Can't Touch This

Can't Touch This: unconditional tamper evidence from short keys

Bart van der Vecht[1], Xavier Coiteux-Roy[2], Boris Škorić[1]

[1]Eindhoven University of Technology [2]Università della Svizzera Italiana, Lugano

Abstract

Storing data on an external server with information-theoretic security, while using a key shorter than the data itself, is impossible. As an alternative, we propose a scheme that achieves information-theoretically secure tamper evidence: The server is able to obtain information about the stored data, but not while staying undetected. Moreover, the client only needs to remember a key whose length is much shorter than the data.

We provide a security proof for our scheme, based on an entropic uncertainty relation, similar to QKD proofs. Our scheme works if Alice is able to (reversibly) randomise the message to almost-uniformity with only a short key. By constructing an explicit attack we show that short-key unconditional tamper evidence cannot be achieved without this randomisability.

1 Introduction

1.1 Delegated Storage

Quantum information processing is markedly different from classical information processing. For instance, performing a measurement on an unknown quantum state typically destroys state information. Furthermore, it is impossible to clone an unknown state by unitary evolution [1]. Such properties are very interesting for security applications, since they provide a certain amount of built-in confidentiality, unclonability and tamper-evidence. Quantum physics also features entanglement of subsystems, which allows for feats like teleportation [2, 3] that have no classical analogue. The laws of quantum physics have been exploited in various security schemes, such as Quantum Key Distribution (QKD) [4, 5, 6], quantum anti-counterfeiting [7], quantum Oblivious Transfer [8, 9], authentication and encryption of quantum states [10, 11, 12], unclonable encryption [13], quantum authentication of PUFs [14, 15], and quantum-secured imaging [16], to name a few. For a recent overview of quantum-cryptographic schemes we refer to [17].

In this paper we look at the problem of *Delegated Storage*. Alice needs to store a large amount of data securely, but she does not have enough storage capacity herself. The typical solution is to encrypt the data and then store it on a remote ('cloud') server Eve. Since Alice has to remember the encryption key, this key is necessarily smaller than the data (otherwise Alice could have just stored the data herself). It is well known that information-theoretic security is possible only when the key is at least as large as the entropy of the data. Hence it is obvious that in Delegated Storage the confidentiality of the data cannot be guaranteed unconditionally, not even using quantum physics. A computationally unbounded Eve will always be able to extract information about the data from the (quantum) ciphertext.

We show that, somewhat surprisingly, it *is* possible in Delegated Storage to get information-theoretic guarantees for a security property other than confidentiality: tamper evidence (tampering detection). We present a quantum Delegated Storage scheme for classical data which makes it impossible for Eve to learn anything about Alice's data without alerting Alice, even if Eve has unbounded powers of (quantum) computation, measurement, storage etc. Our scheme is close in spirit to QKD, and in fact it is useful to imagine Delegated Storage as a sort of QKD where Bob is 'future Alice' who retrieves and decrypts the stored cipherstate, and storage on the server corresponds to travelling qubits. There are some subtle differences with QKD, however, namely

(i) the short encryption key, (ii) the availability of the ciphertext at the moment when Eve attacks the qubits, and (iii) Bob's inability to send any message to Alice. These subtle differences conspire to necessitate a security proof that differs nontrivially from QKD security proofs, though many well known ingredients can be re-used.

1.2 Related work

Several works have appeared on the topic of provable deletion of remotely stored data. Coiteux-Roy and Wolf [18] introduced the task of Delegated Storage and provable deletion with a short-key requirement for both tasks. However, they did not settle the question whether unconditional tamper evidence is achievable. Independently, Broadbent and Islam [19] achieved information-theoretic security for provable deletion using keys that are as long as the message.
Lütkenhaus, Marwah and Touchette [20] use a form of Delegated Storage to store a fully-randomised bit commitment on temporarily trusted servers, with the possibility of recall. They don't require a short key in their definition and use a key as long as the message in their protocol.
The verification process in Delegated Storage involves the measurement of a quantum state by the verifier; the prover has to send this quantum state to the verifier. This is different from Provable Deletion protocols and from Molina, Vidick and Watrous's tickets variant [21] of Wiesner's quantum money, where the stored data is quantum but the communication between the prover and the verifier is classical during the verification phase.

1.3 Contributions and outline

- We define *Correctness*, *Security* and *Usefulness* for Delegated Storage. Correctness means that, in case of low disturbance of the stored quantum states, Alice should not get alerted and should be able to recover the message. Security means that Eve cannot learn a non-negligible amount of information about the stored message without alerting Alice. (This definition *does* allow Eve to learn the full message while alarming Alice.) Usefulness means that Alice's locally stored data is smaller than the remotely stored message.
- We present CAN'TTOUCHTHIS, our Delegated Storage scheme. As a first step Alice derives, in a reversible way, an almost-uniform string m from the message μ. Our scheme requires that this randomisation step is possible without the introduction of long keys; hence the entropy of μ must be sufficiently high to allow for using an extractor, or Alice must know the distribution of μ with sufficient accuracy in order to apply compression-based randomisation techniques. Then, Alice extracts a one-time pad from a random string x; the x is encoded into qubits. She computes a ciphertext by masking m with the one-time pad. She stores the ciphertext and the qubits on the server. In between the qubits that contain x there are 'trap' qubits in random positions. When Alice recovers the stored data, she inspects these trap states to see if they have changed.
- We prove that our scheme satisfies the Correctness and Security properties. If ℓ is the message length, then asymptotically $n = \frac{\ell}{1-h(\beta)}$ qubits are required[1], and Alice has to remember a syndrome of (asymptotic) size $\ell\frac{h(\beta)}{1-h(\beta)}$; the syndrome is the main 'key' that she has to store locally. CAN'TTOUCHTHIS allows the message to be longer than the key only when $1-2h(\beta) > 0$. This inequality is familiar in Quantum Key Distribution, where it represents the condition for having positive key rate without two-way communication.
- We propose a method for recursively applying CAN'TTOUCHTHIS. The syndrome is not stored locally, but using CAN'TTOUCHTHIS. The effect is that Alice has to remember a shorter key; asymptotically the number of qubits stored on the server is $n \to \frac{\ell}{1-2h(\beta)}$. This expression too is familiar from QKD, where it stands for the number of qubits required to generate a key of length ℓ.

[1] β is the tolerated bit error rate in the traps. h is the binary entropy function.

- Our scheme needs a preprocessing step to reversibly transform the message μ into an almost-uniform string m which then serves as the 'message' in the quantum part of the protocol. We show that this need for a uniform input is not a deficiency of our scheme or our proof technique, but in fact a fundamental requirement. We introduce an attack called Support which tries to determine one bit: whether the plaintext is the one with the highest a-priori probability. We consider delegated storage in general *without preprocessing* and lowerbound the advantage that Support yields as a function of the key length and the min-entropy of the plaintext. This lower bound serves as a kind of 'no go' theorem: In the case of a low min-entropy distribution that is not known to Alice, our bound implies that the Security property cannot be achieved with a short key.
- We propose two ways in which to achieve a reduced form of tamper evidence in case of the 'no go' situation mentioned above. (i) Introducing a temporary computational assumption; (ii) secret sharing over multiple servers, with the temporary assumption that they are not all colluding.

The outline is as follows. In Section 2 we introduce notation and list useful definitions and lemmas. The security definition is given in Section 3. In Section 4 we describe Can'tTouchThis, and in Section 5 we do the security analysis. Section 6 discusses parameter settings and the recursive scheme. In Section 7 we prove the 'no go' result for low-minentropy distributions that are not known to Alice. In Section 8 we discuss alternative scheme constructions and weaker schemes in the 'no go' situation.

2 Preliminaries

2.1 Notation and terminology

Sets are written in calligraphic font. Classical Random Variables (RVs) are denoted with capital letters, and their realisations with lowercase letters. The expectation with respect to X is denoted as $\mathbb{E}_x f(x) = \sum_{x\in\mathcal{X}} \Pr[X=x]f(x)$. The notation $X \sim P$ means that X has distribution P. We then write $P(x) = \Pr[X=x]$. The statistical distance between two RVs $X, Y \in \mathcal{X}$, with $X \sim P$ and $Y \sim Q$, is given by $\Delta(X,Y) = \frac{1}{2}\sum_{x\in\mathcal{X}} |P(x) - Q(x)|$.

