# Secure Authentication from a Weak Key, Without Leaking Information

Niek J. Bouman and Serge Fehr

Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands
{n.j.bouman,serge.fehr}@cwi.nl

**Abstract.** We study the problem of authentication based on a weak key in the information-theoretic setting. A key is weak if its min-entropy is an arbitrary small fraction of its bit length. This problem has recently received considerable attention, with different solutions optimizing different parameters. We study the problem in an extended setting, where the weak key is a one-time *session key* that is derived from a public source of randomness with the help of a (potentially also weak) *long-term* key. Our goal now is to authenticate a message by means of the weak session key in such a way that (nearly) no information on the long-term key is leaked. Ensuring privacy of the long-term key is vital for the long-term key to be re-usable. Previous work has not considered such a privacy issue, and previous solutions do not seem to satisfy this requirement.
We show the existence of a practical four-round protocol that provides message authentication from a weak session key and that avoids non-negligible leakage on the long-term key. The security of our scheme also holds in the quantum setting where the adversary may have limited quantum side information on the weak session key. As an application of our scheme, we show the existence of an identification scheme in the bounded quantum storage model that is secure against a man-in-the-middle attack and that is truly password-based: it does not need any high entropy key, in contrast to the scheme proposed by Damgård *et al.*

## 1 Introduction

### 1.1 The Problem

We consider the problem of achieving authentic communication over a public channel that might be under the control of an active adversary. We study this problem in the information-theoretic setting, i.e. we assume the adversary to be computationally unbounded.

Specifically, we consider the following scenario. Alice and Bob share a *long-term* key $W$. When needed, Alice and Bob can extract a weak *session key* $X_W$ from an auxiliary source of randomness with the help of $W$. It should be guaranteed by the property of the auxiliary source that a potential adversary Eve who does not know $W$ has limited information on the weak session key $X_W$. This is formalized by requiring that $H_{\min}(X_W|WE) \geq k$ for some parameter $k$, where $E$ denotes Eve's side information. Examples of where this scenario occurs naturally are the bounded storage model, where $W$ determines which part of the huge string to read, or the quantum setting, where $W$ determines in which basis to measure some quantum state.

The goal now is to authenticate a message $\mu$ from Alice to Bob with the help of the weak session key $X_W$, in such a way that (1) Eve cannot tamper with $\mu$ without being detected, and (2) Eve learns (nearly) no information on the long-term key $W$. We stress that property (2) is vital for Alice and Bob to be able to re-use $W$. Note that once Alice and Bob can do message authentication with a weak key, then they can also do key agreement, simply by doing standard randomness extraction where the seed for the extractor is communicated in an authentic way.

We want to emphasize that, by assumption, every new session key $X_W$ for the same long-term key $W$ contains fresh randomness, provided by the auxiliary source. Therefore, the goal above does not contradict the well-known impossibility result of re-using an authentication key without refreshing. Also note that we do not specify how exactly the auxiliary source of randomness produces $X_W$ from $W$; on the contrary, we want security no matter how $X_W$ is obtained, as long as $X_W$ contains enough min-entropy (given the adversary's information and $W$).

### 1.2 Related Work

Let $n$ be the bitsize of the key (in our case, the session key) and $k$ its min-entropy (in bits). It was proved by Dodis and Wichs [9] that non-interactive authentication is impossible when $k \leq n/2$, even when the parties have access to local non-shared randomness, which we will assume. For a good overview of earlier work on the case where $k > n/2$, we refer to [9].

The first protocol for interactive authentication from arbitrarily weak keys is due to Renner and Wolf [15]. It requires $\Theta(\ell)$ rounds of interaction to authenticate an $\ell$-bit message. In [9], an authentication protocol from arbitrarily weak keys is described that only needs two rounds of interaction, which is optimal (in terms of the number of rounds). Chandran *et al.* [2] focus on minimizing entropy loss and describe a privacy amplification protocol that is optimal with respect to entropy loss (up to constant factors). Their construction needs a linear number of rounds (linear in the security parameter).

The case where Alice and Bob share highly-correlated, but possibly unequal keys – the "fuzzy" case – is addressed in [16] and improved upon by Kanukurthi and Reyzin [11], but also covered by [9] and [2].

We stress that none of these works address the case where the weak key is obtained from a long-term key and where security of the long-term key needs to be guaranteed.

### 1.3 Our Contributions

We propose a new four-round protocol for message authentication with a weak session key $X_W$. We prove that our protocol satisfies *security* and *long-term key* *privacy*, meaning that the adversary Eve cannot tamper with the authenticated message without being detected, nor does she learn any (non-negligible amount of) information on the long-term key $W$. Our proofs also apply in the quantum setting, where Eve's bounded knowledge on $X_W$ may be in the form of a quantum state.

We also discuss how our techniques can be applied in the fuzzy case, where there are some errors between Alice and Bob's weak session keys. Finally, we outline how our scheme can be used to improve an existing password-based identification scheme in the bounded-quantum-storage model (more details on this application are given below).

### 1.4 Application

Our main application is to password-based identification in the bounded quantum storage model, as proposed by Damgård *et al.* [4]. Two identification schemes were proposed in [4], *Q-ID*, which is only secure against dishonest Alice or Bob, and *Q-ID*$^+$, which is also secure against against a man-in-the-middle (MITM) attack. However, only *Q-ID* is truly password-based; in *Q-ID*$^+$, Alice and Bob, in addition to the password, also need to share a high-entropy key. By incorporating our new techniques into *Q-ID*$^+$, we show the existence of a truly password-based identification scheme in the bounded-quantum-storage model with security against MITM attacks.

Based on *Q-ID*$^+$, Damgård *et al.* also propose an *authenticated* quantum key distribution scheme in the bounded quantum storage model, which, in contrast to standard quantum key distribution schemes, does not require authenticated communication but has the authentication "built in".[1] Our relaxation on the required key material in *Q-ID*$^+$ also affects their authenticated quantum key distribution scheme and circumvents the need for a high entropy key. As a result, we obtain a truly password-based authenticated quantum key distribution scheme in the bounded-quantum-storage model.

### 1.5 Organization of the Paper

The paper is structured as follows. In Section 2 we introduce notation, give some standard definitions and introduce the security definition that our authentication protocol should fulfill. Then, in Section 3, we describe an existing authentication protocol that we use as a basis for our protocol. We also explain there why this existing protocol does not fulfill our security definition, and we discuss some steps how we extend that protocol. This ultimately leads to our own protocol `AUTH`, which is introduced in Section 4. In the same section, we present an important lemma that is used in the security proof to deal with a certain circularity issue. Section 5 consists of the proofs for security and privacy. In Section 6 we argue

[1] Furthermore, in contrast to using standard quantum key distribution in combination with standard authentication, in the authenticated quantum key distribution scheme the authentication keys can be re-used.

that our authentication protocol can also be used in the fuzzy case and finally Section 7 discusses our application. Most results related to instantiating our protocol can be found in the full version of this paper [1].

## 2 Notation and Preliminaries

We prove security of our scheme in the presence of a *quantum* adversary with *quantum* side information, and below we introduce some suitable notations. However, we stress that most of the notation and the proofs can also be understood from a purely classical information-theoretical point of view.

The *state* of a quantum system $X$ is given by a *density matrix* $\rho_X$, i.e., a positive-semidefinite trace-1 matrix acting on some Hilbert space $\mathcal{H}_X$. We denote the set of all such matrices, acting on $\mathcal{H}_X$, by $\mathcal{P}(\mathcal{H}_X)$. In the special case where $\rho_X$ is diagonal, $X$ is called *classical*, and in this case we can understand $X$ as a random variable, where its distribution $P_X$ is given by the diagonal entries of $\rho_X$. In this case, we tend to slightly abuse notation and write $X \in \mathcal{X}$ to indicate that the range of the random variable $X$ is $\mathcal{X}$.

If $X$ is part of a bi-partite system $XE$, then $X$ is called classical if the density matrix $\rho_{XE}$ of $XE$ is of the form $\rho_{XE} = \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_{E|X=x}$, where $P_X$ is a probability distribution, $\{|x\rangle\}_x$ forms an orthonormal basis of $\mathcal{H}_X$, and $\rho_{E|X=x} \in \mathcal{P}(\mathcal{H}_E)$. In this case, $X$ can be understood as random variable, and system $E$ is in state $\rho_{E|X=x}$ exactly if $X$ takes on the value $x$. We therefore sometimes also speak of a random variable $X$ and a quantum system $E$. To simplify notation, we often write $\rho_E^x$ instead of $\rho_{E|X=x}$. Readers that are unfamiliar with quantum information can safely think of $E$ as being classical as well, in which case the $\rho_{E|X=x}$'s are all diagonal, with the probabilities of the conditional distributions $P_{E|X}(\cdot|x)$ as diagonal entries.

