# A largely self-contained and complete security proof for quantum key distribution

Marco Tomamichel[1] and Anthony Leverrier[2]

[1]Centre for Quantum Software and Information, University of Technology Sydney, Australia

[2]Inria Paris, France

In this work we present a security analysis for quantum key distribution, establishing a rigorous tradeoff between various protocol and security parameters for a class of entanglement-based and prepare-and-measure protocols. The goal of this paper is twofold: 1) to review and clarify the state-of-the-art security analysis based on entropic uncertainty relations, and 2) to provide an accessible resource for researchers interested in a security analysis of quantum cryptographic protocols that takes into account finite resource effects. For this purpose we collect and clarify several arguments spread in the literature on the subject with the goal of making this treatment largely self-contained.

More precisely, we focus on a class of prepare-and-measure protocols based on the Bennett-Brassard (BB84) protocol as well as a class of entanglement-based protocols similar to the Bennett-Brassard-Mermin (BBM92) protocol. We carefully formalize the different steps in these protocols, including randomization, measurement, parameter estimation, error correction and privacy amplification, allowing us to be mathematically precise throughout the security analysis. We start from an operational definition of what it means for a quantum key distribution protocol to be secure and derive simple conditions that serve as sufficient condition for secrecy and correctness. We then derive and eventually discuss tradeoff relations between the block length of the classical computation, the noise tolerance, the secret key length and the security parameters for our protocols. Our results significantly improve upon previously reported tradeoffs.

## 1  Introduction

Quantum key distribution (QKD) is a cryptographic task that allows two distant parties, Alice and Bob, to exchange secret keys and communicate securely over an untrusted quantum channel, provided they have access to an authenticated classical channel. The first QKD protocol, BB84, was proposed by Bennett and Brassard [1] more than three decades ago and the last 30 years have witnessed staggering experimental advances, making QKD the first quantum information technology. With the advent of quantum information theory, Ekert [2] offered a fruitful new perspective on quantum key distribution by casting it in terms of quantum entanglement and Bell nonlocality and it was quickly noted that the original BB84 protocol can be seen in this light as well [3]. This new perspective was particularly useful for the development of formal security proofs of QKD.

Formalizing the intuitive security arguments accompanying the first protocols has proven to be challenging. Early proofs by Lo and Chau [4], Shor and Preskill [5], and Mayers [6] successfully attacked the problem in the asymptotic limit of infinitely many exchanged quantum signals (and unbounded classical computing power). A later work by Koashi [7] first brought to light that security can be certified using an entropic form [8] of Heisenberg's uncertainty principle [9]. However, these works all lacked a convincing treatment of the security tradeoff in a more realistic regime where the number of exchanged signals and the classical computing power (i.e., the length of the bit strings computations are performed on) are necessarily finite, and the final secret key is thus also of finite length. As absolute security is no longer feasible, the first and most crucial question arising in this finite regime is how to properly define approximate security of a cryptographic protocol. Following developments in classical cryptography, Renner [10] extended the concept of composable security to quantum key distribution and established a first security proof for finite key lengths. This security analysis essentially established a tradeoff between different parameters of a quantum key distribution protocol:

Marco Tomamichel: marco.tomamichel@uts.edu.au, http://www.marcotom.info

Anthony Leverrier: anthony.leverrier@inria.fr, https://who.paris.inria.fr/Anthony.Leverrier

**Block length:** During the run of the protocol the two honest parties, typically called Alice and Bob, prepare, send and measure quantum signals and store the results of these measurements in bit strings that they store on their respective (classical) computers. These bit strings will eventually be used to check for the presence of an interfering eavesdropper and to compute a secret key. As these strings get longer it gets easier to guard against eavesdroppers and secret keys can be extracted more efficiently. On the other hand there are limitations on the length because we would like to start producing a secret key as early as possible[1] and because computations on longer strings get more and more difficult. The length of the bit strings used in the protocol is called the *block length*. Current experimental and commercial implementations of quantum key distribution typically work with block lengths of the order $10^5$–$10^6$ whereas block lengths of the order $10^7$–$10^8$ can be achieved, but require an extreme stability of the system during the several hours required to collect the data [11], [12].

**Key length:** The *secret key length* is the length of the secret key (in bits) that is extracted from a single block of measurement data. An ideal secret key is a uniformly random bit string perfectly correlated between Alice and Bob and independent of any information the eavesdropper might have collected after the run of the protocol. The ratio of the key length to the block length is a key performance indicator of quantum key distribution systems.

**Security parameters:** Modern security definitions for quantum key distribution rely on approximate indistinguishability from an ideal protocol, which ensures that the resulting key can be safely used in any other (secure) application. In this case, the ideal protocol either has the two parties produce an ideal secret key or an abort flag that indicates that no secret key can be extracted, either because an eavesdropper is present, or more mundanely because the quantum channel is too noisy. The *security parameter* is the distinguishing advantage between the real and the ideal protocols, given by the diamond norm distance between the two protocols. Evidently we would like this parameter to be very small and while there is no consensus on what value it should have we will take it to be $10^{-10}$ for our numerical examples.

**Robustness:** According to the notion of security discussed above, a quantum key distribution protocol can be perfectly secure and completely useless because it always aborts. As an additional requirement we thus impose that the protocol succeeds with high probability when the quantum channel is subject to noise below a specific (and realistic) threshold. This describes the *robustness* of the protocol against noise, that is, the probability that the protocol returns a nontrivial key for a given noise level. The noise model used should capture the dynamics of the quantum channel in the expected field operation; however, the exact specification of the noise model — and whether the noise is caused by an eavesdropper or just by the undisturbed operation of the channel — is independent of any security considerations and can thus be treated independently. The robustness, and more specifically the values of the channel parameters for which the robustness goes to zero, is an important figure of merit to compare the expected performance of various protocols.

The tradeoffs between these parameters have been significantly improved since Renner's proof [10], in particular by Tomamichel et al. [13] and Hayashi and Tsurumaru [14], so that the proofs are now sufficiently tight to provide security for realistic implementations of quantum key distribution. The present analysis will mostly follow the approach in the former paper [13].

So what justifies us revisiting this problem here? Firstly, we believe that presenting a complete and rigorous security proof in a single article will make the topic of finite size security more accessible to researchers in quantum cryptography. Secondly, thanks to some improvements in the technical derivation and a steamlining of the analysis, our proof yields significantly stronger tradeoff relations between security and performance parameters. It is worth noting here that strengthening theoretical tradeoff relations of a QKD protocol has rather direct implications for practical implementations as it allows for the generation of more secure key at the same noise level without any changes to the hardware. Thirdly, although all the necessary technical ingredients and conceptual insights are present in the literature, we were not able to find a concise security proof for any QKD protocol that satisfies the following two stringent criteria:

1. The protocol is able to extract a composably secure key for reasonable parameters (i.e. realistic noise levels, security parameters and block lengths that can be handled with state-of-the-art computer hardware).

2. The protocol and all the assumptions on the physical devices used in the protocol are completely specified and all aspects of the protocol are formalized, including the randomness that is required and all the communication transcripts that are produced.

---

[1]This is most relevant for implementations that suffer from a low rate of measurement events, e.g. entanglement-based implementations.

The second point may appear trivial — but we believe the absence of a complete formalization of all aspects of a protocol found in many research papers presents a major obstacle in verifying the proofs and learning about the security of quantum key distribution and quantum cryptography in general. It is in fact common in much of the present literature to fully formalize some aspects of a security proof while keeping other aspects vague and informal — and this has lead to various misconceptions.

Let us illuminate this issue with an example. It is often argued (e.g. in [15]) that collective attacks (where the eavesdropper attacks every quantum signal exchanged between Alice and Bob in the same way) are optimal for the eavesdropper using symmetry and de-Finetti arguments [16, 17]. To get such symmetry it is at some point or another used that measurements are performed in a random basis or that a random subset of raw key bits are used for parameter estimation. However, complete security proofs also must allow for the protocol to abort in case certain thresholds are not met, and one is then left to analyze the state of the system conditioned on the fact that the protocol does not abort. This conditioning will in general introduce correlations between the state held by Alice and Bob and the seeds used to choose the measurement bases and parameter estimation subset, violating the strict symmetry assumptions. Even if these correlations are weak in typical cases, they prove problematic since they allow the eavesdropper to slightly influence Alice and Bob's measurement devices[2], something which is usually explicitly forbidden in security proofs for quantum key distribution with trusted devices. Hence, many simple arguments based on symmetry or independent randomness simply do not go through without modification when the security proof is put under a microscope.

As mentioned above, the early asymptotic proofs fail with Point 1. Moreover, while Renner's analysis [10] gives bounds for finite keys, these are not sufficient to pass Point 1 since the bounds are not strong enough for realistic key lengths.[3] More recent security proofs by Tomamichel et al. [13] and Hayashi and Tsurumaru [14] satisfy Point 1, but they are not fully formalized and thus do not satisfy Point 2.[4] In fact, our requirements in Point 2 are very stringent and we are not aware of any security proof that has met this level of rigor, except arguably Renner's thesis [10]. A recent security proof for the one-sided device-independent setting [20] satisfies Point 2 but provides a key rate that is not optimal asymptotically.

**Results and Outline.**   In the present paper, we give a rigorous and largely self-contained security proof for QKD that satisfies the two conditions above. The proof is based on the security analysis in [13] and uses an entropic uncertainty relation [21] as its main ingredient. A few additional technical results and modifications of previous results are needed. We hope that our proof is accessible to all researchers interested in the security of quantum cryptography. As such, our treatment does not require the reader to have prior knowledge of various tricks and security reductions in quantum cryptography, but presumes a solid understanding of the mathematical foundations of quantum information theory.

The remainder of this manuscript is structured as follows. First we introduce necessary notation and concepts in Section 2. In Part I we analyze a class of simple entanglement-based protocols reminiscent of the BBM92 protocol [3]. Section 3 formally describes the class of protocols we are using (see also Table 2). Section 4 formally introduces our security definitions and claims. We discuss our results in Section 5 and provide a detailed security proof in Section 6. In Part II we move on to a prepare-and-measure protocol that is essentially equivalent to BB84. Section 7 formally introduces this class of protocols (see also Table 5). We discuss our results in Section 8 and provide in Section 9 the security reduction from the prepare-and-measure protocol to the entanglement-based protocol.

## 2   Formalism and notation

We will summarize some concepts necessary for our formal security proofs here, assuming that the reader is familiar with the mathematical foundations of quantum information theory. We refer to [22] for a comprehensive introduction into this mathematical toolkit. Sections 2.1–2.4 are necessary for understanding our main exposition whereas the concepts introduced in Sections 2.5 will be employed only in the security proof.

---

[2]An example is the following attack in the prepare-and-measure version of BB84: for each state sent by Alice, Eve randomly chooses to either measure it in the computational basis, or do nothing. Such an attack would only create errors when Bob measures in the Hadamard basis, and the conditioning on not aborting would therefore slightly favor measurement choices for Bob biased toward the computational basis.

[3]Unfortunately, de Finetti reductions often do not provide good bounds in practice and at least $10^5$ or $10^6$ uses of the quantum channel are typically required for the key rate to effectively become nonzero [15, 17, 18].

[4]For example, a small flaw in the formalization of the protocol in [13] has recently been pointed out by Pfister *et al.* [19]. While this does not suggest that the security guarantees in [13] are invalid, it is a stark reminder that Point 2 is often not taken seriously enough.

## 2.1 Quantum systems, states and metrics

Individual *quantum systems* and the corresponding finite-dimensional Hilbert spaces are denoted by capital letters. The dimension of the system $A$ is denoted by $|A|$. A *joint quantum system $AB$* is defined via the tensor product of the corresponding Hilbert spaces of $A$ and $B$. We use $[m]$ to denote the set $\{1, 2, \ldots, m\}$ and use $A_{[m]}$ to denote a joint quantum system comprised of quantum systems $A_1 A_2 \ldots A_m$. Similarly, if the subscript is a subset of $[m]$, we just refer to the subsystems in the subset. Let us also introduce the notation $\Pi_{m,k} := \{\pi \subset [m] : |\pi| = k\}$, the set of subsets of $[m]$ of size $k$.

We write $\mathcal{S}(A)$ to denote normalized states on $A$, i.e., positive semi-definite operators acting on the Hilbert space $A$ with unit trace. A state is called pure if it the corresponding operator has rank 1. We will employ the trace distance between states, which is defined as

$$\|\rho - \sigma\|_{\mathrm{tr}} := \sup_P \mathrm{tr}\left\{P(\rho - \sigma)\right\}, \tag{1}$$

where $P$ ranges over projectors, i.e. positive semi-definite operators with eigenvalues in $\{0, 1\}$. In particular, we have $\|\rho - \sigma\|_{\mathrm{tr}} = \frac{1}{2}\|\rho - \sigma\|_1$, where $\|\cdot\|_1$ denotes the Schatten 1-norm. The trace distance has an immediate physical interpretation [23]: for two states with trace distance $\varepsilon$, the maximum probability of distinguishing them with a single measurement equals $\frac{1}{2}(1 + \varepsilon)$.

We also collect positive semi-definite operators with trace norm not exceeding 1 on $A$ in the set $\mathcal{S}_\bullet(A)$, and call them sub-normalized states.[5] Sub-normalized states will be very convenient for technical reasons as they allow us to represent the state of quantum systems and classical events simultaneously. The following metric is very useful when dealing with sub-normalized states:

**Definition 1** (Purified distance). For $\rho_A, \sigma_A \in \mathcal{S}_\bullet(A)$, we define the *generalized fidelity*,

$$F(\rho_A, \sigma_A) := \left( \mathrm{tr}\left\{ \sqrt{\sqrt{\rho_A} \sigma_A \sqrt{\rho_A}} \right\} + \sqrt{1 - \mathrm{tr}\{\rho_A\}} \sqrt{1 - \mathrm{tr}\{\sigma_A\}} \right)^2, \tag{2}$$

and the *purified distance*, $P(\rho_A, \sigma_A) := \sqrt{1 - F(\rho_A, \sigma_A)}$.

The purified distance is a metric on sub-normalized states and satisfies [24, Lemma 2]

$$P(\rho_A, \sigma_A) \geq P\big(\mathcal{F}(\rho_A), \mathcal{F}(\sigma_A)\big) \tag{3}$$

for every completely positive (CP) trace non-increasing map $\mathcal{F}$. This means in particular that the distance contracts when we apply a quantum channel to both states. An important property of the purified distance [22, Corollary 3.1] is that for any two states $\rho_A, \sigma_A$ and any extension $\rho_{AB}$ of $\rho_A$, there exists an extension $\sigma_{AB}$ with $P(\rho_{AB}, \sigma_{AB}) = P(\rho_A, \sigma_B)$. (This property is not true in general for the trace distance.) Moreover, it is related to the trace distance as follows [24, Lemma 6]:

$$\left\| \rho_A - \sigma_A \right\|_{\mathrm{tr}} + \frac{1}{2}\left| \mathrm{tr}\{\rho_A\} - \mathrm{tr}\{\sigma_A\} \right| \leq P(\rho_A, \sigma_A). \tag{4}$$

## 2.2 Classical registers and events

We model discrete random variables by finite-dimensional quantum systems, called *registers*, with a fixed orthonormal basis. For example, let $X \in \mathcal{X}$ be a random variable with probability law $x \mapsto P_X(x)$. Then we write the corresponding quantum state as

$$\rho_X = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x|_X, \tag{5}$$

where $\{|x\rangle\}_{x \in \mathcal{X}}$ is an orthonormal basis of the space $X$. Conversely, we write $\Pr[X = x]_\rho = P_X(x)$.

More generally, the classical register might be correlated with a quantum system $A$, and this is modeled using *classical-quantum (cq) states*:

$$\rho_{XA} = \sum_{x \in \mathcal{X}} P_X(x) |x\rangle\langle x|_X \otimes \rho_{A|X=x}, \tag{6}$$

---

[5]The disk in the subscript of $\mathcal{S}_\bullet$ symbolizes the unit disk in trace norm.

where we use $\rho_{A|X=x}$ to denote the quantum state on $A$ conditioned on the register $X$ taking the value $x$. We also write $\Pr[X = x]_\rho = \mathrm{tr}\{|x\rangle\langle x|_X \rho_{XA}\} = P_X(x)$. This convention is extended to arbitrary events defined on a classical register $X$, i.e. if $\Omega : \mathcal{X} \to \{0, 1\}$ is an *event*, we write

$$\Pr[\Omega]_\rho = \sum_{x \in \mathcal{X}} P_X(x)\Omega(x) \quad \text{and} \quad \rho_{XA\wedge\Omega} = \sum_{x \in \mathcal{X}} P_X(x)\Omega(x) |x\rangle\langle x|_X \otimes \rho_{A|X=x}, \tag{7}$$

a state that is generally sub-normalized. We will also write $\rho_{XA|\Omega} = \Pr[\Omega]_\rho^{-1} \rho_{XA\wedge\Omega}$ for the conditional state. For any event $\Omega : \mathcal{X} \to \{0, 1\}$ we denote its complement on $\mathcal{X}$ by $\neg\Omega$.

## 2.3 Quantum channels and measurements

A *quantum channel* $\mathcal{E} : A \to B$ is a completely positive trace-preserving (CPTP) map that maps operators on $A$ to operators on $B$. Prime examples of quantum channels are the trace, denoted tr, and the partial trace over system $A$, denoted $\mathrm{tr}_A$. We will encounter the diamond distance between CPTP maps, which we here define as

$$\|\mathcal{E} - \mathcal{F}\|_\diamond := \sup_{\rho_{AC} \in \mathcal{S}(AC)} \|\mathcal{E}(\rho_{AC}) - \mathcal{F}(\rho_{AC})\|_{\mathrm{tr}}, \tag{8}$$

where the optimization goes over joint states on $A$ and an auxiliary system $C$, and we can assume without loss of generality that $|C| \leq |A|$. The diamond distance also inherits the physical interpretation of the trace distance: for two quantum channels with diamond distance $\varepsilon$, the maximum probability of distinguishing them by preparing a state on the input system and an ancilla system and then measuring the joint system after applying the channel equals $\frac{1}{2}(1 + \varepsilon)$.

A *generalized measurement* on $A$ is a set of linear operators $\{M_A^x\}_{x \in \mathcal{X}}$ such that

$$\sum_{x \in \mathcal{X}} (M_A^x)^\dagger (M_A^x) = 1_A, \tag{9}$$

where $1_A$ denotes the identity operator on $A$. A measurement on $A$ can be represented as a CPTP map $\mathcal{M}_{A \to X}$ that maps states on a quantum system $A$ to measurement outcomes stored in a register $X$. The measurement in (9) applied to a bipartite state $\rho_{AB}$ yields

$$\mathcal{M}_{A \to X} : \rho_{AB} \mapsto \sigma_{XB} = \sum_{x \in \mathcal{X}} |x\rangle\langle x|_X \otimes \mathrm{tr}_A \left\{ M_A^x \rho_{AB} (M_A^x)^\dagger \right\}, \tag{10}$$

where $\sigma_{XB}$ is now a (normalized) classical-quantum state. Finally, let $f : \mathcal{X} \to \mathcal{Y}$ be a function acting on two sets $\mathcal{X}$ and $\mathcal{Y}$. We denote by $\mathcal{E}_f : X \to XY$ the corresponding CPTP map

$$\mathcal{E}_f[\cdot] = \sum_{x \in \mathcal{X}} |f(x)\rangle_Y |x\rangle\langle x|_X \cdot |x\rangle\langle x|_X \langle f(x)|_Y \tag{11}$$

that acts on general quantum states. Note that we defined the map $\mathcal{E}_f$ such that the input register $X$ is kept intact and the operation is deterministic and invertible.