Bitwise XOR of binary strings is written as '$\oplus$'. For the first ℓ bits of the string s we write $s_{[1:\ell]}$. The Hamming weight of s is denoted as $|s|$. The notation 'log' stands for the logarithm with base 2. The function h is the binary entropy function $h(p) = p\log\frac{1}{p} + (1-p)\log\frac{1}{1-p}$.

The Kronecker delta is denoted as δ_{ab}. We will speak about 'the bit error rate β of a quantum channel'. This is defined as the probability that a classical bit x, sent by Alice embedded in a qubit, arrives at Bob's side as the flipped value $\bar{x}$.

For quantum states we use Dirac notation. The notation 'tr' stands for trace. $H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$ is the Hadamard matrix. Let A be a matrix with eigenvalues λ_i. The 1-norm of A is written as $\|A\|_1 = \mathrm{tr}\sqrt{A^\dagger A} = \sum_i |\lambda_i|$. The trace norm is $\|A\|_{\mathrm{tr}} = \frac{1}{2}\|A\|_1$. Quantum states with non-italic label 'A', 'B' and 'E' indicate the subsystem of Alice/Bob/Eve.

Consider classical variables X, Y and a quantum system under Eve's control that depends on X and Y. The combined classical-quantum state is $\rho^{XY\mathrm{E}} = \mathbb{E}_{xy}\, |xy\rangle\langle xy| \otimes \rho^{\mathrm{E}}_{xy}$. The state of a subsystem is obtained by tracing out all the other subspaces, e.g. $\rho^{Y\mathrm{E}} = \mathrm{tr}_X \rho^{XY\mathrm{E}} = \mathbb{E}_y\, |y\rangle\langle y| \otimes \rho^{\mathrm{E}}_y$, with $\rho^{\mathrm{E}}_y = \mathbb{E}_x\, \rho^{\mathrm{E}}_{xy}$. The fully mixed state on Hilbert space $\mathcal{H}_A$ is denoted as χ^A.

We define the rate of a quantum communication protocol as the number of message bits communicated per sent qubit.

2.2 Definitions and lemmas

Definition 2.1 (Rényi entropy) *Let $\mathcal{X}$ be a discrete set. Let $X \in \mathcal{X}$ be a classical variable. Let $\alpha \in (0,1) \cup (1,\infty)$. The Rényi entropy of order α is denoted as $\mathsf{H}_\alpha(X)$ and is defined as*

$$\mathsf{H}_\alpha(X) = \frac{-1}{\alpha-1}\log\sum_{x\in\mathcal{X}}(\Pr[X=x])^\alpha. \tag{1}$$

Definition 2.2 (Smooth Rényi entropy) *Let $X \sim P$ be a discrete classical variable. Let $\alpha \in (0,1) \cup (1,\infty)$. Let $\varepsilon \geq 0$. The ε-smooth Rényi entropy of order α is denoted as $\mathsf{H}^{\varepsilon}_{\alpha}(X)$ and is defined as*

$$\mathsf{H}^{\varepsilon}_{\alpha}(X) = \max_{Y \sim Q,\ Q \in B^{\varepsilon}(P)} \mathsf{H}_{\alpha}(Y), \tag{2}$$

where $B^{\varepsilon}(P)$ is a sub-normalised vicinity of P such that for $Q \in B^{\varepsilon}(P)$ it holds that $\sum_{x \in \mathcal{X}} Q(x) \geq 1 - \varepsilon$ and $\forall_{x \in \mathcal{X}}\ Q(x) \leq P(x)$.

Definition 2.3 (Smooth min-entropy) *Let $\rho^{X\mathrm{E}}$ be a state where X is classical. The ε-smooth min-entropy of X given E is denoted as $\mathsf{H}^{\varepsilon}_{\min}(X|\mathrm{E})_{\rho}$ and is defined as*

$$\mathsf{H}^{\varepsilon}_{\min}(X|\mathrm{E})_{\rho} = \sup_{\tau:\ \|\tau - \rho\|_1 \leq \varepsilon} \mathsf{H}_{\min}(X|\mathrm{E})_{\tau}. \tag{3}$$

Definition 2.4 (Smooth max-entropy) *Let $\rho^{X\mathrm{B}}$ be a state where X is classical. The ε-smooth max-entropy of X given B is denoted as $\mathsf{H}^{\varepsilon}_{\max}(X|\mathrm{B})_{\rho}$ and is defined as*

$$\mathsf{H}^{\varepsilon}_{\max}(X|\mathrm{B})_{\rho} = \inf_{\sigma:\ \|\sigma - \rho\|_1 \leq \varepsilon} \mathsf{H}_{\max}(X|\mathrm{B})_{\sigma}. \tag{4}$$

Definition 2.5 (Extractor) *Let $f : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell}$ be a function. Let $R \in \{0,1\}^d$ be a uniformly random seed. Let $U \in \{0,1\}^{\ell}$ be a uniform RV. The function f is called a (k,ε)-extractor if*

$$\mathsf{H}_{\min}(X) \geq k \implies \Delta\big(f(X,R), U\big) \leq \varepsilon. \tag{5}$$

Definition 2.6 (Strong extractor) *Let $f : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell}$ be a function. Let $R \in \{0,1\}^d$ be a uniformly random seed. Let $U \in \{0,1\}^{\ell}$ be a uniform RV. The function f is called a (k,ε) strong extractor if*

$$\mathsf{H}_{\min}(X) \geq k \implies \Delta\big(Rf(X,R), RU\big) \leq \varepsilon. \tag{6}$$

Definition 2.7 (Quantum-proof strong extractor) *Let $f : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^{\ell}$ be a function. Let $R \in \{0,1\}^d$ be a uniformly random seed. Let $X \in \{0,1\}^n$ be a classical RV and let $\rho^{X\mathrm{E}}$ be a classical-quantum system comprising the classical X entangled with a quantum system 'E'. Let $Z = f(X,R)$. The function f is called a quantum-proof (k,ε) strong extractor if*

$$\mathsf{H}_{\min}(X|E)_{\rho} \geq k \implies \left\|\rho^{ZR\mathrm{E}} - \chi^{Z} \otimes \chi^{R} \otimes \rho^{\mathrm{E}}\right\|_{\mathrm{tr}} \leq \varepsilon. \tag{7}$$

Definition 2.8 (Universal hash) *A family of hash functions $\mathcal{F} = \{f : \mathcal{X} \to \mathcal{T}\}$ is called two-universal (or universal) if for all distinct pairs $x, x' \in \mathcal{X}$ it holds that $\Pr_{f \in \mathcal{F}}[f(x) = f(x')] = 1/|\mathcal{T}|$. Here the probability is over random $f \in \mathcal{F}$.*

A universal hash function is a strong extractor.

Lemma 2.9 (Leftover Hash Lemma [22]) *Let $X \in \mathcal{X}$ be a random variable. Let $f : \mathcal{R} \times \mathcal{X} \to \{0,1\}^{\ell}$ be a universal hash function. Let $U \in \{0,1\}^{\ell}$ be a uniform variable. Then*

$$\ell \leq \max_{\eta \in [0,\varepsilon)} \Big[\mathsf{H}^{\eta}_{2}(X) + 2 - \log \frac{1}{\varepsilon(\varepsilon - \eta)}\Big] \quad \implies \quad \Delta(Rf(R,X), RU) \leq \varepsilon. \tag{8}$$

Definition 2.10 (Pairwise independent hash) *A family of hash functions $\mathcal{F} = \{f : \mathcal{X} \to \mathcal{T}\}$ is called pairwise independent (a.k.a. 2–independent or strongly universal) [23] if for all distinct pairs $x, x' \in \mathcal{X}$ and all pairs $y, y' \in \mathcal{T}$ it holds that $\Pr_{f \in \mathcal{F}}[f(x) = y \wedge f(x') = y'] = |\mathcal{T}|^{-2}$. Here the probability is over random $f \in \mathcal{F}$.*

A pairwise independent hash is also a universal hash.

