ROBUST CRYPTOGRAPHY IN THE NOISY-QUANTUM-STORAGE MODEL

CHRISTIAN SCHAFFNER[a]

Centrum Wiskunde & Informatica (CWI)
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

BARBARA TERHAL[b]

IBM, Watson Research Center
P.O. Box 218, Yorktown Heights, NY, USA

STEPHANIE WEHNER[c]

Caltech, Institute for Quantum Information
1200 E California Blvd, Pasadena CA 91125, USA

It was shown in [42] that cryptographic primitives can be implemented based on the assumption that quantum storage of qubits is noisy. In this work we analyze a protocol for the universal task of oblivious transfer that can be implemented using quantum-key-distribution (QKD) hardware in the practical setting where honest participants are unable to perform noise-free operations. We derive trade-offs between the amount of storage noise, the amount of noise in the operations performed by the honest participants and the security of oblivious transfer which are greatly improved compared to the results in [42]. As an example, we show that for the case of depolarizing noise in storage we can obtain secure oblivious transfer as long as the quantum bit-error rate of the channel does not exceed 11% *and* the noise on the channel is strictly less than the quantum storage noise. This is optimal for the protocol considered. Finally, we show that our analysis easily carries over to quantum protocols for secure identification.

1 Introduction

The noisy-quantum-storage model [42] is based on the assumption that it is difficult to store quantum states. Based on current practical and near-future technical limitations, we assume that any state placed into quantum storage is affected by noise. At the same time the model assumes that preparation, transmission and measurement of simple unentangled quantum states can be performed with much lower levels of noise. The present-day technology of quantum key distribution with photonic qubits demonstrates this contrast between a relatively simple technology for preparation/transmission/measurement versus a limited capability for quantum storage.

---

[a]c.schaffner@cwi.nl

[b]bterhal@gmail.com

[c]wehner@caltech.edu

Table 1. Summary of previous results and the results in this paper. The allowed quantum bit-error rate (QBR) is the maximum effective error-rate on the actions of the honest parties below which we can prove the security of the cryptographic scheme.

	Allowed QBR	Secure 1-2 OT	Secure Identification
PRL [42]	None	Yes	No
Unpublished [41]	2.9%	Yes	No
This work	11% (optimal)	Yes	Yes

Almost all interesting cryptographic tasks are impossible to realize without any restrictions on the participating players, neither classically nor with the help of quantum information, see e.g. [27, 30, 29, 28, 31]. It is therefore an important task to come up with a cryptographic model which restricts the capabilities of adversarial players and in which these tasks become feasible. It turns out that all such two-party protocols can be based on a simple primitive called 1-2 Oblivious Transfer (1-2 OT) [24, 19], first introduced in [44, 35, 16]. In 1-2 OT, the sender Alice starts off with two bit strings S_0 and S_1, and the receiver Bob holds a choice bit C. The protocol allows Bob to retrieve S_C in such a way that Alice does not learn any information about C (thus, Bob cannot simply ask for S_C). At the same time, Alice must be ensured that Bob only learns S_C, and no information about the other string $S_{\overline{C}}$ (thus, Alice cannot simply send him both S_0 and S_1). A 1-2 OT protocol is called unconditionally secure when neither Alice nor Bob can break these conditions, even when given unlimited resources.

2   Results

In this work we focus on the setting where the honest parties are unable to perform perfect operations and experience errors themselves, where we analyze individual-storage attacks. These honest-party errors can be modeled as bit-errors on an effective channel connecting the honest parties. In unpublished work, we have shown that for the case of depolarizing noise in storage, security can be obtained if the actions of the honest parties are noisy but their error rate does not exceed 2.9% [41]. This threshold is too low to be of any practical value. In particular, this result left open the question whether security can be obtained in a real-life scenario.

Using a very different analysis, we are now able to show that in the setting of individual-storage attacks 1-2 oblivious transfer and secure identification can be achieved in the noisy-storage model with depolarizing storage noise, as long as the quantum bit-error rate of the channel does not exceed 11% and the noise on the channel is strictly less than the noise during quantum storage. This is optimal for the protocol considered.

Our result is of great practical significance, since it paves the way to achieve security in a real-life implementation. Our main new Theorems 4 and 6 relate the security of the 1-2 OT protocol to an uncertainty lower bound on the conditional Shannon entropy. In order to prove these theorems, we need to relate the Shannon entropy to the smooth min-entropy and establish several new properties of the smooth min-entropy, see Section 3.2.1.

We evaluate the uncertainty lower bounds on the conditional Shannon entropy in the practically-interesting case of depolarizing noise resulting in Theorems 5 and 7. From this analysis we obtain the clear-cut result that, depending on the amount of storage noise, the

adversary's optimal storage attack is to either store the incoming state as is, or to measure it immediately in one of the two BB84 bases.

2.1 The Noisy-Quantum-Storage Model and Individual-Storage Attacks

The noisy-storage model assumes that any quantum state that is placed into quantum storage is affected by some noise described by a quantum operation $\mathcal{N}$. Practically, noise can arise as a result of transferring the qubit onto a different physical carrier, for example the transfer of a photonic qubit onto an atomic ensemble or atomic state. In addition, a quantum state will undergo noise once it has been transferred into 'storage' if such quantum memory is not 100% reliable.

In principle, one may like to prove security against an adversary that can perform any operation on the incoming quantum states. Here however we analyze the restricted case where the adversary Bob performs *individual-storage* attacks. More precisely, Bob's actions are of the following form as depicted in Figure 1.

1. Bob may choose to (partially) measure (a subset of) his qubits immediately upon reception using an error-free *product* measurement, i.e., when he receives the jth qubit, he may apply any measurement $\mathcal{P}_j$ of his choosing.

2. In addition, he can store each incoming qubit, or post-measurement state from a prior partial measurement, *separately* and wait until he gets additional information from Alice (at Step 3 in Protocol 1). During storage, the jth qubit is thereby affected by some noise described by a quantum operation $\mathcal{N}_j$ acting independently on each qubit. Note that such quantum operation $\mathcal{N}_j$ could come about from encoding an incoming qubit into an error-correcting code and decoding it right before receiving Alice's additional information.

3. Once Bob obtains the additional information he may perform an arbitrary coherent measurement $\mathcal{M}$ on his stored qubits and stored classical data.

We would like to note that we can also derive security if we would allow Bob to initially perform *any*, non-product, destructive measurement on the incoming qubits. By destructive we mean that there is no post-measurement quantum data left. The reason is that we have previously shown in Lemma 2 in [42], that destructive *product* measurements are optimal for Bob if he is not allowed to keep any post-measurement information. Hence this optimality of product measurements reduces such more general destructive measurement model to our model of individual-storage attacks. Measurements in present-day technology with single photon qubits in which photons are detected, are in fact always destructive, hence our model includes many realistic attacks. Intuitively, using entangling operations between the incoming qubits should be of little help in either extracting more information from these independent, uncorrelated, BB84 qubits or in better preserving these qubits against noise when the noise is extremely low and more is lost than gained by measuring some qubits right away and using part of the newly freed space to encode the remaining qubits. Of course this remains to be proven (see also Conclusion). What *can* help is to entangle an incoming qubit individually with ancilla qubits in order to store the incoming qubit in an encoded or other more robust
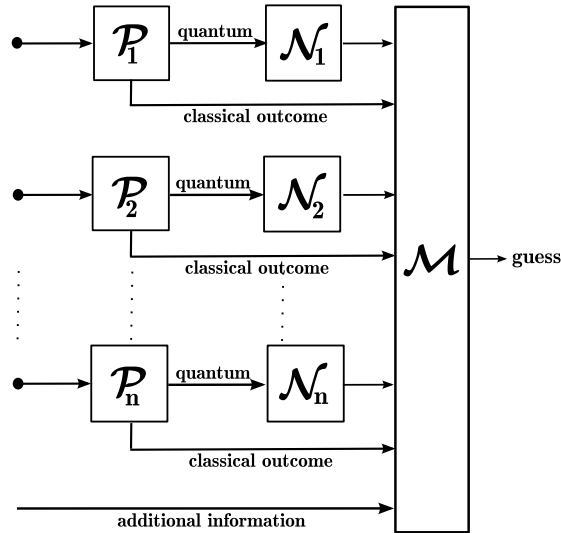
Fig. 1. Individual-Storage Attacks

form. This attack is covered in our model as an effective noisy operation $\mathcal{N}_j$ on incoming qubit j.

In the following, we use the quantum operation $\mathcal{S}_i$ to denote the combined quantum operations of Bob's initial (partial) measurement and the noise.

2.2   Related work

Our model is closely related to the bounded-quantum-storage model, which assumes that the adversary has a limited amount of quantum memory at his disposal [11, 12]. Within this 'bounded-quantum-storage model' OT can be implemented securely as long as a dishonest receiver Bob can store at most $n/4 - O(1)$ qubits coherently, where n is the number of qubits transmitted from Alice to Bob. This approach assumes an explicit limit on the physical number of qubits (or more precisely, the rank of the adversary's quantum state). However, at present we do not know of any practical physical situation which enforces such a limit for quantum information. As was pointed out in [39, 14], the original bounded-quantum-storage analysis applies in the case of noise levels which are so large such that the dishonest player's quantum storage has an effective noise-free Hilbert space with dimension at most $2^{n/4}$. The advantage of our model is that we can evaluate the security parameters of a protocol explicitly in terms of the strength of the noise, even when the noise rate is very low.

Precursors of the idea of basing cryptographic security on storage-noise are already present in [3], but no rigorous analysis was carried through in that paper. We furthermore note that our security proof does not exploit the noise in the communication channel (which has been done in the classical setting to achieve cryptographic tasks, see e.g. [9, 10, 8]), but is solely based on the fact that the dishonest receiver's quantum storage is noisy. A model based on classical noisy storage is akin to the setting of a classical noisy channel, if the operations are

noisy, or the classical bounded-storage model, both of which are difficult to enforce in practice. Another technical limitation has been considered in [38] where a bit-commitment scheme was shown secure under the assumption that the dishonest committer can only measure a limited amount of qubits coherently. Our analysis differs in that we can in fact allow any coherent destructive measurement at the end of the protocol.

2.3   Outline

In Section 3, we introduce some notation and the necessary technical tools. In Section 4, we define the security of 1-2 OT, present the protocol and prove its security in the case when honest players do not experience noise. In Section 5 we then consider the example of depolarizing storage noise explicitly. The lengthy proof of Theorem 5 can be found in Appendix B. In Section 6 we show how to obtain security if the honest players are unable to perform perfect quantum operations. Finally, we point out in Section 7 how our analysis carries over to other protocols.

3   Preliminaries

We start by introducing the necessary definitions, tools and technical lemmas that we need in the remainder of this text.

3.1   Basic Concepts

We use $\in_R$ to denote the uniform random choice of an element from a set. We further use $x_{|\mathcal{T}}$ to denote the string $x = x_1, \ldots, x_n$ restricted to the bits indexed by the set $\mathcal{T} \subseteq \{1, \ldots, n\}$. For a binary random variable C, we denote by $\overline{C}$ the bit different from C.

Let $\mathcal{B}(\mathcal{H})$ denote the set of all bounded operators on a finite-dimensional Hilbert space $\mathcal{H}$. Let $\mathcal{P}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ denote the subset of positive semi-definite Hermitian operators on $\mathcal{H}$, and let $\mathcal{S}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$ denote the subset of all quantum states on $\mathcal{H}$, i.e. $\rho \in \mathcal{S}(\mathcal{H})$ iff $\rho \in \mathcal{B}(\mathcal{H})$ with $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$. $\mathrm{Tr}_A : \mathcal{B}(\mathcal{H}_{AB}) \to \mathcal{B}(\mathcal{H}_B)$ is the partial trace over system A. We denote by id_A the identity operator on system A. Let $|0\rangle_+, |1\rangle_+, |0\rangle_\times := (|0\rangle_+ + |1\rangle_+)/\sqrt{2}, |1\rangle_\times := (|0\rangle_+ - |1\rangle_+)/\sqrt{2}$ denote the BB84-states corresponding to the encoding of a classical bit into the computational or Hadamard basis, respectively.

Classical-Quantum States   A *cq-state* ρ_{XE} is a state that is partly classical, partly quantum, and can be written as

$$\rho_{XE} = \sum_{x \in \mathcal{X}} P_X(x)|x\rangle\langle x| \otimes \rho_E^x \,.$$

Here, X is a classical random variable distributed over the finite set $\mathcal{X}$ according to distribution P_X, $\{|x\rangle\}_{x \in \mathcal{X}}$ is a set of orthonormal states and the register E is in state ρ_E^x when X takes on value x.

Distance measures   The L_1-norm of an operator $A \in \mathcal{B}(\mathcal{H})$ is defined as $\|A\|_1 := \mathrm{Tr}\sqrt{A^\dagger A}$. The fidelity between two quantum states ρ, σ is defined as $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1$. For pure states it takes on the easy form $F(|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|) = |\langle\phi|\psi\rangle|$. The related quantity $C(\rho, \sigma) := \sqrt{1 - F^2(\rho, \sigma)}$ is a convenient distance measure on normalized states [1]. It is invariant

under purifications and equals the trace distance for pure states, i.e. $C(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \sqrt{1 - |\langle\psi|\phi\rangle|^2} = \frac{1}{2}|||\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|||_1$.

Non-uniformity   We can say that a quantum adversary has little information about X if the distribution P_X given his quantum state is close to uniform. Formally, this distance is quantified by the *non-uniformity* of X given $\rho_E = \sum_x P_X(x)\rho_E^x$ defined as

$$d(X|E) := \frac{1}{2}\left\|\operatorname{id}_X/|\mathcal{X}| \otimes \rho_E - \sum_x P_X(x)|x\rangle\langle x| \otimes \rho_E^x\right\|_1 . \tag{1}$$

Intuitively, $d(X|E) \le \varepsilon$ means that the distribution of X is ε-close to uniform even given ρ_E, i.e., ρ_E gives hardly any information about X. A simple property of the non-uniformity which follows from its definition is that it does not change given independent information. Formally,

$$d(X|E, D) = d(X|E) \tag{2}$$

for any cqq-state of the form $\rho_{XED} = \rho_{XE} \otimes \rho_D$.

