$P_{\tilde{M}|\tilde{C}}(\perp|c) = 1 - \sum_{m\in\mathcal{X}} P_{\tilde{M}|\tilde{C}}(m|c) = 1 - \frac{|\mathcal{X}|}{|\mathcal{Y}|}$$

for every $c \in \mathcal{Y}$. Hence $I(\tilde{M};\tilde{C}) = 0$. □

This last lemma is needed in the proof of Theorem 5.1.

Lemma C.2. *For $i \in [n]$, let $0 \le a_i \le 1$ and have weighted average $\sum_i w_i a_i = a$, where $0 \le w_i \le 1$ and $\sum_i w_i = 1$. Then*

$$\sum_{i=1}^{n} w_i |a_i - a| \le 2 \min\{\sqrt{a}, \sqrt{1-a}\}.$$

Proof. Without loss of generality, let $a \le \frac{1}{2}$. If $a > \frac{1}{2}$, set $a_i^{\text{new}} := 1 - a_i^{\text{old}}$ for all i, which leaves $|a_i - a|$ unchanged.

Define $\mathcal{I} := \{i \in [n] : a_i \ge \sqrt{a}\}$. Then

$$\sum_{i=1}^{n} w_i a_i \ge \sum_{i \in \mathcal{I}} w_i \sqrt{a},$$

hence $\sum_{i \in \mathcal{I}} w_i \le \sqrt{a}$. We then have

$$\begin{aligned}
\sum_{i=1}^{n} w_i |a_i - a| &= \sum_{i \in \mathcal{I}} w_i |a_i - a| + \sum_{i \in [n] \setminus \mathcal{I}} w_i |a_i - a| \\
&\le \sum_{i \in \mathcal{I}} w_i (1 - a) + \sum_{i \in [n] \setminus \mathcal{I}} w_i \max\{\sqrt{a} - a, a\} \\
&\le \sqrt{a}(1 - a) + (1 - \sqrt{a}) \max\{\sqrt{a} - a, a\} \\
&= 2\sqrt{a}.
\end{aligned}$$

□

References

[1] Claude Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, 1949.

[2] Mark N. Wegman and Larry Carter. New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences*, 22(3):265–279, 1981.

[3] Danny Dolev, Cynthia Dwork, and Moni Naor. Non-malleable cryptography. *SIAM Journal on Computing*, 30(2):391–437, 2000. A preliminary version appeared at STOC '91. [doi:10.1137/S0097539795291562].

[4] Douglas R. Stinson. Universal hashing and authentication codes. *Designs, Codes and Cryptography*, 4(3):369–380, 1994. A preliminary version appeared at CRYPTO '91. [doi:10.1007/BF01388651].

[5] Mihir Bellare, Anand Desai, David Pointcheval, and Phillip Rogaway. Relations among notions of security for public-key encryption schemes. In *Proceedings of the 18th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '98*, pages 26–45. Springer, 1998. [doi:10.1007/BFb0055718].

[6] Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Construction of a non-malleable encryption scheme from any semantically secure one. In *Proceedings of the 26th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '06*, pages 271–289. Springer, 2006. [doi:10.1007/11818175_16].

[7] Mihir Bellare and Amit Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In *Proceedings of the 19th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '99*, pages 519–536. Springer, 1999. [doi:10.1007/3-540-48405-1_33].

[8] Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Relations among notions of non-malleability for encryption. In *Proceedings of 13th International Conference on the Theory and Application of Cryptology and Information Security, ASIACRYPT '07*, pages 519–535. Springer, 2007. [doi:10.1007/978-3-540-76900-2_32].

[9] Jonathan Katz and Moti Yung. Characterization of security notions for probabilistic private-key encryption. *Journal of Cryptology*, 19(1):67–95, January 2006. [doi:10.1007/s00145-005-0310-8].

[10] Goichiro Hanaoka, Junji Shikata, Yumiko Hanaoka, and Hideki Imai. Unconditionally secure anonymous encryption and group authentication. *The Computer Journal*, 49(3):310–321, 2006. A preliminary version appeared at Asiacrypt '02. [doi:10.1093/comjnl/bxh149].

[11] Goichiro Hanaoka. Some information theoretic arguments for encryption: Non-malleability and chosen-ciphertext security (invited talk). In *Third International Conference on Information Theoretic Security, ICITS 2008*, pages 223–231. Springer, 2008. [doi:10.1007/978-3-540-85093-9_21].

[12] Luke McAven, Reihaneh Safavi-Naini, and Moti Yung. Unconditionally secure encryption under strong attacks. In *9th Australasian Conference on Information Security and Privacy, ACISP 2004*, pages 427–439. Springer, 2004. [doi:10.1007/b98755].

[13] Goichiro Hanaoka, Yumiko Hanaoka, Manabu Hagiwara, Hajime Watanabe, and Hideki Imai. Unconditionally secure chaffing-and-winnowing: A relationship between encryption and authentication. In *16th International Symposium on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, pages 154–162. Springer, 2006. [doi:10.1007/11617983_15].

[14] Jürgen Bierbrauer, Thomas Johansson, Gregory Kabatianskii, and Ben Smeets. On families of hash functions via geometric codes and concatenation. In *Proceedings of the 13th Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '93*, pages 331–342. Springer, 1994. [doi:10.1007/3-540-48329-2_28].

[15] Mihir Bellare and Amit Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. Cryptology ePrint Archive, Report 2006/228, 2006. Full version of [7]. [IACR e-print: 2006/228].

[16] Douglas R. Stinson. *Cryptography: Theory and Practice, Second Edition.* Chapman & Hall/CRC, 2002.

[17] Douglas R. Stinson. On the connections between universal hashing, combinatorial designs and error-correcting codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 2(52), 1995.

[18] Andris Ambainis, Jan Bouda, and Andreas Winter. Non-malleable encryption of quantum information. *Journal of Mathematical Physics*, 50(4):042106, 2009. [doi:10.1063/1.3094756, arXiv:0808.0353].

[19] Christopher Portmann and Keisuke Tanaka. Information-theoretic secrecy with access to decryption oracles. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E94-A(7):1585–1590, 2011.