Lemma 2.11 *(See [24]) Let $F : \{0,1\}^{\nu} \times \{0,1\}^{\nu} \to \{0,1\}^{\nu}$ be given by $F(w,x) = w \cdot x$, where the multiplication is in GF(2^{ν}). Let $\ell \leq \nu$. Let $\Phi : \{0,1\}^{\nu} \times \{0,1\}^{\nu} \to \{0,1\}^{\ell}$ be constructed as $\Phi(w,x) \overset{\mathrm{def}}{=} F(w,x)[1:\ell]$, i.e. the first ℓ bits. Then Φ is a pairwise independent hash.*

Lemma 2.12 *(Theorem 1.5 from [25]) For any $\alpha \in (0,1)$, $\varepsilon > 0$ and any integers n, k satisfying $k \geq \log n + (\log \frac{1}{\varepsilon})^{1+\alpha}$ there exists a quantum-proof (k, ε) strong extractor that gives an output of length $(1-\alpha)k$ and needs a seed of length $\mathcal{O}(\log \frac{n}{\varepsilon})$.*

Lemma 2.13 (Entropic uncertainty relation for BB84 bases) *(See [26]) Let ρ^{ABE} be any state of the three-partite system ABE, where the subsystem A consists of n qubits. Let $X \in \{0,1\}^n$ be the outcome of a measurement on A in the standard basis. Let $X' \in \{0,1\}^n$ be the outcome of a measurement on A in the Hadamard basis. Then*

$$\mathsf{H}^{\varepsilon}_{\min}(X|\mathrm{E})_{\rho} + \mathsf{H}^{\varepsilon}_{\max}(X'|\mathrm{B})_{\rho} \geq n. \tag{9}$$

Lemma 2.14 *(Lemma 6 in [27]) Let $z \in \{0,1\}^{n+r}$. Let $\mathcal{I} \subset [n+r]$, with $|\mathcal{I}| = r$, be a uniformly distributed random variable representing a choice of r out of $n+r$ positions. Then*

$$\Pr\Big[\sum_{i\in\mathcal{I}} z_i \leq r\beta \;\;\wedge\;\; \sum_{i\notin\mathcal{I}} z_i \geq n(\beta+\nu)\Big] \leq e^{-2\nu^2 r \frac{nr}{(n+r)(r+1)}}. \tag{10}$$

3 Attacker model and security definition

We adopt the attacker model that is customary in QKD. No information leaks from Alice's lab, i.e. there are no side channels. Eve has unlimited (quantum) computational resources and is able to perform any measurement allowed by theory. All noise on the quantum channel is considered to be caused by Eve.
Alice draws her message from a certain probability distribution. We will consider three scenarios,

1. **Fully randomised**. From Eve's point of view, the message is uniformly distributed.
2. **Weakly randomised**. From Eve's point of view, the message is not uniformly distributed. However, the distribution has certain favourable properties, and Alice has sufficient knowledge of it to construct an almost-uniform string from the message, by using e.g. an extractor (Lemma 2.9), prefix coding techniques (Section 4.1) or a combination.
3. **Non-randomised**. None of the above apply.

We will show that Delegated Storage can be achieved in the 1st and 2nd scenario, while for a special case of the 3rd scenario we will prove a 'no-go' theorem.
Security proofs are often given for the EPR-based version of a protocol. We will follow the same approach. In this section we define, in the EPR setting, what is meant by 'security' for a Delegated Storage protocol.

The semantics.
The classical variables in the protocol can be abstractly grouped as: The message M, the set of keys K, the data R that Alice has to remember apart from the keys, the transcript T (classical data stored on the server), the modified transcript T' retrieved by Alice, a binary flag $\Omega \in \{0,1\}$ indicating `accept` (1) or `reject` (0), and the reconstructed message $\hat{M}$. The input to the protocol consists of EPR states and the classical M,K. The final output is a quantum-classical state $\rho^{M\hat{M}T\Omega\mathrm{E}}$ containing the classical subsystems M, $\hat{M}$, T, T', Ω and Eve's quantum side information. We denote Eve's system as 'E'. The output state can be written as $\rho^{M\hat{M}TT'\Omega\mathrm{E}} = \rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=0]} + \rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1]}$, with $\operatorname{tr}\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=0]} = \Pr[\Omega=0]$ and $\operatorname{tr}\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1]} = \Pr[\Omega=1]$. Furthermore we write $\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=0]} = \Pr[\Omega=0]_{\rho}\rho^{M\hat{M}TT'\mathrm{E}|\Omega=0}$ and similarly for $\omega = 1$.

Correctness.
Correctness consists of two parts. (i) If Eve behaves honestly, then $\omega = 1$ with overwhelming probability. (ii) If $\omega = 1$ then $\hat{m} = m$ with overwhelming probability.

Security.
We say that the Delegated Storage protocol is ε-secure if the following statement holds.

$$\Big\|\rho^{MTT'\mathrm{E}}_{[\omega=1]} - \mathop{\mathbb{E}}_{m} |m\rangle\langle m| \otimes \rho^{TT'\mathrm{E}}_{[\omega=1]}\Big\|_1 \leq \varepsilon. \tag{11}$$

Eq.(11) can be read as: "If $\Pr[\Omega = 1]$ is negligible then we are making no demands. If $\Pr[\Omega = 1]$ is non-negligible then we demand that M is decoupled from Eve". Note that security properties formulated in terms of the 1-norm (or trace norm) are composable with other (sub-)protocols.

Usefulness.
Let the message space be $\mathcal{M}$ and the key space $\mathcal{K}$. We define the *usefulness* parameter $U \leq 1$ as

$$U = \frac{|\mathcal{M}| - |\mathcal{K}|}{|\mathcal{M}|}. \tag{12}$$

It represents Alice's (relative) gain in the amount of data that she has to store locally.

4 Our Delegated Storage protocol CAN'TTOUCHTHIS

4.1 Design considerations: message randomisation

The potential messages μ that Alice may store come from a message space $\mathcal{M}$. The probability distribution may be far from uniform on $\mathcal{M}$. We will see later on that our scheme requires the stored message to be close to uniform, and that near-uniformity is in fact a necessary condition for Delegated Storage in general. For non-uniform messages one runs into the problem that the ciphertext, which is visible to Eve when she attacks the quantum state, causes leakage about the data contained in the qubits.

Alice hence needs to transform μ into an almost-uniform string M while remembering only a limited amount of information for recovery purposes. She does this in two steps. First she applies a prefix code to losslessly compress μ. The codeword is padded with random bits so that every $\mu \in \mathcal{M}$ that has nonzero probability of occurring is transformed into a string M_0 of fixed length ℓ_0. The ℓ_0 depends only on the probability distribution. It is the length of the longest codeword[2] in the prefix code. The fact that the code is a *prefix code* ensures that the start of the padding can be recognized. Hence, the randomisation comes 'for free': Alice does not have to remember the padding bits.

The second step is to apply an invertible strong extractor on M_0. (See Lemma 2.11). This maps $M_0 \in \{0,1\}^{\ell_0}$ to $M \in \{0,1\}^{\ell}$. If one is willing to tolerate non-uniformity ε_0 then, according to Lemma 2.9, the extractable randomness is

$$\ell \overset{\text{def}}{=} \max_{\eta \in [0,\varepsilon_0)} \left[\mathsf{H}_2^{\eta}(M_0) + 2 - \log \frac{1}{\varepsilon_0(\varepsilon_0 - \eta)} \right]. \tag{13}$$

Alice needs to store locally $\ell_0 - \ell$ secret bits in order to later recover M_0 from M. Once all this is in place, Alice applies a Delegated Storage method that can straightforwardly be proven secure when the message is uniform.

Note that the prefix code method makes sense only if Alice has a reasonably precise knowledge of the distribution P of μ. The step with the strong extractor requires less knowledge: only a correct estimate of $\mathsf{H}_2(M_0)$ is required. Such an estimate is feasible e.g. when P is drawn (in a potentially unknown way) from an ensemble of distributions which each have a known lower bound on the H_2 entropy.

An example of prefix coding plus padding is shown below, for a rather extreme distribution. Here the result M_0 is practically uniform, but this will not be the case in general. Example 4.1 shows that it is sometimes possible to randomise the input very effectively 'for free' even when the H_2-entropy is very low; the randomisation succeeds because Alice knows the distribution of μ.