The distance between two states $\rho_X, \sigma_X \in \mathcal{P}(\mathcal{H}_X)$ is measured by their *trace distance* $\frac{1}{2}\|\rho_X - \sigma_X\|_1$, where $\|\cdot\|_1$ is the $L_1$ norm.[2] In case of classical states, i.e., $\rho_X$ and $\sigma_X$ correspond to distributions $P_X$ and $Q_X$, the trace distance coincides with the statistical distance $\frac{1}{2}\sum_x |P_X(x) - Q_X(x)|$.

In the following definitions, we consider a bi-partite system $XE$ with classical $X$. $X$ is said to be *random and independent* of $E$ if $\rho_{XE} = \rho_U \otimes \rho_E$, where $\rho_U$ is the fully mixed state on $\mathcal{H}_X$ (i.e., $U$ is classical and - as random variable - uniformly distributed). In case of classical $E$, this is equivalent to $P_{XE} = P_U \cdot P_E$ (in the sense that $P_{XE}(x,e) = P_U(x) \cdot P_E(e)\ \forall x, e$). The following definition measures how far away $XE$ is from such an ideal situation.

**Definition 1 (Distance to Uniform).** *The* distance to uniform *of $X$ given $E$ is defined as*

$$d(X|E) := \tfrac{1}{2}\|\rho_{XE} - \rho_U \otimes \rho_E\|_1 .$$

If also $E$ is classical, then $d(X|E)$ simplifies to

$$d(X|E) = \tfrac{1}{2}\sum_{x,e} |P_{XE}(x,e) - P_U(x)P_E(e)| = \sum_e P_E(e)\, \tfrac{1}{2}\sum_x \big|P_{X|E}(x|e) - P_U(x)\big| .$$

[2] Defined by $\|A\|_1 := \mathrm{trace}(\sqrt{A^\dagger A})$, where $A^\dagger$ denotes the Hermitian transpose.

It is not too hard to show that for a tri-partite system $XYE$ with classical $X$ and $Y$

$$d(X|YE) = \sum_{y\in\mathcal{Y}} P_Y(y)\, d(X|E, Y{=}y).$$

From this, the following lemma follows immediately.

**Lemma 1.** *For any $y$: $d(X|E, Y{=}y) \leq d(X|YE)/P_Y(y)$.*

**Definition 2 (Guessing Probability).** *The* guessing probability *of $X$ given $E$ is defined as*

$$\mathsf{Guess}(X|E) := \sup_{\{M_x\}_x} \sum_x P_X(x)\mathrm{tr}(M_x\, \rho_E^x),$$

*where the supremum is over all POVMs $\{M_x\}_x$ on $\mathcal{H}_E$.*

In case also $E$ is classical, $\mathsf{Guess}(X|E)$ simplifies to the standard average guessing probability

$$\mathsf{Guess}(X|E) = \sum_e P_E(e) \max_x P_{X|E}(x|e).$$

**Definition 3 (Min-Entropy).** *The min-entropy of $X$ given $E$ is defined as*

$$H_{\min}(X|E) := -\log \mathsf{Guess}(X|E).$$

This definition coincides with the definition introduced by Renner [13], as shown by [12]; in case of a classical $E$, it coincides with the classical definition of conditional min-entropy (see e.g. [7]).

**Definition 4.** *A function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(k,\varepsilon)$-strong extractor, if for any bipartite quantum system $XE$ with classical $X$ and with $H_{\min}(X|E) \geq k$, and for a uniform and independent seed $Y$, we have*

$$d\big(\mathsf{Ext}(X,Y)\big|YE\big) \leq \varepsilon\,.$$

Note that we find "extractor against quantum adversaries" a too cumbersome terminology; thus we just call $\mathsf{Ext}$ a (strong) extractor, even though it is a stronger notion than the standard notion of a (strong) extractor. When necessary, we distinguish between the two notions by saying that an extractor is or is not *secure against quantum side information.*

A well-known example of a strong extractor (that is secure against quantum side information) is a two-universal hash function $h : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^q$. Indeed, for any $XE$ with classical $X$, and for $Y$ an independent seed, uniformly distributed on $\{0,1\}^d$ privacy amplification [14] guarantees that

$$d(h(X,Y)|YE) \leq \frac{1}{2}\sqrt{2^{q-H_{\min}(X|YE)}} = \frac{1}{2}\sqrt{2^q\,\mathsf{Guess}(X|YE)}.$$

### 2.1 Security Definition

In the scope of this paper, an authentication protocol is understood as a classical protocol between two parties Alice and Bob. Alice inputs a message $\mu$ and a weak session key $X_W$, and Bob inputs a message $\mu'$ and the same session key $X_W$. At the end of the protocol, Bob announces a Boolean decision whether to "accept" or "reject". The weak session key $X_W$ may depend arbitrarily on a long-term key $W$. During the execution of the protocol, an adversary Eve has full control over the communication between Alice and Bob.

We require the protocol to fulfill the following formal definition.

**Definition 5.** *Let $E_\circ, E$ denote Eve's respective a priori and a posteriori quantum systems, where the latter includes Bob's decision on whether to accept or reject. A $(n, k, m, \delta, \varepsilon)$ message authentication protocol with long-term-key privacy is defined to satisfy the following properties:*

CORRECTNESS: *If there is no adversary Eve present, then for any message $\mu \in \{0,1\}^m$ and $\mu' = \mu$, and for any (distribution of the) key $X_W \in \{0,1\}^n$, Bob accepts with certainty.*

SECURITY: *If $H_{\min}(X_W|WE_\circ) > k$, then for any $\mu, \mu' \in \{0,1\}^m$ with $\mu \neq \mu'$, the probability that Bob accepts is at most $\delta$.*

LONG-TERM-KEY PRIVACY: *If $\rho_{WE_\circ} = \rho_W \otimes \rho_{E_\circ}$ and $H_{\min}(X_W|WE_\circ) > k$, then*

$$\frac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \leq \varepsilon\,.$$

## 3 The Dodis-Wichs Authentication Scheme

Here, we describe a slightly modified version of the two-round message authentication protocol due to Dodis and Wichs [9]. Our construction will be based on this protocol. We start by giving a few definitions that are crucial for the understanding of the protocol by Dodis and Wichs.

**Definition 6 (Epsilon Look-Aheadness).** *Let $t, \ell$ be positive integers. Let $A := (A_1, \ldots, A_t)$ and $B := (B_1, \ldots, B_t)$ be random variables over $(\{0,1\}^\ell)^t$, and let $E$ be a quantum system. For all $i \in \{0, \ldots, t-1\}$ let $\varepsilon_i$ be defined as*

$$\varepsilon_i := d\big(A_{i+1} \ldots A_t \big| B_1 \ldots B_i E\big)\,.$$

*The ordered pair $(A, B)$ is* $\varepsilon$-look-ahead conditioned on $E$ *if $\varepsilon \geq \max_i \varepsilon_i$.*

**Definition 7 (Look-Ahead Extractor).** $\mathsf{laExt} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ *is called a $(k, \varepsilon)$-*look-ahead extractor *if for any random variable $X \in \{0,1\}^n$ and quantum system $E$ with $H_{\min}(X|E) \geq k$ the following holds. Let $S \in \{0,1\}^d$ be a independent and uniformly distributed seed, and let $\tilde{S} \in \{0,1\}^d$ be adversarially chosen given $S$ and $E$; this may involve a (partial) measurement of $E$, resulting in the new state $E'$. Then, the ordered pair $(R, \tilde{R})$ where $R = (R_1, \ldots, R_t) :=$ $\mathsf{laExt}(X; S)$ and $\tilde{R} = (\tilde{R}_1, \ldots, \tilde{R}_t) := \mathsf{laExt}(X; \tilde{S})$ is $\varepsilon$-look-ahead conditioned on $S$, $\tilde{S}$ and $E'$.*

Informally, a look-ahead extractor has the property that even if the adversary is allowed to modify the seed, when given the first $i$ blocks of the key that is extracted using the modified seed, the remaining blocks of the key that is extracted using the correct seed still look random.