## 2.4 Universal hash functions

Universal hashing is used (at least[6]) twice in the analysis of the quantum key distribution protocol: first in the error correction step to ensure the correctness of the protocol (Theorem 2), and then in the privacy amplification procedure to guarantee the secrecy of the final key.

**Definition 2** (Universal$_2$ Hashing)**.** Let $\mathcal{H} = \{h\}$ be a family of functions from $\mathcal{X}$ to $\mathcal{Z}$. The family $\mathcal{H}$ is said to be *universal$_2$* if $\Pr[H(x) = H(x')] = \frac{1}{|\mathcal{Z}|}$ for any pair of distinct elements $x, x' \in \mathcal{X}$, when $H$ is chosen uniformly at random in $\mathcal{H}$.

In this work we do not need to specify any particular family of hash functions, and it suffices to note that such families of functions always exist if $|\mathcal{X}|$ and $|\mathcal{Z}|$ are powers of 2. (See, e.g., [25, 26].)

---

[6]Universal hashing is also used to provide authentication of the classical channel, but we will not discuss this issue here.

## 2.5 Conditional entropies

Conditional entropies measure the amount of uncertainty present in a random variable from the perspective of an observer with access to correlated side information. Here we are particularly interested in observers that have access to a quantum system that serves as side information, for example the eavesdropper's memory after interfering with the quantum communication during the run of a quantum key distribution protocol. The most common measure of entropy is the Shannon or von Neumann entropy, defined as $H(X)_\rho := -\sum_{x \in \mathcal{X}} P_X(x) \log P_X(x)$. However, while this entropy has various operational interpretations in the asymptotic limit of infinite repetitions of an information processing task, it is insufficient to describe finite size effects. On the other hand, smooth min- and max-entropy allow us to capture such finite size effects and share many properties with the von Neumann entropy. We will not need the full generality of the smooth entropy formalism here and instead refer to [22] for a comprehensive introduction.

Min- and max-entropy are natural generalizations of conditional Rényi entropies [27] to the quantum setting and were first proposed by Renner [10] and König et al. [28]. The conditional min-entropy captures how difficult it is for an observer with access to quantum side information to guess the content of a classical register. For a bipartite cq state $\rho_{XB} \in \mathcal{S}(AB)$, we define

$$p_{\text{guess}}(X|B)_\rho = \sup_{\{E_B^x\}} \sum_{x \in \mathcal{X}} \Pr[X = x]_\rho \operatorname{tr}\left\{ E_B^x \, \rho_{B|X=x} \left( E_B^x \right)^\dagger \right\}, \tag{12}$$

where the optimization goes over all generalized measurements on $B$.

The conditional min-entropy for a cq state is then defined as $H_{\min}(X|B)_\rho := -\log p_{\text{guess}}(X|B)_\rho$. For later convenience we introduce the measure more generally for any bipartite, potentially sub-normalized, state:

**Definition 3** (Min-entropy). For any bipartite sub-normalized state $\rho_{AB} \in \mathcal{S}_\bullet(AB)$, we define

$$H_{\min}(A|B)_\rho := \sup\left\{ \lambda \in \mathbb{R} : \exists \sigma_B \in \mathcal{S}(B) \text{ such that } \rho_{AB} \leq 2^{-\lambda} \operatorname{id}_A \otimes \sigma_B, \right\}. \tag{13}$$

Showing equivalence between this definition and the special case of cq states in (12) involves semidefinite-programming duality [28] and is outside the scope of this work. We will also encounter the max-entropy, which is a natural dual of the min-entropy in the following sense:

**Definition 4** (Max-entropy). For any bipartite sub-normalized state $\rho_{AB} \in \mathcal{S}_\bullet(AB)$, we define

$$H_{\max}(A|B)_\rho := -H_{\min}(A|C)_\rho, \tag{14}$$

where $\rho_{ABC}$ is any pure state with $\operatorname{tr}_C\{\rho_{ABC}\} = \rho_{AB}$.

The max-entropy is a measure of the size of the support of $X$. In particular, we have the following ordering of unconditional entropies:

$$H_{\min}(X)_\rho \leq H(X)_\rho \leq H_{\max}(X)_\rho \leq \log \big| \{x \in X : \Pr[X = x]_\rho > 0\} \big|, \tag{15}$$

which is a consequence of the monotonicity of the Rényi entropies [27] in the order parameter. Here and throughout this article log denotes the binary logarithm.

We will need a slight generalization of the concepts of conditional min- and max-entropy, which takes into account a ball of states close to $\rho_{AB}$ in terms of the purified distance introduced in the previous section.

**Definition 5** (Smooth Entropies). For $\rho_{AB} \in \mathcal{S}_\bullet(AB)$ and $\varepsilon \in \left[0, \sqrt{\operatorname{tr}(\rho_{AB})}\right)$, we define

$$H_{\min}^\varepsilon(A|B)_\rho := \sup_{\substack{\tilde{\rho}_{AB} \in \mathcal{S}_\bullet(AB), \\ P(\tilde{\rho}_{AB}, \rho_{AB}) \leq \varepsilon}} H_{\min}(A|B)_{\tilde{\rho}}, \qquad H_{\max}^\varepsilon(A|B)_\rho := \inf_{\substack{\tilde{\rho}_{AB} \in \mathcal{S}_\bullet(AB), \\ P(\tilde{\rho}_{AB}, \rho_{AB}) \leq \varepsilon}} H_{\max}(A|B)_{\tilde{\rho}}. \tag{16}$$

In the above definitions we can replace the supremum and infimum with a maximum and minimum, respectively. Roughly speaking, the smooth conditional min-entropy of $X$ given $B$ approximates how much randomness that is uniform for an observer with access to $B$ can be extracted from $X$. (This will be made formal when discussing the Leftover Hashing Lemma in Section 6.4.) The smooth entropies inherit the duality relation [24]. For any pure sub-normalized state $\rho_{ABC} \in \mathcal{S}_\bullet(ABC)$, we have

$$H_{\min}^\varepsilon(A|B)_\rho = -H_{\max}^\varepsilon(A|C)_\rho. \tag{17}$$

The smooth entropies also satisfy a data-processing inequality (DPI) [24, Theorem 18]. For any cq state $\rho_{XB}$ and any completely positive trace-preserving map $\mathcal{E}_{B\to C}$, we have

$$H_{\min}^{\varepsilon}(X|B)_{\rho} \leq H_{\min}^{\varepsilon}(X|C)_{\mathcal{E}(\rho)}, \quad \text{and} \quad H_{\max}^{\varepsilon}(X|B)_{\rho} \leq H_{\max}^{\varepsilon}(X|C)_{\mathcal{E}(\rho)}. \tag{18}$$

This expresses our intuition that performing any processing of the side information can at most increase our uncertainty about $X$. Moreover, we need a simple chain rule [29, Lemma 11], which states that

$$H_{\min}^{\varepsilon}(A|BX)_{\rho} \geq H_{\min}^{\varepsilon}(A|B)_{\rho} - \log|X| \tag{19}$$

where $X$ is a (classical) register of dimension $|X|$. This corroborates our intuition that an additional bit of side information on $X$ cannot decrease our uncertainty about $X$ by more than one bit.

We have defined all these quantities for sub-normalized states so that we can easily treat restrictions to events. Let $\rho_{AXBY} \in \mathcal{S}(ABXY)$ be classical on $X$ and $Y$ and let $\Omega : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be an event. Then we denote by $H_{\min}(AX \wedge \Omega|BY)_{\rho}$ the conditional min-entropy evaluated for the state $\rho_{AXBY \wedge \Omega}$. Similarly, $H_{\min}(AX \wedge \Omega|B)_{\rho}$ denotes the conditional min-entropy evaluated for the marginal $\rho_{AXB \wedge \Omega} = \mathrm{tr}_Y\{\rho_{AXBY \wedge \Omega}\}$. These states are in general sub-normalized. The same notational convention is used for (smooth) min- and max-entropy.

# Part I
# Entanglement-based protocol

## 3 Formal description of the entanglement-based protocol

We first focus on class of simple entanglement-based QKD protocols. We give an overview of the protocols in Table 2. Section 3.1 discusses the assumptions that go into our model, Section 3.2 presents the protocol parameters, and the detailed mathematical description of the individual steps follows in Section 3.3.

Let us emphasize that by *simple* protocols, we mean that we restrict our attention to protocols where the *sifting* procedure is essentially given for free, meaning that Alice and Bob are assumed to initially share a quantum state on which all the measurements are performed. Relatedly, we also do not allow for strategies where the measurement settings are biased towards a specific value.[7] As discussed in Part II, the sifting procedure can be analyzed separately under certain assumptions.

### 3.1 Assumptions of our model

Every mathematical model of physical reality requires some assumptions, and in cryptography it is important to discuss these assumptions since if they are not met by an implementation then the security guarantees derived here are also not applicable to this implementation.

**Finite-dimensional quantum systems:** We assume that Alice's and Bob's relevant quantum degrees of freedom can be effectively represented on a finite-dimensional Hilbert space. (This requirement is not strictly necessary to show security but allows us to circumvent some technical pitfalls.)

**Sealed laboratories:** We assume that the laboratories of Alice and Bob are spatially separated. This allows us to model joint quantum systems $AB$ shared between Alice and Bob as tensor products of respective local Hilbert spaces $A$ and $B$. Moreover, an easily overlooked (an in practice hard to ensure) assumption we need is that we control exactly what information is released from Alice and Bob's laboratory.

**Random seeds:** We assume that Alice has access to uniform randomness (uniformly random seeds). In practice, the seeds can be produced by a trusted quantum random number generator in Alice's lab.[8]

---

[7]Such strategies are advocated in the literature in order to increase the secret key rate by minimizing the cost of sifting [30].

[8]See, for instance, [31] for an analysis of realistic quantum random number generators explaining how randomness originating from quantum processes can be turned into ideal seeds.

| | |
|---|---|
| $\varnothing, \checkmark$ | Symbols for abort and passing, respectively |
| $M_{A_i}^{\phi,x}$ | Measurement operator acting on register $A_i$ with setting $\phi$ and outcome $x$ |
| $c_i$ | Parameter quantifying the quality of the measurement on register $A_i$ |
| $\bar{c}$ | Parameter quantifying the overall (average) quality of measurements on $A$ |
| $m$ | Total number of quantum systems shared and measured by Alice and Bob |
| pe | Parameter estimation scheme: $\text{pe} = \{k, \delta\}$ |
| $k$ | Length (in bits) of the raw key used for parameter estimation |
| $n = m - k$ | Length (in bits) of the raw key used for key distillation |
| $\delta$ | Threshold for the parameter estimation test |
| test | Test function used in the parameter estimation step |
| ec | Error correcting scheme: $\text{ec} = \{t, r, \text{synd}, \text{corr}, \mathcal{H}_{\text{ec}}\}$ |
| $r$ | Length (in bits) of the error correction syndrome |
| $t$ | Length (in bits) of the hash used for verification in the error correcting scheme |
| synd | Function computing the error syndrome |
| corr | Function that calculate the corrected string |
| $\mathcal{H}_{\text{ec}}$ | Universal$_2$ family of hash functions used in the error correcting scheme |
| pa | Privacy amplification scheme: $\text{pa} = \{\ell, \mathcal{H}_{\text{pa}}\}$ |
| $\ell$ | Length (in bits) of the final key |
| $\mathcal{H}_{\text{pa}}$ | Universal$_2$ family of hash functions used in the privacy amplification scheme |
| $A, B$ | Alice's and Bob's initial quantum system |
| $E$ | Eve's quantum memory |
| $\mathcal{M}_{A \to X \mid S}$ | Measurement map applied on register $A$ with setting $S$ and storing the result in register $X$ |
| $V, W$ | Register for Alice's and Bob's classical bits used for parameter estimation |
| $X, Y$ | Register for Alice's and Bob's classical bits used for key distillation |
| $Z$ | Register for Alice's syndrome during error correction |
| $T$ | Register containing the hash of Alice's raw key during error correction |
| $K_A, K_B$ | Register for Alice's and Bob's final keys |
| $S^\Phi$ | Seed for the choice of the measurement bases in the idealized protocol |
| $S^\Pi$ | Seed for the choice of the random subset $\pi \in \Pi_{m,k}$ used for parameter estimation |
| $S^\Xi$ | Seed for the choice of the measurement bases for the subsystems used for parameter estimation |
| $S^\Theta$ | Seed for the choice of the measurement bases for the subsystems used for key distillation |
| $S^{H_{\text{ec}}}$ | Seed for the choice of the hash function used in the error correction test |
| $S^{H_{\text{pa}}}$ | Seed for the choice of the hash function used in the privacy amplification step |
| $S$ | Register corresponding to all the seeds, $S = (S^\Phi, S^\Pi, S^\Xi, S^\Theta, S^{H_{\text{pe}}}, S^{H_{\text{ec}}}, S^{H_{\text{pa}}})$. |
| $F^{\text{pe}}$ | Flag for the parameter estimation test |
| $F^{\text{ec}}$ | Flag for the error correction test |
| $F$ | Register corresponding to all the flags, $F = (F^{\text{pe}}, F^{\text{ec}})$ |
| $C^V$ | Transcript of the register $V$ sent during parameter estimation |
| $C^Z, C^T$ | Transcripts of the registers $Z$ and $T$ sent during error correction |
| $C$ | Register containing all the communication transcripts, $C = (C^V, C^Z, C^T)$ |
| $\rho$ | Quantum state before any measurement took place |
| $\tau$ | Quantum state after the registers used for parameter estimation have been measured |
| $\sigma$ | Quantum state once Alice and Bob's quantum registers have been entirely measured |
| $\omega$ | Quantum state describing the final output of the protocol |

Table 1: Overview of the nomenclature and notation used in Part I.

**Authenticated communication channel:** We assume that Alice and Bob share an authenticated public (classical) communication channel. Everything that is communicated over this channel will be in the public domain and is thus treated as an output of the protocol. The authentication of the classical channel can be obtained with information-theoretic security by tagging every classical message [26]. A more detailed discussion of authentication for QKD is beyond the scope of the present work, and the interested reader is referred to Portmann and Renner [32, Appendix D].

**Deterministic detection:** We further assume that Alice and Bob's measurement devices always output a valid outcome, either 0 or 1. This is unrealistic in practice since it is often the case that detectors will not detect the quantum system (due to losses or imperfect detection efficiency). A simple fix is then to flip a coin and use the resulting bit as the measurement output. Unfortunately, this solution artificially decreases the robustness of the protocol beyond what is usually tolerable in a practical setting. Another much more practical solution consists in discarding these "no detection" events, but this should be done with care and requires

extra-assumptions about the measurement devices to prevent various types of side-channel attacks such as that of Lydersen et al. [33]. We will discuss this solution in more detail in Part II of this work.

**Commuting measurements:** The block length is given by a protocol parameter, $m$, which will be discussed in Section 3.2. For Alice and Bob to run a protocol with block length $m$, we need to assume that both can perform up to $m$ measurements (with either one of two possible settings) on their share of the quantum state in such a way that the order in which they do these measurements does not affect the resulting measurement outcome distribution. This is a standard assumption in the model of trusted measurement devices (by opposition to device-independent cryptography) and ensures that there are no memory effects in the measurement devices.

More formally, we assume that Alice's and Bob's share of the quantum system can be decomposed into $m$ individual quantum systems, $A \equiv A_{[m]} = A_1 A_2 \ldots A_m$ and $B \equiv B_{[m]} = B_1 B_2 \ldots B_m$ and that the measurements can be represented as operators acting on the individual subsystems. We model Alice's $i$-th measurement with setting $\phi \in \{0,1\}$ by a binary generalized measurement $\{M_{A_i}^{\phi,x}\}_{x \in \{0,1\}}$ acting on subsystem $A_i$. The index $x$ ranges over the two possible outcomes of Alice's measurement. Analogously, Bob's $i$-th measurement with setting $\phi \in \{0,1\}$ is a binary generalized measurement $\{M_{B_i}^{\phi,y}\}_{y \in \{0,1\}}$ acting on subsystem $B_i$. The index $y$ ranges over the two possible outcomes of Bob's measurement.

**Complementarity of Alice's measurements:** The exact description of the measurement devices will not be relevant for our derivations. However, we will need to assume that Alice's measurements are sufficiently complementary, a property that is encapsulated by the average overlap, $\bar{c}(m,n)$ that we introduce next. Let $m$ be the block length and $n$ the number of bits used for key extraction (see Section 3.2 for a discussion of the protocol parameters). Let us define

$$c(\{M^x\}_x, \{N^y\}_y) := \max_{x,y \in \{0,1\}} \left\| M^x (N^y)^\dagger \right\|_\infty^2, \quad \text{and} \quad c_i := c\left( \{M_{A_i}^{0,x}\}_x, \{M_{A_i}^{1,y}\}_y \right). \tag{20}$$

In an ideal physical implementation of the protocol with complementary measurements (for example in the computational and Hadamard basis), we would have $c_i = \frac{1}{2}$ for all $i \in [m]$. In realistic implementations, its value will be larger. We assume that there exists a reliable upper bound on $c_i$. More precisely, we assume that

$$\bar{c}(m,n) := \max_{\pi \in \Pi_{m,n}} \left( \prod_{i \in \pi} c_i \right)^{\frac{1}{n}} \leq \bar{c}. \tag{21}$$

We always have $c_i \in [0,1]$ and the condition $\bar{c} < 1$ is necessary to ensure secrecy. As long as the commuting measurement assumption holds, the parameter $\bar{c}$ can in principle be measured directly in an experiment — even if the operators $\{M_{A_i}^{\phi,x}\}_{\phi,x}$ are unknown.[9]

For Bob we do not need to assume a bound on the complementarity parameter. (We only need to assume that the measurements commute, as described in the previous item.)

## 3.2 Protocol parameters and overview

From an information theoretic and mathematical point of view, an entanglement-based QKD protocol is simply a completely positive trace-preserving (CPTP) map composed of local operations and classical communication (LOCC) that takes a bipartite state $\rho_{AB}$ as an input and either aborts or returns two classical binary strings, the keys, which should ideally be identical and independent of the knowledge of any third party having access to a purifying system of $\rho_{AB}$ and to the transcript of the communication performed by the protocol.

We consider protocols $\texttt{qkd\_eb}_{m,\text{pe,ec,pa}}$ that are parametrized by the block length, $m$, and the sub-protocols for parameter estimation, error correction and privacy amplification, respectively denoted by pe, ec, and pa.

- The block length, $m \in \mathbb{N}$, determines the number of individual quantum systems that are shared between Alice and Bob, and thus available to them for parameter estimation and key extraction.

- Parameter estimation is characterized by a tuple $\text{pe} = \{k, \delta\}$, where $k \in \mathbb{N}, k \leq m$ determines the number of quantum systems used for parameter estimation and $\delta \in (0, \frac{1}{2})$ is the tolerated error rate. Let us for later convenience also define $n := m - k$ to denote the number of quantum systems used for key generation.

---

[9]See, e.g., [34, 35] for an estimation strategy based on Bell tests.

<div style="border:1px solid">

$(K_A, K_B, S, C, F) = \texttt{qkd\_eb}_{m,\text{pe,ec,pa}}\left(\rho_{AB}\right)$:

**Input:** Alice and Bob are given a state $\rho_{AB}$, where $A \equiv A_{[m]}$ and $B \equiv B_{[m]}$ are comprised of $m$ quantum systems each.

**Randomization:** They agree on a random string $\Phi \in \{0,1\}^m$, a random subset $\Pi \in \Pi_{m,k}$, and random hash functions $H_{\text{ec}} \in \mathcal{H}_{\text{ec}}$ as well as $H_{\text{pa}} \in \mathcal{H}_{\text{pa}}$. The corresponding uniformly random seeds are denoted $S = (S^\Phi, S^\Pi, S^{H_{\text{ec}}}, S^{H_{\text{pa}}})$.