3.2   *Entropic Quantities*

Throughout this paper we use a number of entropic quantities. The *binary-entropy* function is defined as $h(p) := -p\log p - (1-p)\log(1-p)$, where $\log$ denotes the logarithm base 2 throughout this paper. The *von Neumann entropy* of a quantum state ρ is given by

$$\mathrm{H}(\rho) := -\operatorname{Tr}(\rho\log\rho) .$$

For a bipartite state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$, we use the shorthand

$$\mathrm{H}(A|B) := \mathrm{H}(\rho_{AB}) - \mathrm{H}(\rho_B)$$

to denote the *conditional von Neumann entropy* of the state ρ_{AB} given the quantum state $\rho_B = \operatorname{Tr}_A(\rho_{AB}) \in \mathcal{S}(\mathcal{H}_B)$. Of particular importance to us are the following quantities introduced by Renner [36]. Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$. Then the *conditional min-entropy of ρ_{AB} relative to B* is defined by the following semi-definite program

$$\mathrm{H}_\infty(A|B)_\rho := -\log \min_{\substack{\sigma_B \in \mathcal{P}(\mathcal{H}_B) \\ \rho_{AB} \le \operatorname{id}_A \otimes \sigma_B}} \operatorname{Tr}(\sigma_B) .$$

For a cq-state ρ_{XE} one can show [26] that the conditional min-entropy is the (negative logarithm of the) guessing probability [d]

$$\mathrm{H}_\infty(X|E)_\rho = -\log \mathrm{P}_{\mathrm{guess}}(X|E)_\rho , \tag{3}$$

where $\mathrm{P}_{\mathrm{guess}}(X|E)_\rho$ is defined as the maximum success probability of guessing X by measuring the E-register of ρ_{XE}. Formally, for any (not necessarily normalized) cq-state ρ_{XE}, we define

$$\mathrm{P}_{\mathrm{guess}}(X|E)_\rho := \sup_{\{M_x\}} \sum_x P_X(x)\operatorname{Tr}(M_x\rho_E^x) ,$$

---

[d]Such an "operational meaning" of conditional min-entropy can also be formulated for general qq-states [26].

where the supremum ranges over all positive-operator valued measurements (POVMs) with measurement elements $\{M_x\}_{x \in \mathcal{X}}$, i.e. $M_x \geq 0$ and $\sum_x M_x = \mathsf{id}_E$. If all information in E is classical, we recover the fact that the classical min-entropy is the negative logarithm of the average maximum guessing probability.

In our proofs we also need smooth versions of these entropic quantities. The idea is to no longer consider the min-entropy of a fixed state ρ_{AB}, but take the supremum over the min-entropy of states $\hat{\rho}_{AB}$ which are close to ρ_{AB}, and which may have considerably larger min-entropy. In a cryptographic setting, we are often not interested in the min-entropy of a concrete state ρ_{AB}, but in the maximal min-entropy we can get from states in the neighborhood of ρ_{AB}, i.e. deviating only slightly from the real situation ρ_{AB}. These smooth quantities have some nice properties which are needed in our security proof. For $\varepsilon \geq 0$, the *ε-smooth min-entropy of ρ_{AB}* is given by

$$\mathrm{H}_\infty^\varepsilon (A|B)_\rho := \sup_{\hat{\rho}_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB})} \mathrm{H}_\infty(A|B)_{\hat{\rho}} ,$$

where $\mathcal{K}^\varepsilon(\rho_{AB}) := \{\hat{\rho}_{AB} \in \mathcal{P}(\mathcal{H}_{AB}) \mid C(\rho_{AB}, \hat{\rho}_{AB}) \leq \varepsilon \text{ and } \mathrm{Tr}(\hat{\rho}_{AB}) \leq 1)\}$. If the quantum states ρ are clear from the context, we drop the subscript of the entropies.

3.2.1   Properties of The Conditional Smooth Min-Entropy

In our security analysis we make use of the following properties of smooth min-entropy. First, we need the chain rule whose simple proof can be found in Appendix A:

Lemma 1 (Chain Rule) *For any ccq-state $\rho_{XYE} \in \mathcal{S}(\mathcal{H}_{XYE})$ and for all $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^\varepsilon (X|YE) \geq \mathrm{H}_\infty^\varepsilon (XY|E) - \log |\mathcal{Y}|,$$

where $|\mathcal{Y}|$ is the alphabet size of the random variable Y.

Secondly, we prove the additivity of the smooth conditional min-entropy (see Appendix A):

Lemma 2 (Additivity) *Let ρ_{AB} and $\rho_{A'B'}$ be two independent qq-states. For $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^{\varepsilon^2} (AA'|BB')_\rho \leq \mathrm{H}_\infty^\varepsilon (A|B) + \mathrm{H}_\infty^\varepsilon (A'|B') .$$

Thirdly, adding a classical register can only increase the smooth min-entropy (see Appendix A):

Lemma 3 (Monotonicity) *For a ccq-state ρ_{XYE} and for all $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^\varepsilon (XY|E) \geq \mathrm{H}_\infty^\varepsilon (Y|E) .$$

At last, we deduce a lower bound on the conditional smooth min-entropy of product states. The following theorem is a straightforward generalization of Theorem 7 in [40] (see also [36, Theorem 3.3.6]) to the case where the states are independently, but not necessarily identically distributed. The theorem states that for a large number of independent states, the conditional

smooth min-entropy can be lower-bounded by the conditional Shannon entropy. We note that it is a common feature of equipartition theorems for classical or quantum information that the assumption of i.i.d. sources can be replaced by the weaker assumption of non-i.i.d. but independent sources (see Appendix A for the proof).

Theorem 1 (adapted from [40]) *For $i = 1, \ldots, n$, let $\rho_i \in \mathcal{S}(\mathcal{H}_{AB})$ be density operators. Then, for any $\varepsilon > 0$,*

$$\mathrm{H}_\infty^\varepsilon \left(A^n | B^n\right)_{\otimes_{i=1}^n \rho_i} \geq \sum_{i=1}^n \left[\mathrm{H}(A_i | B_i)_{\rho_i}\right] - \delta(\varepsilon, \gamma) \sqrt{n} \,,$$

where, for $n \geq \frac{8}{5} \log \frac{2}{\varepsilon^2}$, the error is given by

$$\delta(\varepsilon, \gamma) := 4 \log \gamma \sqrt{\log \frac{2}{\varepsilon^2}}$$

and the single-system entropy contribution by

$$\gamma \leq 2 \max_i \sqrt{\mathrm{rank}(\rho_{A_i})} + 1 \,.$$

For the case of independent cq-states in Hilbert spaces with the same dimensions, we obtain

Corollary 2 *For $i = 1, \ldots, n$, let $\rho_{X_i B_i}$ be cq-states over (copies of) the same space $\mathcal{H}_X \otimes \mathcal{H}_B$. Then for every $\varepsilon > 0$ and $n \geq \frac{8}{5} \log \frac{2}{\varepsilon^2}$,*

$$\mathrm{H}_\infty^\varepsilon \left(X^n | B^n\right)_{\otimes_{i=1}^n \rho_{X_i B_i}} \geq \sum_{i=1}^n \mathrm{H}(X|B)_{\rho_{X_i B_i}} - \delta n \,, \tag{4}$$

where $\delta := \sqrt{\frac{\log(2/\varepsilon^2)}{n}} \, 4 \log(2\sqrt{\dim \mathcal{H}_X} + 1)$.

We use the properties of the smooth min-entropy to prove the following two lemmas. These lemmas show that the (smooth) min-entropy of two independent strings can be split.

Lemma 4 *Let $\varepsilon \geq 0$, and let $\rho_{X_0 E_0}, \rho_{X_1 E_1}$ be two independent cq-states with*

$$\mathrm{H}_\infty^{\varepsilon^2}(X_0 X_1 | E_0 E_1) \geq \alpha \,.$$

Additionally, let S_0, S_1 be classical random variables distributed over $\{0,1\}^\ell$. Then, there exists a random variable $D \in \{0,1\}$ such that $\mathrm{H}_\infty^\varepsilon(X_{\overline{D}} D S_D | E_0 E_1) \geq \alpha/2$.

Proof.   From the additivity of smooth min-entropy (Lemma 2) it follows that we can split the min-entropy as

$$\mathrm{H}_\infty^\varepsilon \left(X_0 | E_0\right) + \mathrm{H}_\infty^\varepsilon \left(X_1 | E_1\right) \geq \mathrm{H}_\infty^{\varepsilon^2} \left(X_0 X_1 | E_0 E_1\right) \geq \alpha \,,$$

and therefore, there exists $D \in \{0,1\}$ such that

$$\mathrm{H}_\infty^\varepsilon \left(X_{\overline{D}} D S_D | E_0 E_1\right) \geq \alpha/2 \,,$$

where we used the monotonicity of smooth min-entropy (Lemma 3).   □

Lemma 5 *Let $\varepsilon \geq 0$. Let $\rho_{XE} = \bigotimes_{i=0}^{m-1} \rho_{X_i E_i}$ be a cq-state consisting of m independent cq-substates such that $H_\infty^{\varepsilon^2}(X_i X_j | E) \geq \alpha$ for all $i \neq j$. Then there exists a random variable V over $\{1, \ldots, m\}$ such that for any $v \in \{1, \ldots, m\}$ with $P[V \neq v] > 0$*

$$H_\infty^\varepsilon (X_v | E_v V, V \neq v) \geq \alpha/2 - \log(m) \,.$$

Proof. Let $V \in \{0, \ldots, m-1\}$ be the index which achieves the minimum of $H_\infty^\varepsilon (X_i | E_i)$, i.e. $H_\infty^\varepsilon (X_V | E_V) = \min_i H_\infty^\varepsilon (X_i | E_i)$. By the additivity of smooth min-entropy (Lemma 2), we have for all $v \neq V$,

$$\alpha \leq H_\infty^{\varepsilon^2}(X_v X_V | E) \leq H_\infty^\varepsilon (X_v | E_v) + H_\infty^\varepsilon (X_V | E_V) \,.$$

It follows that $H_\infty^\varepsilon (X_v | E_v, V \neq v) \geq \alpha/2$. The chain rule (Lemma 1) then leads to the claim.
$\square$

3.3   Tools

We also require the following technical results. This lemma is well-known, see [2] or [32] for a proof.

Lemma 6 (Chernoff's inequality) *Let $X_1, \ldots, X_n$ be identically and independently distributed random variables with Bernoulli distribution, i.e. $X_i = 1$ with probability p and $X_i = 0$ with probability $1 - p$. Then $S := \sum_{i=1}^{n} X_i$ has a binomial distribution with parameters (n, p) and it holds that*

$$\Pr\left[\,|S - pn| > \varepsilon n\,\right] \leq 2e^{-2\varepsilon^2 n} \,.$$

Privacy Amplification   The OT protocol makes use of two-universal hash functions. These hash functions are used for privacy amplification similar as in quantum key distribution. A class $\mathcal{F}$ of functions $f : \{0,1\}^n \to \{0,1\}^\ell$ is called two-universal, if for all $x \neq y \in \{0,1\}^n$ and $f \in \mathcal{F}$ chosen uniformly at random from $\mathcal{F}$, we have $\Pr[f(x) = f(y)] \leq 2^{-\ell}$ [6]. The following theorem expresses how the application of hash functions can increase the privacy of a random variable X given a quantum adversary holding ρ_E, the function F and a classical random variable U:

Theorem 3 ([36, 12]) *Let $\mathcal{F}$ be a class of two-universal hash functions from $\{0,1\}^n$ to $\{0,1\}^\ell$. Let F be a random variable that is uniformly and independently distributed over $\mathcal{F}$, and let ρ_{XUE} be a ccq-state. Then, for any $\varepsilon \geq 0$,*

$$d(F(X) | F, U, E) \leq 2^{-\frac{1}{2}(H_\infty^\varepsilon (X | UE) - \ell) - 1} + \varepsilon \,.$$

4   1-2 Oblivious Transfer

4.1   Security Definition and Protocol

In this section we prove the security of a randomized version of 1-2 OT (Theorem 4) from which we can easily obtain 1-2 OT. In such a randomized 1-2 OT protocol, Alice does not input two strings herself, but instead receives two strings S_0, $S_1 \in \{0,1\}^\ell$ chosen uniformly at

random. Randomized OT (ROT) can easily be converted into OT. After the ROT protocol is completed, Alice uses her strings S_0, S_1 obtained from ROT as one-time pads to encrypt her original inputs $\hat{S}_0$ and $\hat{S}_1$, i.e. she sends an additional classical message consisting of $\hat{S}_0 \oplus S_0$ and $\hat{S}_1 \oplus S_1$ to Bob. Bob can retrieve the message of his choice by computing $S_C \oplus (\hat{S}_C \oplus S_C) = \hat{S}_C$. He stays completely ignorant about the other message $\hat{S}_{\overline{C}}$ since he is ignorant about $S_{\overline{C}}$. The security of a quantum protocol implementing ROT is formally defined in [12] and justified in [17] (see also [43]).

Definition 1 *An ε-secure 1-2 ROT$^\ell$ is a protocol between Alice and Bob, where Bob has input $C \in \{0,1\}$, and Alice has no input.*

- *(Correctness) If both parties are honest, then for any distribution of Bob's input C, Alice gets outputs $S_0, S_1 \in \{0,1\}^\ell$ which are ε-close to uniform and independent of C and Bob learns $Y = S_C$ except with probability ε.*

- *(Security against dishonest Alice) If Bob is honest and obtains output Y, then for any cheating strategy of Alice resulting in her state ρ_A, there exist random variables S_0' and S_1' such that $\Pr[Y = S_C'] \geq 1 - \varepsilon$ and C is independent of S_0', S_1' and ρ_A.*[e]

- *(Security against dishonest Bob) If Alice is honest, then for any cheating strategy of Bob resulting in his state ρ_B, there exists a random variable $D \in \{0,1\}$ such that $d(S_{\overline{D}} | S_D D \rho_B) \leq \varepsilon$.*

For convenience, we choose $\{+, \times\}$ instead of $\{0,1\}$ as domain of Bob's choice bit C. We consider the same protocol for ROT as in [12].

Protocol 1 ([12]) *1-2 ROT$^\ell$*

1. *Alice picks $X \in_R \{0,1\}^n$ and $\Theta \in_R \{+, \times\}^n$. Let $\mathcal{I}_b = \{i \mid \Theta_i = b\}$ for $b \in \{+, \times\}$. At time $t = 0$, she sends $|X_1\rangle_{\Theta_1}, \ldots, |X_n\rangle_{\Theta_n}$ to Bob.*

2. *Bob measures all qubits in the basis corresponding to his choice bit $C \in \{+, \times\}$. He obtains outcome $X' \in \{0,1\}^n$.*

3. *Alice picks two hash functions $F_+, F_\times \in_R \mathcal{F}$, where $\mathcal{F}$ is a class of two-universal hash functions. At the reveal time $t = T_{\text{rev}}$, she sends $\mathcal{I}_+, \mathcal{I}_\times$, $F_+, F_\times$ to Bob. Alice outputs $S_+ = F_+(X_{|\mathcal{I}_+})$ and $S_\times = F_\times(X_{|\mathcal{I}_\times})$.*[f]

4. *Bob outputs $S_C = F_C(X'_{|\mathcal{I}_C})$.*

4.2   Security Analysis

We show in this section that Protocol 1 is secure according to Definition 1, in case the dishonest receiver is restricted to individual-storage attacks.