**Definition 8 (Look-ahead security).** *A family of functions*

$$\{\mathsf{MAC}_k : \{0,1\}^m \to \{0,1\}^s\}$$

*indexed by keys* $k \in (\{0,1\}^\ell)^t$ *is an* $(\varepsilon, \delta)$ look-ahead secure *MAC if for any pair of fixed and distinct messages* $\mu_{\mathrm{A}}, \mu_{\mathrm{B}} \in \{0,1\}^m, \mu_{\mathrm{A}} \neq \mu_{\mathrm{B}}$*, and any ordered pair of random variables* $(K, K') \in (\{0,1\}^\ell)^{2t}$ *satisfying the look-ahead property with parameter* $\varepsilon$ *conditioned on quantum system* $E$*,*

$$\mathsf{Guess}\big(\mathsf{MAC}_K(\mu_{\mathrm{B}}) \,\big|\, \mathsf{MAC}_{K'}(\mu_{\mathrm{A}})E\big) < \delta\,.$$

We are now ready to present the Dodis and Wichs message authentication protocol `DW-MAC`. The protocol we present here is slightly modified in that we assume that Alice has already sent her message $\mu_{\mathrm{A}}$ to Bob, who has received it as $\mu_{\mathrm{B}}$ (possibly $\neq \mu_{\mathrm{A}}$). This modification is for simplicity, and because we do not aim at minimizing the number of rounds. $X_W$ is the weak key, known to both Alice and Bob. The function $\mathsf{laExt}$: $\{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ is a $(k, \varepsilon)$-look-ahead extractor and $\mathsf{MAC}_k : \{0,1\}^m \to \{0,1\}^s$ is a $(\varepsilon, \delta)$ look-ahead secure MAC.

**Protocol** `DW-MAC`

| *Alice*$(X_W, \mu_{\mathrm{A}})$ | | *Bob*$(X_W, \mu_{\mathrm{B}})$ |
|---|---|---|
| | | $R \in_{\mathrm{R}} \{0,1\}^d$ |
| | $\xleftarrow{R}$ | |
| $K := \mathsf{laExt}(X_W; R)$<br>$T_{\mathrm{A}} := \mathsf{MAC}_K(\mu_{\mathrm{A}})$ | | $K := \mathsf{laExt}(X_W; R)$<br>$T_{\mathrm{B}} := \mathsf{MAC}_K(\mu_{\mathrm{B}})$ |
| | $\xrightarrow{T_{\mathrm{A}}}$ | |
| | | accept if: $T_{\mathrm{A}} = T_{\mathrm{B}}$<br>else: abort |

Security of `DW-MAC` follows immediately from the definitions of the underlying building blocks: $\mathsf{laExt}$ ensures that Alice and Bob's versions of the key $K$ satisfy the look-ahead property, and in this case it is guaranteed that $\mathsf{MAC}$ acts as a secure MAC, even when Alice's key was modified.

However, in our setting where we additionally want to maintain privacy of the long-term key $W$, which may arbitrarily depend on $X_W$, `DW-MAC` does not seem to be good enough — unless Eve remains passive. Indeed, if Eve does not manipulate the communicated seed $R$, then by the assumed lower bound on $H_{\min}(X_W|WE)$, it follows that the extracted $K$ on Bob's side is close to random and independent of $W$ (and $E$), and thus $T$ leaks no information on $W$. However, if Eve manipulates the seed $R$ (for instance replaces it by a value of her choice), then there is no guarantee anymore that $K$, and thus $T$, does no leak information on $W$.

Another and more subtle way for Eve to (potentially) learn information on $W$ is by not manipulating the message, i.e., have $\mu_{\mathrm{A}} = \mu_{\mathrm{B}}$, but manipulate the seed $R$ and try to obtain information on $W$ by observing if Bob accepts or not.

### 3.1 Towards Achieving Key-Privacy

We give here some intuition on how we overcome the above privacy issues of DW-MAC with respect to the long-term key $W$. Similarly to our notation $T_{\mathrm{A}}$ and $T_{\mathrm{B}}$ to distinguish between the tag computed by Alice and by Bob, respectively, we write $R_{\mathrm{A}}$ and $R_{\mathrm{B}}$ etc. to distinguish between Alice and Bob's values of $R$ etc., which may be different if Eve actively manipulates communicated messages.

A first approach to prevent leakage through $T_{\mathrm{A}}$ is to one-time-pad encrypt $T_{\mathrm{A}}$. The key for the one-time-pad is extracted by means of a strong extractor $\mathsf{Ext}$ from $X_W$, where Alice chooses the seed:

$$
\begin{array}{lcl}
\quad\textit{Alice} & & \quad\textit{Bob} \\
 & \xleftarrow{\quad R \quad} & \\
S \in_{\mathrm{R}} \{0,1\}^k & & \\
Z := \mathsf{Ext}(X_W;S) & & \\
Q := T_{\mathrm{A}} \oplus Z & \xrightarrow{\quad S,Q \quad} & \\
 & & Z := \mathsf{Ext}(X_W;S) \\
 & & \text{accept if: } Q = T_{\mathrm{B}} \oplus Z
\end{array}
$$

In the above protocol (and also below), we understand $T_{\mathrm{A}}$ and $T_{\mathrm{B}}$ to be computed as in DW-MAC. Note that since it is Alice who chooses the seed $S$ and since $H_{\min}(X_W|WE)$ is lower bounded, $Z_{\mathrm{A}}$ is guaranteed to be (close to) random and independent of $W$ (and $E$), and thus hides all information that $T_{\mathrm{A}}$ might leak on $W$. However, this modification renders the *security* of the scheme invalid. For instance, we cannot exclude that by modifying the seed $S$ appropriately, Eve can enforce $Z_{\mathrm{B}} = T_{\mathrm{B}}$, so that she only needs to send $Q = 0$ to have Bob convinced.

In order to re-gain security while still preventing information to leak through $T_{\mathrm{A}}$, we let Bob choose a random non-zero "multiplier" for the one-time pad key $Z$:

$$
\begin{array}{lcl}
\quad\textit{Alice} & & \quad\textit{Bob} \\
 & \xleftarrow{\quad R \quad} & \\
S \in_{\mathrm{R}} \{0,1\}^k & & C \in_{\mathrm{R}} \{0,1\}^k \setminus \{0\}^k \\
Z := \mathsf{Ext}(X_W;S) & & \\
 & \xrightarrow{\quad S \quad} & \\
 & \xleftarrow{\quad C \quad} & \\
\text{abort if } C = 0 & & \\
Q := T \oplus C \cdot Z & \xrightarrow{\quad Q \quad} & \\
 & & Z := \mathsf{Ext}(X_W;S) \\
 & & \text{accept if: } Q = T_{\mathrm{B}} \oplus C \cdot Z
\end{array}
$$

The multiplication $C \cdot Z$ is to be understood in the corresponding binary field. Leakage through $T_{\rm A}$ is still prevented since a non-zero multiple of a good one-time-pad key is still a good one-time-pad key. Furthermore, for security, we can intuitively argue as follows. Consider a snapshot of an execution of the protocol after $S$ has been communicated. We now give Eve the value $T_{\rm A}$ for free; this only makes her stronger. By the security of the underlying `DW-MAC` scheme, we know that it is hard for Eve to guess $T_{\rm B}$. Now, assuming that there exist two distinct values for $C$ for which Eve can predict the corresponding value $Q_{\rm B} = T_{\rm B} \oplus C \cdot Z_{\rm B}$, it follows immediately that Eve can actually predict $T_{\rm B}$; a contradiction. Hence, there can be at most one value for Bob's choice of $C$ for which Eve can guess $Q_{\rm B}$ reasonably well.

We point out that the above intuitive reasoning involves *rewinding*; this is fine in the classical but fails in the quantum setting (see e.g. [17]). Thus, in our formal security proof where we allow Eve to maintain a quantum state, we have to reason in a different way. As a consequence, in the actual protocol, $Q$ is computed in a slightly different way.

One issue that is still unsolved is that Bob's decision to accept or reject may also leak information on $W$ when $\mu_{\rm A} = \mu_{\rm B}$ and Eve modifies one (or both) of the seeds $R$ and $S$. Note that this is not an issue if $\mu_{\rm A} \neq \mu_{\rm B}$ because then, by the security, Bob rejects with (near) certainty. For instance it might be that changing the first bit of $S$ changes $Z$ or not, depending on what the first bit of $X_W$ is. Thus, by changing the first bit of $S$ and observing Bob's decision, Eve can learn the first bit of $X_W$, which may give one bit of information on $W$. The solution to overcome this problem is intuitively very simple: we use $\mathsf{MAC}$ not only to authenticate the actual message, but also to authenticate the two seeds $R$ and $S$. Then, like in the case $\mu_{\rm A} \neq \mu_{\rm B}$, if Eve changes one of the seeds then Bob's decision is determined to be reject. Note that this modification introduces a circularity: the key $K$, which is used to authenticate the seed $R$ (amongst the message and $S$) is extracted from $X_W$ by means of the seed $R$. However, it turns out that we can deal with this.