**Measurement:** Alice and Bob measure the $m$ quantum systems with the setting $\Phi$. They store the binary measurement outcomes in two strings, the *raw keys*. These are denoted $(X, V)$ and $(Y, W)$ for Alice and Bob, respectively. Here $V, W$ are of length $k$ and correspond to the indices in $\Pi$, whereas $X, Y$ of length $n$ correspond to indices not in $\Pi$.

**Parameter Estimation:** Alice sends $V$ to Bob, the transcript is denoted $C^V$. Bob compares $V$ and $W$. If the fraction of errors exceeds $\delta$, Bob sets the flag $F^{\text{pe}} = \varnothing$ and they abort. Otherwise he sets $F^{\text{pe}} = \checkmark$ and they proceed.

**Error Correction:** Alice sends the syndrome $Z = \text{synd}(X)$ to Bob, with transcript $C^Z$. Bob computes $\hat{X} = \text{corr}(Y, Z)$.

To verify the success of the error correction step, Alice computes the hash $T = H_{\text{ec}}(X)$ of length $t$ and sends it to Bob, with transcript $C^T$. Bob computes $H_{\text{ec}}(\hat{X})$. If it differs from $T$, he sets the flag $F^{\text{ec}} = \varnothing$ and they abort the protocol. Otherwise he sets $F^{\text{ec}} = \checkmark$ and they proceed.

**Privacy Amplification:** They compute keys $K_A = H_{\text{pa}}(X)$ and $K_B = H_{\text{pa}}(\hat{X})$ of length $\ell$.

**Output:** The output of the protocol consists of the keys $K_A$ and $K_B$, the seeds $S = (S^\Phi, S^\Pi, S^{H_{\text{ec}}}, S^{H_{\text{pa}}})$, the transcript $C = (C^V, C^Z, C^T)$ and the flags $F = (F^{\text{pe}}, F^{\text{ec}})$. In case of abort, we assume that all registers are initialized to a predetermined value.

</div>

Table 2: Simple QKD Protocol. The precise mathematical model is to be found in Section 3.3.

- The error correcting scheme is described by a quintuple $\text{ec} = \{t, r, \text{synd}, \text{corr}, \mathcal{H}_{\text{ec}}\}$. Here, $r \in \mathbb{N}$ is the length (in bits) of the error correction syndrome. Moreover, synd and corr are functions of the form $\text{synd} : \{0,1\}^n \to \{0,1\}^r$ and $\text{corr} : \{0,1\}^n \times \{0,1\}^r \to \{0,1\}^n$ used to compute the error syndrome and calculate the corrected string, respectively. We do not need to assume anything about the structure of this code. To fix ideas, let us just note that there exist good error correction codes with $r \approx nh(\delta)$, where $h$ is the binary entropy function and $\delta$ is the number of errors to be corrected.[10]

  Finally, $t \in \mathbb{N}$ is the length (in bits) of the hash used for verification and $\mathcal{H}_{\text{ec}} := \{h_{\text{ec}} : \{0,1\}^n \to \{0,1\}^t\}$ is a universal$_2$ family of hash functions. We will see in Theorem 2 that the size $t$ only depends logarithmically on the targeted correctness parameter.

- Privacy amplification is characterized by a tuple $\text{pa} = \{\ell, \mathcal{H}_{\text{pa}}\}$, where $\ell \in \mathbb{N}$ with $\ell \leq n$ is the length (in bits) of the extracted key and $\mathcal{H}_{\text{pa}} := \{h_{\text{pa}} : \{0,1\}^n \to \{0,1\}^\ell\}$ is a universal$_2$ family of hash functions.

  Note that $\ell$, the length of the final key, is fixed. It is in principle possible to design adaptive protocols where the final key length is chosen after parameter estimation, but this is beyond the scope of this work.

This allows us to define a family of protocols $\texttt{qkd\_eb}_{m,\text{pe,ec,pa}}$ in Table 2. Note that any such protocol is simply a completely positive trace-preserving map that maps bipartite quantum states shared between Alice and Bob onto probability distributions of the classical outputs, and we will define their exact operation in Section 3.3.

## 3.3 Exact mathematical model of the protocol

Here we describe in detail the mathematical model underlying the protocol in Table 2. It is worth emphasizing that the eavesdropper does not appear anywhere in this description, but will of course be required when assessing the security of the protocol as discussed in Section 4.

---

[10]For example, synd could be a linear code described by an $r \times n$ parity check matrix $H$ such that $\text{synd}(x) = Hx$. Moreover, corr can be any decoder, for example the (optimal) maximum likelihood decoder, but also a more practical suboptimal iterative decoder.

**Input:** Alice and Bob are given a state $\rho_{AB}$, where $A = A_{[m]} = A_1 A_2 \ldots A_m$ consists of $m$ quantum systems of arbitrary, finite dimension, $B = B_{[m]} = B_1 B_2 \ldots B_m$ consists of $m$ quantum systems of arbitrary, finite dimension. Note that apart from the above structure, the state $\rho_{AB}$ is fully general. The situation is depicted in Figure 1.
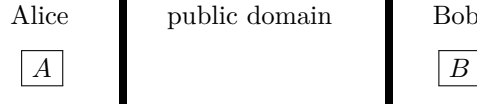


Figure 1: State of the classical and quantum systems at the beginning of the protocol. The initial state is denoted $\rho_{AB}$.

**Randomization:** We model the randomization by random seeds (uniform random variables), shared between Alice and Bob over the public authenticated channel. These random seeds are represented by a quantum state $\rho_S$ which is assumed to be maximally mixed and independent of $\rho_{AB}$. The situation after randomization is depicted in Figure 2. Let us now detail the content of the system $S$.

The first random variable is a random basis choice for each quantum system. This is modeled as a register $S^\Phi$ in the state

$$\rho_{S^\Phi} = \sum_{\phi \in \{0,1\}^m} \frac{1}{2^m} |\phi\rangle\langle\phi|_{S^\Phi}, \tag{22}$$

where $\{|\phi\rangle\}_{\phi \in \{0,1\}^m}$ is an orthonormal basis of the space $S^\Phi$ and $\phi = \phi_{[m]} = (\phi_1, \phi_2, \ldots, \phi_m)$ with $\phi_i \in \{0,1\}$. The total state at the beginning of the protocol is thus of the form $\rho_{AB} \otimes \rho_{S^\Phi}$.

The seed for the choice of the random subset is denoted $S^\Pi$ and is initially in the state

$$\rho_{S^\Pi} = \sum_{\pi \in \Pi_{m,k}} \frac{1}{\binom{m}{k}} |\pi\rangle\langle\pi|_{S^\Pi}, \tag{23}$$

where $\{|\pi\rangle\}_{\pi \in \Pi_{m,k}}$ is an orthonormal basis of the space $S^\Pi$. For any $\pi \in \Pi_{m,k}$, we denote its $k$ elements by $\pi_i$, for $i \in [k]$ and we denote by $\bar\pi \in [m]$ the complement of $\pi$.

At this point we reorder the measurement settings in $S^\Phi$ into two parts: the settings to be used for measuring quantum systems in $\pi$ will be stored in a register $S^\Xi$ and the settings to be used for measuring the remaining $n$ quantum systems in $\bar\pi$ will be stored in a register $S^\Theta$. Formally, we consider the function

$$\mathrm{ro} : \{0,1\}^m \times \Pi_{m,k} \to \{0,1\}^k \times \{0,1\}^n, \quad (\phi, \pi) \mapsto (\phi_\pi, \phi_{\bar\pi}). \tag{24}$$

Since $S^\Phi$ is uniformly random, the resulting state after applying this function and discarding $S^\Phi$ is of the form

$$\rho_{S^\Pi S^\Xi S^\Theta} = \mathrm{tr}_{S^\Phi} \left\{ \mathcal{E}_{\mathrm{ro}}(\rho_{S^\Phi} \otimes \rho_{S^\Pi}) \right\} = \rho_{S^\Pi} \otimes \rho_{S^\Xi} \otimes \rho_{S^\Theta}, \tag{25}$$

where the registers containing $S^\Xi$ and $S^\Theta$ are again uniformly random:

$$\rho_{S^\Xi} = \sum_{\xi \in \{0,1\}^k} \frac{1}{2^k} |\xi\rangle\langle\xi|_{S^\Xi} \quad \text{and} \quad \rho_{S^\Theta} = \sum_{\theta \in \{0,1\}^n} \frac{1}{2^n} |\theta\rangle\langle\theta|_{S^\Theta} \tag{26}$$

for $\xi = \xi_{[k]} = (\xi_1, \xi_2, \ldots, \xi_k)$ and $\theta = \theta_{[n]} = (\theta_1, \theta_2, \ldots, \theta_n)$ with $\xi_i, \theta_i \in \{0,1\}$.

The choice of the hash function in the family $\mathcal{H}_{\mathrm{ec}} = \{h_{\mathrm{ec}} : \{0,1\}^n \to \{0,1\}^t\}$ and the choice of hash function in the family $\mathcal{H}_{\mathrm{pa}} = \{h_{\mathrm{pa}} : \{0,1\}^n \to \{0,1\}^t\}$ are modeled via random seeds

$$\rho_{S^{H_{\mathrm{ec}}}} = \sum_{h \in \mathcal{H}_{\mathrm{ec}}} \frac{1}{|\mathcal{H}_{\mathrm{ec}}|} |h\rangle\langle h|_{S^{H_{\mathrm{ec}}}} \quad \text{and} \quad \rho_{S^{H_{\mathrm{pa}}}} = \sum_{h \in \mathcal{H}_{\mathrm{pa}}} \frac{1}{|\mathcal{H}_{\mathrm{pa}}|} |h\rangle\langle h|_{S^{H_{\mathrm{pa}}}}. \tag{27}$$

**Measurement:** We split the measurement process into two parts, measuring the systems in the set $\pi$ and $\bar\pi$ separately. While this distinction is not relevant for the practical implementation of the protocol, the notation introduced here will be important for the security analysis later. The first measurement concerns the registers in $\pi$, which are used for parameter estimation. For any subset $\pi \in \Pi_{m,k}$, we define a completely positive
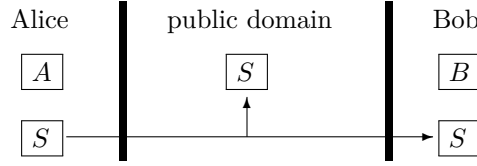
Figure 2: State of the classical and quantum systems after randomization. The state of the total system is $\rho_{AB} \otimes \rho_S$, where $\rho_S = \rho_{S^\Phi} \otimes \rho_{S^\Pi} \otimes \rho_{S^{H_{\text{ec}}}} \otimes \rho_{S^{H_{\text{pa}}}}$.

trace-preserving map $\mathcal{M}^\pi_{A\rightarrow V|S^\Xi} : A_\pi S^\Xi \rightarrow V A_\pi S^\Xi$ where $V = V_{[k]} = V_1 \otimes V_2 \otimes \ldots V_k$ models $k$ binary classical registers storing the measurement outcomes. The map is given by

$$\mathcal{M}^\pi_{A\rightarrow V|S^\Xi}(\cdot) = \sum_{\xi\in\{0,1\}^k} \sum_{v\in\{0,1\}^k} |v\rangle_V \left( M^{\xi,v}_{A_\pi} \otimes |\xi\rangle\langle\xi|_{S^\Xi} \right) \cdot \left( M^{\xi,v}_{A_\pi} \otimes |\xi\rangle\langle\xi|_{S^\Xi} \right)^\dagger \langle v|_V, \quad (28)$$

where $M^{\xi,v}_{A_\pi} := \bigotimes_{i\in[k]} M^{\xi_i,v_i}_{A_{\pi_i}}$. This map measures the $k$ subsystems determined by $\pi$ using the (random) measurement settings stored in the register $S^\Xi$. The results are stored in the classical register $V$, and the post-measurement state remains in the systems $A_\pi$.

Similarly, we define $\mathcal{M}^\pi_{B\rightarrow W|S^\Xi} : B_\pi S^\Xi \rightarrow W B_\pi S^\Xi$ as

$$\mathcal{M}^\pi_{B\rightarrow W|S^\Xi}(\cdot) = \sum_{\xi\in\{0,1\}^k} \sum_{w\in\{0,1\}^k} |w\rangle_W \left( M^{\xi,w}_{B_\pi} \otimes |\xi\rangle\langle\xi|_{S^\Xi} \right) \cdot \left( M^{\xi,w}_{B_\pi} \otimes |\xi\rangle\langle\xi|_{S^\Xi} \right)^\dagger \langle w|_W, \quad (29)$$

where $M^{\xi,w}_{B_\pi} := \bigotimes_{i\in[k]} M^{\xi_i,w_i}_{B_{\pi_i}}$. The two maps $\mathcal{M}^\pi_{A\rightarrow V|S^\Xi}$ and $\mathcal{M}^\pi_{B\rightarrow W|S^\Xi}$ commute since they act on different systems and we write their concatenation as $\mathcal{M}^\pi_{A\rightarrow V|S^\Xi} \circ \mathcal{M}^\pi_{B\rightarrow W|S^\Xi} = \mathcal{M}^\pi_{B\rightarrow W|S^\Xi} \circ \mathcal{M}^\pi_{A\rightarrow V|S^\Xi}$.

So far we have considered $\pi$ to be fixed. The full measurement for parameter estimation instead consults the register $S^\Pi$ and is modeled as a map $\mathcal{M}_{AB\rightarrow VW|S^\Pi S^\Xi} : ABS^\Pi S^\Xi \rightarrow ABVWS^\Pi S^\Xi$ given by

$$\mathcal{M}_{AB\rightarrow VW|S^\Pi S^\Xi}(\cdot) = \sum_{\pi\in\Pi_{m,k}} \mathcal{M}^\pi_{A\rightarrow V|S^\Xi} \circ \mathcal{M}^\pi_{B\rightarrow W|S^\Xi}\left(|\pi\rangle\langle\pi|_{S^\Pi} \cdot |\pi\rangle\langle\pi|_{S^\Pi}\right). \quad (30)$$

The state of the total system after the measurement required for parameter estimation is thus given by

$$\tau_{ABVWS^\Pi S^\Xi S^\Theta} = \mathcal{M}_{AB\rightarrow VW|S^\Pi S^\Xi}(\rho_{ABS^\Pi S^\Xi S^\Theta}) \quad (31)$$

$$= \sum_{\pi\in\Pi_{m,k}} \sum_{\xi\in\{0,1\}^k} \sum_{v,w\in\{0,1\}^k} \frac{1}{2^k\binom{m}{k}} |\pi,\xi\rangle\langle\pi,\xi|_{S^\Pi S^\Xi} \otimes \rho_{S^\Theta} \otimes$$

$$\ldots |v,w\rangle\langle v,w|_{VW} \otimes \left(M^{\xi,v}_{A_\pi} \otimes M^{\xi,w}_{B_\pi}\right)\rho_{AB}\left(M^{\xi,v}_{A_\pi} \otimes M^{\xi,w}_{B_\pi}\right)^\dagger. \quad (32)$$

The second measurement concerns the quantum systems used for extracting the secret key. The corresponding measurement maps are defined analogously to the measurements maps above, but now act on the systems determined by $\bar\pi$, the complement of $\pi$ in $[m]$. We define

$$\mathcal{M}_{A\rightarrow X|S^\Pi S^\Theta}(\cdot) = \sum_{\pi\in\Pi_{m,k}} \sum_{\theta,x\in\{0,1\}^n} |x\rangle_X \left( M^{\theta,x}_{A_{\bar\pi}} \otimes |\pi,\theta\rangle\langle\pi,\theta|_{S^\Pi S^\Theta} \right) \cdot \left( M^{\theta,x}_{A_{\bar\pi}} \otimes |\pi,\theta\rangle\langle\pi,\theta|_{S^\Pi S^\Theta} \right)^\dagger \langle x|_X, \quad (33)$$

$$\mathcal{M}_{B\rightarrow Y|S^\Pi S^\Theta}(\cdot) = \sum_{\pi\in\Pi_{m,k}} \sum_{\theta,y\in\{0,1\}^n} |y\rangle_Y \left( M^{\theta,y}_{B_{\bar\pi}} \otimes |\pi,\theta\rangle\langle\pi,\theta|_{S^\Pi S^\Theta} \right) \cdot \left( M^{\theta,y}_{B_{\bar\pi}} \otimes |\pi,\theta\rangle\langle\pi,\theta|_{S^\Pi S^\Theta} \right)^\dagger \langle y|_Y \quad (34)$$

as well as $\mathcal{M}_{AB\rightarrow XY|S^\Pi S^\Theta} = \mathcal{M}_{A\rightarrow X|S^\Pi S^\Theta} \circ \mathcal{M}_{B\rightarrow Y|S^\Pi S^\Theta}$. It is evident that all measurements $\mathcal{M}$ defined so far mutually commute because they act either on classical registers or on distinct quantum registers. Finally, we define the total measurement map as $\mathcal{M}_{AB\rightarrow VWXY|S^\Pi S^\Xi S^\Theta} := \mathcal{M}_{AB\rightarrow VW|S^\Pi S^\Xi} \circ \mathcal{M}_{AB\rightarrow XY|S^\Pi S^\Theta}$.

Of particular interest is the state of the system after measurement and after we discard the quantum systems.
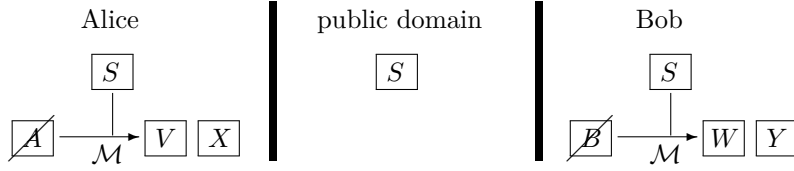
Figure 3: State of the classical and quantum systems during and after measurement. The measurement can be summarized as a CPTP map $\rho_{AB} \otimes \rho_{S^\Phi} \otimes \rho_{S^\Pi} \mapsto \sigma_{VWXYABS^\Phi S^\Pi}$.
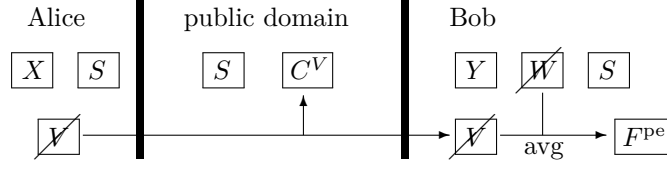


Figure 4: State of the classical and quantum systems during and after parameter estimation. Parameter estimation is summarized as a CPTP map $\sigma_{VW} \mapsto \sigma_{C^V F^{\mathrm{pe}}}$.