Example 4.1 *Consider the following probability distribution on $\mathcal{M} = \{0,1\}^L$. For one string μ_0 it holds that $\Pr[\mu = \mu_0] = \frac{1}{2}$; all other strings have probability $\frac{1/2}{2^L - 1}$. This distribution has $\mathsf{H}_{\min}(\mu) = 1$, collision entropy $\mathsf{H}_2(\mu) = 2 + \mathcal{O}(2^{-L})$ and Shannon entropy $\mathsf{H}(\mu) = \frac{L}{2} + \mathcal{O}(1)$.*

[2] In some extreme cases, such as Example 4.1, it may happen that ℓ_0 is slightly larger than $\log |\mathcal{M}|$, in which case it is not really a compression.

The prefix code with padding is constructed as follows. A string $\mu \neq \mu_0$ is encoded as $(0||\mu) \in \{0,1\}^{L+1}$. The string μ_0 is encoded as '1' followed by L random padding bits. The resulting string $M_0 \in \{0,1\}^{L+1}$ has the probability distribution

$$\Pr[M_0 = (0||x)] = (1 - \delta_{x,\mu_0})\frac{1/2}{2^L - 1}; \qquad \Pr[M_0 = (1||x)] = 2^{-(L+1)} \tag{14}$$

for any $x \in \{0,1\}^L$. It has min-entropy $\mathsf{H}_{\min}(M_0) = \log(2^L - 1) + 1 > L + 1 - 2^{-L}/\ln 2$.

4.2 Protocol steps

Setup phase.

Alice chooses ε_0 and sets ℓ according to (13). She chooses values r, λ, β, ν, ε and α, taking care that the following condition is met,

$$\frac{\ell}{1-\alpha} \geq \log\frac{\ell}{(1-\alpha)[1-h(\beta+\nu)]} + (\log\frac{1}{\varepsilon})^{1+\alpha}. \tag{15}$$

She sets $n = \frac{\ell}{(1-\alpha)[1-h(\beta+\nu)]}$. She chooses a MAC function $\Gamma : \mathcal{K} \times \{0,1\}^* \to \{0,1\}^\lambda$, and a quantum-proof strong extractor $f : \{0,1\}^d \times \{0,1\}^n \to \{0,1\}^\ell$. She chooses an Error-Correcting Code C that is able to deal with error rate $\beta + \nu$. The ECC message length is κ and the codeword length is n. We denote the syndrome function as $\texttt{Syn} : \{0,1\}^n \to \{0,1\}^{n-\kappa}$, and the syndrome decoding as $\texttt{SynDec} : \{0,1\}^{n-\kappa} \to \{0,1\}^n$. Alice has an invertible randomisation function $\texttt{Compress} : \mathcal{M} \to \{0,1\}^{\ell_0}$. This includes the random padding step. The corresponding inverse function is $\texttt{Decompress} : \{0,1\}^{\ell_0} \to \mathcal{M}$, which includes discarding the padding bits.

Message preparation.

Alice has a message $\mu \in \mathcal{M}$. She performs the following steps.

1. $m_0 = \texttt{Compress}(\mu)$.
2. Draw random seed $w \in \{0,1\}^{\ell_0}$. Compute $p = w \cdot m_0$, where the multiplication '$\cdot$' is in GF(2^{ℓ_0}). Parse p as $p = m\|m_\triangle$, with $m \in \{0,1\}^\ell$.

Encryption and storage.

3. Draw random strings $\xi, t \in \{0,1\}^{n+r}$ with $|t| = r$. (The string t defines a subset $\mathcal{T} \subset [n+r]$ of size r which points at the trap locations.) Create strings $v = \xi_\mathcal{T}$ (trap values) and $x = \xi_{[n+r]\setminus\mathcal{T}}$ (payload). Prepare the quantum state $|\Psi\rangle = \bigotimes_{j=1}^{n+r} H^{t_j}|\xi_j\rangle$.
4. Draw random seed $u \in \{0,1\}^d$. Compute syndrome $s = \texttt{Syn}\,x$, one-time-pad $z = f(u,x)$ and ciphertext $c = m \oplus z$.
5. Draw random MAC key $\eta \in \mathcal{K}$. Compute the tag $\theta = \Gamma(\eta, w\|u\|c)$. Store w, u, c, θ and $|\Psi\rangle$ on the server. Remember $\eta, \mathcal{T}, v, s, m_\triangle$. Forget all other variables.

Testing and decryption.

6. Retrieve classical data w', u', c', θ' and state $|\Psi'\rangle$. Set $\omega = 0$. If $\theta' \neq \Gamma(\eta, w'\|u'\|c')$ then abort.
7. In positions $\mathcal{T}$ measure $|\Psi'\rangle$ in the Hadamard basis; in all other positions in the standard basis. The measurement result is $v' \in \{0,1\}^r$ from the trap locations and $x' \in \{0,1\}^n$ from the other locations. If $|v' \oplus v| > \beta r$ then abort.
8. Reconstruct $\hat{x} = x' \oplus \texttt{SynDec}(s \oplus \texttt{Syn}\,x')$. If $\texttt{SynDec}$ fails then abort, else continue. Set $\omega = 1$. Compute $\hat{z} = f(u', \hat{x})$ and $\hat{m} = \hat{z} \oplus c'$.
9. Compute $\hat{m}_0 = (w')^{-1} \cdot (\hat{m}\|m_\triangle)$ and $\hat{\mu} = \texttt{Decompress}(\hat{m}_0)$.

Notation	*Meaning*
β	bit error rate threshold
c	classical ciphertext
d	seed length for quantum-proof extractor
η	MAC key for w, c
f	quantum-proof extractor
Γ	MAC function
H	Hadamard operator
κ	ECC message length
ℓ	output length of strong extractor
ℓ_0	lossless compression length
λ	length of authentication tag
m_0	compressed message
m	output of strong extractor
$m_\triangle$	stored for strong extractor inversion
μ	original message
n	number of qubits that carry a payload
ω	success flag
r	number of trap qubits
s	syndrome of payload x
$\mathtt{Syn}$	syndrome function
t	string indicating the trap positions
$\mathcal{T}$	set of trap positions
θ	tag for w, c
u	seed for the quantum-proof extractor
v	trap values
w	seed for the strong extractor
x	payload in the non-trap positions
z	classical one-time pad

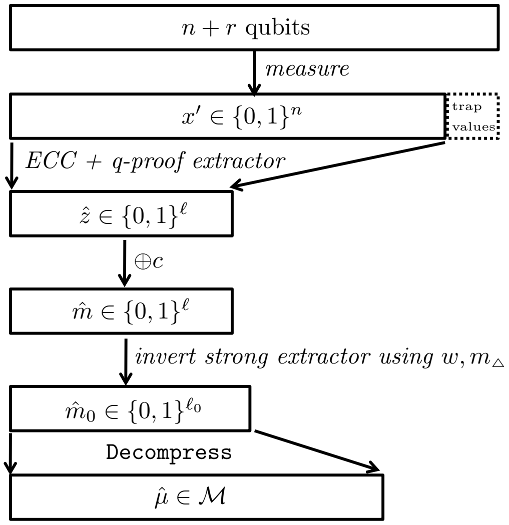
$n + r$ qubits
measure
$x' \in \{0,1\}^n$
trap values
ECC + q-proof extractor
$\hat{z} \in \{0,1\}^\ell$
$\oplus c$
$\hat{m} \in \{0,1\}^\ell$
invert strong extractor using $w, m_\triangle$
$\hat{m}_0 \in \{0,1\}^{\ell_0}$
Decompress
$\hat{\mu} \in \mathcal{M}$

Figure 1: *Visualisation of the variables and the steps in the reconstruction of the message μ.*