---

[e] Existence of the random variables S_0', S_1' has to be understood as follows: given the cq-state ρ_{YA} of honest Bob and dishonest Alice, there exists a cccq-state $\rho_{Y S_0' S_1' A}$ such that tracing out the registers of S_0', S_1' yields the original state ρ_{YA} and the stated properties hold.

[f] If $X_{|\mathcal{I}_b}$ is less than n bits long Alice pads the string $X_{|\mathcal{I}_b}$ with 0's to get an n bit-string in order to apply the hash function to n bits.
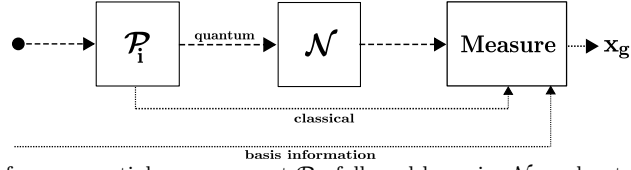
Fig. 2. Bob performs a partial measurement $\mathcal{P}_i$, followed by noise $\mathcal{N}$, and outputs a guess bit x_g depending on his classical measurement outcome, the remaining quantum state, and the additional basis information.

Correctness First of all, note that it is clear that the protocol fulfills its task correctly. Bob can determine the string $X_{|\mathcal{I}_C}$ (except with negligible probability 2^{-n} the set $\mathcal{I}_C$ is non-empty) and hence obtains S_C. Alice's outputs $S_+, S_\times$ are perfectly independent of each other and of C.

Security against Dishonest Alice Security holds in the same way as shown in [12]. As the protocol is non-interactive, Alice never receives any information from Bob at all, and Alice's input strings can be extracted by letting her interact with an unbounded receiver.

Security against Dishonest Bob Proving that the protocol is secure against Bob requires more work. Our goal is to show that there exists a $D \in \{+, \times\}$ such that Bob is completely ignorant about $S_{\overline{D}}$.

Recall that in round i, honest Alice picks $X_i \in_R \{0,1\}$ and $\Theta_i \in_R \{+, \times\}$ and sends $|X_i\rangle_{\Theta_i}$ to dishonest Bob. Bob can subsequently do a partial measurement to obtain the classical outcome K_i and store the remaining quantum state in register E_i which is then subject to noise (see Figure 2). It is important to note that Bob's initial partial measurement does not depend on the basis information Θ. Since we are modeling individual-storage attacks, the overall state (as viewed by Bob) for Alice and Bob right before time T_{rev} is of the form

$$\rho_{X\Theta KE} = \bigotimes_{i=1}^{n} \rho_{X_i\Theta_i K_i E_i},$$

with

$$\rho_{X_i\Theta_i K_i E_i} = \frac{1}{4}\sum_{x_i,\theta_i,k_i} p_{k_i|x_i\theta_i} \underbrace{|x_i\rangle\langle x_i|}_{X_i} \otimes \underbrace{|\theta_i\rangle\langle\theta_i|}_{\Theta_i} \otimes \underbrace{|k_i\rangle\langle k_i|}_{K_i} \otimes \underbrace{\mathcal{N}_i\left(\rho_{x_i\theta_i}^{k_i}\right)}_{E_i}, \qquad (5)$$

where we use X_i to denote Alice's system corresponding to her choice of bit x_i, Θ_i for the system corresponding to her choice of basis θ_i, and K_i and E_i for Bob's systems corresponding to the classical outcome k_i (with probability $p_{k_i|x_i\theta_i}$) of his partial measurement and his remaining quantum system respectively.

It is clear that a dishonest receiver will have some uncertainty about the bit X_i, given that he either measured the register E without the correct basis information and that storage noise occurred on the post-measurement quantum state. To formalize this uncertainty, let us call t an *uncertainty lower bound* on the conditional Shannon entropy if, for all $i = 1, \ldots, n$, we have

$$\mathrm{H}(X_i|\Theta_i K_i E_i) = \mathrm{H}(\rho_{X_i\Theta_i K_i E_i}) - \mathrm{H}(\rho_{\Theta_i K_i E_i}) \geq t. \qquad (6)$$

The parameter t thereby depends on the specific kind of noise in the quantum storage. In Section 5, we evaluate the uncertainty lower-bound t for the case of depolarizing noise.

The following theorem shows that as long as $\ell \lesssim tn/4$, the protocol is secure except with probability ε. Since we are performing 1-out-of-2 oblivious transfer of ℓ-bit strings, ℓ corresponds to the "amount" of oblivious transfer we can perform for a given security parameter ε and number of qubits n. In QKD, ℓ corresponds to the length of the key generated.

Theorem 4 *Protocol 1 is 2ε-secure against a dishonest receiver Bob according to Definition 1, if $n \geq \frac{8}{5} \log \frac{2}{\varepsilon^4}$ and*

$$\ell \leq \frac{1}{4} \left(t - \delta \right) n + \frac{1}{2} - \log \left(\frac{1}{\varepsilon} \right) \, ,$$

where $\delta = 8\sqrt{\log(2/\varepsilon^4)/n}$, and t is the uncertainty lower bound on the conditional Shannon entropy fulfilling Eq. (6).

Proof.   We need to show the existence of a binary random variable D such that $S_{\bar{D}}$ is ε-close to uniform from Bob's point of view. As noted above, the overall state of Alice and Bob before time T_{rev} has a product form. After time T_{rev}, dishonest Bob holds the classical registers Θ, K, the quantum register E as well as classical information about the hash functions $F_+, F_\times$. To prove security, we first lower-bound Bob's uncertainty about X in terms of min-entropy, use Lemma 4 to obtain D and then apply the privacy amplification theorem.

First of all, we know from Corollary 2 that the smooth min-entropy of an n-fold tensor state is roughly equal to n times the von Neumann entropy of its substates. Hence, applying Corollary 2 to our setting with $B_i := \Theta_i K_i E_i$ and $\log(2\sqrt{\dim \mathcal{H}_{X_i}} + 1) = \log(2\sqrt{2} + 1) \leq 2$ we obtain for $n \geq \frac{8}{5} \log \frac{2}{\varepsilon^4}$ that

$$\mathrm{H}_\infty^{\varepsilon^2} (X|\Theta KE) \geq \sum_{i=1}^n \mathrm{H}(X_i|\Theta_i K_i E_i) - \delta n \geq (t - \delta)n \, ,$$

with $\delta = 8\sqrt{\log(2/\varepsilon^4)/n}$. We used Equation (4) in the first inequality and the last follows by Definition (6) of the uncertainty bound t.

For ease of notation, we use X_+ and $X_\times$ to denote $X_{|\mathcal{I}_+}$ and $X_{|\mathcal{I}_\times}$, the sequences of bits X_i where $\Theta_i = +$ and $\Theta_i = \times$, respectively. From $\mathrm{H}_\infty^{\varepsilon^2} (X_+ X_\times|\Theta KE) \geq (t - \delta)n$ and Lemma 4 it follows that $D \in \{+, \times\}$ exists such that

$$\mathrm{H}_\infty^\varepsilon (X_{\bar{D}} D S_D|\Theta KE) \geq (t - \delta)\frac{n}{2} \, .$$

The rest of the security proof is analogous to the derivation in [12]: It follows from the chain rule (Lemma 1) and the monotonicity (Lemma 3) of the smooth min-entropy that

$$\mathrm{H}_\infty^\varepsilon (X_{\bar{D}}|\Theta D S_D KE) \geq \mathrm{H}_\infty^\varepsilon (X_{\bar{D}} D S_D|\Theta KE) - (\ell + 1)$$
$$\geq (t - \delta)\frac{n}{2} - 1 - \ell.$$

The privacy amplification Theorem 3 yields

$$d(F_{\bar{D}}(X_{\bar{D}}) \mid \Theta F_D D S_D KE) \leq 2^{-\frac{1}{2}((t-\delta)\frac{n}{2}+1-2\ell)} + \varepsilon \tag{7}$$

which is smaller than 2ε as long as

$$(t - \delta)\frac{n}{4} + \frac{1}{2} - \ell \geq \log\left(\frac{1}{\varepsilon}\right).$$

from which our claim follows.   □

We note that one can improve on the extractable length ℓ by using a quantum version of Wullschleger's distributed-privacy-amplification theorem [45]. Since this technique is specific to oblivious transfer and does not apply to our extension to the case of secure identification, we do not go into the details here.

5   Example: Depolarizing Noise

In this section, we consider the case when Bob's storage is affected by depolarizing noise as described by the quantum operation

$$\mathcal{N}(\rho) = r\rho + (1 - r)\frac{\mathsf{id}}{2}. \tag{8}$$

Depolarization noise will leave the input state ρ intact with probability r, but replace it with the completely mixed state with probability $1-r$. In order to give explicit security parameters for this setting, our goal is to prove an uncertainty bound t for the conditional von Neumann entropy $\mathrm{H}(X_i|\Theta_i K_i E_i)$ as in Eq. (6). Exploiting the symmetries in the setting, we derive in Appendix B the following result. We drop the index i in this Theorem.

Theorem 5 *Let $\mathcal{N}$ be the depolarizing quantum operation given by Eq. (8) and let $\mathrm{H}(X|\Theta KE)$ be the conditional von Neumann entropy of one qubit. Then*

$$\mathrm{H}(X|\Theta KE) \geq \left\{ \begin{array}{ll} h(\frac{1+r}{2}) & \textit{for } r \geq \hat{r}\,, \\ 1/2 & \textit{for } r < \hat{r}\,, \end{array} \right.$$

where $\hat{r} := 2h^{-1}(1/2) - 1 \approx 0.7798$.

Our result shows that when the probability of retaining the input state ρ is small, $r < 0.7798$, the best attack for Bob is to measure everything right away in the computational basis. For this measurement, we have $\mathrm{H}(X|\Theta KE) \geq 1/2$. If the depolarizing rate is low, i.e. $r \geq 0.7798$, our result says that the best strategy for Bob is to simply store the qubit as is.

Our result may seem contradictory to our previous error trade-off obtained in [41], where Bob's best strategy was to either store the qubit as is or measure it in the *Breidbart basis* depending on the amount of depolarizing noise. Note, however, that the quantity we optimize in this work is the *von Neumann entropy* and not the guessing probability considered in [41]. This phenomenon is similar to the setting of QKD, where Eve's strategy that optimizes her guessing probability is different from the one that optimizes the entropy [18]. In general, the von Neumann entropy is larger than the min-entropy (which corresponds to the guessing probability). Corollary 2 provides the explanation why the von Neumann entropy is the relevant quantity in the setting of individual-storage attacks.

6   Robust Oblivious Transfer

In a practical setting, honest Alice and honest Bob are not able to perform perfect quantum operations or transmit qubits through a noiseless channel. We must therefore modify the ROT protocol to make it robust against noise for the honest parties. The protocol we consider is a small modification of the protocol considered in [39]. The idea is to let Alice send additional error-correcting information which can help honest Bob to retrieve S_C as desired. The main difficulty in the analysis of the extended protocol is the fact that we have to assume a worst-case scenario: If Bob is dishonest, we give him access to a perfect noise-free quantum channel with Alice and he only experiences noise during storage.

We can divide the noise on the channel into two categories, which we consider separately: First, we consider erasure noise (in practice corresponding to photon loss) during preparation, transmission and measurement of the qubits by the honest parties. Let $1 - p_{\mathrm{erase}}$ be the total probability for an honest Bob to measure and detect a photon in the $\{+, \times\}$-basis given that an honest Alice prepares a weak pulse in her lab and sends it to him. The probability p_{erase} is determined, among other things, by the mean photon number in the pulse, the loss on the channel and the quantum efficiency of the detector. In our protocol we assume that the erasure rate p_{erase} is *independent* for every pulse and independent of whether qubits were encoded or measured in the $+$- or $\times$-basis whenever Bob is honest. This assumption is necessary to guarantee the correctness and the security against a *cheating Alice* only. Fortunately, this assumption is well matched with the possible physical implementations of the protocol.

Any other noise source during preparation, transmission and measurement can be characterized as an effective classical noisy channel resulting in the output bits X' that Bob obtains at Step 3 of Protocol 2. For simplicity, we model this compound noise source as a classical binary symmetric channel acting independently on each bit of X. Typical noise sources for polarization-encoded qubits are depolarization during transmission, dark counts in Bob's detector and misaligned polarizing beam-splitters. Let the effective bit-error probability, called the quantum bit-error rate in quantum key distribution, of this binary symmetric channel be $p_{\mathrm{error}} < 1/2$.

6.1   Protocol

In this section we present the modified version of the ROT protocol. Before engaging in the actual protocol, Alice and Bob agree on a small enough security-error probability $\varepsilon > 0$ that they are willing to tolerate. In addition, they determine the system parameters p_{erase} and p_{error} similarly to Step 1 of the protocol in [3]. Furthermore, they agree on a family $\{C_n\}$ of linear error-correcting codes of length n capable of efficiently correcting $n \cdot p_{\mathrm{error}}$ errors [8]. For any string $x \in \{0,1\}^n$, error-correction is done by sending the syndrome information $syn(x)$ to Bob from which he can correctly recover x if he holds an output $x' \in \{0,1\}^n$ obtained by flipping each bit of x independently with probability p_{error}. It is known that for large enough n, the code C_n can be chosen such that its rate is arbitrarily close to $1 - h(p_{\mathrm{error}})$ and the syndrome length (the number of parity check bits) is asymptotically bounded by $|syn(x)| < h(p_{\mathrm{error}})n$ [8]. We assume that the players have synchronized clocks. In each time slot, Alice sends one qubit to Bob.