This is given by a classical state $\sigma_{VWXYS^\Pi S^\Xi S^\Theta}$. This state is of the form

$$\sigma_{VWXYS^\Pi S^\Xi S^\Theta} \tag{35}$$

$$= \mathrm{tr}_{AB}\left(\mathcal{M}_{AB \to VWXY|S^\Pi S^\Xi S^\Theta}(\rho_{ABS^\Pi S^\Xi S^\Theta})\right) \tag{36}$$

$$= \mathrm{tr}_{AB}\left(\mathcal{M}_{AB \to XY|S^\Pi S^\Theta}(\tau_{ABVWS^\Pi S^\Xi S^\Theta})\right) \tag{37}$$

$$= \sum_{\pi \in \Pi_{m,k}} \sum_{\substack{\xi \in \{0,1\}^k \\ \theta \in \{0,1\}^n}} \sum_{\substack{v,w \in \{0,1\}^k \\ x,y \in \{0,1\}^n}} \frac{1}{2^m \binom{m}{k}} |\pi,\xi,\theta\rangle\langle\pi,\xi,\theta|_{S^\Pi S^\Xi S^\Theta} \otimes |v,w,x,y\rangle\langle v,w,x,y|_{VWXY} \otimes$$

$$\ldots \; \mathrm{tr}_{AB}\left\{\left(\widetilde{M}_{A_\pi}^{\xi,v} \otimes \widetilde{M}_{A_{\bar\pi}}^{\theta,x} \otimes \widetilde{M}_{B_\pi}^{\xi,w} \otimes \widetilde{M}_{B_{\bar\pi}}^{\theta,y}\right)\rho_{AB}\right\}, \tag{38}$$

where we write $\widetilde{M}_{A_\pi}^{\xi,v} = \left(M_{A_\pi}^{\xi,v}\right)^\dagger M_{A_\pi}^{\xi,v}$ and analogously introduce $\widetilde{M}_{A_{\bar\pi}}^{\theta,x}$, $\widetilde{M}_{B_\pi}^{\xi,w}$ and $\widetilde{M}_{B_{\bar\pi}}^{\theta,y}$.

The situation after the complete measurement is depicted in Figure 3.

**Parameter estimation:** We model parameter estimation by a test function acting on the registers $V$ and $W$ and creating a binary flag $F^{\mathrm{pe}}$ as follows:

$$\mathrm{test} : \{0,1\}^k \times \{0,1\}^k \to \{\varnothing, \checkmark\}, \quad \mathrm{pe}(v,w) = \begin{cases} \varnothing & \text{if } \sum_{i \in [k]} 1\{v_i \neq w_i\} \geq k\delta, \\ \checkmark & \text{otherwise.} \end{cases} \tag{39}$$

This test can be applied to the states $\tau_{ABVWS^\Pi S^\Xi S^\Theta}$ or $\sigma_{VWXYS^\Pi S^\Xi S^\Theta}$ defined previously. This requires Alice to communicate $V$ to Bob on the authenticated classical channel in order to evaluate the value of $\mathrm{pe}(v,w)$ and the transcript of this communication is stored in the variable $C^V = V$.

We are specifically interested in the state $\tau_{ABVWS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}}} = \mathcal{E}_{\mathrm{pe}}(\tau_{ABVWS^\Pi S^\Xi S^\Theta})$ and the corresponding state conditioned on the outcome $F^{\mathrm{pe}} = \checkmark$, given by

$$\tau_{ABVWS^\Pi S^\Xi S^\Theta|F^{\mathrm{pe}}=\checkmark} = \frac{1}{\Pr[F^{\mathrm{pe}}=\checkmark]_\tau} \sum_{\pi \in \Pi_{m,k}} \sum_{\xi \in \{0,1\}^k} \sum_{\substack{v,w \in \{0,1\}^k \\ \sum_{i=1}^k 1\{v_i \neq w_i\} < k\delta}} \frac{1}{2^k \binom{m}{k}} |\pi,\xi\rangle\langle\pi,\xi|_{S^\Pi S^\Xi} \otimes$$

$$\ldots \; \rho_{S^\Theta} \otimes |v,w\rangle\langle v,w|_{VW} \otimes \left(M_{A_\pi}^{\xi,v} \otimes M_{B_\pi}^{\xi,w}\right)\rho_{AB}\left(M_{A_\pi}^{\xi,v} \otimes M_{B_\pi}^{\xi,w}\right)^\dagger. \tag{40}$$

We will see that this state is crucial for the security analysis in the next section. Finally, we note that $\mathcal{M}_{AB \to XY|S^\Pi S^\Theta}$ and $\mathcal{E}_{\mathrm{pe}}$ commute, and thus in particular we find that

$$\mathrm{tr}_{AB}\left\{\mathcal{M}_{AB \to XY|S^\Pi S^\Theta}(\tau_{ABVWS^\Pi S^\Xi S^\Theta|F^{\mathrm{pe}}=\checkmark})\right\} = \sigma_{VWXYS^\Pi S^\Xi S^\Theta|F^{\mathrm{pe}}=\checkmark}, \quad \text{where} \tag{41}$$

$$\sigma_{VWXYS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}}} = \mathcal{E}_{\mathrm{pe}}(\sigma_{VWXYS^\Pi S^\Xi S^\Theta}). \tag{42}$$

We then relabel $V$ to $C^V$ and keep it around as part of the transcript, while we discard $W$ after performing parameter estimation. The situation after parameter estimation is depicted in Figure 4.

**Error correction:** The error correction part of the protocol is split into two parts. The first part consists of the actual error correction procedure, determined by two functions synd and corr that are executed by Alice and Bob, respectively. We do not assume anything about these functions, but rather check their success in the second part by evaluating hash functions.

First Alice computes a syndrome $Z = \text{synd}(X)$ and sends it to Bob over the public channel. Bob then computes an estimate $\hat{X} = \text{corr}(Y, Z)$, discarding $Y$ in the process.

Alice and Bob then need to check that the decoding procedure succeeded with high probability by comparing hashes of their respective strings $X$ and $\hat{X}$ and abort the protocol if they differ. Alice computes a hash of size $t$ (in bits) of $X$ and sends it to Bob, who computes the corresponding hash for $\hat{X}$. This test is summarized as a classical map ec acting on registers $X$, $\hat{X}$ and $S^{H_{\text{ec}}}$ creating a transcript of the hash value $C^T$ and a binary flag $F^{\text{ec}}$ as follows:

$$\text{ec} : \{0,1\}^n \times \{0,1\}^n \times \mathcal{H}_{\text{ec}} \to \{0,1\}^t \times \{\varnothing, \checkmark\}, \quad (x, \hat{x}) \mapsto \begin{cases} (h_{\text{ec}}(x), \varnothing) & \text{if } h_{\text{ec}}(x) \neq h_{\text{ec}}(\hat{x}), \\ (h_{\text{ec}}(x), \checkmark) & \text{otherwise.} \end{cases} \tag{43}$$

These classical functions are modeled using CPTP maps $\mathcal{E}_{\text{synd}}$, $\mathcal{E}_{\text{corr}}$ and $\mathcal{E}_{\text{ec}}$, respectively. Applying them to the state $\sigma_{XYC^V S^\Pi S^\Xi S^\Theta F^{\text{pe}}}$ yields

$$\sigma_{X\hat{X}C^V C^Z C^T S^\Pi S^\Xi S^\Theta S^{H_{\text{ec}}} F^{\text{pe}} F^{\text{ec}}} = \text{tr}_Y \left\{ \mathcal{E}_{\text{ec}} \circ \mathcal{E}_{\text{corr}} \circ \mathcal{E}_{\text{synd}} \left( \sigma_{XYC^V S^\Pi S^\Xi S^\Theta F^{\text{pe}}} \otimes \rho_{S^{H_{\text{ec}}}} \right) \right\}, \tag{44}$$

where the transcript register $C^Z$ contains the value of the syndrome and $C^T$ the output of Alice's hash. This process is depicted in Figure 5.
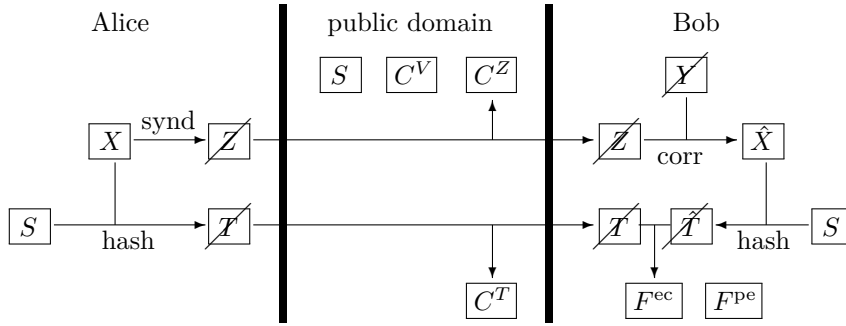


Figure 5: State of the classical and quantum systems during and after error correction. Error correction is summarized as a CPTP map $\sigma_{XY} \otimes \rho_{S^{H_{\text{ec}}}} \mapsto \sigma_{X\hat{X}C^Z F^{\text{ec}}}$.

**Privacy amplification:** Alice and Bob use the seed $H_{\text{pa}}$ to choose a hash function, which they then both apply on their raw key to compute $K_A = H_{\text{pa}}(X)$ and $K_B = H_{\text{pa}}(\hat{X})$, their respective keys. Formally, the privacy amplification map is defined as:

$$\text{pa} : \{0,1\}^n \times \{0,1\}^n \times \mathcal{H}_{\text{pa}} \to \{0,1\}^\ell \times \{0,1\}^\ell, \quad (x, \hat{x}, h_{\text{pa}}) \mapsto (h_{\text{pa}}(x), h_{\text{pa}}(\hat{x})). \tag{45}$$

Denoting by $K_A$ and $K_B$ the respective key spaces of Alice and Bob, the final quantum state is

$$\omega_{K_A K_B CSF} = \text{tr}_{X\hat{X}} \left\{ \mathcal{E}_{\text{pa}}(\sigma_{X\hat{X}CSF} \otimes \rho_{S^{H_{\text{pa}}}}) \right\}. \tag{46}$$

Finally, Bob reveals the status of his flag registers. This final step is depicted in Figure 6.

# 4  Security of the generated key

For a detailed discussion of the security of quantum key distribution, we refer the reader to Portmann and Renner [32]. For our purposes here (consistent with [10] and [32]), we say that our protocol is $\Delta$-secure if it is $\Delta$-close to an ideal protocol in terms of the diamond distance. Note in particular that an ideal protocol is allowed to abort, but it will always output a uniformly random shared key in case it does not. Table 3 gives such an ideal protocol, denoted $\texttt{qkd\_ideal}_{m,\text{pe},\text{ec},\text{pa}}$ designed in such a way that it is close to the original $\texttt{qkd\_eb}_{m,\text{pe},\text{ec},\text{pa}}$ protocol.
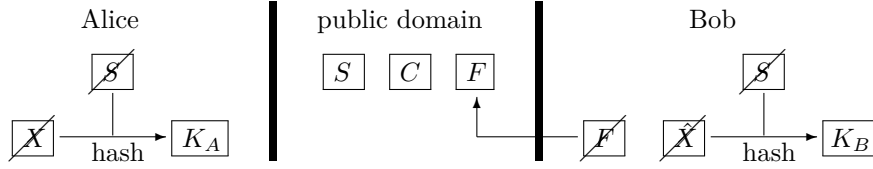
Figure 6: State of the classical and quantum systems during and after privacy amplification. Privacy amplification is a CPTP map $\sigma_{X\hat{X}} \otimes \rho_{S^{H_{\mathrm{pa}}}} \mapsto \omega_{K_A K_B}$. The complete final state is denoted by $\omega_{K_A K_B SCF}$.

---

$(K_A, K_B, S, C, F) = \texttt{qkd\_ideal}_{m,\mathrm{pe,ec,pa}}\left(\rho_{AB}\right):$

**Run protocol:** Set $(K_A, K_B, S, C, F) = \texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}(\rho_{AB})$.

**Output:** If $F^{\mathrm{pe}} = F^{\mathrm{ec}} = \checkmark$, then replace $K_A$ and $K_B$ by an independent and uniformly distributed random string $K$ of length $\ell$, i.e. set $K_A = K_B = K$.

---

Table 3: An ideal QKD protocol that is close to $\texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}$.

In order to show that the protocol is secure, it thus suffices to show that

$$\Delta_{m,\mathrm{pe,ec,pa}} := \left\|\texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}} - \texttt{qkd\_ideal}_{m,\mathrm{pe,ec,pa}}\right\|_\diamond \tag{47}$$

$$= \sup_{\rho_{ABE} \in \mathcal{S}(ABE)} \left\|\texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}(\rho_{ABE}) - \texttt{qkd\_ideal}_{m,\mathrm{pe,ec,pa}}(\rho_{ABE})\right\|_{\mathrm{tr}} \tag{48}$$

is very small for certain choices of parameters $k, n, \delta, \mathrm{ec}$ and pa. In the latter expression $\rho_{ABE}$ is an arbitrary extension of $\rho_{AB}$ to an auxiliary system $E$. Without loss of generality we may take $|E| = |A||B|$, which is sufficient to achieve the supremum. Physically the system $E$ is held by a potential adversary, the eavesdropper. In particular, this assures that $E$ can be assumed finite-dimensional. Hence, we need to show that the trace distance between the protocols' outputs is small for all possible input states $\rho_{ABE}$.

Let us now fix $\rho_{ABE}$ for the moment. The trace distance in (48) can be simplified by noting that the output of $\texttt{qkd\_ideal}$ equals the output of $\texttt{qkd\_eb}$ if the protocol aborts. We find

$$\left\|\texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}(\rho_{ABE}) - \texttt{qkd\_ideal}_{m,\mathrm{pe,ec,pa}}(\rho_{ABE})\right\|_{\mathrm{tr}}$$

$$= \Pr\left[F = (\checkmark,\checkmark)\right]_\omega \cdot \left\|\omega_{K_A K_B SCE|F=(\checkmark,\checkmark)} - \chi_{K_A K_B} \otimes \omega_{SCE|F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}} \tag{49}$$

$$= \left\|\omega_{K_A K_B SCFE \wedge F=(\checkmark,\checkmark)} - \chi_{K_A K_B} \otimes \omega_{SCFE \wedge F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}} \tag{50}$$

where we use $\omega_{K_A K_B SCFE} = \texttt{qkd\_eb}_{k,n,\delta,\mathrm{ec,pa}}(\rho_{ABE})$ and define a perfect key $\chi_{K_A K_B}$ as follows:

$$\chi_{K_A K_B} := \frac{1}{2^\ell} \sum_{k \in \{0,1\}^\ell} |k\rangle\langle k|_{K_A} \otimes |k\rangle\langle k|_{K_B} . \tag{51}$$

Recall that $\omega$ in (50) corresponds to a subnormalized state with trace equal to $\Pr[F = (\checkmark,\checkmark)]_\omega$.

Our goal in the following is to bound (50) or (49) uniformly in $\rho_{ABE}$, which implies an upper bound on (48) as well. In order to do this we will employ the following lemma which allows us to split the norm into two terms corresponding to correctness and secrecy. This has been shown, e.g., in [32, Theorem 4.1], but we provide a proof here for completeness.

**Lemma 1.** *Let $\varepsilon_{ec}, \varepsilon_{pa} \in [0,1)$ be two constants. If, for every state $\rho_{ABE} \in \mathcal{S}(ABE)$ and $\omega_{K_A K_B SCFE} = \texttt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}(\rho_{ABE})$, we have*

$$\Pr[K_A \neq K_B \wedge F^{pe} = F^{ec} = \checkmark]_\omega \leq \varepsilon_{ec} \qquad and \tag{52}$$

$$\left\|\omega_{K_A SCFE \wedge F=(\checkmark,\checkmark)} - \chi_{K_A} \otimes \omega_{SCFE \wedge F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}} \leq \varepsilon_{pa} . \tag{53}$$

*Then, $\Delta_{m,\mathrm{pe,ec,pa}} \leq \varepsilon_{ec} + \varepsilon_{pa}$.*

*Proof.* Let us introduce an auxiliary state $\eta_{K_A K_B SCFE}$ that is equal to $\omega_{K_A K_B SCFE}$ except that we set $K_B = K_A$. Then, applying the triangle inequality to the trace distance in (49) and simplifying the resulting terms,

we find

$$\left\|\omega_{K_A K_B SCE|F=(\checkmark,\checkmark)} - \chi_{K_A K_B} \otimes \omega_{SCE|F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}}$$

$$\leq \left\|\omega_{K_A K_B SCE|F=(\checkmark,\checkmark)} - \eta_{K_A K_B SCE|F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}}$$

$$+ \left\|\eta_{K_A K_B SCE|F=(\checkmark,\checkmark)} - \chi_{K_A K_B} \otimes \omega_{SCE|F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}} \tag{54}$$

$$= \Pr\left[K_A \neq K_B | F = (\checkmark,\checkmark)\right]_\omega + \left\|\eta_{K_A SCE|F=(\checkmark,\checkmark)} - \chi_{K_A} \otimes \omega_{SCE|F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}}. \tag{55}$$

Multiplying this with $\Pr[F = (\checkmark,\checkmark)]_\omega$ as in (49) yields the desired implication. $\qquad\square$

The first condition of the above Lemma 1 ensures that the protocol is $\varepsilon_{ec}$-*correct*, and the second condition ensures that the protocol $\varepsilon_{pa}$-*secret*. If both are satisfied, we say that the protocol is $(\varepsilon_{\mathrm{ec}} + \varepsilon_{pa})$-*secure*. In the security proof we can thus verify the two conditions separately.

# 5   Results and discussion

We will show the following theorems, which essentially give bounds on the security parameters in terms of the protocol parameters. The first theorem establishes correctness of the protocol. Correctness of the protocol is ensured in the error correction step using hash functions, and consequently correctness can be bounded in term of the length $t$ of the hash that is used. The proof is given in Section 6.1.

**Theorem 2.** *Consider the protocol* qkd_eb$_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}$ *in Section 3 with* ec $= \{t,\ldots\}$. *Then for every state* $\rho_{AB} \in \mathcal{S}(AB)$ *and* $\omega_{K_A K_B SCF} = $ qkd_eb$_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}(\rho_{AB})$ *we have*

$$\Pr[K_A \neq K_B \wedge F = (\checkmark,\checkmark)]_\omega \leq \varepsilon_{\mathrm{ec}} := 2^{-t}. \tag{56}$$

The second theorem asserts secrecy. Secrecy is ensured by a combination of the parameter estimation and privacy amplification steps of the protocol, which both introduce an error. There is a tradeoff between these two errors, parametrized by a scalar $\nu$, which ought to be optimized numerically. The proof is given in Sections 6.2–6.4.

**Theorem 3.** *Consider the protocol* qkd_eb$_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}$ *in Section 3 with* pe $= \{k,\delta\}$, ec $= \{t,r,\ldots\}$ *and* pa $= \{\ell,\ldots\}$. *Then, for every state* $\rho_{ABE}$ *and* $\omega_{K_A K_B SCFE} = $ qkd_eb$_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}(\rho_{ABE})$, *we have*

$$\left\|\omega_{K_A SCFE \wedge F=(\checkmark,\checkmark)} - \chi_{K_A} \otimes \omega_{SCFE \wedge F=(\checkmark,\checkmark)}\right\|_{\mathrm{tr}} \leq \inf_{\nu \in (0, \frac{1}{2}-\delta)} \varepsilon_{pe}(\nu) + \varepsilon_{pa}(\nu), \tag{57}$$

*where the error functions are given as*

$$\varepsilon_{pa}(\nu) := \frac{1}{2}\sqrt{2^{-(m-k)\left(\log\frac{1}{c} - h(\delta+\nu)\right) + r + t + \ell}} \quad and \quad \varepsilon_{pe}(\nu) := 2\,e^{-\frac{(m-k)k^2\nu^2}{m(k+1)}} \tag{58}$$

*and* $h(x) := -x \log x - (1-x)\log(1-x)$ *denotes the binary entropy.*

Combining Theorems 2 and 3 we see that total error is thus composed of three components, $\varepsilon_{\mathrm{pe}}$, $\varepsilon_{\mathrm{ec}}$, and $\varepsilon_{\mathrm{pa}}$. Let us take a close look at these errors for the case of large $m$. First, we note that $\varepsilon_{\mathrm{ec}}$ vanishes asymptotically when we choose $t = \log(m)$, or any other slowly growing function of $m$. To make sure that $\varepsilon_{\mathrm{pe}}$ vanishes we choose $k = \sqrt{m}$ and $\nu = \log(m)^{-1}$, for example. For a robust operation at noise level $\delta$ it is necessary (and in theory sufficient) that the error correction leakage satisfies $r \approx (m-k)h(\delta)$. Since $h(\delta+\nu) \approx h(\delta)$ by continuity, we find that $\varepsilon_{\mathrm{pa}}$ vanishes as long as

$$(m-k)\left(\log\frac{1}{c} - 2h(\delta)\right) - \ell - \log(m) \tag{59}$$

is positive and grows in $m$. Since $k$ and $\log m$ become negligible compared to $m$ as $m$ gets large, our protocol thus achieves the asymptotically optimal rate by Devetak and Winter [36], with $\ell/m = \log\frac{1}{c} - 2h(\delta)$.