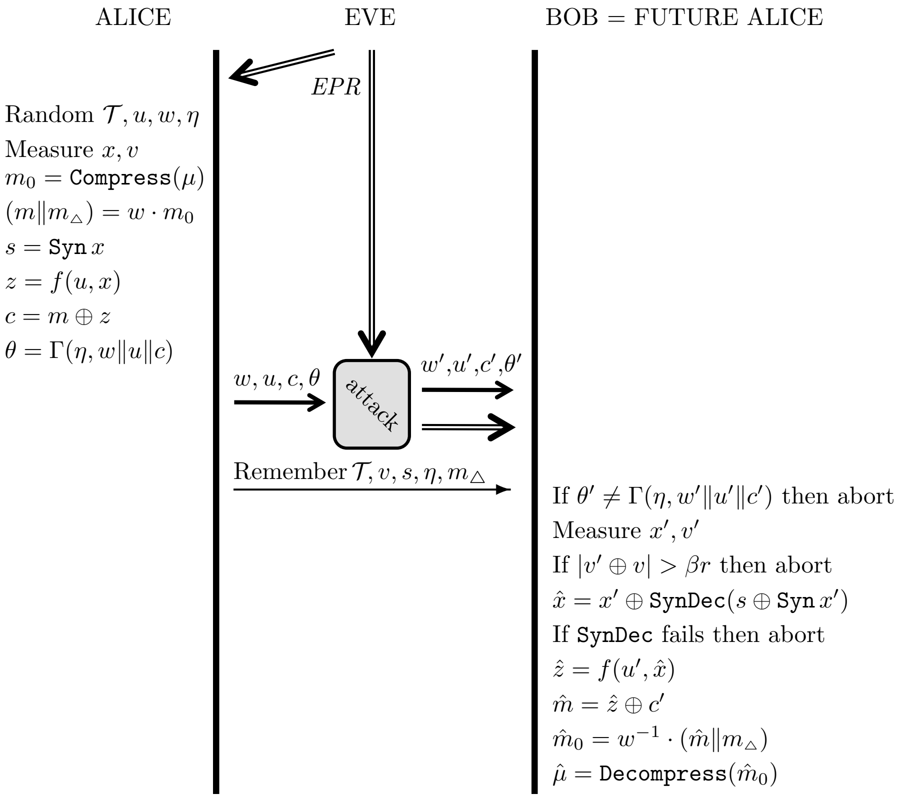
ALICE
EVE
BOB = FUTURE ALICE
EPR
Random $\mathcal{T}, u, w, \eta$
Measure x, v
$m_0 = \texttt{Compress}(\mu)$
$(m \| m_\triangle) = w \cdot m_0$
$s = \texttt{Syn}\, x$
$z = f(u, x)$
$c = m \oplus z$
$\theta = \Gamma(\eta, w \| u \| c)$
w, u, c, θ
attack
w', u', c', θ'
Remember $\mathcal{T}, v, s, \eta, m_\triangle$
If $\theta' \neq \Gamma(\eta, w' \| u' \| c')$ then abort
Measure x', v'
If $|v' \oplus v| > \beta r$ then abort
$\hat{x} = x' \oplus \texttt{SynDec}(s \oplus \texttt{Syn}\, x')$
If $\texttt{SynDec}$ fails then abort
$\hat{z} = f(u', \hat{x})$
$\hat{m} = \hat{z} \oplus c'$
$\hat{m}_0 = w^{-1} \cdot (\hat{m} \| m_\triangle)$
$\hat{\mu} = \texttt{Decompress}(\hat{m}_0)$

Figure 2: *The EPR version of* Can'tTouchThis. *The double lines represent quantum states ($n + r$ qubits).*

5 Security analysis

5.1 EPR version of the protocol

We present the EPR-pair based version of CAN'TTOUCHTHIS (Fig.2). Only the differences with respect to Section 4.2 are listed.
Eve creates $n+r$ EPR-pairs. Of each EPR pair she gives one qubit to Alice and keeps one for herself. In step 3 of the protocol, instead of preparing a state, Alice now measures her j'th qubit in the Hadamard basis if $j \in \mathcal{T}$, and in the standard basis otherwise. The random string ξ is now the result of Alice's measurement.
In step 6 and later, we refer to Alice as 'Bob'.

5.2 Main result

Theorem 5.1 (Main theorem) *Consider the EPR version of* CAN'TTOUCHTHIS *as described in Section 5.1, with parameter values as in Section 4.2. Let the distribution of the message μ and Alice's knowledge about this distribution be such such that the length ℓ in (13) is positive. Let δ be defined as $\delta \overset{\text{def}}{=} \exp\left[-2\nu^2 r \frac{nr}{(n+r)(r+1)}\right]$. The protocol satisfies Correctness and is $(2\cdot 2^{-\lambda}+2\delta+4\varepsilon_0+\varepsilon)$-Secure as defined in Section 3.*

The security of the EPR version implies security of the actual protocol.
Proof of Theorem 5.1.
Correctness.
(i) If Eve behaves honestly then the bitflips in the traps are caused entirely by noise; the parameter β is chosen such that the number of naturally occurring bitflips does not exceed $r\beta$ except with negligible probability (binomial tail).
(ii) If all tests are passed ($\omega = 1$), then the event $\hat{m} \neq m$ could occur in a number of ways: (a) Eve forges a tag. For this the probability is $2^{-\lambda}$, negligible; (b) In the non-trap part there are too many errors to correct (probability $\leq \delta$).
Security.
We have to prove that $\|\rho^{MTT'\mathrm{E}}_{[\omega=1]} - \rho^M \otimes \rho^{TT'\mathrm{E}}_{[\omega=1]}\|_1 \leq 2\cdot 2^{-\lambda} + 2\delta + 4\varepsilon_0 + \varepsilon$. First we note that $\Pr[\Omega = 1|T' \neq T]_\rho \leq 2^{-\lambda}$. This allows us to write $\|\rho^{MTT'\mathrm{E}}_{[\omega=1]} - \rho^M \otimes \rho^{TT'\mathrm{E}}_{[\omega=1]}\|_1 \leq 2\cdot 2^{-\lambda} + \|\rho^{MTT'\mathrm{E}}_{[\omega=1,t'=t]} - \rho^M \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t]}\|_1$. Next we write $\|\rho^{MTT'\mathrm{E}}_{[\omega=1,t'=t]} - \rho^M \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t]}\|_1 \leq \|\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1,t'=t]} - \rho^{M\hat{M}}_{[\omega=1,t'=t]} \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t]}\|_1$. (Taking the $\hat{M}$-trace cannot increase the norm.) We apply Lemma 2.14 to the noise $v' \oplus v \in \{0,1\}^r$ in the trap positions and the noise $x' \oplus x \in \{0,1\}^n$ in the payload positions. Lemma 2.14 gives $\Pr\big[|v' \oplus v| \leq r\beta \;\wedge\; |x' \oplus x| > n(\beta+\nu)\big] \leq \delta$. It follows that $\Pr[\Omega = 1, \hat{M} \neq M] \leq \delta$. This allows us to write $\|\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1,t'=t]} - \rho^{M\hat{M}}_{[\omega=1,t'=t]} \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t]}\|_1 \leq 2\delta + \|\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1,t'=t,\hat{m}=m]} - \rho^{M\hat{M}}_{[\omega=1,t'=t,\hat{m}=m]} \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t,\hat{m}=m]}\|_1$. Next we note that, by Lemma 2.9, the M is ε_0 removed from being uniform. If M were uniform, the transcript T (containing the ciphertext $C = M \oplus Z$) would be statistically decoupled from Z. Hence, the state $\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1,t'=t,\hat{m}=m]}$ is ε_0-close to a state σ where Z does not depend on T. Given C, the M and Z are equivalent. We can write $\|\rho^{M\hat{M}TT'\mathrm{E}}_{[\omega=1,t'=t,\hat{m}=m]} - \rho^{M\hat{M}}_{[\omega=1,t'=t,\hat{m}=m]} \otimes \rho^{TT'\mathrm{E}}_{[\omega=1,t'=t,\hat{m}=m]}\|_1 \leq 4\varepsilon_0 + \|\sigma^{Z\mathrm{E}} - \chi^Z \otimes \sigma^{\mathrm{E}}\|_1$. Finally we have to show that we satisfy the conditions for the existence of a quantum-proof extractor (Lemma 2.12), in order to obtain $\|\sigma^{Z\mathrm{E}} - \chi^Z \otimes \sigma^{\mathrm{E}}\|_1 \leq \varepsilon$. For this we need a lower bound on the min-entropy $\mathsf{H}_{\min}(X|\mathrm{E})_\sigma$. Lemma 2.13 gives $\mathsf{H}_{\min}(X|\mathrm{E})_\sigma \geq n - \mathsf{H}_{\max}(X'|\mathrm{B})_\sigma$. The $\mathsf{H}_{\max}(X'|\mathrm{B})_\sigma$ represents the amount of redundancy information that Bob (who shares noisy EPR pairs with Alice) needs in order to reconstruct a measurement at Alice's side. For the state σ, which is conditioned on $\omega = 1$ and $\hat{m} = m$, this redundancy can be upper bounded [27] as $nh(\beta+\nu)$. This yields $\mathsf{H}_{\min}(X|\mathrm{E})_\sigma \geq n - nh(\beta+\nu)$. By (15) and the setting of n relative to ℓ the conditions for Lemma 2.12 are indeed met. □