Protocol 2 *Robust 1-2 $ROT^\ell(C, T, \varepsilon)$*

1. *Alice picks $X \in_R \{0,1\}^n$ and $\Theta \in_R \{+, \times\}^n$.*

2. *For $i = 1, \ldots, n$: In time slot $t = i$, Alice sends $|X_i\rangle_{\Theta_i}$ as a phase- or polarization-encoded weak pulse of light to Bob.*

3. *In each time slot, Bob measures the incoming qubit in the basis corresponding to his choice bit $C \in \{+, \times\}$ and records whether he detects a photon or not. He obtains some bit-string $X' \in \{0,1\}^m$ with $m \le n$.*

4. *Bob reports back to Alice in which time slots he received a qubit. Alice restricts herself to the set of $m < n$ bits that Bob did not report as missing. Let this set of qubits be S_{remain} with $|S_{\text{remain}}| = m$.*

5. *Let $\mathcal{I}_b = \{i \in S_{\text{remain}} \mid \Theta_i = b\}$ for $b \in \{+, \times\}$ and let $m_b = |\mathcal{I}_b|$. Alice aborts the protocol if either m_+ or $m_\times$ are outside the interval $[(1-p_{\text{erase}}-\varepsilon)n/2, (1-p_{\text{erase}}+\varepsilon)n/2]$. If this is not the case, Alice picks two two-universal hash functions $F_+, F_\times \in_R \mathcal{F}$. At time $t = n + T_{\text{rev}}$, Alice sends $\mathcal{I}_+, \mathcal{I}_\times$, $F_+, F_\times$, and the syndromes $syn(X_{|\mathcal{I}_+})$ and $syn(X_{|\mathcal{I}_\times})$ according to codes of appropriate length m_b to Bob. Alice outputs $S_+ = F_+(X_{|\mathcal{I}_+})$ and $S_\times = F_\times(X_{|\mathcal{I}_\times})$.*

6. *Bob uses $syn(X_{|\mathcal{I}_C})$ to correct the errors on his output $X'_{|\mathcal{I}_C}$. He obtains the corrected bit-string X_{cor} and outputs $S'_C = F_C(X_{\text{cor}})$.*

6.2   Security Analysis

Correctness   By assumption, p_{erase} is independent for every pulse and independent of the basis in which Alice sent the qubits. Thus, by Chernoff's Inequality (Lemma 6), S_{remain} is, except with negligible probability, a random subset of m qubits independent of the value of Θ and such that $(1 - p_{\text{erase}} - \varepsilon)n \le m \le (1 - p_{\text{erase}} + \varepsilon)n$. This implies that in Step 5 the protocol is aborted with a probability only exponentially small in n. The codes are chosen such that Bob can decode except with negligible probability. These facts imply that if both parties are honest, the protocol is correct (i.e. $S_C = S'_C$) with exponentially small probability of error.

Security against Dishonest Alice   Even though in this scenario Bob *does* communicate to Alice, the information about which qubits were erased is (by assumption) independent of the basis in which he measured and thus of his choice bit C. Hence Alice does not learn anything about his choice bit C. Her input strings can be extracted as in the analysis of Protocol 1.

Security against Dishonest Bob   We prove the following:

Theorem 6 *Protocol 2 is secure against a dishonest receiver Bob with error of at most 2ε, if $n \ge \frac{8}{5} \log \frac{2}{\varepsilon^4}$ and*

$$\ell \le (t - \delta - h(p_{\text{error}})) (1 - p_{\text{erase}})\frac{n}{4} - \varepsilon \frac{n}{2} + \frac{1}{2} - \log\left(\frac{1}{\varepsilon}\right), \tag{9}$$

where $\delta = 8\sqrt{\log(2/\varepsilon^4)/((1 - p_{\text{erase}} - \varepsilon)n)}$, and t is the uncertainty bound on the conditional Shannon entropy fulfilling Eq. (6).

Proof (Sketch).   First of all, we note that Bob can always make Alice abort the protocol by reporting back an insufficient number of received qubits. If Alice does not abort the protocol in Step 5, we have that $(1 - p_{\text{erase}} - \varepsilon)n/2 \leq m_+, m_\times \leq (1 - p_{\text{erase}} + \varepsilon)n/2$. We define D as in the security proof of Protocol 1. The security analysis is the same, but we need to subtract the amount of error correcting information $|syn(X_{|\mathcal{I}_{\overline{D}}})|$ from the entropy of the dishonest receiver. If Alice does not abort the protocol in Step 5, we have that $|syn(X_{|\mathcal{I}_{\overline{D}}})| \leq h(p_{\text{error}})(1 - p_{\text{erase}} + \varepsilon)n/2$. Hence,

$$
\begin{aligned}
&\mathrm{H}_\infty^\varepsilon \left(X_{\overline{D}} | \Theta F_D D S_D syn(X_{|\mathcal{I}_{\overline{D}}}) K E \right) \\
&\geq \mathrm{H}_\infty^\varepsilon \left(X_{\overline{D}} D S_D syn(X_{|\mathcal{I}_{\overline{D}}}) | \Theta F_D K E \right) - (\ell + 1) - h(p_{\text{error}})m/2 \\
&\geq (t - \delta)(1 - p_{\text{erase}} - \varepsilon)n/2 - (\ell + 1) - h(p_{\text{error}})(1 - p_{\text{erase}} + \varepsilon)n/2 - 1 - \ell \\
&\geq (t - \delta - h(p_{\text{error}})(1 - p_{\text{erase}})n/2 - \underbrace{(t - \delta + h(p_{\text{error}}))}_{\leq 2} \varepsilon n/2 - 1 - \ell \,,
\end{aligned}
$$

where $(t - \delta + h(p_{\text{error}})) \leq 2$ since $t \leq 1$. Using this inequality to bound the security parameter via the privacy amplification Theorem 3 gives the claimed bound on ℓ, Eq. (9).   □

Remarks   Note that it is only possible to choose a code C that satisfies the stated parameters *asymptotically.* For a real—finite block-length—code, deviations from this asymptotic behavior need to be taken into account. For the sake of clarity we have omitted these details in the analysis above. Secondly, the dishonest parties need to obtain an estimate for p_{error} prior to the protocol. One approach would be to use a worst case estimate based what is possible with present-day technology. Alternatively, one could follow Step 1 of the protocol in [3] as suggested above. However, one needs to analyze this estimation procedure in a practical setting. Thirdly, when weak photon sources are used in this protocol, one needs to analyze the security threat due to the presence of multi-photon emissions which Bob can exploit in photon-number-splitting attacks as in QKD. See [41] for a first discussion of the effect of such attacks.

6.3   *Depolarizing Noise*

As an example, we again consider the security trade-off when Bob's storage is affected by depolarizing noise. It follows directly from Theorems 3, 6 and 5 that

Corollary 7 *Let $\mathcal{N}$ be the depolarizing quantum operation given by Eq. (8). Then the protocol can be made secure (by choosing a sufficiently large n) as long as*

$$
\begin{aligned}
h\left(\frac{1 + r}{2}\right) &> h(p_{\text{error}}) & \text{for } r \geq \hat{r}, \\
1/2 &> h(p_{\text{error}}) & \text{for } r < \hat{r},
\end{aligned}
$$

where $\hat{r} := 2h^{-1}(1/2) - 1 \approx 0.7798.$

Hence, our security parameters are greatly improved from our previous analysis [41]. For $r < \hat{r}$ we can now obtain security as long as the quantum bit error rate $p_{\text{error}} \lesssim 0.11$, compared to 0.029 before. For the case of $r \geq \hat{r}$, we can essentially show security as long as the noise on the channel is strictly less than the noise in Bob's quantum storage. Note that we cannot hope to construct a protocol that is both correct and secure when the noise of the channel exceeds the noise in Bob's quantum storage. However, it remains an open question whether it is possible to construct a protocol or improve the analysis of the current protocol such that security can be achieved even for very small n.

Corollary 7 puts a restriction on the noise rate of the honest protocol. Yet, since our protocols are particularly interesting at short distances (e.g. in the case of secure identification we describe below), we can imagine free-space implementations over very short distances such that depolarization noise during transmission is negligible and the main noise source is due to Bob's honest measurements.

In the near-future, if good photonic memories become available (see e.g. [23, 4, 7, 15, 37, 34] for recent progress), we may anticipate that storing the qubit is a better attack than a direct measurement. Note, however, that we are free in our protocol to stretch the reveal time T_{rev} between Bob's reception of the qubits and his reception of the classical basis information, say, to seconds, which means that one has to consider the overall noise rate on a qubit that is stored for seconds.

In terms of long-term security, *fault-tolerant* photonic computation (e.g., with the KLM scheme [25]) might allow a dishonest Bob to encode the incoming quantum information into a fault-tolerant quantum memory. Such an encoding would guarantee that the effective noise rate in storage can be made arbitrarily small. The encoding of a single unknown state is *not* a fault-tolerant quantum operation however. Hence, even in the presence of a quantum computer, there is a residual storage noise rate due to the unprotected encoding operation. The question of security then becomes a question of a trade-off between this residual noise rate versus the intrinsic noise rate for honest parties. Intuitively, it might be possible to arrange the setting such that tasks of honest players are always technically easier (and/or cheaper) to perform than the ones for dishonest players. Possibly, this intrinsic gap can be exploited for cryptographic purposes. The current paper can be appreciated as a first step in this direction.

7 Extension to Secure Identification

In this section, we like to point out how our model of noisy quantum storage with individual-storage attacks also applies to protocols that achieve more advanced tasks such as secure identification. The protocol from [13] allows a user U to identify him/herself to a server S by means of a personal identification number (PIN). This task can be achieved by securely evaluating the equality function on the player's inputs. In other words, both U and S input passwords W_U and W_S into the protocol and the server learns as output whether $W_U = W_S$ or not. The protocol proposed in [13] is secure against an unbounded user U and a quantum-memory bounded server S in the sense that it is guaranteed that if a dishonest player starts with quantum side information which is uncorrelated with the honest player's password W, the

only thing the dishonest player can do is guess a possible W' and learn whether $W = W'$ or not while not learning anything more than this mere bit of information about the honest user's password W. This protocol can also be (non-trivially) extended to additionally withstand man-in-the-middle attacks.

The security proof against a quantum-memory bounded dishonest server (and man-in-the-middle attacks) relies heavily on the uncertainty relation first derived in [12] and used for proving the security of 1-2 OT. This uncertainty relation guarantees a lower bound on the smooth min-entropy of the encoded string X from the dishonest player's point of view. As we establish a similar type of lower bound (Cor. 2 and Eq. (6)) on the smooth min-entropy in the noisy-storage model, the security proof for the identification scheme (and its extension) translates to our model.

In terms of the proof of Proposition 3.1 of [13], the pair X_i, X_j has essentially $t \cdot d$ bits of min-entropy given Θ, K, and E, where t is the uncertainty lower bound on the conditional Shannon entropy from Eq. (6) and d is the minimal distance of the code used in the identification scheme. Lemma 5 implies that there exists W' (called V in Lemma 5) such that if $W \neq W'$ then X_W has essentially $td/2 - \log(m)$ bits of min-entropy given W, W', Θ, K, E. Privacy amplification then guarantees that $F(X_W)$ is ε'-close to uniform and independent of F, W, W', Θ, K, E, conditioned on $W \neq W'$, where $\varepsilon' = \frac{1}{2}2^{-\frac{1}{2}(td/2-\log(m)-\ell)}$. Security against a dishonest server with noisy quantum storage follows as in [13] for an error parameter ε which is exponentially small in $td - 2\log(m) - 2\ell$.

8   Conclusion

We have obtained improved security parameters for oblivious transfer in the noisy-quantum-storage model. Yet, it remains to prove security against general coherent noisy attacks. The problem with analyzing a coherent attack of Bob described by some quantum operation $\mathcal{S}$ affecting all his incoming qubits is not merely a technical one: one first needs to determine a realistic noise model in this setting. Symmetrizing the protocol as in the proof of QKD [36] and using de Finetti type arguments does not immediately work here. However, one can analyze a specific type of coherent noise, one that essentially corresponds to an eavesdropping attack in QKD. Note that the 1-2 OT protocol can be seen as two runs of QKD interleaved with each other. The strings $f(x_{|\mathcal{I}_+})$ and $f(x_{|\mathcal{I}_\times})$ are then the two keys generated. The noise must be such that it leaves Bob with exactly the same information as the eavesdropper Eve in QKD. In this case, it follows from the security of QKD that the dishonest Bob (learning exactly the same information as the eavesdropper Eve) does not learn anything about the two keys.

Clearly, there is a strong relation between QKD and the protocol for 1-2 OT, and one may wonder whether other QKD protocols can be used to perform oblivious transfer in our model. Intuitively, this is indeed the case, but it remains to evaluate explicit parameters for the security of the resulting protocols.

It will be interesting to extend our results to a security analysis of a noise-robust protocol in a realistic physical setting, where, for example, the use of weak laser pulses allows the possibility of photon-number-splitting attacks. Such a comprehensive security analysis has been carried out in [20] for quantum key distribution.

Acknowledgments

References

1. G. Alexei, N. Langford, and M. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, 2005.
2. N. Alon and J. Spencer. *The Probabilistic Method*. Series in Discrete Mathematics and Optimization. Wiley-Interscience, 2nd edition, 2000.
3. C. H. Bennett, G. Brassard, C. Crépeau, and M.-H. Skubiszewska. Practical quantum oblivious transfer. In *CRYPTO '91: Proceedings of the 11th Annual International Cryptology Conference on Advances in Cryptology*, pages 351–366. Springer-Verlag, 1992.
4. A. D. Boozer, A. Boca, R. Miller, T. E. Northup, and H. J. Kimble. Reversible state transfer between light and a single trapped atom, 2007. quant-ph/0702248.
5. S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
6. J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and System Sciences*, 18:143–154, 1979.
7. T. Chanelière, D.N. Matsukevich, S.D. Jenkins, S.-Y. Lan, T.A.B. Kennedy, and A. Kuzmich. Storage and retrieval of single photons between remote quantum memories. *Nature*, 438:pp. 833–836, 2005.
8. C. Crépeau. Efficient cryptographic protocols based on noisy channels. In *Advances in Cryptology – Proceedings of EUROCRYPT '97*, 1997.
9. C. Crépeau and J. Kilian. Achieving oblivious transfer using weakened security assumptions. In *Proceedings of 29th IEEE FOCS*, pages 42–52, 1988.
10. C. Crépeau, K. Morozov, and S. Wolf. Efficient unconditional oblivious transfer from almost any noisy channel. In *International Conference on Security in Communication Networks (SCN)*, volume 4 of *Lecture Notes in Computer Science*, pages 47–59, 2004.
11. I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the Bounded-Quantum-Storage Model. In *Proceedings of 46th IEEE FOCS*, pages 449–458, 2005.
12. I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner. A tight high-order entropic quantum uncertainty relation with applications. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 360–378. Springer-Verlag, 2007.
13. I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Secure identification and QKD in the bounded-quantum-storage model. In *Advances in Cryptology—CRYPTO '07*, volume 4622 of *Lecture Notes in Computer Science*, pages 342–359. Springer-Verlag, 2007.
14. I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner. Cryptography in the bounded-quantum-storage model. *SIAM Journal on Computing*, 37(6):1865–1890, 2008.
15. M.D. Eisaman, A. André, F. Massou, M. Fleischauer, A.S. Zibrov, and M. D. Lukin. Electromagnetically induced transparency with tunable single-photon pulses. *Nature*, 438:pp. 837–841, 2005.
16. S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *Communications of the ACM*, 28(6):637–647, 1985.
17. S. Fehr and C. Schaffner. Composing quantum protocols in a classical environment. In *Theory of Cryptography—TCC '09*, volume 5444 of *Lecture Notes in Computer Science*, pages 350–367. Springer, 2009.
18. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74:145–195, 2002.
19. O. Goldreich and R. Vainish. How to solve any protocol problem - an efficiency improvement. In *Advances in Cryptology — CRYPTO '87*, volume 293 of *Lecture Notes in Computer Science*,

pages 73–86. Springer, 1988.