# 6   Security proof
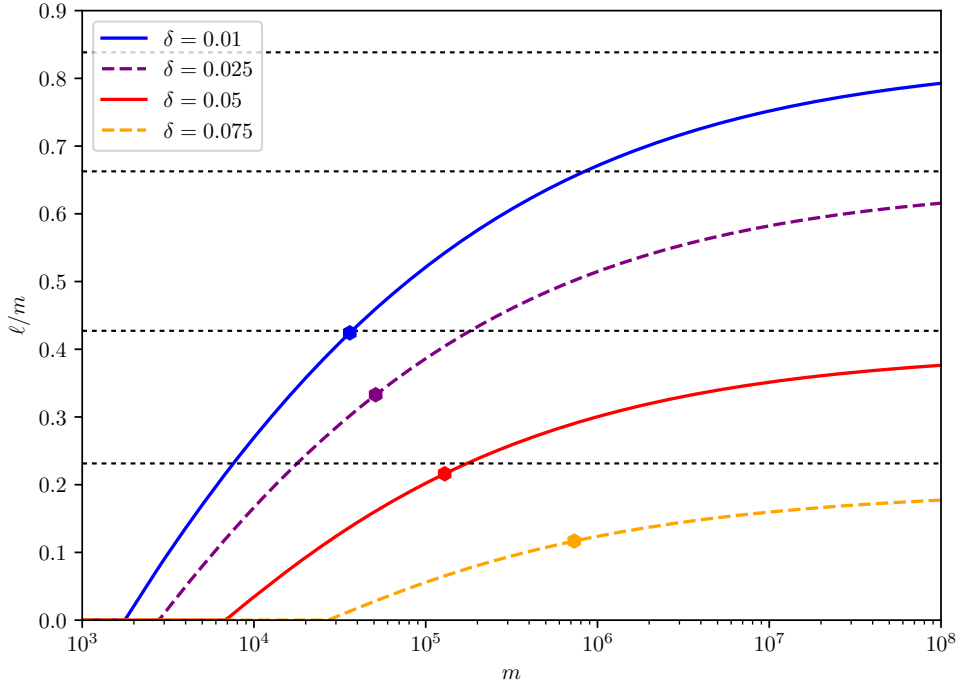
The purpose of this section is to prove Theorems 2 and 3.

Figure 7: This plot shows the maximal secret key rate $\ell/m$ as a function of $m$ for different error thresholds $\delta \in \{0.01, 0.025, 0.05, 0.075\}$, optimized over all protocols. The protocols are required to be $\varepsilon$-secure with $\varepsilon = 10^{-10}$ and the device parameter is assumed to be $\bar{c} = 0.5$. The error correction leakage is approximated to be $r = 1.1(m-k)h(\delta)$, see for instance [37]. (A more detailed approximation that includes finite-size effects was recently given in [38].) All remaining parameters, i.e. $\nu, k$ and $t$, are optimized numerically to maximize $\ell$ (code available online). The dotted horizontal lines show the corresponding asymptotic limit of the key rate for each value of $\delta$, given as $1 - 2h(\delta)$. The markers indicate the points at which the key rate matches $50\%$ of the asymptotic limit.

## 6.1 Error correction: Proof of Theorem 2

We wish to upper bound the probability of the protocol not aborting and outputting distinct final keys for Alice and Bob.

*Proof of Theorem 2.* We consider the following chain of inequalities:

$$\Pr[K_A \neq K_B \wedge F^{\mathrm{pe}} = F^{\mathrm{ec}} = \checkmark]_\omega \leq \Pr[K_A \neq K_B \wedge F^{\mathrm{ec}} = \checkmark]_\omega \tag{60}$$
$$= \Pr[H_{\mathrm{pa}}(X) \neq H_{\mathrm{pa}}(X') \wedge H_{\mathrm{ec}}(X) = H_{\mathrm{ec}}(X')]_\omega \tag{61}$$
$$\leq \Pr[X \neq X' \wedge H_{\mathrm{ec}}(X) = H_{\mathrm{ec}}(X')]_\sigma \tag{62}$$
$$= \Pr[X \neq X']_\sigma \Pr[H_{\mathrm{ec}}(X) = H_{\mathrm{ec}}(X') \mid X \neq X']_\sigma \tag{63}$$
$$\leq \Pr[H_{\mathrm{ec}}(X) = H_{\mathrm{ec}}(X') \mid X \neq X']_\sigma \tag{64}$$
$$\leq |\mathcal{H}_{\mathrm{ec}}|^{-1} = 2^{-t}. \tag{65}$$

The first inequality follows since we ignore the status of the flag $F^{\mathrm{pe}}$. The second inequality is a consequence of the fact $X = X'$ implies $H_{\mathrm{pa}}(X) = H_{\mathrm{pa}}(X')$. The third inequality follows since $\Pr[X \neq X']_\sigma \leq 1$ and the last one by definition of universal$_2$ hashing. $\qquad\square$

## 6.2 Measurements: Uncertainty tradeoff between smooth min- and max-entropy

The crucial bound on the smooth entropy of Alice's measurement outcomes follows by the entropic uncertainty relation, suitably applied. We state the uncertainty relation in a natural form [39, Corollary 7.4].

**Proposition 4.** *Let $\tau_{APRS} \in \mathcal{S}_\bullet(APRS)$ be an arbitrary sub-normalized state with $P$ a classical register, and set $t := \mathrm{tr}\{\tau_{APRS}\}$. Furthermore, let $\varepsilon \in [0, \sqrt{t})$ and let $q$ be a bijective function on $P$ that is a symmetry of*

$\tau_{ABCP}$ *in the sense that* $\tau_{ARS,P=p} = \tau_{ARS,P=q(p)}$ *for all* $p \in P$. *Then, we have*

$$H^\varepsilon_{\min}(X|PR)_\sigma + H^\varepsilon_{\max}(X|PS)_\sigma \geq \log \frac{1}{c_q}, \qquad where \tag{66}$$

*where* $c_q = \max_{p \in P} \max_{x,z \in X} \left\| F_A^{q(p),x} \big( F_A^{p,z} \big)^\dagger \right\|_\infty^2$. *Here,* $\sigma_{XPRS} = \mathcal{M}_{A \to X|P}(\tau_{APRS})$ *for the map*

$$\mathcal{M}_{A \to X|P}\big[\, \cdot \,\big] = \mathrm{tr}_A \left( \sum_{p \in P} \sum_{x \in X} |x\rangle_X \left( |p\rangle\langle p|_P \otimes F_A^{p,x} \right) \cdot \left( |p\rangle\langle p|_P \otimes F_A^{p,x} \right)^\dagger \langle x|_X \right). \tag{67}$$

*and any set (indexed by* $p \in P$*) of generalized measurements* $\{F_A^{p,x}\}_{x \in X}$.

A variant of this uncertainty relation was first shown in [39], based on the techniques introduced in [21]. We provide a full proof of the uncertainty relation in Appendix A for completeness. In the following corollary we apply it to the situation at hand during our protocol.

**Corollary 5.** *Consider the protocol* qkd_eb$_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}$ *in Section 3 with* $\mathrm{pe} = \{k, \ldots\}$ *applied to a state* $\rho_{ABE} \in \mathcal{S}(ABE)$ *and the state* $\sigma_{XYVWS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}} E}$ *as in* (42) *that results after measurement and parameter estimation. Define* $\bar{c}$ *as in* (21). *Then, for* $\varepsilon \in \big[0, \sqrt{\Pr[F^{\mathrm{pe}} = \checkmark]_\sigma}\big)$, *we have*

$$H^\varepsilon_{\min}(X \wedge F^{\mathrm{pe}} = \checkmark | VWS^\Pi S^\Xi S^\Theta E)_\sigma + H^\varepsilon_{\max}(X \wedge F^{\mathrm{pe}} = \checkmark | Y)_\sigma \geq (m-k) \log \frac{1}{\bar{c}}. \tag{68}$$

*Proof.* Consider the state $\tau_{ABVWS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}} E \wedge F^{\mathrm{pe}} = \checkmark}$ defined in (40) and note that it is of the form

$$\tau_{ABVWS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}} E \wedge F^{\mathrm{pe}} = \checkmark} = \tau_{ABVWS^\Pi S^\Xi F^{\mathrm{pe}} E \wedge F^{\mathrm{pe}} = \checkmark} \otimes \rho_{S^\Theta}. \tag{69}$$

This is the state of the system after parameter estimation and after measuring $V$ and $W$, but with the measurement of $X$ and $Y$ (in the basis determined by $S^\Theta$) delayed. In particular we have used the fact that the register $S^\Theta$ has not yet been touched in the protocol, and is thus independent and uniform and independent even after we consider the event $F^{\mathrm{pe}} = \checkmark$.

Let us now apply Proposition 4 to this state. For this purpose we equate $P = S^\Pi S^\Xi S^\Theta$, $R = VWE$, and $S = B$. The symmetry is determined by the map $q : \theta \mapsto \bar{\theta}$ with $\bar{\theta}_i = 1 - \theta_i$, which only acts on $S^\Theta$ and since this system is uniform and in product with the rest of the state trivially satisfies the symmetry condition of the theorem. The measurement map is then simply $\mathcal{M}_{A \to X|S^\Pi S^\Theta}$ and we can calculate

$$c_q = \max_{\pi \in \Pi_{m,k}} \max_{\theta,x,z \in \{0,1\}^n} \left\| M_{A_{\bar{\pi}}}^{\bar{\theta},x} \big(M_{A_{\bar{\pi}}}^{\theta,z}\big)^\dagger \right\|_\infty^2 = \max_{\pi \in \Pi_{m,k}} \left( \prod_{i \in \bar{\pi}} c_i \right) = \bar{c}^n. \tag{70}$$

Proposition 4 applied to our setup thus yields

$$H^\varepsilon_{\min}(X \wedge F^{\mathrm{pe}} = \checkmark | VWES^\Pi S^\Xi S^\Theta)_\sigma + H^\varepsilon_{\max}(X \wedge F^{\mathrm{pe}} = \checkmark | BS^\Pi S^\Xi S^\Theta)_\tau \geq n \log \frac{1}{\bar{c}}. \tag{71}$$

Finally, the statement of the Proposition follows by applying the measurement map $\mathcal{M}_{B \to Y|S^\Pi S^\Theta}$ (and discarding the seed registers) and noting that $H^\varepsilon_{\max}(X \wedge F^{\mathrm{pe}} = \checkmark | BS^\Pi S^\Xi S^\Theta)_\tau \leq H^\varepsilon_{\max}(X \wedge F^{\mathrm{pe}} = \checkmark | Y)_\sigma$ by the data-processing inequality. $\qquad \square$

## 6.3 Parameter estimation: Statistical bounds on smooth max-entropy

This section covers the necessary statistical analysis. This is essentially a variation of the analysis in [13], but requires a new tool, Lemma 7, presented in Section 2.2, as we are finding it clearer to do the analysis with sub-normalized states here. We use the following standard tail bound.

**Lemma 6.** *Consider a set of binary random variables* $Z = (Z_1, Z_2, \ldots, Z_m)$ *with* $Z_i$ *taking values in* $\{0,1\}$ *and* $m = n + k$. *Let* $\Pi \in \Pi_{m,k}$ *be an independent, uniformly distributed random variable. Then,*

$$\Pr\left[ \sum_{i \in \Pi} Z_i \leq k\delta \ \wedge \ \sum_{i \in \bar{\Pi}} Z_i \geq n(\delta + \nu) \right] \leq e^{-2\nu^2 \frac{nk^2}{(n+k)(k+1)}}. \tag{72}$$

Remarkably this bound is valid without any assumption on the distribution of $Z$.

*Proof.* Let $\mu(z) = \frac{1}{m}\sum_{i\in[m]}z_i$. Consider the following sequence of inequalities:

$$\Pr\left[\frac{1}{k}\sum_{i\in\Pi}Z_i \le \delta \ \wedge \ \frac{1}{n}\sum_{i\notin\Pi}Z_i \ge \delta+\nu\right] \le \Pr\left[\frac{1}{n}\sum_{i\in\bar\Pi}Z_i \ge \frac{1}{k}\sum_{i\in\Pi}Z_i+\nu\right] \tag{73}$$

$$= \sum_{z\in\{0,1\}^m}\Pr[Z=z]\Pr\left[\frac{1}{n}\sum_{i\in\bar\Pi}z_i \ge \frac{1}{k}\sum_{i\in\Pi}z_i+\nu\right] \tag{74}$$

$$= \sum_{z\in\{0,1\}^m}\Pr[Z=z]\Pr\left[\frac{1}{n}\sum_{i\in\bar\Pi}z_i \ge \mu(z)+\frac{k\nu}{m}\right]. \tag{75}$$

Here, the first inequality holds since $A \implies B$ implies $\Pr[A] \le \Pr[B]$ for any events $A$ and $B$. The first equality follows from the fact that $\Pi$ is independent of $Z$. The last equality follows by substituting $\sum_{i\in\Pi}z_i = m\mu(z) - \sum_{i\in\bar\Pi}z_i$ and rearranging the terms appropriately.

Now note that the random sums $S_n := \sum_{i\in\bar\Pi}z_i$ can be seen as emanating from randomly sampling without replacement $n$ balls labelled by $z_i \in \{0,1\}$ from a population $z$ with mean $\mu(z)$. Serfling's bound [40, Corollary 1.1] then tells us that

$$\Pr\left[\frac{1}{n}S_n \ge \mu(z)+\frac{k\nu}{m}\right] \le e^{-2n\left(\frac{k\nu}{m}\right)^2\frac{1}{1-f_n^*}} = e^{-2\nu^2\frac{nk^2}{(n+k)(k+1)}}. \tag{76}$$

where we substituted $f_n^* = \frac{n-1}{m}$. It is important to note that this bound is independent of $\mu(z)$. Thus, substituting this back into (75), we conclude the proof. $\qquad\square$

The following lemma ensures that disregarding an unlikely event will not disturb the state too much in terms of the purified distance.

**Lemma 7.** *Let $\rho_{AX} \in \mathcal{S}_\bullet(AX)$ be classical on $X$ and $\Omega : \mathcal{X} \to \{0,1\}$ an event with $\Pr[\Omega]_\rho = \varepsilon < \mathrm{tr}\{\rho_{AX}\}$. Then there exists a sub-normalized state $\tilde\rho_{AX} \in \mathcal{S}_\bullet(AX)$ with $\Pr[\Omega]_{\tilde\rho} = 0$ and $P(\rho_{AX},\tilde\rho_{AX}) \le \sqrt\varepsilon$.*

*Proof.* Set $\xi = \mathrm{tr}\{\rho_{AX}\}$. Let $\tilde\rho_{AX} = \frac{\sin^2(\phi)}{\xi-\varepsilon}\rho_{AX\wedge\neg\Omega}$ for some normalization $\phi \in [0,\frac{\pi}{2}]$ to be determined. Then the generalized fidelity evaluates to

$$\sqrt{F(\rho_{AX},\tilde\rho_{AX})} = \frac{\sin(\phi)}{\sqrt{\xi-\varepsilon}}\mathrm{tr}\left\{\sqrt{\sqrt{\rho_{AX}}\rho_{AX\wedge\neg\Omega}\sqrt{\rho_{AX}}}\right\} + \cos(\phi)\sqrt{1-\xi} \tag{77}$$

$$= \sin(\phi)\sqrt{\xi-\varepsilon} + \cos(\phi)\sqrt{1-\xi}. \tag{78}$$

This expression is maximized for $\tan(\phi) = \sqrt{\frac{\xi-\varepsilon}{1-\xi}}$ and, thus, $\sin(\phi) = \sqrt{\frac{\xi-\varepsilon}{1-\varepsilon}}$ and $\cos(\phi) = \sqrt{\frac{1-\xi}{1-\varepsilon}}$. Substituting this into (78) yields $F(\rho_{AX},\tilde\rho_{AX}) = 1-\varepsilon$, concluding the proof. $\qquad\square$

With this in hand, we wish to bound the smooth max-entropy of the state when passing the parameter estimation test.

**Proposition 8.** *Consider the protocol $\mathtt{qkd\_eb}_{m,\mathrm{pe,ec,pa}}$ in Section 3 with $\mathrm{pe} = \{k,\delta\}$ applied to a state $\rho_{AB} \in \mathcal{S}(AB)$ and the state $\sigma_{XYF^{\mathrm{pe}}}$ in (42) that results after measurement and parameter estimation. For any $\nu \in (0,1)$, we first define*

$$\varepsilon(\nu) := e^{-\frac{(m-k)k^2\nu^2}{m(k+1)}}. \tag{79}$$

*Then, for any $\nu \in (0,\frac{1}{2}-\delta]$ such that $\varepsilon(\nu)^2 < \Pr[F^{\mathrm{pe}}=\checkmark]_\sigma$, the following holds:*

$$H_{\max}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}}=\checkmark|Y)_\sigma \le (m-k)\,h(\delta+\nu) \qquad \textit{where} \qquad h(x) := -x\log x - (1-x)\log(1-x). \tag{80}$$

Intuitively, this result is a consequence of the fact that when we pass the parameter estimation test, conditioned on any particular value of $Y$, the support of $X$ is small as the number of errors (positions where $x_i \ne y_i$) is bounded (with high probability).

*Proof.* We use the shorthand $p = \Pr[F^{\mathrm{pe}} = \checkmark]_\sigma$ and $n := m - k$. Define the event event $\Omega := 1\{\sum_{i\in[n]} 1\{X_i \neq Y_i\} \geq n(\delta + \nu)\}$. We show that the statement in (80) holds when $p > \varepsilon^2$. Using Lemma 6, we find

$$\Pr\left[F^{\mathrm{pe}} = \checkmark \wedge \Omega\right]_\sigma = \Pr\left[\sum_{i\in[k]} 1\{V_i \neq W_i\} \leq k\delta \ \wedge \ \sum_{i\in[n]} 1\{X_i \neq Y_i\} \geq n(\delta + \nu)\right]_\sigma \leq \varepsilon(\nu)^2. \tag{81}$$

This gives an upper bound on the probability of the unlikely coincidence where the parameter estimation test passes with threshold $\delta$ but the fraction of errors between $X$ and $Y$ exceeds the threshold $\delta$ by a constant amount. We now want to remove the above unlikely events from our state $\sigma_{XYF^{\mathrm{pe}}\wedge F^{\mathrm{pe}}=\checkmark}$ by means of smoothing. Lemma 7 allows to do just that, and we introduce the state $\tilde{\sigma}_{XYF^{\mathrm{pe}}}$ that is $\varepsilon(\nu)$-close to $\sigma_{XYF^{\mathrm{pe}}\wedge F^{\mathrm{pe}}=\checkmark}$ in purified distance and satisfies $\Pr[\Omega]_{\tilde\sigma} = 0$. From this we conclude that

$$H_{\max}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|Y)_\sigma \leq H_{\max}(X \wedge F^{\mathrm{pe}} = \checkmark|Y)_{\tilde\sigma} = H_{\max}(X|Y)_{\tilde\sigma}\,, \tag{82}$$

where the last equality is a consequence of the fact that $\tilde\sigma$ is only supported on $F^{\mathrm{pe}} = \checkmark$.