## 4 Main Construction

We now turn to our construction for the message authentication protocol with long-term-key privacy (Definition 5). In the construction, we will use `DW-MAC` as a building block. Informally speaking, the basic idea is to encrypt the authentication tag from `DW-MAC` using a one-time pad, which prevents key leakage. The key for this one-time pad is established in a challenge-response sequence, from a mix of local and shared randomness. Additionally, we use the `DW-MAC` protocol to authenticate some of the extractor seeds that appear in the construction, to prevent key-leakage from Bob's accept/reject decision.

Let $\mathsf{laExt} : \{0,1\}^n \times \{0,1\}^d \to (\{0,1\}^\ell)^t$ be a $(k_K, \varepsilon_K)$ look-ahead extractor. Let $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^v \to \{0,1\}^q$ be a $(k_Z, \varepsilon_Z)$-strong extractor. Let $\mathsf{MAC} : (\{0,1\}^\ell)^t \times (\{0,1\}^m \times \{0,1\}^d \times \{0,1\}^v) \to \{0,1\}^s$ be a $(\epsilon, \lambda + \epsilon)$ look-ahead secure MAC, for any $\epsilon > 0$. Let $X_W$ be the session key, shared among Alice and Bob, and satisfy $H_{\min}(X_W|WE_\circ) > \max(k_K + q, k_Z)$. The "$\oplus$" symbol represents bitwise addition modulo 2. Multiplication, denoted by "$\cdot$", should be understood as multiplication in the corresponding finite field: $\mathrm{GF}(2^s)$ or $\mathrm{GF}(2^q)$. We write $[b]_q$ for the $q$ most significant bits of the bit-string $b$. Protocol `AUTH` is shown below.

**Protocol** `AUTH`

| *Alice*$(X_W, \mu_\mathrm{A})$ | | *Bob*$(X_W, \mu_\mathrm{B})$ |
|---|---|---|
| | | $R \in_\mathrm{R} \{0,1\}^d$ |
| | $\xleftarrow{R}$ | |
| $K := \mathsf{laExt}(X_W; R)$<br>$S \in_\mathrm{R} \{0,1\}^v$ | | $K := \mathsf{laExt}(X_W; R)$ |
| | $\xrightarrow{S}$ | |
| $Z := \mathsf{Ext}(X_W; S)$<br>$T_\mathrm{A} := \mathsf{MAC}_K((\mu_\mathrm{A}, R, S))$ | | $Z := \mathsf{Ext}(X_W; S)$<br>$T_\mathrm{B} := \mathsf{MAC}_K((\mu_\mathrm{B}, R, S))$<br>$U \in_\mathrm{R} \{0,1\}^s, V \in_\mathrm{R} \{0,1\}^q \setminus \{0\}^q$ |
| | $\xleftarrow{U,V}$ | |
| if $U, V = 0$: abort<br>$Q := [U \cdot T_\mathrm{A}]_q \oplus V \cdot Z$ | | |
| | $\xrightarrow{Q}$ | |
| | | accept if: $Q = [U \cdot T_\mathrm{B}]_q \oplus V \cdot Z$<br>else: abort |

In the full version of this paper [1], we show how to instantiate the building blocks (due to space restrictions we have only included the part about instantiating the MAC in the present version, in Appendix A) to obtain a scheme with reasonable parameters. In doing so, we use similar techniques as [9], except that we replace the strong extractors that are part of the look-ahead extractor construction by extractors that are proven secure against quantum side information (by [6]).

Depending on the parameters of an instantiation of `AUTH` and on the bitsize of $\mu_\mathrm{A}$, it might be beneficial, or could even be necessary, to authenticate a hash of the tuple $(\mu_\mathrm{A}, R, S)$, instead of authenticating the tuple itself. In this case, we let Alice choose a small seed for an almost universal hash function and apply $\mathsf{MAC}_K$ to this seed and the hash of the the tuple $(\mu_\mathrm{A}, R, S)$ (with respect to this seed). We will actually make use of this suggested modification when instantiating `AUTH`.

Before going into the security proof for protocol `AUTH`, we resolve here the circularity issue obtained by authenticating the seed $R$ that was used to extract the authentication key $K$.

**Lemma 2.** *Consider a MAC that is $(\epsilon, \lambda + \epsilon)$-look-ahead-secure for any $\epsilon$. Let $K, K', M_\mathrm{A}, M_\mathrm{B}$ be arbitrary random variables and $E$ a quantum state, and let the ordered pair $(K, K') \in (\{0,1\}^\ell)^{2t}$ satisfy the look-ahead property with parameter*

*$\varepsilon$ conditioned on $M_{\mathrm{A}}, M_{\mathrm{B}}, E$ and the event $M_{\mathrm{A}} \neq M_{\mathrm{B}}$. Then,*

$$\mathsf{Guess}\big(\mathsf{MAC}_K(M_{\mathrm{B}}) \,\big|\, \mathsf{MAC}_{K'}(M_{\mathrm{A}}) M_{\mathrm{A}} M_{\mathrm{B}} E, M_{\mathrm{A}} \neq M_{\mathrm{B}}\big) < \lambda + t\varepsilon .$$

*Proof.* We condition on $M_{\mathrm{A}} = m_{\mathrm{A}}$ and $M_{\mathrm{B}} = m_{\mathrm{B}}$ and assume throughout the proof that $m_{\mathrm{A}} \neq m_{\mathrm{B}}$. Because $(K, K')$ may depend on $(M_{\mathrm{A}}, M_{\mathrm{B}})$, conditioning on fixed values for the latter implies that $(K, K')$ is not necessarily $\varepsilon$-look-ahead anymore. Let $\varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}}}$ be the maximum over $i \in [t]$ of the following expression,

$$\varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}},i} := d(K_{i+1} \dots K_t \big| K'_1 \dots K'_i E, M_{\mathrm{A}} = m_{\mathrm{A}}, M_{\mathrm{B}} = m_{\mathrm{B}}).$$

Hence, by Definition 6, $(K, K')$ is $\varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}}}$-look-ahead conditioned on $E$ and the events $M_{\mathrm{A}} = m_{\mathrm{A}}$ and $M_{\mathrm{B}} = m_{\mathrm{B}}$. Note that averaging $\varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}},i}$ over $m_{\mathrm{A}}$ and $m_{\mathrm{B}}$ (conditioned on them being distinct) results in

$$\varepsilon_i = d(K_{i+1} \dots K_t | K'_1 \dots K'_i M_{\mathrm{A}} M_{\mathrm{B}} E, M_{\mathrm{A}} \neq M_{\mathrm{B}}\big) \leq \varepsilon \,.$$

Furthermore, note that by conditioning on fixed and distinct values for $M_{\mathrm{A}}$ and $M_{\mathrm{B}}$, we fulfill the requirements for MAC look-ahead security from Definition 8. I.e. we can conclude that

$$\mathsf{Guess}\big(\mathsf{MAC}_K(M_{\mathrm{B}}) \,\big|\, \mathsf{MAC}_{K'}(M_{\mathrm{A}}) E, M_{\mathrm{A}} = m_{\mathrm{A}}, M_{\mathrm{B}} = m_{\mathrm{B}}\big) < \lambda + \varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}}} .$$

It now follows that

$$\begin{aligned}
&\mathsf{Guess}\big(\mathsf{MAC}_K(M_{\mathrm{B}}) \,\big|\, \mathsf{MAC}_{K'}(M_{\mathrm{A}}) M_{\mathrm{A}} M_{\mathrm{B}} E, M_{\mathrm{A}} \neq M_{\mathrm{B}}\big) \\
&= \sum_{m_{\mathrm{A}},m_{\mathrm{B}}} P_{M_{\mathrm{A}} M_{\mathrm{B}} | M_{\mathrm{A}} \neq M_{\mathrm{B}}}(m_{\mathrm{A}}, m_{\mathrm{B}}) \\
&\qquad \cdot \mathsf{Guess}\big(\mathsf{MAC}_K(M_{\mathrm{B}}) \,\big|\, \mathsf{MAC}_{K'}(M_{\mathrm{A}}) E, M_{\mathrm{A}} = m_{\mathrm{A}}, M_{\mathrm{B}} = m_{\mathrm{B}}\big) \\
&< \sum_{m_{\mathrm{A}},m_{\mathrm{B}}} P_{M_{\mathrm{A}} M_{\mathrm{B}} | M_{\mathrm{A}} \neq M_{\mathrm{B}}}(m_{\mathrm{A}}, m_{\mathrm{B}}) \,(\lambda + \max_{i \in [t]} \varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}},i}) \\
&\leq \lambda + \sum_{m_{\mathrm{A}},m_{\mathrm{B}}} P_{M_{\mathrm{A}} M_{\mathrm{B}} | M_{\mathrm{A}} \neq M_{\mathrm{B}}}(m_{\mathrm{A}}, m_{\mathrm{B}}) \sum_{i \in [t]} \varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}},i} \\
&= \lambda + \sum_{i \in [t]} \sum_{m_{\mathrm{A}},m_{\mathrm{B}}} P_{M_{\mathrm{A}} M_{\mathrm{B}} | M_{\mathrm{A}} \neq M_{\mathrm{B}}}(m_{\mathrm{A}}, m_{\mathrm{B}}) \, \varepsilon_{m_{\mathrm{A}},m_{\mathrm{B}},i} \\
&= \lambda + \sum_{i \in [t]} \varepsilon_i \leq \lambda + \sum_{i \in [t]} \varepsilon = \lambda + t\varepsilon .
\end{aligned}$$

This concludes the proof.