6 Setting the parameters

6.1 Asymptotics

We look at the asymptotic case $\mathsf{H}_2(\mu) \to \infty$. We tune the parameters such that the three terms $4\varepsilon_0$, $2^{-\lambda+1}$, and 2δ in Theorem 5.1 equal ε, where ε is constant.
Consider setting the number of trap states proportional to n, i.e. $r = \zeta n$ for constant ζ. This yields (asymptotically)

$$2\delta \to 2e^{-2n\nu^2 \frac{\zeta}{1+\zeta}}. \tag{16}$$

Hence we may set both ν and ζ to small values, $\nu^2\zeta = \mathcal{O}(\frac{1}{n}\ln\frac{1}{\varepsilon})$. With vanishing ν the value of κ goes to $n - nh(\beta)$.
How much local storage Alice needs.
The main data item that Alice needs to remember (store locally) is the syndrome $s \in \{0,1\}^{n-\kappa}$. Asymptotically the size of the syndrome is $nh(\beta)$ bits, with $n \to \ell\frac{1}{1-h(\beta)}$. The other items are the constant-size MAC key η, the trap locations set of size $\log|\mathcal{T}| \approx nh(\frac{\zeta}{1+\zeta}) \approx n\zeta\log\frac{1}{\zeta}$, the $n\zeta$ trap values, and the $m_\triangle \in \{0,1\}^{\ell_0-\ell}$.
It is difficult to make general statements about the gap $\ell_0 - \ell$. However, for sources like human language the quality of the compression becomes better with increasing message length; the average compressed size (without padding) approaches the Shannon entropy, meaning that M_0 becomes more uniform, and the ratio $(\ell_0 - \ell)/\ell_0$ becomes smaller. Furthermore, for sources that produce i.i.d. symbols it is known that asymptotically the smooth Rényi entropy approaches the Shannon entropy, which has an advantageous effect on (13). For these reasons, we expect the overhead $\ell_0 - \ell$ to be sub-linear in $\log|\mathcal{M}|$.
Delegated storage has positive Usefulness (see Section 3) only if Alice needs to remember fewer than roughly ℓ bits,

$$\ell\frac{h(\beta)}{1-h(\beta)} < \ell, \tag{17}$$

i.e. when the bit error rate is small enough to satisfy $1 - 2h(\beta) > 0$. This threshold is (perhaps unsurprisingly) the same as the BB84 threshold for having a positive QKD rate.

6.2 Recursive application of Can'tTouchThis

The $nh(\beta)$ storage requirement for Alice means that she gains very little from Delegated Storage at large β. The following method improves that. The storage of the syndrome s itself can be Delegated using Can'tTouchThis; then Alice has to remember only a fraction $\frac{h(\beta)}{1-h(\beta)}$ of the original size. This principle can be applied recursively until (asymptotically) Alice's local storage needs are very small compared to ℓ_0. The number of stored qubits is then $\frac{\ell}{1-h(\beta)}[1 + \frac{h(\beta)}{1-h(\beta)} + \{\frac{h(\beta)}{1-h(\beta)}\}^2 + \cdots] = \frac{\ell}{1-2h(\beta)}$; this formula is familiar: it is associated with the key rate of (efficient) QKD, i.e. the number of qubits needed to convey a k_0-bit message one-time-pad-encrypted with a QKD key.

7 Why the message needs to be randomizable

The security proof (Section 5.2) needs the assumption that either the *fully randomised* or *weakly randomised* scenario holds (see Section 3). That leaves the question: Is it *impossible in general* to achieve Delegated Storage in the *non-randomised* scenario, or is it just a quirk of our scheme and/or our proof method?

7.1 The non-randomised scenario

We consider a scenario where the distribution P of the message is not known to Alice and is controlled by Eve. This is essentially the setting of Indistinguishability under Chosen Plaintext Attacks.

Alice's lack of knowledge about P prevents her from applying the prefix-code preprocessing trick, i.e. *any randomisation that she performs comes at the cost of having to remember keys, and hence becomes part of the encryption procedure.*
By explicitly constructing an attack (we call it SUPPORT) we demonstrate that, in this scenario, Delegated Storage with short keys is impossible.
We consider a general protocol, not restricted to the one proposed in Section 4.2. We use the following notation. The plaintext is a random variable $M \in \mathcal{M}$ with distribution P. We write $p_m \overset{\text{def}}{=} \Pr[M = m]$. We denote the highest-probability plaintext as m_*, with probability p_*. Let $\rho(m,k)$ denote the quantum encryption of message m using key k. The $\rho(m,k)$ represents *everything* (quantum and classical) that Alice stores on the server, while k is everything that Alice stores privately. The stored state does not have to be pure. The event that Alice does not notice disturbance is called `acc` ('accept'), and if she notices disturbance `rej` ('reject'). We consider only 'correct' schemes, i.e. decryption succeeds with certainty when there is no attack.

7.2 The SUPPORT Attack

Alice receives m from the distribution P and draws a key k uniformly from $\mathcal{K}$. She creates the encryption $\rho(m,k)$. Eve's task is to determine from the encryption whether $m == m_*$ without causing a `rej`. We set some notation:

- For $y \in \mathcal{M}$ we write $\rho(\neg y, k) \overset{\text{def}}{=} \frac{1}{1-p_y} \sum_{m \neq y} p_m \rho(m,k)$ and $\rho(\mathcal{M}, k) \overset{\text{def}}{=} \sum_{m \in \mathcal{M}} p_m \rho(m,k)$.
- For $m \in \mathcal{M}$, $k \in \mathcal{K}$ let $\Pi_{m,k}$ be the projector onto $\text{span}\,(\rho(m,k))$.
- For $m \in \mathcal{M}$ let $\Pi_{m,\mathcal{K}}$ be the projector onto $\text{span}\left(\sum_{k \in \mathcal{K}} \Pi_{m,k}\right)$.
- For $k \in \mathcal{K}$ let $\Pi_{\mathcal{M},k}$ be the projector onto $\text{span}\left(\sum_{m \in \mathcal{M}} \Pi_{m,k}\right)$. For the correctness of the decryption the orthogonality property $\Pi_{\mathcal{M},k} = \sum_{m \in \mathcal{M}} \Pi_{m,k}$ must hold.
- Let $\mathrm{I}_{\mathcal{M},\mathcal{K}}$ be the (identity) projector onto $\text{span}\left(\sum_{m \in \mathcal{M}} \Pi_{m,\mathcal{K}}\right)$.

Definition 7.1 *The attack* SUPPORT *proceeds as follows: Eve applies, on the quantum state she receives from Alice, the projective measurement* $\{\Pi_{m_*,\mathcal{K}},\ \mathrm{I}_{\mathcal{M},\mathcal{K}} - \Pi_{m_*,\mathcal{K}}\}$. *If she obtains* $\Pi_{m_*,\mathcal{K}}$*, she guesses* m_**; otherwise, she guesses* $\neg m_*$.

The projective measurement is potentially very complex.

7.3 SUPPORT breaks the security in the non-randomised scenario

We denote by WIN the event that Eve guesses correctly whether $m == m_*$.

Lemma 7.2 *For* $m = m_*$ *the* SUPPORT *attack causes* WIN *and* `acc`.