20. D. Gottesman, H.-K. Lo, N. Lutkenhaus, and J. Preskill. Security of quantum key distribution with imperfect devices. *Quant. Inf. Comp*, 5:325–360, 2004. quant-ph/0212066.

21. M. Hayashi. *Quantum Information - An introduction*. Springer, 2006.

22. R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

23. B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurasek, and E. S. Polzik. Experimental demonstration of quantum memory for light. *Nature*, 432:pp. 482–485, 2004.

24. J. Kilian. Founding cryptography on oblivious transfer. In *Proceedings of 20th ACM STOC*, pages 20–31, 1988.

25. E. Knill, R. Laflamme, and G. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, 2001.

26. R. König, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Transactions on Information Theory*, 2009. arxiv:0807.1338

27. H-K. Lo. Insecurity of quantum secure computations. *Physical Review A*, 56:1154, 1997.

28. H-K. Lo and H. F. Chau. Is quantum bit commitment really possible? *Physical Review Letters*, 78:3410, 1997.

29. H-K. Lo and H.F. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. In *Proceedings of PhysComp96*, 1996. quant-ph/9605026.

30. D. Mayers. The trouble with quantum bit commitment. quant-ph/9603015, 1996.

31. D. Mayers. Unconditionally secure quantum bit commitment is impossible. *Physical Review Letters*, 78:3414–3417, 1997.

32. R. Motwani and R. Prabhakar. *Randomized Algorithms*. Cambridge University Press, 1995.

33. M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

34. T. B. Pittman and J. D. Franson. Cyclical quantum memory for photonic qubits. *Phys. Rev. A*, 66(6):062302, Dec 2002.

35. M. Rabin. How to exchange secrets by oblivious transfer. Technical report, Aiken Computer Laboratory, Harvard University, 1981. Technical Report TR-81.

36. R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005. quant-ph/0512258.

37. W. Rosenfeld, S. Berner, J. Volz, M. Weber, and H. Weinfurter. Remote preparation of an atomic quantum memory. *Physical Review Letters*, 98:0505004, 2007.

38. L. Salvail. Quantum bit commitment from a physical assumption. In *Proceedings of CRYPTO'98*, volume 1462 of *Lecture Notes in Computer Science*, pages 338–353, 1998.

39. C. Schaffner. *Cryptography in the Bounded-Quantum-Storage Model*. PhD thesis, University of Aarhus, 2007. arxiv:0709.0289.

40. M. Tomamichel, R. Colbeck, and R. Renner. A fully quantum asymptotic equipartition property. arxiv:0811.1221, 2008.

41. S. Wehner, C. Schaffner, and B. Terhal. Cryptography from noisy photonic storage. arxiv:0711.2895, 2007.

42. S. Wehner, C. Schaffner, and B. M. Terhal. Cryptography from noisy storage. *Physical Review Letters*, 100(22):220502, 2008.

43. S. Wehner and J. Wullschleger. Security in the bounded-quantum-storage model. In *ICALP 2008*, 2008.

44. S. Wiesner. Conjugate coding. *Sigact News*, 15(1), 1983.

45. J. Wullschleger. Oblivious-transfer amplification. In *Advances in Cryptology — EUROCRYPT '07*, Lecture Notes in Computer Science. Springer-Verlag, 2007.

Appendix A

In this Appendix we provide the technical proofs of the Lemmas and the Theorem in Section 3.2.1. We restate the claims for convenience.

Proof of Lemma 1 (Chain Rule)

Lemma A.1 (Chain Rule) *For any ccq-state $\rho_{XYE} \in \mathcal{S}(\mathcal{H}_{XYE})$ and for all $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^\varepsilon (X|YE) \geq \mathrm{H}_\infty^\varepsilon (XY|E) - \log |\mathcal{Y}|,$$

where $|\mathcal{Y}|$ is the alphabet size of the random variable Y.

Proof. For $\varepsilon = 0$, it follows from Eq. (3) that we need to show that

$$\mathrm{P}_{\mathrm{guess}}(XY|E) \geq \mathrm{P}_{\mathrm{guess}}(X|YE) \cdot \frac{1}{|\mathcal{Y}|} . \tag{A.1}$$

For a given value y, let $\{M_x^y\}_x$ be the POVM on register E which optimally guesses X given Y. A particular strategy of guessing X *and* Y from E is to guess a value of y uniformly at random from $\mathcal{Y}$ and subsequently measure E with the POVM $\{M_x^y\}_x$. The success probability of this strategy is exactly the r.h.s of (A.1). Clearly, the optimal guessing probability $\mathrm{P}_{\mathrm{guess}}(XY|E)$ can only be better than this particular strategy. For $\varepsilon > 0$, let $\hat{\rho}_{XYE} \in \mathcal{K}^\varepsilon(\rho_{XYE})$ be the state in the ε-ball around ρ_{XYE} that maximizes the min-entropy $\mathrm{H}_\infty^\varepsilon (XY|E)$. The technique from Remark 3.2.4 in [36] can be used to show that $\hat{\rho}_{XYE}$ is a ccq-state. By the derivation above for $\varepsilon = 0$, we obtain that

$$\mathrm{P}_{\mathrm{guess}}(XY|E)_{\hat{\rho}} \geq \mathrm{P}_{\mathrm{guess}}(X|YE)_{\hat{\rho}} \cdot \frac{1}{|\mathcal{Y}|}$$

$$\geq \min_{\tilde{\rho}_{XYE} \in \mathcal{K}^\varepsilon(\rho_{XYE})} \mathrm{P}_{\mathrm{guess}}(X|YE)_{\tilde{\rho}} \cdot \frac{1}{|\mathcal{Y}|} ,$$

which proves the lemma by taking the negative logarithms and using Eq. (3). $\qquad\square$

Proof of Lemma 2 (Additivity)

To show additivity of the smooth min-entropy we will employ semidefinite programming, where we refer to [5] for in-depth information. Here, we will use semidefinite programming in the language of [26] to express the primal and dual optimization problem given by parameters $c \in \mathcal{V}_1$ and $b \in \mathcal{V}_2$ in vector spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$. We will optimize over variables $v_1 \in K_1$ and $v_2 \in K_2$, where $K_1 \subset \mathcal{V}_1$ and $K_2 \subset \mathcal{V}_2$ are convex cones in the respective vector spaces. In our application below, these will simply be the cones of positive-semidefinite matrices. We can then write

$$\gamma^{\mathrm{primal}} = \min_{\substack{v_1 \geq 0 \\ Av_1 \geq b}} \langle v_1, c \rangle_1 \qquad \text{and} \qquad \gamma^{\mathrm{dual}} = \max_{\substack{v_2 \geq 0 \\ A^* v_2 \leq c}} \langle b, v_2 \rangle_2, \tag{A.2}$$

where $A : \mathcal{V}_1 \to \mathcal{V}_2$ is a linear map defining the particular problem we wish to solve. We use $A^* : \mathcal{V}_2 \to \mathcal{V}_1$ to denote its dual map satisfying

$$\langle Av_1, v_2 \rangle_2 = \langle v_1, A^* v_2 \rangle_1 \qquad \text{for all } v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2 .$$

Note that we have $\gamma^{\mathrm{primal}} \geq \gamma^{\mathrm{dual}}$ by weak duality. In this case our SDPs will be strongly feasible, giving us $\gamma^{\mathrm{primal}} = \gamma^{\mathrm{dual}}$ known as strong duality. Our proof is based on the same

idea as [41, Lemma 2] applied to the smoothed setting. We thank Robert König for allowing us to include the following.

Lemma A.2 (Additivity (König and Wehner)) *Let ρ_{AB} and $\rho_{A'B'}$ be two independent qq-states. For $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^{\varepsilon^2}(AA'|BB') \leq \mathrm{H}_\infty^\varepsilon(A|B) + \mathrm{H}_\infty^\varepsilon(A'|B') \, .$$

Proof.   In order to prove additivity, it is important to realize that the smooth conditional min-entropy can be written as semi-definite program:

$$
\begin{aligned}
\mathrm{H}_\infty^\varepsilon(A|B) &= \max_{\hat\rho_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB})} \mathrm{H}_\infty(A|B)_{\hat\rho} \\
&= \max_{\hat\rho_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB})} -\log \min_{\substack{\sigma_B \geq 0 \\ \rho_{A\hat B C} \geq 0 \\ \mathsf{id}_A \otimes \sigma_B \geq \hat\rho_{AB}}} \mathrm{Tr}(\sigma_B) & \text{(A.3)} \\
&= -\log \min_{\substack{\sigma_B \geq 0 \\ \rho_{A\hat B C} \geq 0 \\ \hat\rho_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB}) \\ \mathsf{id}_A \otimes \sigma_B \geq \hat\rho_{AB}}} \mathrm{Tr}(\sigma_B) \, . & \text{(A.4)}
\end{aligned}
$$

where $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ throughout. Let $|\psi\rangle_{ABC}$ be a purification of ρ_{AB}. Then, all states $\hat\rho_{AB} \in \mathcal{K}^\varepsilon(\rho_{AB})$ can be obtained by an extension $\hat\rho_{ABC} \geq 0$ such that $\mathrm{Tr}(\hat\rho_{ABC}) \leq 1$, and $\mathrm{Tr}(\hat\rho_{ABC}|\psi_{ABC}\rangle\langle\psi_{ABC}|) \geq 1 - \delta$ with $\delta = \varepsilon^2$. Therefore, we can write

$$\mathrm{H}_\infty^\varepsilon(A|B) = -\log \min_{\substack{\mathrm{Tr}(\hat\rho_{ABC}|\psi_{ABC}\rangle\langle\psi_{ABC}|) \geq 1-\delta \\ 1 \geq \mathrm{Tr}(\hat\rho_{ABC}) \\ \mathsf{id}_A \otimes \sigma_B \geq \hat\rho_{AB}}} \mathrm{Tr}(\sigma_B) \, ,$$

where the minimum is taken over all $\sigma_B \in \mathcal{P}(\mathcal{H}_B)$ and $\hat\rho_{ABC} \in \mathcal{P}(\mathcal{H}_{ABC})$, which is a semi-definite program (SDP). Our goal will be to determine the dual of this semidefinite program which will then allow us to put an upper bound on the smooth min-entropy as desired.

We now first show how to convert the primal of this semidefinite program into the form of Eq. (A.2). Let $\mathcal{V}_1 = \mathrm{Herm}(\mathcal{H}_B) \oplus \mathrm{Herm}(\mathcal{H}_{ABC})$ where $\mathrm{Herm}(\mathcal{H})$ is the (real) vector space of Hermitian operators on $\mathcal{H}$. Let $K_1 \subset \mathcal{V}_1$ be the cone of positive semi-definite operators. Let $c = \mathsf{id}_B \oplus 0_{ABC}$ where 0_{ABC} is the zero-operator on $\mathcal{H}_{ABC}$. Let the inner product be defined as $\langle v_1, v_1' \rangle_1 = \mathrm{Tr}(v_1^\dagger v_1')$. Note that this allows us to express our objective function as

$$\langle \sigma_B \oplus \hat\rho_{ABC}, c \rangle_1 = \mathrm{Tr}(\sigma_B).$$

It remains to rewrite the constraints in the appropriate form. To this end, we need to define $\mathcal{V}_2 = \mathbb{R} \oplus \mathbb{R} \oplus \mathrm{Herm}(\mathcal{H}_A) \oplus \mathrm{Herm}(\mathcal{H}_B)$, $K_2 \subset \mathcal{V}_2$ the cone of positive semi-definite operators and take the inner product to have the same form $\langle v_2, v_2' \rangle_2 = \mathrm{Tr}(v_2^\dagger v_2')$. We then let $b \in \mathcal{V}_2$ be given as

$$b = (1 - \delta) \oplus (-1) \oplus 0_{AB} \, ,$$

and define the map

$$A(\sigma_B \oplus \hat\rho_{ABC}) = \mathrm{Tr}(\hat\rho_{ABC}|\psi_{ABC}\rangle\langle\psi_{ABC}|) \oplus (-\mathrm{Tr}(\hat\rho_{ABC})) \oplus (\mathsf{id}_A \otimes \sigma_B - \hat\rho_{AB}) \, .$$

Note that $v_1 = \sigma_B \oplus \hat{\rho}_{ABC} \geq 0$ and $A(v_1) \geq b$ now exactly represent our constraints.

We now use this formalism to find the dual. Note that we may write any $v_2 \in \mathcal{V}_2$ with $v_2 \geq 0$ as $v_2 = r \oplus s \oplus Q_{AB}$ where $Q_{AB} \in \mathcal{P}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $r, s \in \mathbb{R}$. To find the dual map A^* note that

$$\langle Av_1, v_2 \rangle_2 = r\text{Tr}(\hat{\rho}_{ABC}|\psi_{ABC}\rangle\langle\psi_{ABC}|) - s\text{Tr}(\hat{\rho}_{ABC}) + \text{Tr}(Q_{AB}(\text{id}_A \otimes \sigma_B - \hat{\rho}_{AB}))$$
$$= r\text{Tr}(\hat{\rho}_{ABC}|\psi_{ABC}\rangle\langle\psi_{ABC}|) - s\text{Tr}(\hat{\rho}_{ABC}) + \text{Tr}(Q_B\sigma_B) - \text{Tr}((Q_{AB} \otimes \text{id}_C)\hat{\rho}_{ABC}),$$

and we therefore have

$$A^*(v_2) = (0_B \oplus r|\psi_{ABC}\rangle\langle\psi_{ABC}|) - (s\text{id}_{ABC}) + (Q_B \oplus 0_{ABC}) - (0_B \oplus Q_{AB} \otimes \text{id}_C) ,$$

which is all we require using Eq. (A.2). To find a more intuitive interpretation of the dual note that $A^*(v_2) \leq c$ is equivalent to

$$\text{id}_B \geq Q_B , \tag{A.5}$$
$$Q_{AB} \otimes \text{id}_C \geq r|\psi_{ABC}\rangle\langle\psi_{ABC}| - s\text{id}_{ABC} , \tag{A.6}$$

and $\langle b, v_2 \rangle_2 = r(1 - \delta) - s$. The dual can thus be written as

$$\gamma^{\text{dual}} = \max_{\substack{r \geq 0, s \geq 0 \\ \text{id}_B \geq Q_B \\ Q_{AB} \otimes \text{id}_C \geq r|\psi_{ABC}\rangle\langle\psi_{ABC}| - s\text{id}_{ABC}}} r(1 - \delta) - s .$$