## 5 Proofs of Security and Privacy

In this section we show that protocol `AUTH` fulfills the properties listed in Definition 5. First of all, note that it is easy to see from the protocol description that the correctness property is satisfied, we do not elaborate further on this here.

Throughout the proofs, let $E_\circ$ be Eve's quantum side information before executing AUTH. $E_i$, where $i \in \{1,\dots,4\}$, represents Eve's (quantum) side information after the $i$th round of communication, and hence includes the communicated random variables up to this $i$th round. $E$ represents Eve's side information after executing AUTH, including Bob's decision to accept or reject ($E_4$ does not include this decision). Furthermore, like in Section 3.1, we write $R_\mathrm{A}$ and $R_\mathrm{B}$ etc. for Alice and Bob's respective values for $R$ etc.

**Theorem 1 (Security).** *Assuming that $H_{\min}(X_W|WE_\circ) > k_K + q$, Protocol* AUTH *fulfills the security property defined in Definition 5 with*

$$\delta \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K)}.$$

In fact, we will prove a slightly stronger statement than the security statement, which will be of use also in the proof of the key privacy statement. Let $M_\mathrm{A} := (\mu_\mathrm{A}, R_\mathrm{A}, S_\mathrm{A})$ and $M_\mathrm{B} := (\mu_\mathrm{B}, R_\mathrm{B}, S_\mathrm{B})$. We will prove that in protocol AUTH, if $H_{\min}(X_W|WE_\circ) > k_K + q$, and conditioned on the event $M_\mathrm{A} \neq M_\mathrm{B}$, Bob rejects except with probability

$$\delta' \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}])}.$$

Note that this expression reduces to the simpler expression of Theorem 1 when proving security, because in that case $\mu_\mathrm{A} \neq \mu_\mathrm{B}$ (by Definition 5) which implies that $\Pr[M_\mathrm{A} \neq M_\mathrm{B}] = 1$.

*Proof.* Consider the phase in protocol AUTH after the second round of communication. Assume that $Z_\mathrm{A}$ and $T_\mathrm{A}$ are given to the adversary (this will only make her stronger). Let $K_\mathrm{A} := \mathsf{laExt}(X_W; R_\mathrm{A})$ and $K_\mathrm{B} := \mathsf{laExt}(X_W; R_\mathrm{B})$.

From the chain rule, and by subsequently using that $R_\mathrm{B}$ and $S_\mathrm{A}$ are sampled independently, it follows that

$$H_{\min}(X_W|Z_\mathrm{A}WE_2) \geq H_{\min}(X_W|WE_2) - q \geq H_{\min}(X_W|WE_\circ) - q.$$

By assumption on the parameters, i.e. $H_{\min}(X_W|WE_\circ) > k_K + q$, it follows that $(K_\mathrm{B}, K_\mathrm{A})$ is $\varepsilon_K$-look-ahead conditioned on $Z_\mathrm{A}, W$ and $E_2$. In order to apply Lemma 2, we additionally condition on the event $M_\mathrm{A} \neq M_\mathrm{B}$. By Lemma 1, it is guaranteed that $\varepsilon_K$ grows at most by a factor $1/\Pr[M_\mathrm{A} \neq M_\mathrm{B}]$ as a result of this conditioning. We now apply Lemma 2 and conclude that

$$\mathsf{Guess}(T_\mathrm{B}|T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B}) \leq \lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}].$$

The next step is to view $Q_\mathrm{B} := [U_\mathrm{B} \cdot T_\mathrm{B}]_q \oplus V_\mathrm{B} \cdot Z_\mathrm{B}$ as the output of a strong extractor, with seed $(U_\mathrm{B}, V_\mathrm{B})$. Indeed, it is straightforward to verify that $h : \{0,1\}^s \times \{0,1\}^q \times \{0,1\}^s \times \{0,1\}^q \to \{0,1\}^q$, which maps $(t,z,u,v)$ to $[u \cdot t]_q \oplus v \cdot z$, is a universal hash function (with random seed $(u,v)$). Thus, we can apply privacy amplification. One subtlety is that in protocol AUTH, $V_\mathrm{B}$ is random in $\{0,1\}^q \setminus \{0\}^q$, rather than in $\{0,1\}^q$. However, this affects the overall state by at most an additive term $2^{-q}$, and thus, by triangle inequality, the distance-to-uniform by at most $2 \cdot 2^{-q}$:

$$\begin{aligned} &d(Q_\mathrm{B}|U_\mathrm{B}V_\mathrm{B}T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B}) \\ &\quad \leq \tfrac{1}{2}\sqrt{2^q \mathsf{Guess}(T_\mathrm{B}Z_\mathrm{B}|T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B})} + 2 \cdot 2^{-q} \\ &\quad \leq \tfrac{1}{2}\sqrt{2^q \mathsf{Guess}(T_\mathrm{B}|T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B})} + 2 \cdot 2^{-q} \\ &\quad \leq \tfrac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}])} + 2 \cdot 2^{-q}. \end{aligned}$$

Finally, we have that

$$\begin{aligned} \delta' &= \mathsf{Guess}\big(Q_\mathrm{B}\big|Q_\mathrm{A}WE_3, M_\mathrm{A} \neq M_\mathrm{B}\big) \\ &\leq \mathsf{Guess}\big(Q_\mathrm{B}\big|U_\mathrm{B}V_\mathrm{B}T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B}\big) \\ &\leq 2^{-q} + d(Q_\mathrm{B}|U_\mathrm{B}V_\mathrm{B}T_\mathrm{A}Z_\mathrm{A}WE_2, M_\mathrm{A} \neq M_\mathrm{B}) \\ &\leq 3 \cdot 2^{-q} + \tfrac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K / \Pr[M_\mathrm{A} \neq M_\mathrm{B}])}. \end{aligned}$$

**Theorem 2 (Long-Term-Key Privacy).** *Assuming that* $H_{\min}(X_W|WE_\circ) > \max(q+k_K, k_Z)$*, Protocol* `AUTH` *fulfills the long-term-key privacy property defined in Definition 5 with*

$$\varepsilon \leq 6 \cdot 2^{-q} + \sqrt{2^q\big(\lambda + t\,\varepsilon_K\big)} + \varepsilon_K + 2\,\varepsilon_Z.$$

*Proof.* We first prove that none of the messages exchanged during the protocol leaks information about $W$. Then, we show that in our protocol Bob's decision on whether to accept or reject neither leaks information about $W$.

Because $R_\mathrm{B}$ is sampled independently of $X_W$, and by the chain rule, it follows that

$$H_{\min}(X_W|WE_1[U_\mathrm{A} \cdot T_\mathrm{A}]_q) \geq H_{\min}(X_W|WE_\circ) - q.$$

By assumption on the parameters in the statement of the proposition, i.e. that $H_{\min}(X_W|WE_\circ) > q + k_Z$, and by the properties of $\mathsf{Ext}$ it follows that

$$d(Z_\mathrm{A}|WE_2[U_\mathrm{A} \cdot T_\mathrm{A}]_q) \leq d(Z_\mathrm{A}|S_\mathrm{A}WE_1[U_\mathrm{A} \cdot T_\mathrm{A}]_q) \leq \varepsilon_Z.$$

By the fact that $U_\mathrm{B}$ and $V_\mathrm{B}$ are sampled independently, the following also holds

$$d(Z_\mathrm{A}|WE_3[U_\mathrm{A} \cdot T_\mathrm{A}]_q) \leq \varepsilon_Z.$$

Then, by security of the one-time pad, by the fact that Eve cannot gain information on $W$ by computing $Q_\mathrm{B}$, and by assumption that $\rho_{WE_\circ} = \rho_W \otimes E_\circ$,

$$\tfrac{1}{2}\|\rho_{WE_4} - \rho_W \otimes \rho_{E_4}\|_1 \leq \tfrac{1}{2}\|\rho_{WE_3Q_\mathrm{A}} - \rho_W \otimes \rho_{E_3Q_\mathrm{A}}\|_1 \leq \varepsilon_Z.$$

This completes the first part of the proof.