Proof: In the case that Alice encrypts $m = m_*$, no matter the actual key k used, the SUPPORT measurement $\Pi_{m_*,\mathcal{K}}$ leaves the state unchanged and correctly yields the measurement result m_*. □

Lemma 7.3 *For the overall* `acc` *probability we have* $\Pr[\texttt{acc}] \geq p_*$.

Proof: $\Pr[\texttt{acc}] = \Pr[M = m_*]\Pr[\texttt{acc}|M = m_*] + \Pr[M \neq m_*]\Pr[\texttt{acc}|M \neq m_*] \geq p_* \Pr[\texttt{acc}|M = m_*] = p_*$. □
We introduce the following notation. Let P be a distribution on $\mathcal{M}$, and let π be a permutation on $\mathcal{M}$. The permuted distribution is denoted as $\pi(P)$.

Lemma 7.4 *Consider the scenario described in Section 7.1. Let* P *be a distribution on* $\mathcal{M}$. *There exists a permutation* π *on* $\mathcal{M}$ *such that the* SUPPORT *attack has* $\Pr_{M \sim \pi(P)}[\text{WIN}|M \neq m_*] \geq 1 - \frac{|\mathcal{K}|}{|\mathcal{M}|}$.

Proof: We use $\max_\pi \Pr_{M\sim\pi(P)}[\mathrm{WIN}|M \neq m_*] \geq \mathbb{E}_\pi \Pr_{M\sim\pi(P)}[\mathrm{WIN}|M \neq m_*]$. In the derivation below the effect of $\mathbb{E}_\pi$ is a uniform choice of m_*. For all $k \in \mathcal{K}$ we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\pi} \Pr_{M\sim\pi(P)}[\mathrm{WIN}|M \neq m_*, k] &= 1 - \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \mathrm{tr}\, \Pi_{m_*,\mathcal{K}} \cdot \rho(\neg m_*, k) \\
&= 1 - \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \mathrm{tr}\, \Pi_{m_*,\mathcal{K}} \cdot \frac{1}{1-p_*} \sum_{m\neq m_*} p_* \rho(m, k) \\
&= 1 - \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \frac{1}{1-p_*} [\mathrm{tr}\, \Pi_{m_*,\mathcal{K}} \cdot \rho(\mathcal{M}, k) - p_*] && (18) \\
&\geq 1 - \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \mathrm{tr}\, \Pi_{m_*,\mathcal{K}} \cdot \rho(\mathcal{M}, k) && (19) \\
&\geq 1 - \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \sum_{k'} \mathrm{tr}\, \Pi_{m_*,k'} \cdot \rho(\mathcal{M}, k) && (20) \\
&= 1 - \sum_{k'} \frac{1}{|\mathcal{M}|} \sum_{m_*\in\mathcal{M}} \mathrm{tr}\, \Pi_{m_*,k'} \cdot \rho(\mathcal{M}, k) && (21) \\
&\overset{\text{correctness}}{=} 1 - \sum_{k'} \frac{1}{|\mathcal{M}|} \mathrm{tr}\, \Pi_{\mathcal{M},k'} \cdot \rho(\mathcal{M}, k) && (22) \\
&\geq 1 - \frac{|\mathcal{K}|}{|\mathcal{M}|}. && (23)
\end{aligned}$$

The equality (18) follows from the definitions of $\rho(\mathcal{M}, k)$, $\rho(\neg m_*, k)$ and the fact that $\mathrm{tr}\,[\Pi_{m_*,\mathcal{K}}$ $\rho(m_*, k)] = 1$. The inequality (19) follows from $\mathrm{tr}\,(\cdots) \leq 1$. We get (20) from $\mathrm{span}(\sum_{k'} \Pi_{m_*,k'}) \leq \sum_{k'} \Pi_{m_*,k'}$. □

Lemma 7.5 *Consider the scenario described in Section 7.1. Let P be a distribution on $\mathcal{M}$. There exists a permutation π on $\mathcal{M}$ such that*

$$\Pr_{M\sim\pi(P)}[\mathsf{acc}|M \neq m_*] \leq \Pr_{M\sim\pi(P)}[\mathrm{WIN} \wedge \mathsf{acc}|M \neq m_*] + \frac{|\mathcal{K}|}{|\mathcal{M}|}. \tag{24}$$

Proof: We have $\Pr[\mathrm{WIN}\wedge\mathsf{acc}|M \neq m_*] = \Pr[\mathrm{WIN}|M \neq m_*] - \Pr[\mathrm{WIN}\wedge\neg\mathsf{acc}|M \neq m_*]$. Applying Lemma 7.4 gives

$$\begin{aligned}
\exists_\pi \quad \Pr_{M\sim\pi(P)}[\mathrm{WIN} \wedge \mathsf{acc}|M \neq m_*] &\geq 1 - \frac{|\mathcal{K}|}{|\mathcal{M}|} - \Pr_{M\sim\pi(P)}[\mathrm{WIN} \wedge \neg\mathsf{acc}\,|M \neq m_*] && (25) \\
&\geq 1 - \frac{|\mathcal{K}|}{|\mathcal{M}|} - \Pr_{M\sim\pi(P)}[\neg\mathsf{acc}|M \neq m_*] && (26) \\
&= \Pr_{M\sim\pi(P)}[\mathsf{acc}|M \neq m_*] - \frac{|\mathcal{K}|}{|\mathcal{M}|}\,. && (27)
\end{aligned}$$

□

Proposition 7.6 *Consider the scenario described in Section 7.1. Let P be a distribution on $\mathcal{M}$. There exists a permutation π on $\mathcal{M}$ such that the* Support *attack has the following advantage in guessing the bit $[m == m_*]$ correctly while staying unnoticed by Alice.*

$$\Pr_{M\sim\pi(P)}[\mathrm{WIN}|\mathsf{acc}] - p_* \;\geq\; p_*(1-p_*)(1 - \frac{|\mathcal{K}|}{|\mathcal{M}|}). \tag{28}$$

Proof: We have

$$\begin{aligned}
\Pr[\mathrm{WIN}|\mathsf{acc}] &= \Pr[\mathrm{WIN} \wedge \mathsf{acc}]/\Pr[\mathsf{acc}] && (29) \\
&= \frac{p_* \cdot \Pr[\mathrm{WIN} \wedge \mathsf{acc}|M = m_*] + (1-p_*) \cdot \Pr\,[\mathrm{WIN} \wedge \mathsf{acc}|M \neq m_*]}{p_* \cdot \Pr[\mathsf{acc}|M = m_*] + (1-p_*) \cdot \Pr[\mathsf{acc}|M \neq m_*]} && (30) \\
&= \frac{p_* + (1-p_*) \cdot \Pr\,[\mathrm{WIN} \wedge \mathsf{acc}|M \neq m_*]}{p_* + (1-p_*) \cdot \Pr[\mathsf{acc}|M \neq m_*]} && (31)
\end{aligned}$$

Note that $\Pr[\mathtt{acc}] \neq 0$ by Lemma 7.3, allowing the division by $\Pr[\mathtt{acc}]$ in the first line. Applying Lemma 7.5 yields

$$\begin{aligned}
\exists_\pi \quad \Pr_{M\sim\pi(P)}[\mathrm{WIN}|\mathtt{acc}] \quad &\geq \quad \frac{p_* + (1-p_*)\Pr_{M\sim\pi(P)}\,[\mathrm{WIN}\wedge\mathtt{acc}|M\neq m_*]}{p_* + (1-p_*)\{\Pr_{M\sim\pi(P)}[\mathrm{WIN}\wedge\mathtt{acc}|M\neq m_*] + \frac{|\mathcal{K}|}{|\mathcal{M}|}\}} \qquad (32)\\
&\geq \quad \frac{p_*}{p_* + (1-p_*)\frac{|\mathcal{K}|}{|\mathcal{M}|}} = \frac{p_*}{1-(1-p_*)(1-\frac{|\mathcal{K}|}{|\mathcal{M}|})} \qquad (33)\\
&\geq \quad p_* + p_*(1-p_*)(1-\frac{|\mathcal{K}|}{|\mathcal{M}|}). \qquad (34)
\end{aligned}$$

□

For the non-randomised scenario, Proposition 7.6 and Lemma 7.3 prove that the security property (11) cannot be achieved with a short key. Consider a distribution P such that $p_* \gg 1/|\mathcal{M}|$. In order to make Eve's advantage (28) negligible it is necessary to make the key length almost equal to the message length; this negates all the advantages of delegated storage.