We now use the dual formulation to upper bound the smooth min-entropy of the combined state $\rho_{AB} \otimes \rho_{A'B'}$ and parameter $\tilde{\delta}$ by finding a lower bound to the dual semidefinite program. Let $\tilde{\gamma}(\tilde{\delta})$ denote the optimal solution of the dual of the SDP for the combined state for error $\tilde{\delta}$. For each individual state, we may solve the above SDP, where we let Q_{AB}, r and s denote the optimal solution for state ρ_{AB} with parameter δ and optimal value $\gamma(\delta)$, and let $Q_{A'B'}, r'$ and s' denote the optimal solution for state $\rho_{A'B'}$ with parameter δ' and optimal value $\gamma(\delta')$. We now use these solutions to construct a solution (not necessarily the optimal one) for the combined state $\rho_{AB} \otimes \rho_{A'B'}$. Let $\tilde{Q} = Q_{AB} \otimes Q_{A'B'}$, $\tilde{r} = rr'$ and $\tilde{s} = rs'(1-\delta) + sr'(1-\delta') - ss'$. Note that $rs' \geq 0$ and $r'(1 - \delta') - s' \geq 0$ for the optimal r', s' and hence

$$\tilde{r} \geq 0 , \ \tilde{s} \geq 0 ,$$
$$\text{id}_{BB'} \geq \tilde{Q}_{BB'} ,$$
$$\tilde{Q}_{AA'BB'} \otimes \text{id}_{CC'} \geq (r|\psi_{ABC}\rangle\langle\psi_{ABC}| - s\,\text{id}_{ABC}) \otimes (r'|\psi_{A'B'C'}\rangle\langle\psi_{A'B'C'}| - s\,\text{id}_{A'B'C'})$$
$$\geq \tilde{r}|\psi_{ABC}\rangle\langle\psi_{ABC}| \otimes |\psi_{A'B'C'}\rangle\langle\psi_{A'B'C'}| - \tilde{s}\,\text{id}_{ABC} \otimes \text{id}_{A'B'C'} ,$$

and thus $\tilde{Q}$ is indeed a feasible solution for the combined problem. Choosing $\tilde{\delta}$ as

$$\tilde{\delta} = \delta + \delta' - \delta\delta'$$

we have

$$\tilde{\gamma}(\tilde{\delta}) \geq \tilde{r}(1 - \tilde{\delta}) - \tilde{s} = \gamma(\delta)\gamma'(\delta') .$$

We hence obtain

$$\mathrm{H}_\infty^{\sqrt{\tilde{\delta}}}(\tilde{A}|\tilde{B}) \leq \mathrm{H}_\infty^{\sqrt{\delta}}(A|B) + \mathrm{H}_\infty^{\sqrt{\delta'}}(A'|B') \ .$$

For $\delta = \delta'$, we have

$$\tilde{\delta} = 2\delta - \delta^2 \geq \delta^2 \ .$$

Putting everything together we thus have

$$\mathrm{H}_\infty^{\delta}(\tilde{A}|\tilde{B}) \leq \mathrm{H}_\infty^{\sqrt{\delta}}(A|B) + \mathrm{H}_\infty^{\sqrt{\delta}}(A'|B') \ ,$$

from which the result follows since $\delta = \varepsilon^2$. $\hfill \square$

Proof of Lemma 3 (Monotonicity)

Lemma A.3 (Monotonicity) *For a ccq-state ρ_{XYE} and for all $\varepsilon \geq 0$, it holds that*

$$\mathrm{H}_\infty^\varepsilon (XY|E) \geq \mathrm{H}_\infty^\varepsilon (Y|E) \ .$$

Proof.   For $\varepsilon = 0$, the lemma follows from Eq. (3), that is, guessing XY from E is harder than guessing only Y from E and therefore, $\mathrm{P}_{\mathrm{guess}}(XY|E) \leq \mathrm{P}_{\mathrm{guess}}(Y|E)$.

For $\varepsilon > 0$ the idea behind the argument is similar. Let the maximum in $\mathrm{H}_\infty^\varepsilon (Y|E)$ be achieved by a density matrix $\hat{\rho}_{YE}$, i.e. $\mathrm{H}_\infty^\varepsilon (Y|E) = \mathrm{H}_\infty(Y|E)_{\hat{\rho}}$ such that $C(\rho_{YE}, \hat{\rho}_{YE}) \leq \varepsilon$ and $\mathrm{Tr}(\hat{\rho}_{YE}) \leq 1$. Remark 3.2.4 in [36] shows that $\hat{\rho}_{YE}$ is a cq-state. We can express this min-entropy in terms of the guessing probability, Eq. (3), and thus

$$\mathrm{H}_\infty^\varepsilon (Y|E)_{\hat{\rho}} = -\log \mathrm{P}_{\mathrm{guess}}(Y|E)_{\hat{\rho}_{YE}} \leq -\log \mathrm{P}_{\mathrm{guess}}(XY|E)_{\hat{\rho}_{XYE}} \ , \tag{A.7}$$

where $\hat{\rho}_{XYE}$ is any ccq-state which has $\hat{\rho}_{YE}$ as its reduced state, i.e $\mathrm{Tr}_X(\hat{\rho}_{XYE}) = \hat{\rho}_{YE}$. Now we would like to show that one can choose an extension $\hat{\rho}_{XYE}$ such that $C(\rho_{XYE}, \hat{\rho}_{XYE}) = \sqrt{1 - F(\rho_{XYE}, \hat{\rho}_{XYE})^2} \leq \varepsilon$ and $\mathrm{Tr}(\hat{\rho}_{XYE}) \leq 1$. If we can determine such an extension, we can upper-bound the r.h.s. in Eq. (A.7) by $\mathrm{H}_\infty^\varepsilon (XY|E)$ which is the supremum of $-\log \mathrm{P}_{\mathrm{guess}}(XY|E)$ over states in the ε-neighborhood of ρ_{XYE}. This would prove the Lemma.

Let $|\Psi\rangle_{XYEC}$ be a purification of ρ_{XYE} and hence also a purification of ρ_{YE}. By Uhlmann's theorem (see e.g. [33]), we have for the fidelity $F(\rho_{YE}, \hat{\rho}_{YE})$ between ρ_{YE} and $\hat{\rho}_{YE}$ that

$$F(\rho_{YE}, \hat{\rho}_{YE}) = \max_{|\Psi'\rangle_{XYEC}} |\langle\Psi|\Psi'\rangle| := F(|\Psi\rangle\langle\Psi|, |\hat{\Psi}\rangle\langle\hat{\Psi}|) \ ,$$

where $|\hat{\Psi}\rangle_{XYEC}$ is the purification of $\hat{\rho}_{YE}$ achieving the maximum. The monotonicity property of the fidelity under taking the partial trace gives

$$F(|\Psi\rangle\langle\Psi|, |\hat{\Psi}\rangle\langle\hat{\Psi}|) \leq F(\mathrm{Tr}_C(|\Psi\rangle\langle\Psi|), \mathrm{Tr}_C(|\hat{\Psi}\rangle\langle\hat{\Psi}|)) = F(\rho_{XYE}, \hat{\rho}_{XYE}) \ ,$$

where $\hat{\rho}_{XYE} := \mathrm{Tr}_C(|\hat{\Psi}_{XYEC}\rangle\langle\hat{\Psi}_{XYEC}|)$. Hence

$$\sqrt{1 - \varepsilon^2} \leq F(\rho_{YE}, \hat{\rho}_{YE}) \leq F(\rho_{XYE}, \hat{\rho}_{XYE}) \ , \tag{A.8}$$

and therefore, $C(\rho_{XYE}, \hat{\rho}_{XYE}) \leq \varepsilon$. If $\mathrm{Tr}(\hat{\rho}_{XYE}) > 1$, it follows that also $\mathrm{Tr}(\hat{\rho}_{YE}) > 1$ which contradicts the assumption. Therefore, it must be the case that $\mathrm{Tr}(\hat{\rho}_{XYE}) \leq 1$.

It remains to show that $\hat{\rho}_{XYE}$ is a ccq-state. Because of

$$
\begin{aligned}
F(\rho_{YE}, \hat{\rho}_{YE}) &= F(|\Psi\rangle\langle\Psi|, |\hat{\Psi}\rangle\langle\hat{\Psi}|) \\
&\leq F(\mathrm{Tr}_C(|\Psi\rangle\langle\Psi|), \mathrm{Tr}_C(|\hat{\Psi}\rangle\langle\hat{\Psi}|)) = F(\rho_{XYE}, \hat{\rho}_{XYE}) \leq F(\rho_{YE}, \hat{\rho}_{YE}),
\end{aligned}
$$

these quantities are all equal and in particular, we could do a measurement on the X-register of $\hat{\rho}_{XYE}$ without increasing the fidelity. Hence, we can assume the optimal purification $|\hat{\Psi}_{XYEC}\rangle\langle\hat{\Psi}_{XYEC}|$ is such that $\hat{\rho}_{XYE}$ is a ccq-state. $\qquad\square$

Proof of Theorem 1

Theorem A.1 *For $i = 1, \ldots, n$, let $\rho_i \in \mathcal{S}(\mathcal{H}_{AB})$ be density operators. Then, for any $\varepsilon > 0$,*

$$
\mathrm{H}_\infty^\varepsilon (A^n|B^n)_{\otimes_{i=1}^n \rho_i} \geq \sum_{i=1}^n [\mathrm{H}(A_i|B_i)_{\rho_i}] - \delta(\varepsilon, \gamma)\sqrt{n},
$$

where, for $n \geq \frac{8}{5} \log \frac{2}{\varepsilon^2}$, the error is given by

$$
\delta(\varepsilon, \gamma) := 4 \log \gamma \sqrt{\log \frac{2}{\varepsilon^2}}
$$

and the single-system entropy contribution by

$$
\gamma \leq 2 \max_i \sqrt{\mathrm{rank}(\rho_{A_i})} + 1.
$$

Proof. The proof is analogous to the proof of Theorem 7 in [40]. For convenience, we point out where their proof needs to be adapted. We need the following definitions. Let $\mathcal{H}'_{AB}$ be a copy of $\mathcal{H}_{AB}$ and let $|\gamma\rangle := \sum_i |i\rangle \otimes |i\rangle$ be the unnormalized fully entangled state on $\mathcal{H}_{AB} \otimes \mathcal{H}'_{AB}$. Define the purification $|\phi\rangle := (\sqrt{\rho_{AB}} \otimes \mathrm{id}_{AB})|\gamma\rangle$ of ρ_{AB} and let $1 < \alpha \leq 2$, $\beta := \alpha - 1$, and $X := \rho_{AB} \otimes (\mathrm{id}_A \otimes \rho_B^{-1})^T$. The conditional α-entropy is defined as $\mathrm{H}_\alpha(A|B)_{\rho|\sigma} := \frac{1}{1-\alpha} \log \mathrm{Tr}(\rho_{AB}^\alpha (\mathrm{id}_A \otimes \sigma_B)^{1-\alpha})$. The authors of [40] prove the following lower bound

$$
\mathrm{H}_\alpha(A|B)_{\rho|\rho} \geq H(A|B)_\rho - \frac{1}{\beta \ln 2}\langle\phi|r_\beta(X)|\phi\rangle, \tag{A.9}
$$

where $r_\beta(t) := t^\beta - \beta \ln t - 1$.

Let $\hat{\rho} = \rho_{AB}^1 \otimes \ldots \otimes \rho_{AB}^n$. Then, as in Equation (27) of [40], we have

$$
\begin{aligned}
\mathrm{H}_\infty^\varepsilon (A^n|B^n)_{\hat{\rho}} &\geq \mathrm{H}_\infty^\varepsilon (A^n|B^n)_{\hat{\rho}|\hat{\rho}} \geq \mathrm{H}_\alpha(A^n|B^n)_{\hat{\rho}|\hat{\rho}} - \frac{1}{\beta} \log \frac{2}{\varepsilon^2} \\
&= \sum_{i=1}^n \mathrm{H}_\alpha(A|B)_{\rho^i|\rho^i} - \frac{1}{\beta} \log \frac{2}{\varepsilon^2} \\
&\geq \sum_{i=1}^n \left(\mathrm{H}(A|B)_{\rho^i} - \frac{1}{\beta \ln 2}\langle\phi|r_\beta(X^i)|\phi\rangle\right) - \frac{1}{\beta} \log \frac{2}{\varepsilon^2},
\end{aligned}
$$

$$\tag{A.10}$$

where we used (A.9) in the last step.

Let us define the single-system entropy contributions $\gamma^i := \langle\phi|\sqrt{X^i} + 1/\sqrt{X^i} + \mathsf{id}|\phi\rangle$ of which we know that they are all ≥ 3 and let $\gamma_{\max}$ be the largest of them. By choosing an appropriate $\mu \geq 0$ such that

$$\beta = \frac{1}{2\mu\sqrt{n}} \leq \sqrt{\frac{5}{8}}\,\frac{1}{2\log\gamma_{\max}} \leq \min\left\{\frac{1}{4}, \frac{1}{2\log\gamma_{\max}}\right\}\,,$$

we can bound

$$\frac{1}{\beta\ln 2}\langle\phi|r_\beta(X^i)|\phi\rangle \leq \frac{2}{\mu\sqrt{n}}\log^2(\gamma^i) \leq \frac{2}{\mu\sqrt{n}}\log^2(\gamma_{\max})$$

Therefore, we can further lower bound (A.10) as

$$\mathrm{H}^\varepsilon_\infty\left(A^n|B^n\right)_{\hat\rho} \geq \sum_{i=1}^n H(A|B)_{\rho^i} - \sum_{i=1}^n \frac{2}{\mu\sqrt{n}}\log^2(\gamma_{\max}) - 2\mu\sqrt{n}\log\frac{2}{\varepsilon^2}$$

$$\geq \sum_{i=1}^n H(A|B)_{\rho^i} - 2\sqrt{n}\left(\frac{1}{\mu}\log^2(\gamma_{\max}) + \mu\log\frac{2}{\varepsilon^2}\right)\,.$$

and the rest of the derivation goes as after Equation (28) in [40].

In order to obtain the upper bound on γ, we notice that $H_{1/2}(A|B)_{\rho|\rho} \leq H_{1/2}(A)_\rho \leq H_0(A)_\rho = \log(\mathrm{rank}(\rho_A))$.   □

Appendix B

In this appendix, we use the symmetries inherent in our problem to prove Theorem 5 in a series of steps.

Theorem B.1 *Let $\mathcal{N}$ be the depolarizing quantum operation given by Eq. (8) and let $\mathrm{H}(X|\Theta KE)$ be the conditional von Neumann entropy of one qubit. Then*

$$\mathrm{H}(X|\Theta KE) \geq \left\{\begin{array}{ll} h(\frac{1+r}{2}) & \text{for } r \geq \hat{r}\,, \\ 1/2 & \text{for } r < \hat{r}\,, \end{array}\right.$$

where $\hat{r} := 2h^{-1}(1/2) - 1 \approx 0.7798$.