It remains to show that Bob's decision to accept or reject cannot leak (a substantial amount of) information about $W$. To show this, we make the following case distinction. In case $\mu_\mathrm{A} \neq \mu_\mathrm{B}$, the security proof applies and Bob rejects except with probability $\delta \leq 3 \cdot 2^{-q} + \frac{1}{2}\sqrt{2^q(\lambda + t\,\varepsilon_K)}$. It now immediately follows that

$$\tfrac{1}{2}\|\rho_{WE_4} - \rho_{WE}\|_1 \leq \delta, \quad \text{and} \quad \tfrac{1}{2}\|\rho_W \otimes \rho_{E_4} - \rho_W \otimes \rho_E\|_1 \leq \delta.$$

Hence, in case $\mu_{\text{A}} \neq \mu_{\text{B}}$ (by the triangle inequality),

$$\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \leq \varepsilon_Z + 2\delta.$$

We now turn to the case $\mu_{\text{A}} = \mu_{\text{B}}$ and we analyze for two disjoint events. Conditioned on $M_{\text{A}} \neq M_{\text{B}}$, the strengthened version of the security statement applies, i.e.

$$\delta' \leq 3 \cdot 2^{-q} + \tfrac{1}{2}\sqrt{2^q\big(\lambda + t\,\varepsilon_K / \Pr[M_{\text{A}} \neq M_{\text{B}}]\big)},$$

and again by applying the triangle inequality, we obtain

$$\tfrac{1}{2}\|\rho_{WE|M_{\text{A}}\neq M_{\text{B}}} - \rho_W \otimes \rho_{E|M_{\text{A}}\neq M_{\text{B}}}\|_1 \leq \varepsilon_Z + 2\delta'.$$

Secondly, we analyze for the event $M_{\text{A}} = M_{\text{B}}$. Nevertheless, we start this analysis without conditioning on $M_{\text{A}} = M_{\text{B}}$. (We'll condition on this event later in the proof.) Since $S_{\text{A}}$ is sampled at random and independently of $X_W$, and since $H_{\min}(X_W|WE_\circ) > k_Z$, it follows that

$$d(Z_{\text{A}}|S_{\text{A}}WE_\circ) < \varepsilon_Z\,.$$

By the chain rule (and the independent choice of $S_{\text{A}}$),

$$H_{\min}(X_W|Z_{\text{A}}WE_2) \geq H_{\min}(X_W|WE_\circ) - q > k_K\,,$$

and thus

$$d(K_{\text{B}}|R_{\text{B}}Z_{\text{A}}S_{\text{A}}WE_\circ) < \varepsilon_K\,.$$

From the above, and the independent choices of $R_{\text{B}}$ and $S_{\text{A}}$, it follows that

$$\tfrac{1}{2}\|\rho_{K_{\text{B}}Z_{\text{A}}R_{\text{B}}S_{\text{A}}WE_\circ} - \rho_U \otimes \rho_{U'} \otimes \rho_{R_{\text{B}}} \otimes \rho_{S_{\text{A}}} \otimes \rho_W \otimes \rho_{E_\circ}\|_1 \leq \varepsilon_K + \varepsilon_Z.$$

where $\rho_U$ is the fully mixed state on $\mathcal{H}_{K_{\text{B}}}$ and $\rho_{U'}$ is the fully mixed state on $\mathcal{H}_{Z_{\text{A}}}$, and therefore that

$$\tfrac{1}{2}\|\rho_{K_{\text{B}}Z_{\text{A}}WE_2} - \rho_U \otimes \rho_{U'} \otimes \rho_W \otimes \rho_{E_2}\|_1 \leq \varepsilon_K + \varepsilon_Z.$$

We now condition on $M_{\text{A}} = M_{\text{B}}$. Note that conditioned on this event, $K_{\text{A}} = K_{\text{B}}$ and $Z_{\text{A}} = Z_{\text{B}}$, and therefore, from here on, we omit the subscripts for these random variables and simply write $K$ and $Z$. From Lemma 1 (noting that whether the event $M_{\text{A}} = M_{\text{B}}$ holds is determined by $E_2$), we get

$$\tfrac{1}{2}\|\rho_{KZWE_2|M_{\text{A}}=M_{\text{B}}} - \rho_U \otimes \rho_{U'} \otimes \rho_W \otimes \rho_{E_2|M_{\text{A}}=M_{\text{B}}}\|_1 \leq \frac{\varepsilon_K + \varepsilon_Z}{\Pr[M_{\text{A}} = M_{\text{B}}]}.$$

$U_{\text{B}}$ and $V_{\text{B}}$ are chosen uniformly at random and independent of the rest (and also independently of the event $M_{\text{A}} = M_{\text{B}}$). Furthermore, since $E$ is computed from $(KZE_4)$ alone, it follows that

$$\tfrac{1}{2}\|\rho_{WE|M_{\text{A}}=M_{\text{B}}} - \rho_W \otimes \rho_{E|M_{\text{A}}=M_{\text{B}}}\|_1 \leq \frac{\varepsilon_K + \varepsilon_Z}{\Pr[M_{\text{A}} = M_{\text{B}}]}.$$

We now combine the analyses for the two disjoint events, and conclude that in case $\mu_\mathrm{A} = \mu_\mathrm{B}$,

$$\begin{aligned}
&\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \\
&\quad \le \Pr[M_\mathrm{A} \neq M_\mathrm{B}]\, \tfrac{1}{2}\|\rho_{WE|M_\mathrm{A}\neq M_\mathrm{B}} - \rho_W \otimes \rho_{E|M_\mathrm{A}\neq M_\mathrm{B}}\|_1 \\
&\qquad\quad + \Pr[M_\mathrm{A} = M_\mathrm{B}]\, \tfrac{1}{2}\|\rho_{WE|M_\mathrm{A}=M_\mathrm{B}} - \rho_W \otimes \rho_{E|M_\mathrm{A}=M_\mathrm{B}}\|_1 \\
&\quad = \Pr[M_\mathrm{A} \neq M_\mathrm{B}]\,(\varepsilon_Z + 2\delta') + \varepsilon_K + \varepsilon_Z \\
&\quad \le \Pr[M_\mathrm{A} \neq M_\mathrm{B}] \left[\varepsilon_Z + 6\cdot 2^{-q} + \sqrt{2^q\big(\lambda + t\,\varepsilon_K/\Pr[M_\mathrm{A}\neq M_\mathrm{B}]\big)}\right] + \varepsilon_K + \varepsilon_Z \\
&\quad \le 6\cdot 2^{-q} + \sqrt{2^q\big(\lambda + t\,\varepsilon_K\big)} + \varepsilon_K + 2\,\varepsilon_Z .
\end{aligned}$$

Note that we have computed two upper bounds on $\frac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1$, for two distinct cases: $\mu_\mathrm{A} \neq \mu_\mathrm{B}$ and $\mu_\mathrm{A} = \mu_\mathrm{B}$. Obviously, the weaker (larger) upper bound holds in both cases, and we finally conclude that

$$\tfrac{1}{2}\|\rho_{WE} - \rho_W \otimes \rho_E\|_1 \le 6\cdot 2^{-q} + \sqrt{2^q\big(\lambda + t\,\varepsilon_K\big)} + \varepsilon_K + 2\,\varepsilon_Z .$$

## 6 The Fuzzy Case

Up to here, we assumed a scenario where Alice and Bob share *identical* copies of the session key $X_W$. Let us now consider the "fuzzy" case, where Alice and Bob hold keys that are only close in some sense, but not necessarily equal. This kind of scenario naturally arises when Alice and Bob obtain their session keys in the presence of noise. For simplicity and with our application (Section 7) in mind, we use the Hamming distance to measure closeness between keys.

Consider the following simple approach. Let Bob's key be called $X_W$. Before executing the authentication scheme, Bob sends some error correcting information (like the syndrome of $X_W$ with respect to some error correcting code) to Alice, so that she can correct the errors in her key, $X'_W$, or vice versa. Unfortunately, Eve may of course also modify this error-correcting information, so that Alice might not correct $X'_W$ correctly, in which case our scheme is not guaranteed to work. However, as proved in [9], this approach *does* work if one uses alternating-extraction-based instantiations of look-ahead extractors. For this solution to work it is important that both $X_W$ and $X'_W$ have sufficient min-entropy, and that Bob sends the error correcting information to Alice (i.e. the error-correction information must be sent in the same direction as the seed for the look-ahead extractor). The same holds in our setting where Eve is allowed to have quantum side information.