It remains to show that $H_{\max}(X|Y)_{\tilde\sigma} \leq nh(\delta + \nu)$. Using the expansion of the conditional max-entropy in [39, Sec. 4.3.2], we find

$$H_{\max}(X|Y)_{\tilde\sigma} = \log\left(\sum_{y\in\{0,1\}^n} \Pr[Y = y]_{\tilde\sigma}\, 2^{H_{\max}(X|Y=y)_{\tilde\sigma}}\right) \tag{83}$$

$$\leq \max_{\substack{y\in\{0,1\}^n \\ \Pr[Y=y]_{\tilde\sigma}>0}} H_{\max}(X|Y = y)_{\tilde\sigma} \tag{84}$$

$$\leq \max_{\substack{y\in\{0,1\}^n \\ \Pr[Y=y]_{\tilde\sigma}>0}} \log\left|\left\{x \in \{0,1\}^n : \Pr[X = x|Y = y]_{\tilde\sigma} > 0\right\}\right| \tag{85}$$

$$= \max_{y\in\{0,1\}^n} \log\left|\left\{x \in \{0,1\}^n : \Pr[X = x \wedge Y = y]_{\tilde\sigma} > 0\right\}\right|. \tag{86}$$

In the ultimate inequality we used that the (unconditional) Rényi entropy is upper bounded by the logarithm of the distribution's support [27]. Furthermore, since $\Pr[\Omega]_{\tilde\sigma} = 0$, we have

$$\left|\left\{x \in \{0,1\}^n : \Pr[X = x \wedge Y = y]_{\tilde\sigma} > 0\right\}\right| \leq \sum_{x\in\{0,1\}^n} 1\left\{\sum_{i\in[n]} 1\{x_i \neq y_i\} < n(\delta + \nu)\right\} \tag{87}$$

$$= \sum_{e\in\{0,1\}^n} 1\left\{\sum_{i=1}^n e_i < n(\delta + \nu)\right\} \tag{88}$$

$$= \sum_{\lambda=0}^n \binom{n}{\lambda} 1\{\lambda < n(\delta + \nu)\} = \sum_{\lambda=0}^{\lfloor n(\delta+\nu)\rfloor} \binom{n}{\lambda}. \tag{89}$$

Here, in order to derive (88) we reparameterize $e_i = x_i$ xor $y_i$, indicating if there is an error at the $i$-th position. Finally, in (89) we substitute $\lambda = \sum_{i=1}^n e_i$, the total number of errors. The inequality $\sum_{\lambda=0}^{\lfloor n(\delta+\nu)\rfloor} \binom{n}{\lambda} \leq 2^{nh(\delta+\nu)}$ for $\delta + \nu \leq 1/2$ (see, e.g., [41, Sec. 1.4]) then concludes the proof. $\qquad\square$

## 6.4  Privacy amplification: Proof of Theorem 3

The last main ingredient of our proof is a so-called Leftover Hashing Lemma. It ensures that if the smooth min-entropy of $X$ given some side information $B$ is large, then we can extract randomness from $X$ that is independent of $B$. The Leftover Hashing Lemma is, up to a slight change of the definition of the smooth min-entropy, due to Renner [10, Corollary 5.6.1]. The proof of this exact statement is provided in Appendix B for the convenience of the reader.

**Proposition 9.** *Let $\sigma_{XD} \in \mathcal{S}_\bullet(XD)$ be classical on $X$ and $\varepsilon \in \left[0, \sqrt{\mathrm{tr}(\sigma_{XD})}\right)$. Let $\mathcal{H}$ be a universal$_2$ family of hash functions from $\mathcal{X} = \{0,1\}^n$ to $\mathcal{K} = \{0,1\}^\ell$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{H}|}\sum_{h\in\mathcal{H}} |h\rangle\langle h|_{S^H}$. Then,*

$$\|\omega_{KS^HD} - \chi_K \otimes \omega_{S^HD}\|_{\mathrm{tr}} \leq \frac{1}{2} 2^{-\frac{1}{2}(H_{\min}^\varepsilon(X|D)_\sigma - \ell)} + 2\varepsilon \tag{90}$$

*where $\chi_K = \frac{1}{2^\ell}\mathrm{id}_K$ is the fully mixed state and $\omega_{KS^HD} = \mathrm{tr}_X\left(\mathcal{E}_f(\sigma_{XD} \otimes \rho_{S^H})\right)$ for the function $f : (x,h) \mapsto h(x)$ that acts on the registers $X$ and $S^H$.*

The following technical lemma allows us to bound the smooth conditional min-entropy restricted to events in terms of the unrestricted entropy.

**Lemma 10.** *Let $\rho_{ABXY} \in \mathcal{S}_\bullet(ABXY)$ be classical on $X$ and $Y$ and let $\Omega : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be an event with $\Pr[\Omega]_\rho > 0$. Then, for $\varepsilon \in \left[0, \sqrt{\Pr[\Omega]_\rho}\right)$, we have*

$$H_{\min}^\varepsilon(AX \wedge \Omega|BY)_\rho \geq H_{\min}^\varepsilon(AX|BY)_\rho \quad and \quad H_{\min}^\varepsilon(AX \wedge \Omega|B)_\rho \geq H_{\min}^\varepsilon(AX|B)_\rho \tag{91}$$

*Proof.* Let us start with the first inequality. By definition of the smooth conditional min-entropy there exists a sub-normalized state $\tilde{\rho}_{ABXY} \in \mathcal{S}_\bullet(ABXY)$ and a state $\sigma_{BY} \in \mathcal{S}(BY)$ such that

$$\tilde{\rho}_{ABXY} \leq 2^{-H_{\min}^\varepsilon(AX|BY)_\rho} \mathrm{id}_{AX} \otimes \sigma_{BY} \quad and \quad P(\tilde{\rho}_{ABXY}, \rho_{ABXY}) \leq \varepsilon. \tag{92}$$

Without loss of generality [22, Lemma 6.6] we can assume that $\tilde{\rho}_{ABXY}$ is classical on $X$ and $Y$. As such, we have $\tilde{\rho}_{ABXY \wedge \Omega} \leq \tilde{\rho}_{ABXY}$ and $P(\tilde{\rho}_{ABXY \wedge \Omega}, \rho_{ABXY \wedge \Omega}) \leq P(\tilde{\rho}_{ABXY}, \rho_{ABXY}) \leq \varepsilon$ by the monotonicity of the purified distance under trace non-increasing maps. The desired inequality then follows by definition of the smooth min-entropy evaluated for the state with the event $\Omega$.

The second inequality follows similarly. By definition of the smooth conditional min-entropy there exists a sub-normalized state $\tilde{\rho}_{ABX} \in \mathcal{S}_\bullet(ABX)$ and a state $\sigma_B \in \mathcal{S}(B)$ such that

$$\tilde{\rho}_{ABX} \leq 2^{-H_{\min}^\varepsilon(AX|B)_\rho} \mathrm{id}_{AX} \otimes \sigma_B \quad and \quad P(\tilde{\rho}_{ABX}, \rho_{ABX}) \leq \varepsilon. \tag{93}$$

As discussed previously, this implies in particular the existence of an extension $\tilde{\rho}_{ABXY} \in \mathcal{S}_\bullet(ABXY)$ that satisfies $P(\tilde{\rho}_{ABXY}, \rho_{ABXY}) \leq \varepsilon$. Without loss of generality we can assume that $\tilde{\rho}_{ABXY}$ is classical on $X$ and $Y$. (To see this, note that pinching in the computational basis on $Y$ would indeed only decrease the distance between the $\tilde{\rho}_{ABXY}$ and $\rho_{ABXY}$, leaving the latter state invariant.) On the state $\tilde{\rho}_{ABXY}$ we can now define the restriction on the event $\Omega$ and find

$$\tilde{\rho}_{ABXY \wedge \Omega} \leq \tilde{\rho}_{ABXY} \implies \tilde{\rho}_{ABX \wedge \Omega} = \mathrm{tr}_Y\{\tilde{\rho}_{ABXY \wedge \Omega}\} \leq \tilde{\rho}_{ABX}. \tag{94}$$

Finally, we proceed in the same fashion as for the first inequality to show that $P(\tilde{\rho}_{ABX \wedge \Omega}, \rho_{ABX \wedge \Omega}) \leq \varepsilon$ and conclude the proof. □

The next proposition builds on Corollary 5 and Proposition 8 and the above Leftover Hashing Lemma to establish the secrecy of the key.

**Proposition 11.** *Let $\rho_{ABE} \in \mathcal{S}(ABE)$. Consider the protocol $\mathtt{qkd\_eb}_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}$ in Section 3 with $\mathrm{pe} = \{k, \delta\}$, $\mathrm{ec} = \{t, r, \ldots\}$ and $\mathrm{pa} = \{\ell, \ldots\}$ and the state $\omega_{K_A K_B SCFE} = \mathtt{qkd\_eb}_{m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}(\rho_{ABE})$. Define $\varepsilon(\nu)$ as in (79). Then, for any $\nu \in (0, \frac{1}{2} - \delta]$ such that $\varepsilon(\nu)^2 < \Pr[F = (\checkmark, \checkmark)]_\sigma$, the following holds:*

$$\left\| \omega_{K_A SCFE \wedge F=(\checkmark,\checkmark)} - \chi_{K_A} \otimes \omega_{SCFE \wedge F=(\checkmark,\checkmark)} \right\|_{\mathrm{tr}} \leq \frac{1}{2} 2^{-\frac{1}{2} g(\nu)} + 2\varepsilon(\nu). \tag{95}$$

*where $g(\nu) := (m - k)\left(\log \frac{1}{c} - h(\delta + \nu)\right) - r - t - \ell$.*

Note that in the above proposition, in order to ensure that the smooth min-entropy is always well-defined, we restricted ourselves to the case where the success probabillity exceeds the squared smoothing parameter, i.e. we required that $\varepsilon(\nu)^2 < \Pr[F = (\checkmark, \checkmark)]$. The case where the success probability is small will be handled separately in Corollary 12 below.

*Proof.* We use $n = m - k$. Since $\Pr[F = (\checkmark\checkmark)]_\sigma \leq \Pr[F^{\mathrm{pe}} = \checkmark]_\sigma$ the condition of Proposition 8 is satisfied and we find that $H_{\max}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|Y)_\sigma \leq nh(\delta + \nu)$ for the state $\sigma_{XYVWS^\Pi S^\Xi S^\Theta F^{\mathrm{pe}} E}$ as in (42) that results after measurement and parameter estimation. Combining this with Corollary 5 yields

$$H_{\min}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|VWS^\Pi S^\Xi S^\Theta E)_\sigma \geq nq, \tag{96}$$

where we introduced the shorthand $q = \log \frac{1}{c} - h(\delta + \nu)$.

Our goal is to translate this in a condition on the state $\sigma_{X\hat{X}C^V C^Z C^T S^\Pi S^\Xi S^\Theta S^{H_{\mathrm{ec}}} F^{\mathrm{pe}} F^{\mathrm{ec}} E}$ as in (44) that results after error correction. The following chain of inequalities holds:

$$nq \leq H_{\min}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|S^\Pi S^\Xi S^\Theta C^V E)_\sigma \tag{97}$$

$$\leq H_{\min}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|S^\Pi S^\Xi S^\Theta C^V C^Z E)_\sigma + r \tag{98}$$

$$= H_{\min}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|S^\Pi S^\Xi S^\Theta S^{H_{\mathrm{ec}}} C^V C^Z E)_{\sigma \otimes \rho} + r \tag{99}$$

$$\leq H_{\min}^{\varepsilon(\nu)}(X \wedge F^{\mathrm{pe}} = \checkmark|S^\Pi S^\Xi S^\Theta S^{H_{\mathrm{ec}}} C^V C^Z C^T E)_\sigma + r + t \tag{100}$$

$$\leq H_{\min}^{\varepsilon(\nu)}\left(X \wedge F^{\mathrm{pe}} = \checkmark \wedge F^{\mathrm{ec}} = \checkmark \middle| S^\Pi S^\Xi S^\Theta S^{H_{\mathrm{ec}}} C^V C^Z C^T E\right)_\sigma + r + t. \tag{101}$$

21

The first inequality follows by relabeling $V$ to $C^V$ and discarding $W$, an instance of the data-processing inequality. The transcript register $C^Z$ contains the syndrome sent from Alice to Bob and the inequality (98) follows by the chain rule in (19), and the fact that $\log|C^Z| = r$. The register $S^{H_{ec}}$ in the state $\rho_{SH_{ec}}$ is independent of the other registers. The register $C^T$ contains the hash of the raw key $X$ of size $\log|C^T| = t$ leading to the penultimate inequality. In the last step we used Lemma 10.

Summarizing $S' = (S^\Pi, S^\Xi, S^\Theta, S^{H_{ec}})$ as well as $C = (C^V, C^Z, C^T)$, and $F = (F^{pe}, F^{ec})$ as usual, we can thus more compactly write this as

$$H_{\min}^{\varepsilon(\nu)}(X \wedge F = (\checkmark, \checkmark) | S'CE)_\sigma \geq nq - r - t. \tag{102}$$

Proposition 9 applied with this bound then immediately yields the desired inequality. $\qquad \square$

The above proposition suffers from the assumption $\varepsilon(\nu)^2 < \Pr[F = (\checkmark, \checkmark)]_\omega$. However, the other case can easily be dealt with. Indeed, the trace distance $\left\| \omega_{K_A SCFE \wedge F = (\checkmark, \checkmark)} - \chi_{K_A} \otimes \omega_{SCFE \wedge F = (\checkmark, \checkmark)} \right\|_{tr}$ is upper bounded by the the trace of both states, i.e. the probability $\Pr[F = (\checkmark, \checkmark)]_\omega$. The inequalities $\Pr[F = (\checkmark, \checkmark)]_\sigma \leq \varepsilon(\nu)^2 \leq \varepsilon(\nu)$ then yield the following corollary, which is equivalent to Theorem 3.

**Corollary 12.** *Consider the setup of Proposition 11. For any $\nu \in (0, \frac{1}{2} - \delta]$ the following holds:*

$$\left\| \omega_{K_A SCFE \wedge F = (\checkmark, \checkmark)} - \chi_{K_A} \otimes \omega_{SCFE \wedge F = (\checkmark, \checkmark)} \right\|_{tr} \leq \frac{1}{2}\sqrt{2^{-(m-k)\left(\log\frac{1}{\bar{c}} - h(\delta + \nu)\right) + r + t + \ell}} + 2\varepsilon(\nu). \tag{103}$$

# Part II
# Prepare-and-measure protocol

| | |
|---|---|
| $\mathcal{N}_{A \to B}$ | Quantum channel between Alice and Bob |
| $\mathcal{P}_{\varnothing \to A | RS^{\Phi_A}}$ | Preparation map that returns a state in register $A$ depending on the settings $R, S^{\Phi_A}$ |
| $M$ | Number of states sent by Alice in the prepare-and-measure version |
| $\Omega$ | Subset of $[M]$ for which Bob obtains a conclusive measurement result |
| $\Sigma$ | Subset of $m$ indices where Alice and Bob's settings agree and Bob obtained a conclusive outcome |
| ro | Reordering map used in the sifting step. |
| $R$ | Register for Alice's raw key in the prepare-and-measure protocol |
| $U$ | Register for Bob's measurement results in the prepare-and-measure protocol |
| $S^{\Phi_A}$ | Seed for the choice of Alice's measurement bases in the prepare-and-measure protocol |
| $S^{\Phi_B}$ | Seed for the choice of Bob's measurement bases in the prepare-and-measure protocol |
| $S$ | Register corresponding to all the seeds that Alice communicates to Bob after state distribution $S = (S^\Pi, S^\Xi, S^\Theta, S^{H_{pe}}, S^{H_{ec}})$ |
| $F^{si}$ | Flag for the sifting procedure in the prepare-and-measure protocol |
| $F$ | Register corresponding to all the flags, $F = (F^{si}, F^{pe}, F^{ec})$ |
| $C^\Omega, C^\Sigma$ | Transcripts of the registers $\Omega$ and $\Sigma$ sent during sifting |
| $C$ | Register containing all the communication transcripts, $C = (C^\Omega, C^\Sigma, C^V, C^Z, C^T)$ |

Table 4: Additional nomenclature and notation used in Part II. See also Table 1

## 7 Formal description of the prepare-and-measure protocol

Here we discuss a prepare-and-measure (PM) protocol for QKD, denoted `qkd_pm`$_{M,m,pe,ec,pa}$, which is essentially equivalent to BB84 [1], and prove that its security follows from that of the entanglement-based protocol considered in Part I, provided that some *additional assumptions* are made.

Section 7.2 provides the details of the protocol described in Table 5, for the steps where it differs from the entanglement-based protocol. We describe the additional assumptions on the preparation and measurement devices in 7.1 and present a mathematical model of the protocol in 7.3.

## 7.1  Additional assumptions on preparation and measurement devices

The physical equipment of Alice an Bob is modified compared to that of the entanglement-based protocol considered in Part I. Indeed, the main point of implementing a prepare-and-measure protocol is that it is no longer required for Alice and Bob to share an entangled state, a task that remains very challenging if the two parties are a few tens of kilometers apart, which is typical in realistic scenarios where one wants to distribute secret keys at the scale of a metropolitan area. In the prepare-and-measure setup Alice and Bob do not start with an entangled state but instead have access to a quantum channel from Alice to Bob.

The assumption on finite-dimensional quantum systems, sealed laboratories, random seeds and authenticated communication channel discussed in Section 3.1 still apply. The assumption of sealed laboratories in particular implies that the quantum channel between Alice and Bob models all quantum communication leaving Alice's lab. However, we will replace the assumption of commuting measurements, deterministic detection and the complementarity of Alice's measurements.

**Alice's preparation:**  In every round, indexed by $i \in [M]$, Alice's preparation device takes two bits as input: $\phi \in \{0,1\}$ describing a basis choice and $x \in \{0,1\}$ describing the bit value within each basis. It produces a quantum state $\rho_{A_i}^{\phi,x}$, ideally corresponding to one of the four BB84 states. The commuting measurement assumption for Alice is replaced with the requirement that these states do not depend on the preparations in previous or later rounds (which is already ensured by our notation). Our next assumption is that the state $\rho_{A_i}^{\phi,x}$ does not leak any information about the basis choice $\phi$, i.e., we require that

$$\sum_{x \in \{0,1\}} \rho_{A_i}^{0,x} = \sum_{x \in \{0,1\}} \rho_{A_i}^{1,x}. \tag{104}$$

Moreover, instead of the complementarity assumption on the measurements, we require that the prepared states are sufficiently complementary. More precisely, let us define

$$c'(\{\rho^x\}_x, \{\sigma^y\}_y) := \max_{x,y} \left\| \sqrt{\rho^x} X^{-1} \sqrt{\sigma^y} \right\|_\infty^2, \quad \text{for states satisfying} \quad \sum_x \rho^x = \sum_y \sigma^y := X. \tag{105}$$

In case $X$ has not full support we take the generalized inverse (on its support) in the above definition. Our second assumption on Alice's preparation is that

$$c_i' := c'\left(\left\{\rho_{A_i}^{0,x}\right\}_x, \left\{\rho_{A_i}^{1,y}\right\}_y\right) \le \bar{c}' \tag{106}$$

for all $i \in [M]$ and some constant $\bar{c}' < 1$. The constant $\bar{c}'$ is closely related to the constant $\bar{c}$ that described the complementarity of Alice's measurement in the entanglement-based protocol, as we will see in Corollary 15.