Setting the Stage

In order to prove the theorem, we find Bob's strategy which minimizes $\mathrm{H}(X|\Theta KE)$ as a function of the depolarizing noise parameter r. As depicted in Figure 2, in each round the dishonest receiver Bob receives one of the four possible BB84 states $\rho_{x\theta}$ at random. On such state he may then perform any (partial) measurement M given by measurement operators $M = \{F_k\}$ such that $\sum_k F_k^\dagger F_k = \mathsf{id}$. For clarity of notation, we do not use a subscript to indicate the round i as in the Figure. We denote by E the register containing the renormalized post-measurement state

$$\rho_{x\theta}^{k,M} = \frac{F_k\rho_{x,\theta}F_k^\dagger}{p_{k|x\theta}},$$

to which the depolarizing quantum operation $\mathcal{N}$ is applied. Here

$$p_{k|x\theta}^M = \text{Tr}(F_k \rho_{x\theta} F_k^\dagger)$$

is the probability to measure outcome k when given state $\rho_{x\theta}$. We omit the superscript M if it is clear which measurement is used. Note that we may write

$$p_{x\theta k}^M = \frac{1}{4} p_{k|x\theta}^M,$$

$$\sum_x p_{x\theta k}^M = p_{\theta k}^M = \frac{1}{4} \text{Tr}\left(F_k \left(\rho_{0\theta} + \rho_{1\theta}\right) F_k^\dagger\right) = \frac{1}{4} \text{Tr}(F_k F_k^\dagger),$$

and

$$p_{x|\theta k}^M = \frac{p_{k|x\theta}^M}{4 p_{\theta k}^M}.$$

Here we have used the fact that Alice chooses the basis and bit in each round uniformly and independently at random.

First of all, note that for a cq-state $\rho_{YE} = \sum_y P_Y(y)|y\rangle\langle y| \otimes \rho_y^E$, the von Neumann entropy can be expanded as

$$\text{H}(YE) = \text{H}(Y) + \sum_y P_Y(y)\,\text{H}(\rho_y^E).$$

Using this expansion, we can write

$$\text{H}(X|\Theta KE)_M$$

$$= \text{H}(X\Theta KE)_M - \text{H}(\Theta KE)_M \tag{B.1}$$

$$= \text{H}(X\Theta K)_M + \sum_{x\theta k} p_{x\theta k}^M\,\text{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right) - \text{H}(\Theta K)_M - \sum_{\theta k} p_{\theta k}^M\,\text{H}\left(\mathcal{N}\left(\sum_x p_{x|\theta k}^M \rho_{x\theta}^{k,M}\right)\right)$$

$$= \text{H}(X|\Theta K)_M + \sum_{x\theta k} p_{x\theta k}^M\,\text{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right) - \sum_{\theta k} p_{\theta k}^M\,\text{H}\left(\sum_x p_{x|\theta k}^M \mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right). \tag{B.2}$$

We use the notation $\text{H}(X|\Theta KE)_M$ to emphasize that we consider the conditional von Neumann entropy when Bob performed a partial measurement M. In the following, we use the shorthand

$$\text{B}(M) := \text{H}(X|\Theta KE)_M.$$

Using Symmetries to Reduce Degrees of Freedom

Our goal is to minimize $\text{B}(M)$ over all possible measurements $M = \{F_k\}$ as a function of r. We proceed in three steps. First, we simplify our problem considerably until we are left with a single Hermitian measurement operator over which we need to minimize the entropy. Second, we show that the optimal measurement operator is diagonal in the computational basis. And finally, we show that depending on the amount of noise, this measurement operator is either proportional to the identity, or proportional to a rank one projector.

First, we prove a property of the function $B(M)$ for a composition of two measurements. Intuitively, the following statement uses the fact that if we choose one measurement with probability α and another measurement with probability β our average success probability is the average of the success probabilities obtained via the individual measurements:

Claim 1 *Let $F = \{F_k\}_{k=1}^{f}$ and $G = \{G_k\}_{k=f+1}^{f+g}$ be two measurements. Then, for $0 \leq \alpha \leq 1$ and a combined measurement $M = \alpha F + (1-\alpha)G := \{\sqrt{\alpha}F_k\}_{k=1}^{f} \cup \{\sqrt{1-\alpha}G_k\}_{k=f+1}^{f+g}$, we have*

$$\mathrm{B}(\alpha F + (1-\alpha)G) = \alpha\,\mathrm{B}(F) + (1-\alpha)\,\mathrm{B}(G)\,.$$

Proof.   Let $F = \{F_k\}_{k=1}^{f}$ and $G = \{G_k\}_{k=1}^{g}$ be measurements, $0 \leq \alpha \leq 1$ and let $M := \{\sqrt{\alpha}F_k\}_{k=1}^{f} \cup \{\sqrt{1-\alpha}G_k\}_{k=f+1}^{f+g}$.

It is easy to verify that we have the following relations for $1 \leq k \leq f$: $p_{x\theta k}^{M} = \alpha p_{x\theta k}^{F}$, $p_{x|\theta k}^{M} = \frac{\alpha p_{x\theta k}^{F}}{\alpha p_{\theta k}^{F}} = p_{x|\theta k}^{F}$, $p_{k|x\theta}^{M} = \alpha p_{k|x\theta}^{F}$ and $\rho_{x\theta}^{k,M} = \frac{\alpha F_k \rho_{x\theta} F_k^{\dagger}}{p_{k|x\theta}^{M}} = \frac{F_k \rho_{x\theta} F_k^{\dagger}}{p_{k|x\theta}^{F}} = \rho_{x\theta}^{k,F}$ and analogously for $f+1 \leq k \leq f+g$.

We consider the three summands in Eq. (B.2) separately. For the first term we get

$$\mathrm{H}(X|\Theta K)_M = \sum_{\theta k} p_{\theta k}^{M} h\left(p_{0|\theta k}^{M}\right)$$

$$= \sum_{\theta}\sum_{k=1}^{f} \alpha p_{\theta k}^{F} h\left(p_{0|\theta k}^{F}\right) + \sum_{\theta}\sum_{k=f+1}^{f+g} (1-\alpha) p_{\theta k}^{G} h\left(p_{0|\theta k}^{G}\right)$$

$$= \alpha\,\mathrm{H}(X|\Theta K)_F + (1-\alpha)\,\mathrm{H}(X|\Theta K)_G\,.$$

For the second term, we obtain

$$\sum_{x\theta k} p_{x\theta k}^{M}\,\mathrm{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right)$$

$$= \alpha\sum_{x\theta}\sum_{k=1}^{f} p_{x\theta k}^{F}\,\mathrm{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,F}\right)\right) + (1-\alpha)\sum_{x\theta}\sum_{k=f+1}^{f+g} p_{x\theta k}^{G}\,\mathrm{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,G}\right)\right)\,.$$

The third term yields

$$\sum_{\theta k} p_{\theta k}^{M}\,\mathrm{H}\left(\sum_{x} p_{x|\theta k}^{M}\mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right)$$

$$= \alpha\sum_{\theta}\sum_{k=1}^{f} p_{\theta k}^{F}\,\mathrm{H}\left(\sum_{x} p_{x|\theta k}^{F}\mathcal{N}\left(\rho_{x\theta}^{k,F}\right)\right) + (1-\alpha)\sum_{\theta}\sum_{k=f+1}^{f+g} p_{\theta k}^{G}\,\mathrm{H}\left(\sum_{x} p_{x|\theta k}^{G}\mathcal{N}\left(\rho_{x\theta}^{k,G}\right)\right)\,.$$

$\square$

We can now make a series of observations.

Claim 2 *Let $M = \{F_k\}$ and $G = \{\mathsf{id}, X, Z, XZ\}$. Then for all $g \in G$ we have $\mathrm{B}(M) = \mathrm{B}(gMg^{\dagger})$.*

Proof.   First of all, note that for all $g \in G$, g can at most exchange the roles of 0 and 1. That is, we can perform a bit flip before the measurement which we can correct for afterwards by applying classical post-processing. Furthermore, since $g \in G$ is Hermitian and unitary we have

$$p_{\theta k}^{gMg^{\dagger}} = \frac{1}{4}\mathrm{Tr}(gF_kg^{\dagger}(\rho_{0\theta}+\rho_{1\theta})g^{\dagger}F_k^{\dagger}g) = \frac{1}{4}\mathrm{Tr}(F_kF_k^{\dagger}) = p_{\theta k}^{M}\,,$$

and hence there exists a bijection $f : \{0,1\} \to \{0,1\}$ such that

$$p_{x|\theta k}^{gMg^\dagger} = p_{f(x)|\theta k}^M \,.$$

Again, we consider the three summands in Eq. (B.2) separately. For the first term, observe that $h(p_{0|\theta k}^M) = h(p_{1|\theta k}^M)$.

$$\mathrm{H}(X|\Theta K)_{gMg^\dagger} = \sum_{\theta k} p_{\theta k}^{gMg^\dagger} h\left(p_{0|\theta k}^{gMg^\dagger}\right) = \sum_{\theta k} p_{\theta k}^M h\left(p_{0|\theta k}^M\right) = \mathrm{H}(X|\Theta K)_M \,.$$

To analyze the second term, note that we can write

$$p_{x\theta k}^{gMg^\dagger} = p_{f(x)\theta k}^M \,,$$

and for depolarizing noise $\mathcal{N}\left(U\rho U^\dagger\right) = U\mathcal{N}(\rho)U^\dagger$, in addition the von Neumann entropy itself is invariant under unitary operations $\mathrm{H}(g\mathcal{N}(\rho)g^\dagger) = \mathrm{H}(\mathcal{N}(\rho))$. Putting everything together, we obtain

$$\sum_{x\theta k} p_{x\theta k}^{gMg^\dagger} \, \mathrm{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,gMg^\dagger}\right)\right) = \sum_{x\theta k} p_{x\theta k}^M \, \mathrm{H}\left(\mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right).$$

By a similar argument, we derive the equality for the third term

$$\sum_{\theta k} p_{\theta k}^{gMg^\dagger} \, \mathrm{H}\left(\sum_x p_{x|\theta k}^{gMg^\dagger} \mathcal{N}\left(\rho_{x\theta}^{k,gMg^\dagger}\right)\right) = \sum_{\theta k} p_{\theta k}^M \, \mathrm{H}\left(\sum_x p_{x|\theta k}^M \mathcal{N}\left(\rho_{x\theta}^{k,M}\right)\right).$$

$\square$

Claim 3 *Let $G = \{\mathrm{id}, X, Z, XZ\}$. There exists a measurement operator F such that the minimum of $\mathrm{B}(M)$ over all measurements M is achieved by a measurement proportional to $\{gFg^\dagger \mid g \in G\}$.*

Proof. Let $M = \{F_k\}$ be a measurement. Let $K = |M|$ be the number of measurement operators. Clearly, $\hat{M} = \{\hat{F}_{g,k}\}$ with

$$\hat{F}_{g,k} = \frac{1}{2} g F_k g^\dagger \,,$$

is also a quantum measurement since $\sum_{g,k} \hat{F}_{g,k}^\dagger \hat{F}_{g,k} = \mathrm{id}$. It follows from Claims 1 and 2 that $\mathrm{B}(M) = \mathrm{B}(\hat{M})$. Define operators

$$N_{g,k} = \frac{1}{\sqrt{2\mathrm{Tr}(F_k^\dagger F_k)}} g F_k g^\dagger \,.$$

Note that

$$\sum_{g \in G} N_{g,k} = \frac{1}{\sqrt{2\mathrm{Tr}(F_k^\dagger F_k)}} \sum_{u,v \in \{0,1\}} X^u Z^v F_k^\dagger F_k Z^v X^u = \mathrm{id} \,.$$

(see for example Hayashi [21]). Hence $M_k = \{N_{g,k}\}$ is a valid quantum measurement. Now, note that $\hat{M}$ can be obtained from $M_1, \ldots, M_K$ by averaging. Hence, by Claim 1 we have

$$\mathrm{B}(M) = \mathrm{B}(\hat{M}) \geq \min_k \mathrm{B}(M_k).$$

Let M^* be the optimal measurement. Clearly, $m = \mathrm{B}(M^*) \geq \min_k \mathrm{B}(M_k^*) \geq m$ by the above and Claim 2 from which the present claim follows. $\qquad\square$

Finally, we note that we can restrict ourselves to optimizing over positive semi-definite (and hence Hermitian) matrices only.

Claim 4 *Let F be a measurement operator and $M^F = \{gFg^\dagger | g \in G\}$ the associated measurement. Then there exists a Hermitian operator $\hat{F}$ such that $\mathrm{B}(M^F) = \mathrm{B}(M^{\hat{F}})$.*

Proof. Let $F^\dagger = \hat{F}U$ be the polar decomposition of $F^\dagger$, where $\hat{F}$ is positive semi-definite and U is unitary [22, Corollary 7.3.3]. Evidently, since the trace is cyclic, all probabilities remain the same. Using the invariance of the von Neumann entropy and the depolarizing quantum operation under unitaries, the claim follows. $\qquad\square$

Note that Claim 3 also gives us that we have at most 4 measurement operators. Wlog, we take the measurement outcomes to be labeled $1, 2, 3, 4$ and measurement operators $F_1 = F, F_2 = XFX, F_3 = ZFZ, F_4 = XZFZX$. Our final observation is the following easy claim.