One subtlety is that the error correcting information must not leak information about $W$, to preserve the privacy property. Exactly this problem is addressed in [8], and is generalized to the quantum setting in [10]. Note that it is straightforward to upper bound the min-entropy loss in $X_W$ (and $X'_W$) due to error correction: by the chain rule this is at most the bitsize of the error-correction information.

## 7 Application: Password-Based Identification in the Bounded Quantum Storage Model

Our main application is to password-based identification in the bounded quantum storage model. Damgård *et al.* proposed in [4] two password-based identification schemes, *Q-ID* and*Q-ID*$^+$. The former is truly password based but does not protect against a man-in-the-middle attack, whereas the latter is secure against a man-in-the-middle attack but is not truly password-based, because the "User" $U$ and "Server" $S$ need to additionally share a secret high-entropy key.[3] We sketch here how our authentication scheme leads to a truly password-based identification scheme in the bounded quantum storage model with security against man-in-the-middle attacks.

The idea of *Q-ID* and *Q-ID*$^+$ is as follows. $U$ sends $n$ BB84 qubits $H^\theta|x\rangle = H^{\theta_1}|x_1\rangle \otimes \cdots \otimes H^{\theta_n}|x_n\rangle$ to $S$, who measures them in basis $c(w) \in \{0,1\}^n$, where $w$ is the common password and $c$ is some appropriate code with large minimal distance $d$. Then, $U$ announces the basis $\theta \in \{0,1\}^n$ used for the BB84 qubits. This allows $U$ and $S$ to compute the string $x_w$ consisting of all the positions of $x$ with $\theta_i = c(w)_i$, i.e., where $U$ and $S$ used the same basis. Then, $U$ needs to convince $S$ that he indeed knows (the same) string $x_w$. Damgård *et al.* show a way to do this which is guaranteed to not leak any information on $w$ to a potentially dishonest $U$ or $S$. Security against a dishonest $U$ holds unconditionally, whereas security against a dishonest $S$ holds in the bounded quantum storage model (where $S$ is assumed to have limited quantum storage). At the core of the latter proof is a lower bound on the min-entropy of $x_w$ from the dishonest server's point of view, which follows from the uncertainty relation from [5].

To make the protocol secure against man-in-the-middle attacks, some way is needed to protect the (classical and quantum) communication against tampering. In order to detect tampering with the communicated qubits, $U$ and $S$ choose a random sample of the qubits and verify that on those no tampering took place. In order to detect tampering with the classical communication, Damgård *et al.* propose to use a so-called extractor MAC. Such a MAC is similar to a standard information-theoretic MAC, and as such requires a high-entropy key, but is also an extractor. The way in which this extractor MAC is used in *Q-ID*$^+$ allows to re-use the high-entropy key.

### 7.1 Our Approach

Our approach of obtaining security against man-in-the-middle attacks without a high-entropy key is now simply to do the authentication of the classical communication by applying protocol `AUTH` of Section 4, using $x_w$ as weak session key. Our privacy property guarantees that the authentication does not leak information on the password $w$. We stress that previous schemes for authentication based on weak keys would (potentially) leak here information on $w$.

[3] The high entropy key is only needed to protect against a man-in-the-middle attack, security against dishonest $U$ and $S$ only relies on the password and holds even if the dishonest party knows the high entropy key.

There are a couple of subtleties to be taken care of with our approach. If the quantum communication is noisy (which it is in realistic scenarios) or if the man-in-the-middle attacker modifies some of the qubits (but few enough so that he is not detected) or $\theta$, then $U$ and $S$'s versions of $x_w$ are not identical. Thus, we are in the fuzzy case. As discussed in Section 6, this is not a problem as long as the error-correcting information is sent from Bob to Alice, which means from $S$ to $U$ in the identification setting, and as long as we have lower bounds on both $U$ and $S$'s versions of $x_w$ (from the attacker's point of view). The first requirement is easily taken care of, we just perform the error correction in the required direction; from $S$ to $U$. In order to guarantee that both versions of $x_w$ have sufficient min-entropy (the analysis of Damgård *et al.* only guarantees min-entropy in $U$'s version), we modify the scheme as follows. Instead of measuring the BB84 qubits in basis $c(w)$, $S$ measures them in a *random* basis $\hat{\theta}$ and announces the difference $r = c(w) \oplus \hat{\theta}$. Then, $U$ and $S$ update the code $c$ by shifting every code word by $r$, so that with respect to the updated code $c'$, $S$ has actually measured the BB84 qubits in basis $c'(w)$. This trick has also been used in [3], though for a different reason, and has no real effect on the analysis of the scheme. However, as we show below, it enables us to argue that also $S$'s version of $x_w$ has lower-bounded min-entropy, and therefore the authentication of the classical messages is guaranteed to work, which implies security of our password-based identification scheme.

Recall that security against a dishonest $U$ or a dishonest $S$ requires that the dishonest party can exclude at most one possibility for the password $w$ (in one execution of the attack); indeed, this is the best we can hope for, because the dishonest party can always try to guess $w$. For password-based man-in-the-middle security, we require that the attacker can exclude at most *two* possibilities for the password. Again, this is the best we can hope for, because in a man-in-the-middle attack, the attacker can (but of course does not have to) individually attack $U$ and $S$, and in both attacks he can try to guess $w$. This is the man-in-the-middle security that we achieve with our scheme.

We first outline our scheme below and then argue (informally) why it is secure. From here, we use upper case letters for the random variables that describe the values $x, \theta, w$, etc. in a (purified) execution of the protocol. It follows from the analysis of *Q-ID*$^+$ in [4] (which still applies under the shifted-codeword modification outlined above) that there exists a $W'$ (independent of $W$) such that unless $W' = W$, there is min-entropy in $X$ restricted to $I_W$ from Eve's point of view.

For $X'$ we reason as follows. Consider two possibilities for $W$; say $w_1$ and $w_2$. We focus on the positions where $c(w_1) \neq c(w_2)$ (which will be the same positions when replacing $c$ by $c'$). We now fix $\theta$; the following will hold for any choice of $\theta$ (chosen by $U$). From the uncertainty relation of [5] it follows that, approximately,

$$H_{\min}(X'_{12}|\hat{\Theta}) \geq d/2,$$

where $X'_{12}$ is the restriction of $X'$ to the positions where $c(w_1) \neq c(w_2)$, and remember that $d$ represents the minimum distance of $c$. Because $X'$ and $\hat{\Theta}$ are independent of $W$, and, in turn, $R$ is determined by $\hat{\Theta}$ and $W$, we have that

1. $U$ picks $x, \theta \in_R \{0,1\}^n$ and sends the $n$-qubit state $H^\theta|x\rangle$ to $S$.
2. $S$ picks $\hat{\theta} \in_R \{0,1\}^n$ and measures $H^\theta|x\rangle$ in basis $\hat{\theta}$. Let $x'$ be the outcome. $S$ computes and sends $r := \hat{\theta} \oplus c(w)$ to $U$. We define $c'(w) := c(w) \oplus r$ and $I_w := \{i : \theta_i = c'(w)_i\}$.
3. $U$ sends $\theta$ and $f \in_R \mathcal{F}$ to $S$.
4. $S$ picks $g \in \mathcal{G}$, $j \in_R \mathcal{J}$ and a random subset $T \subset \{1, \ldots, n\}$ of size $\ell$, computes $s := syn_j(x'|_{I_w})$ and $test' := x'|_T$, and sends $g$, $j$, $s$ and $T$ to $U$.
5. $U$ sets $test := x|_T$, recovers $x'|_{I_w}$ from $x|_{I_w}$ with the help of $s$, and sends $test$ and $z := f(x'|_{I_w}) \oplus g(w)$ to $S$.
6. Using weak key $x'|_{I_w}$, $U$ authenticates all communicated classical messages, i.e. $r, \theta, f, g, j, s, T, test, z$, using AUTH, towards $S$.
7. $S$ accepts if and only if (1) AUTH accepts, (2) $test$ coincides with $test'$ wherever the bases coincide (up to some allowed noise level), and (3) $z = f(x'|_{I_w}) \oplus g(w)$.

$H_{\min}(X'_{12}|\hat{\Theta}WR) \geq d/2$. To additionally condition on Eve's quantum system $E$, we apply the storage bound $q$ and conclude that $H_{\min}(X'_{12}|\hat{\Theta}WRE) \geq d/2 - q$. Since $F$ is independently chosen, we may additionally condition on $F$.