In an ideal implementation of the BB84 protocol, the states $\rho^{\phi,r}$ would be single-qubit states given by

$$\rho^{0,0} = |0\rangle\langle0|, \quad \rho^{0,1} = |1\rangle\langle1|, \quad \rho^{1,0} = |+\rangle\langle+|, \quad \rho^{1,1} = |-\rangle\langle-|. \tag{107}$$

These states obviously satisfy the assumption in (104) and it is easy to verify that $\bar{c}' = \frac{1}{2}$ is a valid bound.

We should note that our first assumption is rather strong and for instance does not allow us to assess the security of popular implementations of the BB84 protocol relying on a weak-coherent-state encoding. Since single-photon sources remain expensive and imperfect today, it is indeed tempting to encode each qubit with two polarization modes and replace single-photons by phase-randomized weak-coherent states[11]. For such an implementation, the four BB84 states become linearly independent, and Eq. (104) cannot hold. It is well-known that such implementations are sensitive to "photon-number-splitting" attacks but that solutions exist to restore their security, for instance with the help of decoy states [42]. While we believe that our framework could accommodate such modifications (see for instance [43, 44]), we do not address this issue here.

**Bob's measurement:**  As for the simple protocol of Part I, we require that Bob's quantum system can be decomposed as $B \equiv B_{[M]} = B_1 B_2 \dots B_M$. Bob's measurement device is similar to that of the simple protocol of Part I, but the measurement operators now need to be specified for indices in $[M]$ and allow for an additional outcome, '∅', corresponding to an inconclusive result. Such an inconclusive result can for instance occur when no detector clicked (photon loss) or when more than 1 detector clicked (dark counts)[12].

---

[11]If the single-photon polarization qubit states are given by (107), then the four encoded BB84 states are $\tau^{i,j} = e^{-\alpha^2} \bigoplus_{k=0}^{\infty} \frac{\alpha^{2k}}{k!} (\rho^{i,j})^{\otimes k}$ for $i, j \in \{0,1\}$, where $\alpha > 0$ is the amplitude of the coherent states.

[12]Indeed, in most experiments, the measurement device is usually implemented with the help of two single-photon detectors, and a conclusive measurement outcome, 0 or 1, will correspond to which detector clicked while inconclusive outcomes occur if none or both of the detectors clicked.

For any $i \in [M]$, we model Bob's measurement on subsystem $B_i$ with setting $\phi \in \{0,1\}$ by a ternary generalized measurement $\{M_{B_i}^{\phi,z}\}_{z \in \{0,1,\varnothing\}}$ acting on $B_i$. The index $z$ ranges over the two conclusive outcomes, 0 and 1, of Bob's measurement as well as the inconclusive outcomes $\varnothing$. We require that

$$(M_{B_i}^{0,\varnothing})^\dagger (M_{B_i}^{0,\varnothing}) = (M_{B_i}^{1,\varnothing})^\dagger (M_{B_i}^{1,\varnothing}), \tag{108}$$

meaning that the element corresponding to an inconclusive result coincides for both measurements. As will be formalized in Lemma 16, this implies that Bob's measurement map can be interpreted as a two-step process first deciding whether the result is conclusive or not, and then, in the former case, proceeding with the ideal measurement considered in Part I. While this assumption seems quite reasonable for photon detector working in the few photons regime, it usually fails to apply when avalanche photo diodes are accessed in the linear mode, and this was precisely the origin of the "blinding attack" of [33].

In any realistic implementation, Alice and Bob would need to be synchronized so that Bob can keep track of which system he is currently measuring. This is especially relevant in high-loss regimes where Bob's detectors would not click most of the time. Such a synchronization procedure can be realized classically, provided that both players have access to an authenticated channel. For simplicity, we ignore this synchronization issue in our model.

## 7.2 Protocol parameters and overview

The protocol $\texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}$ is parametrized as $\texttt{qkd\_eb}_{m,\text{pe},\text{ec},\text{pa}}$, but with one extra parameter:

- The number of individual states prepared and sent through the quantum channel by Alice, $M \in \mathbb{N}$. We require that $M \geq m$ and for optical implementations a typical value for $M$ is $\frac{2m}{\eta}$ where $\eta$ is the overall transmittance of the optical channel between Alice and Bob.

## 7.3 Exact mathematical model of the protocol

Here we describe in detail the mathematical model corresponding to the protocol in Table 5, for the steps where it differs from the simple protocol of Part I.

**Input:** The realistic protocol $\texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}$ we consider is a prepare-and-measure protocol, and the role of the input is now played by an (arbitrary) quantum channel $\mathcal{N}_{A \to B}$ between Alice and Bob. Here $A = A_{[M]}$ and $B = B_{[M]}$. This situation is depicted in Figure 8.

As before, we make the assumption that the input and output of the quantum channel are finite-dimensional. This arguably appears quite restrictive since any complete description of the optical channel would involve infinite-dimensional Fock spaces, with the idea that each of the $M$ systems prepared by Alice corresponds to two polarization modes for instance. However, we point out that we do not require any explicit upper bound on the dimension of the Hilbert spaces occurring in the protocol, and that any *physical* state necessarily has a bounded energy, which means that it can be arbitrarily well approximated by a quantum state in a finite-dimensional Hilbert space of sufficiently large dimension.
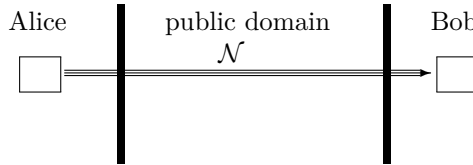


Figure 8: (Input.) Alice and Bob have access to a quantum channel: $\mathcal{N}: A \to B$.

**Randomization:** The random seeds are modeled similarly as for the idealized version of the protocol. Here, the seed $S^\Phi$ corresponding to identical measurement settings is not provided directly. Instead, Alice and Bob initially choose independently two strings $\Phi_A, \Phi_B \in \{0,1\}^M$, and it will later be the role of the sifting procedure to produce a set of identical measurement settings $\Phi$. The random choice of the strings $\Phi_A, \Phi_B$ is modeled by two registers $S^{\Phi_A}, S^{\Phi_B}$ in the state

$$\rho_{S^{\Phi_A}} \otimes \rho_{S^{\Phi_B}} = \frac{1}{4^M} \sum_{\phi_A, \phi_B \in \{0,1\}^M} |\phi_A\rangle\langle\phi_A|_{S^{\Phi_A}} \otimes |\phi_B\rangle\langle\phi_B|_{S^{\Phi_B}}, \tag{109}$$

24

$(K_A, K_B, S, C, F) = \texttt{qkd\_pm}_{M,m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}(\mathcal{N}_{A \to B})$:

**Input:** Alice and Bob have access to a quantum channel $\mathcal{N}_{A \to B} : A \to B$ where $A \equiv A_{[M]}$ and $B \equiv B_{[M]}$ are comprised of $M$ quantum systems.

**Randomization:** Alice and Bob respectively choose two random strings $\Phi_A, \Phi_B \in \{0,1\}^M$. Alice also chooses a random string $R \in \{0,1\}^M$. These private seeds are denoted $S^{\Phi_A}, S^{\Phi_B}$ and $R$. Finally, similarly as in the simple protocol, Alice chooses a random subset $\Pi \in \Pi_{m,k}$, and random hash functions $H_{\mathrm{ec}} \in \mathcal{H}_{\mathrm{ec}}$ as well as $H_{\mathrm{pa}} \in \mathcal{H}_{\mathrm{pa}}$. These uniformly random seeds are denoted $S = (S^\Pi, S^{H_{\mathrm{ec}}}, S^{H_{\mathrm{pa}}})$.

**State Preparation:** Alice prepares a quantum state $\rho_A^{R,\Phi_A}$, encoding the string $R$ in the measurement basis corresponding to $\phi_A$.

**State Distribution:** Alice sends the state $\rho_A^{R,\Phi_A}$ through the quantum channel $\mathcal{N}$ and Bob receives the output state $\rho_B^{R,\Phi_A} = \mathcal{N}(\rho_A^{R,\Phi_A})$. (In practice, Alice would send the systems one by one, and use the quantum channel $M$ times.)

**Measurement:** Bob measures the $M$ quantum systems with the setting $\Phi_B$, and stores his ternary measurement outcomes in a string $U \in \{0, 1, \varnothing\}^M$, where $\varnothing$ denotes an inconclusive measurement result. He also computes the set $\Omega \subseteq 2^{[M]}$ of indices corresponding to conclusive measurements (In practice, Bob would start measuring the systems as soon as he receives them.)

**Randomness distribution:** Bob publicly announces both the value of $\Phi_B$ and $\Omega$. The corresponding transcripts are denoted by $C^{\Phi_B}$ and $C^\Omega$, respectively. Alice sends the value of the seeds $S = (S^\Pi, S^{H_{\mathrm{ec}}}, S^{H_{\mathrm{pa}}})$ to Bob through the authenticated public channel.

**Sifting:** If it exists, Alice publicly announces a set $\Sigma \subseteq \Omega$, with transcript $C^\Sigma$ of cardinality $m$, such that $\Phi_A$ and $\Phi_B$ coincide on $\Sigma$, and sets the flag $F^{\mathrm{si}} = \checkmark$. Otherwise, she sets $F^{\mathrm{si}} = \varnothing$ and they abort. The respective binary substrings $R'$ and $U'$ of $R$ and $U$ restricted to $\Sigma$ become the raw keys. As in the idealized protocol, they are then reordered and denoted $(X, V)$ and $(Y, W)$ for Alice and Bob, respectively. Here $V, W$ are of length $k$ and correspond to the indices in $\Pi$, whereas $X, Y$ of length $n$ correspond to indices not in $\Pi$.

**Parameter Estimation:** Alice sends $V$ to Bob, the transcript is denoted $C^V$. Bob compares $V$ and $W$. If the fraction of errors exceeds $\delta$, Bob sets the flag $F^{\mathrm{pe}} = `\varnothing`$ and they abort. Otherwise he sets $F^{\mathrm{pe}} = `\checkmark`$ and they proceed.

**Error Correction:** Alice sends the syndrome $Z = \mathrm{synd}(X)$ to Bob, with transcript $C^Z$. Bob computes $\hat{X} = \mathrm{corr}(Y, Z)$.

Alice computes the hash $T = H_{\mathrm{ec}}(X)$ of length $t$ and sends it to Bob, with transcript $C^T$. Bob computes $H_{\mathrm{ec}}(\hat{X})$. If it differs from $T$, he sets the flag $F^{\mathrm{ec}} = `\varnothing`$ and they abort the protocol. Otherwise he sets $F^{\mathrm{ec}} = `\checkmark`$ and they proceed.

**Privacy Amplification:** They compute keys $K_A = H_{\mathrm{pa}}(X)$ and $K_B = H_{\mathrm{pa}}(\hat{X})$ of length $\ell$.

**Output:** The output of the protocol consists of the keys $K_A$ and $K_B$, the seeds $(S^{\Phi_B}, S^\Pi, S^{H_{\mathrm{ec}}}, S^{H_{\mathrm{pa}}})$, the transcript $C = (C^\Omega, C^\Sigma, C^V, C^Z, C^H)$ and the flags $F = (F^{\mathrm{si}}, F^{\mathrm{pe}}, F^{\mathrm{ec}})$. In case of abort, we assume that all registers are initialized to a predetermined value.

Table 5: Realistic Prepare-and-Measure QKD Protocol $\texttt{qkd\_pm}_{M,m,\mathrm{pe},\mathrm{ec},\mathrm{pa}}$. The precise mathematical model is described in Section 7. This protocol differs from the entanglement-based protocol in several points: in particular, the input now corresponds to the quantum channel $\mathcal{N}$ between Alice and Bob.

where $\{|\phi_A\rangle\}, \{|\phi_B\rangle\}$ are orthonormal bases of $S^{\Phi_A}$ and $S^{\Phi_B}$, respectively.

Another difference with the simple protocol of Part I is that Alice also has access to a register $R$ that she will use to choose which state to prepare. This register is modeled similarly as the other seeds as a maximally mixed state:

$$\rho_R = \frac{1}{2^M} \sum_{r \in \{0,1\}^M} |r\rangle\langle r|_R, \tag{110}$$

where $\{|r\rangle\}$ is an orthonormal basis of $R$. The other random seeds $\rho_{S^\Pi}, \rho_{S^{H_{ec}}}, \rho_{S^{H_{pa}}}$ are identical to the idealized version. This situation is depicted in Figure 9.
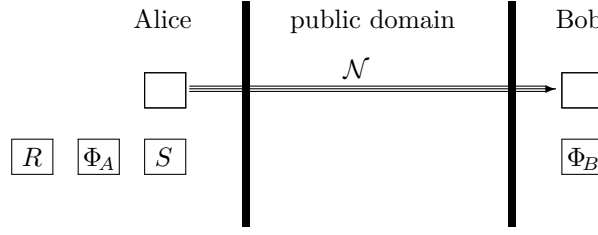


Figure 9: (Randomization.) Alice prepares a public seed $S$ that will later be communicated to Bob through an authenticated classical channel. Alice and Bob also prepare private seeds: $R, \Phi_A$ for Alice, $\Phi_B$ for Bob.

Note that while in the simple protocol all the seeds are communicated publicly during a step of the protocol, it is crucially not the case here for the seed $R$, from which the final key could be immediately inferred. We will see that it is not necessary to communicate the seed $S^{\Phi_A}$ to Bob since the sifting procedure is performed by Alice.

In a practical implementation, the various random seeds, except for $S^{\Phi_B}$ would be initially prepared by Alice, and only communicated to Bob when needed (except for $R$ and $S^{\Phi_A}$). In particular, one should wait until the state distribution is over before communicating the value of the chosen subset for parameter estimation or of the various hash functions.

**State preparation:** Alice prepares a quantum state on $M$ systems $A \equiv A_{[M]}$ using the map

$$\mathcal{P}_{\varnothing \to A | R S^{\Phi_A}}(\cdot) = \sum_{r, \phi \in \{0,1\}^M} \left( |r\rangle\langle r|_R \otimes |\phi\rangle\langle\phi|_{S^{\Phi_A}} \right) \cdot \left( |r\rangle\langle r|_R \otimes |\phi\rangle\langle\phi|_{S^{\Phi_A}} \right) \otimes \rho_A^{r,\phi} \tag{111}$$

where $\rho_A^{r,\phi} = \bigotimes_{i=1}^M \rho_{A_i}^{r_i, \phi_i}$. Applying this map to the seeds in registers $R$ and $S^{\Phi_A}$ results in the state

$$\rho_{R S^{\Phi_A} A} = \frac{1}{4^M} \sum_{r, \phi \in \{0,1\}^M} |r\rangle\langle r|_R \otimes |\phi\rangle\langle\phi|_{S^{\Phi_A}} \otimes \rho_A^{r,\phi}. \tag{112}$$
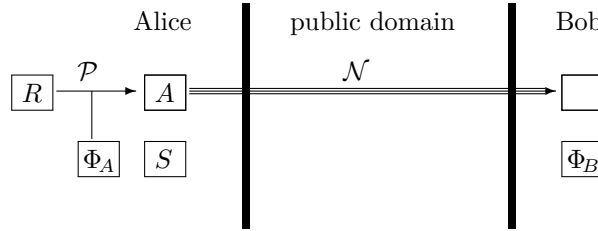
This situation is depicted in Figure 10.



Figure 10: (State Preparation.) Using her private seeds $R$ and $\Phi_A$, Alice prepares system $A$.

**State distribution:** Alice sends her register $A$ to Bob through the quantum channel $\mathcal{N} : A \to B$. The state shared by Alice and Bob is given by:

$$\rho_{R S^{\Phi_A} B} = \mathcal{N}_{A \to B}(\rho_{R S^{\Phi_A} A}) \tag{113}$$

$$= \frac{1}{4^M} \sum_{r, \phi \in \{0,1\}^M} |r\rangle\langle r|_R \otimes |\phi\rangle\langle\phi|_{S^{\Phi_A}} \otimes \rho_B^{r,\phi}, \tag{114}$$

where we defined $\rho_B^{r,\phi} = \mathcal{N}\left( \rho_A^{r,\phi} \right)$. This situation is depicted in Figure 11.
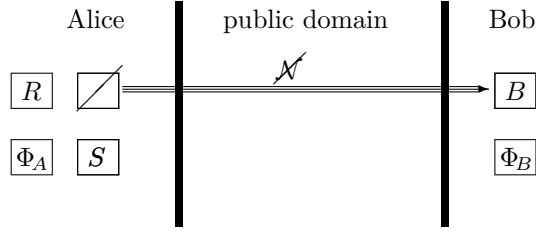
Figure 11: (State Distribution.) Alice sends the system $A$ through the quantum channel $\mathcal{N}$ and Bob receives system $B$.
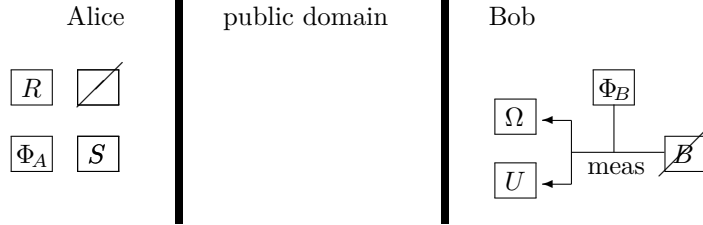


Figure 12: (Bob's Measurement.) Bob measures the quantum system $B$ using the measurement settings given by his private randomness $\Phi_B$ and obtains two classical strings stored in registers $\Omega$ and $U$.

**Measurement:** Bob measures each of his $M$ quantum systems in the basis corresponding to $\Phi_B$ and stores his measurement outcomes, either 0, 1, or $\varnothing$ in the case of inconclusive outcomes, in a classical register $U$ taking values in $\{0, 1, \varnothing\}^M$. The measurement map $\mathcal{M}_{B \to U\Omega|S^{\Phi_B}}$ is defined as:

$$\mathcal{M}_{B \to U\Omega|S^{\Phi_B}}(\cdot) = \sum_{\phi \in \{0,1\}^M} \sum_{u \in \{0,1,\varnothing\}^M} |u, \omega\rangle_{TC^\Omega} \left( M_B^{\phi,u} \otimes |\phi\rangle\langle\phi|_{S^{\Phi_B}} \right) \cdot \left( M_B^{\phi,u} \otimes |\phi\rangle\langle\phi|_{S^{\Phi_B}} \right)^\dagger \langle u, \omega|_{UC^\Omega}, \quad (115)$$

where $\omega = \omega(u)$ is the subset of $[M]$ where $u$ takes values in $\{0, 1\}$, namely

$$\omega(u) := \{i \in [M] : u_i \neq \varnothing\}. \quad (116)$$

The state of the total system after Bob's measurement is given by

$$\sigma_{RUC^\Omega BS^{\Phi_A}S^{\Phi_B}} = \frac{1}{8^M} \sum_{r,\phi_A,\phi_B \in \{0,1\}^M} \sum_{u \in \{0,1,\varnothing\}^M} |r, u, \omega, \phi_A, \phi_B\rangle\langle r, u, \omega, \phi_A, \phi_B|_{RUC^\Omega S^{\Phi_A}S^{\Phi_B}} \otimes M_B^{\phi_B,u} \rho_B^{r,\phi_A} \left( M_B^{\phi_B,u} \right)^\dagger,$$
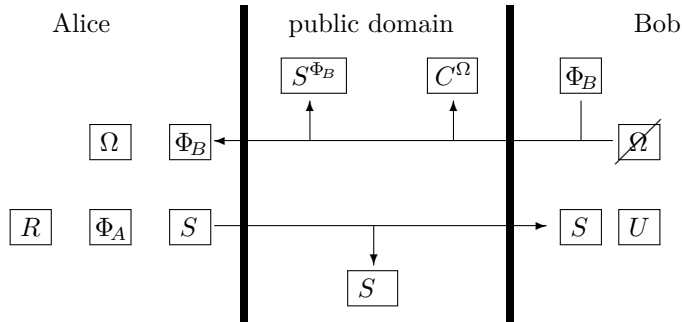
$$(117)$$

and the situation is depicted in Figure 12.