Claim 5 *For any linear operator F on Hilbert space $\mathcal{H}$ and any state $|\phi\rangle \in \mathcal{H}$ such that $F|\phi\rangle \neq 0$, it holds that the operator $P := \frac{F|\phi\rangle\langle\phi|F^\dagger}{\mathrm{Tr}(F|\phi\rangle\langle\phi|F^\dagger)}$ is a projector with $\mathrm{rank}(P) = 1$.*

Proof. Notice that $|\phi\rangle\langle\phi|F^\dagger F|\phi\rangle\langle\phi| = \mathrm{Tr}(F|\phi\rangle\langle\phi|F^\dagger)|\phi\rangle\langle\phi|$. Thus

$$PP = \frac{F|\phi\rangle\langle\phi|F^\dagger F|\phi\rangle\langle\phi|F^\dagger}{\mathrm{Tr}(F|\phi\rangle\langle\phi|F^\dagger)^2} = \frac{F|\phi\rangle\langle\phi|F^\dagger}{\mathrm{Tr}(F|\phi\rangle\langle\phi|F^\dagger)} = P.$$

As $F|\phi\rangle \neq 0$ we have that $\mathrm{rank}(F|\phi\rangle\langle\phi|F^\dagger) = 1$. $\qquad\square$

Exploiting our observations, we can considerably simplify the expression $B(M)$ to be minimized:

Lemma B.1 *Let $\mathrm{B}(M)$ be defined as above. Then*

$$\min_M \mathrm{B}(M) = \min_F C(F),$$

where the minimization is taken over Hermitian operators $F \in \mathbb{C}^{2\times 2}$ and $C(F)$ is defined as

$$C(F) = \frac{1}{2}\left(h\left(2\,\mathrm{Tr}\left(F\rho_{0+}F\right)\right) + h\left(2\,\mathrm{Tr}\left(F\rho_{0\times}F\right)\right)\right) + h\left(\frac{1+r}{2}\right) - \mathrm{H}\left(\mathcal{N}\left(2F^2\right)\right). \qquad (\mathrm{B}.3)$$

Proof. First of all, note that

$$p_{\theta k} = p_{0\theta k} + p_{1\theta k} = \frac{1}{4}\mathrm{Tr}\left(F_k\left(\rho_{0\theta} + \rho_{1\theta}\right)F_k\right) = \frac{1}{4}\mathrm{Tr}(F^2),$$

which is independent of k. Thus we have

$$\frac{1}{2} = p_\theta = \sum_{k=1}^{4} p_{\theta k} = \text{Tr}(F^2),$$

and hence $p_{\theta k} = \frac{1}{8}$. Furthermore, as in the proof of Claim 2, there exists a bijection $f : \{0,1\} \to \{0,1\}$ such that

$$p_{x|\theta k} = \frac{p_{x\theta k}}{p_{\theta k}} = \frac{\text{Tr}\left(F_k \rho_{x\theta} F_k\right)/4}{1/8} = 2\,\text{Tr}\left(F_k \rho_{x\theta} F_k\right) = 2\,\text{Tr}\left(F \rho_{f(x)\theta} F\right) .$$

Note again that $h(p_{0|\theta k}^M) = h(p_{1|\theta k}^M)$. We then obtain for the first term

$$\begin{aligned}
\text{H}(X|\Theta K) &= \sum_{\theta k} p_{\theta k} h(p_{0|\theta k}) \\
&= \sum_{\theta k} \frac{1}{8} h(2\,\text{Tr}(F_k \rho_{0\theta} F_k)) \\
&= \frac{1}{2} \left(h(2\,\text{Tr}(F\rho_{0+}F)) + h(2\,\text{Tr}(F\rho_{0\times}F)) \right) .
\end{aligned}$$

For the second term, we need to evaluate $\text{H}(\mathcal{N}(\rho_{x\theta}^k))$. It follows from Claim 5 that if $p_{x\theta k} > 0$, the normalized post-measurement state $\rho_{x\theta}^k$ has eigenvalues 0 and 1. Applying the depolarizing quantum operation to such rank 1 state gives an entropy $\text{H}(\mathcal{N}(\rho_{x\theta}^k)) = h((1+r)/2)$, independent of the state. Thus the second term becomes

$$\sum_{x\theta k} p_{x\theta k}\, \text{H}\left(\mathcal{N}\left(\rho_{x\theta}^k\right)\right) = \sum_{x\theta k} p_{x\theta k} h\left(\frac{1+r}{2}\right) = h\left(\frac{1+r}{2}\right) .$$

For the third term, we use that for $0 \le \alpha \le 1$, it holds that $\mathcal{N}(\alpha\rho + (1-\alpha)\sigma) = \alpha\mathcal{N}(\rho) + (1-\alpha)\mathcal{N}(\sigma)$. Hence,

$$\begin{aligned}
p_{0|\theta k}\mathcal{N}\left(\rho_{0\theta}^k\right) + p_{1|\theta k}\mathcal{N}\left(\rho_{1\theta}^k\right) &= \mathcal{N}\left(p_{0|\theta k}\rho_{0\theta}^k + p_{1|\theta k}\rho_{1\theta}^k\right) \\
&= \mathcal{N}\left(2\,\text{Tr}\left(F_k\rho_{0\theta}F_k^\dagger\right)\frac{F_k\rho_{0\theta}F_k^\dagger}{\text{Tr}\left(F_k\rho_{0\theta}F_k^\dagger\right)} + 2\,\text{Tr}\left(F_k\rho_{1\theta}F_k^\dagger\right)\frac{F_k\rho_{1\theta}F_k^\dagger}{\text{Tr}\left(F_k\rho_{1\theta}F_k^\dagger\right)}\right) \\
&= \mathcal{N}\left(2F_k(\rho_{0\theta} + \rho_{1\theta})F_k^\dagger\right) \\
&= U_k\mathcal{N}\left(2F^2\right)U_k^\dagger,
\end{aligned}$$

where $U_k \in G$. The third term then yields

$$\sum_{\theta k} p_{\theta k}\, \text{H}\left(\sum_x p_{x|\theta k}\mathcal{N}\left(\rho_{x\theta}^k\right)\right) = \text{H}\left(\mathcal{N}(2F^2)\right) .$$

These arguments prove the Lemma. $\qquad\square$

F is Diagonal in the Computational Basis

Now that we have simplified our problem considerably, we are ready to perform the actual optimization. We first show that we can take F to be diagonal in the computational (or Hadamard) basis.

Claim 6 *Let $F \in \mathbb{C}^{2 \times 2}$ be the Hermitian operator that minimizes $\mathrm{C}(F)$ as defined by Eq. (B.3). Then $F = \alpha |\phi\rangle\langle\phi| + \beta (\mathsf{id} - |\phi\rangle\langle\phi|)$ for some $\alpha, \beta \in \mathbb{R}$ and pure state $|\phi\rangle$ lying in the XZ plane of the Bloch sphere. (i.e. $\mathrm{Tr}(FY) = 0$).*

Proof. Since F is a Hermitian on a 2-dimensional space, we may express F as

$$F = \alpha |\phi\rangle\langle\phi| + \beta |\phi^\perp\rangle\langle\phi^\perp| \,,$$

for some state $|\phi\rangle$ and real numbers α, β. We first of all note that from $\sum_k F_k F_k = \mathsf{id}$, we obtain that

$$\sum_k \mathrm{Tr}(F_k F_k) = \sum_{g \in \{\mathsf{id}, X, Z, XZ\}} \mathrm{Tr}(gFgg^\dagger Fg^\dagger) = 4 \mathrm{Tr}(F^2) = \mathrm{Tr}(\mathsf{id}) = 2 \,,$$

and hence $\mathrm{Tr}(F^2) = \alpha^2 + \beta^2 = 1/2$. Furthermore, using that $|\phi\rangle\langle\phi| + |\phi^\perp\rangle\langle\phi^\perp| = \mathsf{id}$ gives

$$F = \alpha |\phi\rangle\langle\phi| + \beta (\mathsf{id} - |\phi\rangle\langle\phi|) \,, \tag{B.4}$$

with $\beta = \sqrt{1/2 - \alpha^2}$. Hence without loss of generality, we can consider $0 \leq \alpha \leq 1/\sqrt{2}$. The eigenvalues of $2F^2$ are $2\alpha^2$ and $1 - 2\alpha^2$. Hence, the third term of $\mathrm{C}(F)$ becomes $\mathrm{H}(\mathcal{N}(2F^2)) = h(2r\alpha^2 + (1 - r)/2)$ which does not depend on $|\phi\rangle$. We want to minimize

$$\min_F \frac{1}{2} (h(2 \mathrm{Tr}(F\rho_{0+}F)) + h(2 \mathrm{Tr}(F\rho_{0\times}F))) + h((1 + r)/2) - h(2r\alpha^2 + (1 - r)/2) \,. \tag{B.5}$$

We first parametrize the state $|\phi\rangle$ in terms of its Bloch vector

$$|\phi\rangle\langle\phi| = \frac{\mathsf{id} + \hat{x}X + \hat{y}Y + \hat{z}Z}{2} \,.$$

Since $|\phi\rangle$ is pure we can write $\hat{y} = \sqrt{1 - \hat{x}^2 - \hat{z}^2}$. Note that we may wlog assume that $0 \leq \hat{x}, \hat{z} \leq 1$, since the remaining three measurement operators are given by XFX, ZFZ, and $XZFZX$. A small calculation shows that for the encoded bit $x \in \{0, 1\}$

$$2 \mathrm{Tr}(F\rho_{x+}F) = \frac{1}{2} \left(1 + (-1)^x (4\alpha^2 - 1)\hat{z}\right) \,,$$

and similarly

$$2 \mathrm{Tr}(F\rho_{x\times}F) = \frac{1}{2} \left(1 + (-1)^x (4\alpha^2 - 1)\hat{x}\right) \,.$$

Our goal is to show that for every $0 \leq \alpha \leq 1/\sqrt{2}$, the function

$$f(\hat{z}) := h(2 \mathrm{Tr}(F\rho_{x+}F))$$

is non-increasing on the interval $0 \leq \hat{z} \leq 1$. First of all, note that $f(\hat{z}) = 1$ for $\alpha = 1/2$. We now consider the case of $\alpha \neq 1/2$. A simple computation (using Mathematica) shows that when differentiating f with respect to $\hat{z}$ we obtain

$$f'(\hat{z}) = \frac{\partial}{\partial \hat{x}} f(\hat{z}) = \frac{1}{2}(1 - 4\alpha^2) \log \left(\frac{2}{1 + \hat{z} - 4\alpha^2 \hat{z}} - 1 \right),$$

$$f''(\hat{z}) = \frac{\partial^2}{\partial \hat{x}} f(\hat{z}) = \frac{(1 - 4\alpha^2)^2}{\ln 2(\hat{z}^2(1 - 4\alpha^2)^2 - 1)}$$

Hence the function has one maximum at $\hat{z} = 0$ with $f(0) = 1$. Since $0 \leq \alpha \leq 1/\sqrt{2}$ and $\alpha \neq 1/2$ we also have that $(1 - 4\alpha^2)^2 \leq 1$ and hence $f''(\hat{z}) \leq 0$ everywhere and f is concave (though not strictly concave). Thus $f(\hat{z})$ is decreasing with $\hat{z}$.

Since we have $\hat{x}^2 + \hat{z}^2 + \hat{y}^2 = 1$ we can thus conclude that in order to minimize C(F), we want to choose $\hat{x}$ and $\hat{z}$ as large as possible and thus let $\hat{y} = 0$ from which the claim follows. $\square$

We can immediately extend this analysis to find

Claim 7 *Let F be the operator that minimizes C(F), and write F as in Eq. B.4. Then*

$$|\phi\rangle = g|0\rangle \,,$$

for some $g \in \{\mathsf{id}, X, Z, XZ\}$.

Proof. By Claim 6, we can rewrite our optimization problem as

$$\begin{aligned} \text{minimize} \quad & (f(\hat{x}) + f(\hat{z}))/2 + h((1 + r)/2) - h(2r\alpha^2 + (1 - r)/2) \\ \text{subject to} \quad & \hat{x}^2 + \hat{z}^2 = 1 \\ & 0 \leq \hat{x} \leq 1 \\ & 0 \leq \hat{z} \leq 1 \,. \end{aligned}$$

By using Lagrange multipliers we can see that for an extreme point we must have either $\hat{x} = \hat{z} = 1/\sqrt{2}$ or $\hat{x} = 0, \hat{z} = 1$ or $\hat{z} = 0, \hat{x} = 1$. From the definition of f above we can see that to minimize the expression, we want to choose the latter, from which the claim follows. $\square$

Optimality of the Trivial Strategies

We have shown that without loss of generality F is diagonal in the computational basis. Hence, we have only a single parameter left in our optimization problem. We must optimize over all operators F of the form

$$F = \alpha|\phi\rangle\langle\phi| + \sqrt{1/2 - \alpha^2}|\phi^\perp\rangle\langle\phi^\perp| \,,$$

where we may take $|\phi\rangle$ to be $|0\rangle$ or $|1\rangle$. Our aim is to show that either F is the identity, or $F = |\phi\rangle\langle\phi|$ depending on the value of r.

Claim 8 *Let F be the operator that minimizes C(F), and let $r_0 := 2h^{-1}\left(\frac{1}{2}\right) - 1$. Then $F = c\,\mathsf{id}$ (for some $c \in \mathbb{R}$) for $r \geq r_0$, and $F = |\phi\rangle\langle\phi|$ for $r < r_0$, where*

$$|\phi\rangle = g|0\rangle \,,$$

for some $g \in \{\mathsf{id}, X, Z, XZ\}$.

Proof.   We can plug $x = 0$ and $z = 1$ in the expressions in the proof of our previous claim. Thus our goal is to minimize

$$t(r, \alpha) := \frac{1}{2} \left(1 + g(1, \alpha) \right) + h \left(\frac{1 + r}{2} \right) - g(r, \alpha) \,,$$

where

$$g(r, \alpha) := h \left(\frac{1 + r}{2} - 2\alpha^2 r \right) \,.$$

Differentiating g with respect to α gives us

$$\frac{\partial}{\partial \alpha} g(r, \alpha) = 4\alpha r \left(\log \left(\frac{1 + r}{2} - 2\alpha^2 r \right) - \log \left(\frac{1 - r}{2} + 2\alpha^2 r \right) \right) \,,$$

with which we can easily differentiate t with respect to α as

$$\frac{\partial}{\partial \alpha} t(r, \alpha) = \frac{1}{2} \frac{\partial}{\partial \alpha} g(1, \alpha) - \frac{\partial}{\partial \alpha} g(r, \alpha) \,.$$

We can calculate

$$\lim_{\alpha \to 0} \frac{\partial}{\partial \alpha} t(r, \alpha) = 0$$

and

$$\frac{\partial}{\partial \alpha} t(r, 1/2) = 0 \,.$$

We thus have two extremal points. By computing the second derivative which is equal to $8(2r^2 - 1)/\ln 2$ at the point $\alpha = 1/2$, we can see that as r grows from 0 to 1, the second extreme point switches from a maximum to a minimum at $r = 1/\sqrt{2}$. Our goal is thus to determine for which r we have $t(r, 0) \leq t(r, 1/2)$. Note that shortly after the transition point $r = 1/\sqrt{2}$, we do obtain two additional maxima, but since we are interested in finding the minimum they do not contribute to our analysis. By plugging in the definition for t from above, we have that $t(r, 0) \leq t(r, 1/2)$ iff

$$\frac{1}{2} \leq h \left(\frac{1 + r}{2} \right) \,,$$

or in other words iff

$$2h^{-1} \left(\frac{1}{2} \right) - 1 \leq r \,,$$

as promised.   □

We conclude that Bob's optimal strategy, –the one which minimizes $H(X|\Theta KE)$–, is an extremal strategy, that is, he either measures his qubit in the computational basis, or he stores the qubit as is. This is the content of Theorem 5. We believe that a similar analysis can be done for the dephasing quantum operation, by first symmetrizing the noise by applying a rotation over $\pi/4$ to the input states.