Theorem 7.7 *For all $U > 0$ there exists $\varepsilon > 0$ such that there exists no U-useful, ε-secure Delegated Storage protocol in the non-randomised scenario.*

Proof. By Proposition 7.6 the attacker's advantage is lower bounded by $p_*(1-p_*)U$, with U constant as a function of the message size. In the non-randomised scenario, the attacker controls p_*, so Alice has no way to reduce the attacker's advantage below $p_*(1-p_*)U$. (E.g. increasing the message length, which typically improves security, does not help here.) The fixed advantage does not allow ε to be decreased indefinitely. □

8 Discussion

Various alternative constructions are of course possible. If confidentiality is required in case of a `reject`, Alice can classically encrypt the message, before or after randomisation, or as part of the randomisation. The confidentiality will not be information-theoretic since the encryption key has to be short.

A different way to improve the `reject`-case confidentiality is to do secret sharing of the message between two or more servers. Then the plaintext is compromised only if (i) all the retrievals are `reject`s; and (ii) all servers collude.

Encryption can be used for a different purpose as well. Consider the scenario discussed in Section 7, i.e. Alice is unable to randomise the message μ 'for free'. Let Alice use a classical cipher F to create ciphertext $\nu = F_k(\mu)$, where k is a short key. This gives Alice a benefit: from Eve's point of view, the ν *temporarily* looks random (until Eve breaks F), enabling Alice to apply CAN'TTOUCHTHIS with ν as the message to be stored. Thus, the tamper evidence in this scenario is computational instead of information-theoretic; still a feat that cannot be accomplished classically. Furthermore, Eve works under a time limit. She is forced break F before Alice retrieves the data.

Secret sharing over multiple servers can achieve this message-randomisation purpose too, and the servers only need to be prohibited from colluding *during* the protocol, because the security granted by tamper evidence is unaffected by the servers sharing information *after* the verification phase.

An interesting aspect of our results is that unconditional tamper evidence for the randomised version of delegated storage does not imply unconditional tamper evidence for its non-randomised version. In fact, Theorem 5.2 and Theorem 7.7 respectively prove that, given a requirement of non-zero usefulness for at least some large-enough messages, the former is possible while the latter is not. In contrast, the security of non-randomised Oblivious Transfer [28] has been shown, under many sequential composability scenarios [29, 30, 31], to be reducible to the security of Randomised Oblivious Transfer. Similarly, Quantum Key Distribution has been shown [32] secure in the universally composable sense: In order to communicate non-random messages, the random keys that QKD distributes can be securely correlated with subsequent information.

An interesting extension would be to apply CAN'TTOUCHTHIS to *quantum information*. The trap qubits would work in a similar way; the mask z would become a key for Quantum-One-Time-Pad encrypting Alice's quantum information.

Acknowledgements

We thank Daan Leermakers, Dave Touchette, and Thomas Vidick for useful discussions and references. Part of this work was funded by Fonds de Recherche du Québec – Nature et Technologies (FRQNT) and the Swiss National Science Foundation (SNF).

References

[1] W.K. Wootters and W.H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.

[2] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, 1993.

[3] D.N. Matsukevich and A. Kuzmich. Quantum state transfer between matter and light. *Science*, 306(5696):663–666, 2004.

[4] C.H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. *IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[5] A.K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661 – 663, 1991.

[6] D. Gottesman and J. Preskill. *Quantum Information with Continuous Variables*, chapter Secure quantum key exchange using squeezed states, pages 317–356. Springer, 2003. arXiv:quant-ph/0008046v2.

[7] C.H. Bennett, G. Brassard, S. Breidbard, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In *CRYPTO*, pages 267–275, 1982.

[8] I.B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the bounded quantum-storage model. In *IEEE Symposium on Foundations of Computer Science*, page 449, 2005.

[9] C. Schaffner. Simple protocols for oblivious transfer and secure identification in the noisy-quantum-storage model. *Phys. Rev. A*, 82:032308, 2010.

[10] H. Barnum, C. Crépeau, D. Gottesman, A. Smith, and A. Tapp. Authentication of quantum messages. In *IEEE Symposium on Foundations of Computer Science*, pages 449–458, 2002. Full version at `http://arxiv.org/abs/quant-ph/0205128`.

[11] P.O. Boykin and V. Roychowdhury. Optimal encryption of quantum bits. *Phys.Rev.A*, 67:042317, 2003.

[12] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf. Private quantum channels. In *Annual Symposium on Foundations of Computer Science*, pages 547–553, 2000.

[13] D. Gottesman. Uncloneable encryption. *Quantum Information and Computation*, 3(6):581–602, 2003.

[14] B. Škorić. Quantum Readout of Physical Unclonable Functions. *International Journal of Quantum Information*, 10(1):1250001:1–31, 2012.

[15] S.A. Goorden, M. Horstmann, A.P. Mosk, B. Škorić, and P.W.H. Pinkse. Quantum-Secure Authentication of a physical unclonable key. *Optica*, 1(6):421–424, 2014.

[16] M. Malik, O.S. Magaña-Loaiza, and R.W. Boyd. Quantum-secured imaging. *Appl.Phys.Lett.*, 101:241103, 2012.

[17] A. Broadbent and C. Schaffner. Quantum cryptography beyond quantum key distribution. *Des. Codes Cryptogr.*, 78:351–382, 2016.

[18] X. Coiteux-Roy and S. Wolf. Proving erasure. In *IEEE International Symposium on Information Theory (ISIT) 2019*, pages 832–836, 2019.

[19] A. Broadbent and R. Islam. Quantum encryption with certified deletion, 2019. `https://arxiv.org/abs/1910.03551`.

[20] N. Lütkenhaus, A.S. Marwah, and D. Touchette. Erasable bit commitment from temporary quantum trust, 2019. `https://export.arxiv.org/pdf/1910.13949`.

[21] A. Molina, T. Vidick, and J. Watrous. Optimal counterfeiting attacks and generalizations for Wiesners quantum money. In *Conference on Quantum Computation, Communication, and Cryptography*, pages 45–64. Springer, 2012.

[22] E. Verbitskiy, P. Tuyls, C. Obi, B. Schoenmakers, and B. Škorić. Key extraction from general nondiscrete signals. *IEEE Transactions on Information Forensics and Security*, 5(2):269–279, 2010.

[23] M.N. Wegman and J.W. Carter. New hash functions and their use in authentication and set equality. *Journal of computer and system sciences*, 22:265–279, 1981.

[24] Y. Dodis, 2013. Lecture notes. `https://cs.nyu.edu/courses/spring14/CSCI-GA.3220-001/lecture1.pdf`.

[25] K.-M. Chung, G. Cohen, T. Vidick, and X. Wu. Quantum-proof extractors: optimal up to constant factors, 2016. `https://arxiv.org/abs/1605.04194v1`.

[26] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover hashing against quantum side information. *IEEE Transactions on Information Theory*, 57(8):5524–5535, 2011.

[27] M. Tomamichel and A. Leverrier. A largely self-contained and complete security proof for quantum key distribution. *Quantum*, 1:14, 07 2017.

[28] D. Beaver. Precomputing oblivious transfer. In *Annual International Cryptology Conference*, pages 97–109. Springer, 1995.

[29] C. Crépeau, G. Savvides, C. Schaffner, and J. Wullschleger. Information-theoretic conditions for two-party secure function evaluation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 538–554. Springer, 2006.

[30] S. Wehner and J. Wullschleger. Composable security in the bounded-quantum-storage model. In *International Colloquium on Automata, Languages, and Programming*, pages 604–615. Springer, 2008.

[31] S. Fehr and C. Schaffner. Composing quantum protocols in a classical environment. In *Theory of Cryptography Conference*, pages 350–367. Springer, 2009.

[32] M. Ben-Or, M. Horodecki, D. Leung, D. Mayers, and J. Oppenheim. The universal composable security of quantum key distribution. In *Theory of Cryptography Conference*, pages 386–406. Springer, 2005.