Let $\tilde{\Theta}$ be the (possibly) adversarially modified version of $\theta$, which is sent in step 3. The adversary obtains $\tilde{\Theta}$ as a quantum measurement on $FRE$, so we may condition on $\tilde{\Theta}$ instead of $E$ without lowering the bound. Now, because of the conditioning on $\hat{\Theta}\tilde{\Theta}WFR$, we can replace $X'_{12}$ by the pair $X'_1X'_2$ in the min-entropy bound, where $X'_1$ consists of the positions where $\tilde{\Theta} = c'(w_1)$ and similarly $X'_2$. (The entropy cannot decrease by not restricting to the positions $c(w_1) \neq c(w_2)$ anymore.) Thus,

$$H_{\min}(X'_1X'_2|\hat{\Theta}\tilde{\Theta}WFR) \geq d/2 - q,$$

and therefore in particular

$$H_{\min}(X'_1X'_2|\tilde{\Theta}FR) \geq d/2 - q.$$

This holds for any $w_1$ and $w_2$, so that the entropy splitting lemma [4] implies the existence of $W''$ (independent of $W$), so that unless $W'' = W$, there is lower-bounded min-entropy in $X'$ restricted to $I_W$ from Eve's point of view after step 3. Note that at the point where $X'|_{I_W}$ is actually used to run protocol AUTH, Eve will have obtained additional information (i.e. during step 4 and 5). However, it is not hard to upper-bound the min-entropy loss in $X'|_{I_W}$ due to this additional information, so that with the right choice of parameters there is still lower bounded min-entropy.

We have argued that both $X|_{I_W}$ and $X'|_{I_W}$, or, respectively $X'_W$ and $X_W$ in the terminology of Section 6, have lower-bounded min-entropy from Eve's point of view. Furthermore, in our proposed identification scheme above, $S$ sends the error-correcting information to $U$. Together, this guarantees the security of AUTH when applied in the fuzzy case. Although in the original protocol (*Q-ID*$^+$) the error-correction information is sent in the other direction, reversing this direction is allowed because the authentication makes sure that no message is modified.

Now, security follows from the analysis of *Q-ID*$^+$ [4] (as well as from [3] regarding the shifted-codeword modification).

## Acknowledgment

We would like to thank Krzysztof Pietrzak for interesting discussions and valuable comments regarding this work.

## References

1. Bouman, N.J., Fehr, S.: Secure authentication from a weak key, without leaking information (full version). Cryptology ePrint Archive (2011), http://eprint.iacr.org/2011/034
2. Chandran, N., Kanukurthi, B., Ostrovsky, R., Reyzin, L.: Privacy amplification with asymptotically optimal entropy loss. In: STOC '10: Proceedings of the 42nd ACM symposium on theory of computing. pp. 785–794. ACM (2010)
3. Damgård, I., Fehr, S., Lunemann, C., Salvail, L., Schaffner, C.: Improving the security of quantum protocols via commit-and-open. In: Advances in Cryptology - CRYPTO '09. pp. 408–427. Lecture Notes in Computer Science, Springer (2009)
4. Damgård, I., Fehr, S., Salvail, L., Schaffner, C.: Secure identification and QKD in the bounded-quantum-storage model. In: Advances in Cryptology - CRYPTO '07. Lecture Notes in Computer Science, vol. 4622, pp. 342–359. Springer (2007)
5. Damgård, I.B., Fehr, S., Renner, R., Salvail, L., Schaffner, C.: A tight high-order entropic quantum uncertainty relation with applications. In: Advances in Cryptology - CRYPTO '07. pp. 360–378. Lecture Notes in Computer Science, Springer (2007)
6. De, A., Portmann, C., Vidick, T., Renner, R.: Trevisan's extractor in the presence of quantum side information. arXiv (2009), http://arxiv.org/abs/0912.5514
7. Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. SIAM J. Comput. 38(1), 97–139 (2008)
8. Dodis, Y., Smith, A.: Correcting errors without leaking partial information. In: STOC '05: Proceedings of the 37th annual ACM symposium on theory of computing. pp. 654–663. ACM (2005)
9. Dodis, Y., Wichs, D.: Non-malleable extractors and symmetric key cryptography from weak secrets. In: STOC '09: Proceedings of the 41st annual ACM symposium on theory of computing. pp. 601–610 (2009)
10. Fehr, S., Schaffner, C.: Randomness extraction via delta-biased masking in the presence of a quantum attacker. In: Theory of Cryptography - TCC '08. pp. 465–481. Lecture Notes in Computer Science (2008)
11. Kanukurthi, B., Reyzin, L.: Key agreement from close secrets over unsecured channels. In: Joux, A. (ed.) Advances in Cryptology - EUROCRYPT '09, Lecture Notes in Computer Science, vol. 5479, pp. 206–223. Springer (2009)
12. König, R., Renner, R., Schaffner, C.: The operational meaning of min- and max-entropy. IEEE Transactions on Information Theory 55(9), 4337–4347 (2009)
13. Renner, R.: Security of Quantum Key Distribution. Ph.D. thesis, ETH Zürich (Switzerland) (2005)

14. Renner, R., König, R.: Universally composable privacy amplification against quantum adversaries. In: Kilian, J. (ed.) Theory of Cryptography - TCC '05, Lecture Notes in Computer Science, vol. 3378, pp. 407–425. Springer (2005)
15. Renner, R., Wolf, S.: Unconditional authenticity and privacy from an arbitrarily weak secret. In: Boneh, D. (ed.) Advances in Cryptology - CRYPTO '03. Lecture Notes in Computer Science, vol. 2729, pp. 78–95. Springer (2003)
16. Renner, R., Wolf, S.: The exact price for unconditionally secure asymmetric cryptography. In: Cachin, C., Camenisch, J. (eds.) Advances in Cryptology - EUROCRYPT '04, Lecture Notes in Computer Science, vol. 3027, pp. 109–125. Springer (2004)
17. Van De Graaf, J.: Towards a formal definition of security for quantum protocols. Ph.D. thesis, Univ. de Montreal (Quebec, Canada) (1998)

## A Security and Instantiation of MAC

To construct a MAC with look-ahead security, we adopt the construction given in [9]. Because our look-ahead security definition, Definition 8, is slightly weaker than the one given in [9] (in that both $\mu_\mathrm{A}$ and $\mu_\mathrm{B}$ are fixed), we obtain a better security parameter, as argued below.

With respect to a different aspect, the requirement on the MAC for our construction is somewhat stronger, because we need a "universal" MAC which is $(\epsilon, \lambda + \epsilon)$-look-ahead secure *for any* $\epsilon \geq 0$ (and some $\lambda$). (This requirement stems from the proof of Lemma 2.) It turns out that the construction from [9] in the light of our weaker security definition does satisfy this property.

**Proposition 1.** *For any positive integers $m$ and $\ell$, there exists a family of functions $\{\mathsf{MAC}_k : \{0,1\}^m \to \{0,1\}^s\}$, indexed by keys $k \in (\{0,1\}^\ell)^t$, that is $(\epsilon, 2^{-\ell} + \epsilon)$ look-ahead secure for any $\epsilon > 0$, where $t = 4m$ and $s = 2m\ell$.*

The proof of the statement that $\mathsf{MAC}_k$ is $(\epsilon, 2^{-\ell} + \epsilon)$ look-ahead secure for any $\epsilon > 0$ largely follows the proof of Lemma 15 Appendix E.3 of [9] (and still applies in the quantum setting). However, our modification (of fixing both $\mu_\mathrm{A}$ and $\mu_\mathrm{B}$ before executing `DW-MAC`) overcomes the need for a union bound over all possible messages $\mu_\mathrm{B}$ in that original proof, and hence saves us a factor of $2^m$.

For completeness, we very briefly describe the idea of the construction here. $\mathsf{MAC}_k(\mu)$ outputs some of the blocks $k_i$ of the key $k = (k_1, \ldots, k_t)$; where the choice of this subset is determined by $\mu$. Furthermore, the construction guarantees that for any two distinct messages $\mu$ and $\mu'$, there exists an index $i_\circ < t$ such that $\mathsf{MAC}_k(\mu)$ outputs more blocks $k_i$ with $i > i_\circ$ than $\mathsf{MAC}_k(\mu')$ does. From the look ahead property, it follows that given $k'_1, \ldots, k'_{i_\circ}$, the remaining blocks $k_{i_\circ+1}, \ldots, k_t$ are (close to) random. Then, from the choice of $i_\circ$ and from the chain rule we conclude that when given $\mathsf{MAC}_{k'}(\mu')$, the tag $\mathsf{MAC}_k(\mu)$ still contains at least (nearly) $\ell$ bits of min-entropy.