Figure 13: (Randomness distribution.) Bob communicates the both the value of $\Phi_B$ and $\Omega$ to Alice with the authenticated classical channel. Alice publicly announces the value of the various seeds, $S = (S^\Pi, S^{H_{ec}}, S^{H_{pa}})$.

**Randomness distribution:** Bob publicly announces the content of the register $S^{\Phi_B}$ together with the description $C^\Omega$ of the set $\omega$ of indices corresponding to conclusive measurement results. This situation is depicted in Figure 13. Alice publicly announces the value of the various seeds, $S = (S^\Pi, S^{H_{ec}}, S^{H_{pa}})$.

**Sifting:** Alice applies the *sifting map*, a classical map 'sift' defined as follows

$$\text{sift} : \begin{cases} \{0,1\}^M \times \{0,1\}^M \times 2^{[M]} & \to & \Pi_{M,m} \times \{\varnothing, \checkmark\} \\ (\phi_A, \phi_B, \omega) & \mapsto & (\Sigma, F^{\text{si}}) \end{cases} \tag{118}$$

where $\Sigma$ is either the first subset of $\Omega$ of cardinality $m$ in the lexicographic order where $\phi_A$ and $\phi_B$ coincide, if such a set exists, or else it is set to a dummy value, for instance $[m]$. In the first case, the flag $F^{\text{si}}$ is set to $\checkmark$, otherwise it is set to $\varnothing$ and the protocol aborts. The output of the sifting map immediately allows Alice and Bob to compute the value of the seed $S^\Phi$ which is simply the restriction of either $S^{\Phi_A}$ or $S^{\Phi_B}$ to the indices in $\Sigma$. This classical map is lifted to give a CPTP map $\mathcal{E}_{\text{si}} = S^{\Phi_A} S^{\Phi_B} C^\Omega \to C^\Sigma C^\Omega F^{\text{si}} S^{\Phi_A} S^{\Phi_B}$. This situation is depicted in Figure 14.
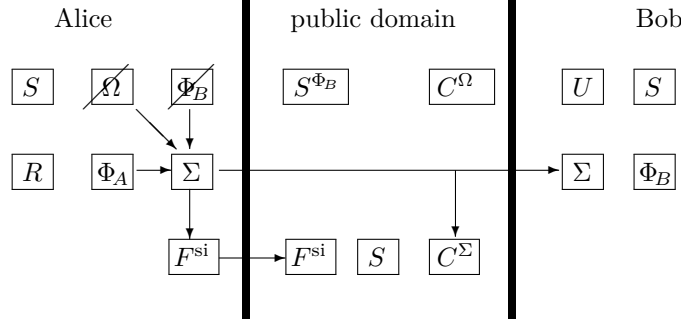


Figure 14: (Sifting $1/2$.) Using her private randomness $\Phi_A$ together with the registers $\Omega$ and $\Phi_B$ sent by Bob, Alice computes the set $\Sigma$ when it exists and sets the value of the flag $F^{\text{si}}$. Both values are then publicly announced.

We then define a CPTP map $\mathcal{E}_{\text{di}}$ that discards the systems in $R$, $U$ and $\Phi_A$ which do not correspond to the subset $\Sigma$ and put the remaining systems in registers denoted $R'$, $U'$ and $\Phi$ of size $m$, respectively.

$$\mathcal{E}_{\text{di}} : C^\Sigma R U S^{\Phi_A} \to R' U' C^\Phi C^\Sigma. \tag{119}$$
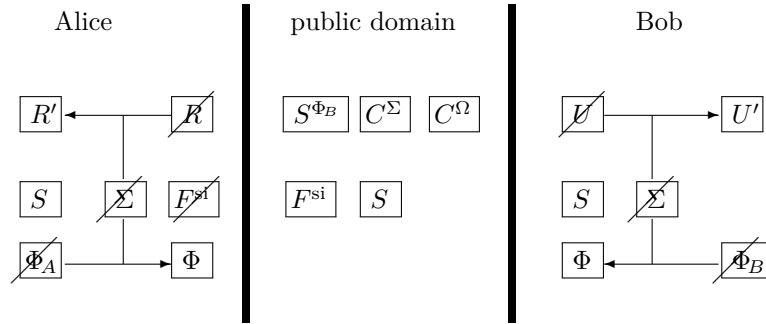
This situation is depicted in Figure 15.



Figure 15: (Sifting $2/2$.) Using $\Sigma$, Alice (resp. Bob) discards the $M - m$ irrelevant systems of $R$ (resp. $U$) and stores the $m$ remaining ones in $R'$ (resp. $U'$). Both Alice and Bob further compute $\Phi$ from $\Sigma$ and either $\Phi_A$ or $\Phi_B$.

Finally, similarly as in the simple protocol of Part I, Alice and Bob use the content of register $\Pi$ to reorder their raw keys, which become $(V, X)$ for Alice and $(W, Y)$ for Bob. The situation here is similar to that obtained after measurement in the simple protocol (compare Figures 16 and 3), with the addition of the registers $C^\Sigma$, $C^\Omega$, $S^{\Phi_B}$ and $F^{\text{si}}$ now available in the public domain.

**Remaining steps:** The remaining steps are as in the entanglement-based QKD protocol presented in Part I.

## 8 Results and Discussion

The security proof should establish that for any input channel $\mathcal{N}_{A \to B}$ given to $\texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}$, either the protocol outputs secret identical keys, or else it aborts. In the same spirit as the entanglement-based version of

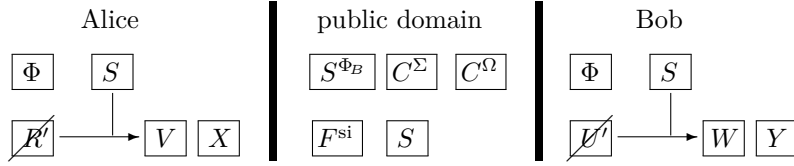Figure 16: (Reordering.) Using the content of $S^\Pi$, Alice reorder their raw keys, $R'$ and $U'$, which become respectively $(X, V)$ and $(Y, W)$.

Part I, we define the security parameter

$$\Delta_{M,\text{si},\text{pe},\text{ec},\text{pa}} := \sup_{\mathcal{N}_{A\to BE}} \left\| \texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}(\mathcal{N}_{A\to BE}) - \texttt{qkd\_ideal}_{M,k,n,\delta,\text{ec},\text{pa}}(\mathcal{N}_{A\to BE}) \right\|_{\text{tr}}, \tag{120}$$

where again $\texttt{qkd\_ideal}_{M,k,n,\delta,\text{si},\text{ec},\text{pa}}$ is defined analogously to the entanglement-based case and simply replaces the output of $\texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}(\mathcal{N}_{A\to BE})$ with a perfect key in case the protocol does not abort. Here, the channels $\mathcal{N}_{A\to BE}$ have an additional output that goes to an eavesdropper, and it again suffices to consider maps where $E$ is finite-dimensional. Establishing security thus boils down to showing that this trace distance is small for all such channels.

Our strategy is to show that the realistic protocol is equivalent to applying the idealized QKD protocol on a virtual quantum state $\rho_{AB}$ independent from the uniformly distributed random seed $S^\Phi$ for the measurement settings. If this holds, then the security proof of Part I for the simple protocol applies, and establishes the security of the prepare-and-measure protocol. For this, we need to make explicit assumptions about (i) the state preparation on Alice's side to make sure that no basis information is leaked and (ii) the measurement device on Bob's side to ensure that the invalid measurement results do not depend on the measurement basis.

Under the assumptions in Section 7.1, we show that the prepare-and-measure QKD protocol is secure.

**Theorem 13.** *Let* $m, \text{pe}, \text{ec}, \text{pa}$ *be such that the protocol* $\texttt{qkd\_eb}_{m,\text{pe},\text{ec},\text{pa}}$ *in Section 3 is $\varepsilon$-secure with device parameter* $\bar{c} = \bar{c}'$. *Then* $\texttt{qkd\_pm}_{M,m,\text{pe},\text{ec},\text{pa}}$ *is also $\varepsilon$-secure.*

As will be shown in Section 9, the security of the prepare-and-measure protocol is a consequence of that of the simple protocol studied in Part I, provided that some additional assumptions are made.

When assessing the performance of the protocol, however, two modifications appear. First, the device parameter ($\bar{c}'$ instead of $\bar{c}$) needs to be defined differently since it is no longer a function of the measurement device of Alice, but rather of her preparation device. In an ideal implementation, it is still expected to be equal to $\frac{1}{2}$, as was discussed in Section 7.1. The more important difference is due to the sifting procedure. Indeed, the definition of the *secret key rate* should now be modified to mean the ration between the key length $\ell$ and the number $M$ of individual states prepared and sent by Alice (instead of the number $m$ in the simple entanglement-based protocol). The means that the secret key rate achieved with the prepare-and-measure protocol is given by

$$\frac{\ell}{M} = \frac{m}{M} \cdot \frac{\ell}{m}. \tag{121}$$

The sifting procedure that we have considered here (and described in Section 7.3) is not optimized to maximize the secret key rate (or equivalently the ratio $\frac{m}{M}$), but rather to simplify the analysis as much as possible. Better sifting procedures are discussed in [19] and could involve not fixing the value of $M$ in advance for instance.

A typical experiment would be characterized by a given overall transmittance $\eta$ of the optical channel, meaning that approximately $\eta M$ photons will be detected by Bob, or in other words, that the expected value of $|\Omega|$ is $\eta M$. Given that Alice and Bob's measurement bases will coincide on expectation 50% of the time, we therefore expect that

$$\frac{m}{M} \approx \frac{\eta}{2} \tag{122}$$

holds asymptotically.

In particular, the secret key rate of the prepare-and-measure protocol is then expected to be equal to $\frac{\eta}{2}$ times the secret key rate of the simple protocol displayed in Figure 7.

## 9 Security reduction

In Sections 9.1 and 9.2, we discuss some implications of the device assumptions made in Section 7.1. The security reduction to the simple protocol will be addressed in Section 9.3.

## 9.1 Preparation: Assumptions on Alice's device

Consider the four states $\{\rho_{A_i}^{x,\phi}\}_{x,\phi\in\{0,1\}}$ created by Alice in round $i$ of the protocol for some $i \in [M]$. Since these states adhere to the assumptions stated in (104) and (106), the following lemma is applicable:

**Lemma 14.** *Let $\{\rho_A^{\phi,x}\}_{\phi,x} \subset \mathcal{S}(A)$ where $x$ and $\phi$ are taken from discrete sets. Moreover, let $\{p_x^\phi\}_x$ be a probability distribution for each $\phi$ such that*

$$\sum_x p_x^\phi \rho_A^{\phi,x} = \sum_x p_x^{\phi'} \rho_A^{\phi',x} \quad \text{for all } \phi \text{ and } \phi'. \tag{123}$$

*Then there exists a state $\tau_{AA'} \in \mathcal{S}(AA')$ where $A' \equiv A$ and a generalized measurement $\{M_{A'}^{\phi,x}\}_x$ on $A'$ for each $\phi$ such that*

$$p_x^\phi \rho_A^{\phi,x} = \mathrm{tr}_{A'}\left[\tau_{AA'} \, \mathrm{id}_A \otimes \left(M_{A'}^{\phi,x}\right)^\dagger M_{A'}^{\phi,x}\right] \quad \text{and} \quad c\left(\{M_{A'}^{\phi,x}\}_x, \{M_{A'}^{\phi',x}\}_x\right) = c'\left(\{\rho_A^{\phi,x}\}_x, \{\rho_A^{\phi',x}\}_x\right). \tag{124}$$

*Proof.* We will explicitly construct the state and measurement as follows. First, let us introduce $\tau_A := \sum_x p_x^\phi \rho_A^{\phi,x}$ and choose $\tau_{AA'}$ as its purification on $A'$. Now we choose

$$M_{A'}^{\phi,x} := \sqrt{p_x^\phi} \left(\left(\rho_A^{\phi,x}\right)^{\frac{1}{2}}\right)^T \left(\tau_A^{-\frac{1}{2}}\right)^T \tag{125}$$

where the transpose is taken with regards to the Schmidt basis of $\tau_{AA'}$. Let us first verify that this constitutes a generalized measurement. Indeed, for every $\phi$, we find

$$\sum_x \left(M_{A'}^{\phi,x}\right)^\dagger M_{A'}^{\phi,x} = \sum_x p_x^\phi \left(\tau_A^{-\frac{1}{2}}\right)^T \left(\rho_A^{\phi,x}\right)^T \left(\tau_A^{-\frac{1}{2}}\right)^T = \mathrm{id}_{A'}. \tag{126}$$

Let us now verify the conditions in (124). Since $\tau_{AA'} = |\tau_{AA'}\rangle\langle\tau_{AA'}|$ purifies $\tau_A$, we find

$$\mathrm{id}_A \otimes M_{A'}^{\phi,x} |\tau_{AA'}\rangle = \sqrt{p_x^\phi} \, \mathrm{id}_A \otimes \left(\left(\rho_A^{\phi,x}\right)^{\frac{1}{2}}\right)^T \left(\tau_A^{-\frac{1}{2}}\right)^T |\tau_{AA'}\rangle = \sqrt{p_x^\phi} \, |\rho_{AA'}^{\phi,x}\rangle, \tag{127}$$

where $\rho_{AA'}^{\phi,x} = |\rho_{AA'}^{\phi,x}\rangle\langle\rho_{AA'}^{\phi,x}|$ purifies $\rho_A^{\phi,x}$. The first equality readily follows. The second equality can be confirmed by consulting the definitions of $c$ and $c'$ in (20) and (105), respectively. $\qquad\square$

These assumptions on Alice's preparation allow us to replace the state preparation by a measurement performed on a virtual extension of the average prepared state.

**Corollary 15.** *If the two assumptions in (104) and (106) on Alice's preparation device hold, then for each $i \in [M]$ there exists a state $\rho_{A_i A_i'}$ where $A_i' \equiv A_i$ and generalized measurements on $A_i'$ of the form prescribed in Section 3.1 such that $c_i \leq c_i'$, and, in particular, $\bar{c} \leq \bar{c}'$. Combining all these measurements into a map $\mathcal{M}_{A' \to R|S^{\Phi_A}}$ acting on $A' \equiv A_{[M]}'$, we further have*

$$\rho_{AR\Phi_A} = \mathcal{M}_{A' \to R|S^{\Phi_A}}\left(\rho_{AA'} \otimes \rho_{\Phi_A}\right). \tag{128}$$

## 9.2 Assumption on Bob's measurement:

The objective of the assumption made on Bob's measurement is to show that when the sifting procedure succeeds, the resulting register $S^\Phi$ is independent of the state shared by Alice and Bob, similarly as in the entanglement-based protocol of Part I.

**Lemma 16.** *If Bob's measurement satisfies (108), that is, $(M_{B_i}^{0,\varnothing})^\dagger(M_{B_i}^{0,\varnothing}) = (M_{B_i}^{1,\varnothing})^\dagger(M_{B_i}^{1,\varnothing})$ for all $i \in [M]$, then the measurement map $\mathcal{M}_{B \to U\Omega|S^{\Phi_B}}$ can be decomposed as*

$$\mathcal{M}_{B \to U\Omega|S^{\Phi_B}} = \mathcal{M}_{B \to U|S^{\Phi_B}\Omega} \circ \mathcal{M}_{B \to B\Omega}. \tag{129}$$

*Proof.* Define for each $i \in [M]$ operators $M_{B_i}^{\checkmark}$ and $M_{B_i}^{\varnothing}$ satisfying

$$(M_{B_i}^{\checkmark})^\dagger(M_{B_i}^{\checkmark}) = \sum_{z\in\{0,1\}} (M_{B_i}^{0,z})^\dagger(M_{B_i}^{0,z}) = \sum_{z\in\{0,1\}} (M_{B_i}^{1,z})^\dagger(M_{B_i}^{1,z}) \tag{130}$$

$$(M_{B_i}^{\varnothing})^\dagger(M_{B_i}^{\varnothing}) = (M_{B_i}^{0,\varnothing})^\dagger(M_{B_i}^{0,\varnothing}) = (M_{B_i}^{1,\varnothing})^\dagger(M_{B_i}^{1,\varnothing}). \tag{131}$$

Note in particular that $(M_{B_i}^\checkmark)^\dagger(M_{B_i}^\checkmark) + (M_{B_i}^\varnothing)^\dagger(M_{B_i}^\varnothing) = \mathrm{id}_{B_i}$. The generalized measurement $\mathcal{M}_{B\to B\Omega}$ is then described by

$$\mathcal{M}_{B\to B\Omega} : \rho_B \mapsto \rho_{B\Omega} = \sum_{c\in\{\checkmark,\varnothing\}^M} |c\rangle\langle c|_\Omega \otimes M_B^c \rho_B (M_B^c)^\dagger, \tag{132}$$

with $M_B^c := \bigotimes_{i=1}^M M_{B_i}^{c_i}$. Up to relabeling, the string $c$ contained in register $\Omega$ describes the subset $\omega$ describing the indices for which the measurement was conclusive. Indeed $\omega(c) = \{i \in [M] : c_i \neq \varnothing\} = \{i \in [M] : c_i = \checkmark\}$. It will be convenient in the following to write $M_B^\omega$ instead of $M_B^c$, and put the value $\omega = \omega(c)$ in register $\Omega$. Let us now define the second measurement, $\mathcal{M}_{B\to U|S^\Phi B\Omega}$, characterized by operators $\{M_{B_i}^{u,\phi,c}\}_{u\in\{0,1,\varnothing\}}$ for each $i \in [M]$. In the case where the register $\Omega_i$ is set to $\varnothing$ (or equivalently that the set $\omega$ does not contain index $i$), we choose

$$M_{B_i}^{\varnothing,\phi,\varnothing} = \mathrm{id}_{B_i}, \quad M_{B_i}^{0,\phi,\varnothing} = 0, \quad M_{B_i}^{1,\phi,\varnothing} = 0 \tag{133}$$

which means that the measurement is essentially trivial and simply reveals the content of register $\Omega_i$. To address the case where the register is set to $\checkmark$, indicating that a conclusive outcome should be obtained, we choose operators $M_{B_i}^{0,\phi,\checkmark}$ and $M_{B_i}^{1,\phi,\checkmark}$ given by

$$M_{B_i}^{0,\phi,\checkmark} = M_{B_i}^{0,\phi}(M_{B_i}^\checkmark)^{-1/2}, \quad M_{B_i}^{1,\phi,\checkmark} = M_{B_i}^{1,\phi}(M_{B_i}^\checkmark)^{-1/2}. \tag{134}$$

It is immediate to check that these define valid generalized measurements and that Eq. (129) is satisfied. $\qquad\square$

**Lemma 17.** *If the assumption of Eq. (108) on Bob's measurement holds, then for any state $\rho_{A'BE}$, define*

$$\rho_{A''B'C^\Sigma C^\Omega S^\Phi EF^{\mathrm{si}}} = \mathcal{E}_{\mathrm{di}} \circ \mathcal{E}_{\mathrm{si}} \circ \mathcal{M}_{B\to B\Omega}(\rho_{A'BE} \otimes \rho_{S^{\Phi_A}} \otimes \rho_{S^{\Phi_B}}). \tag{135}$$

*Then, the state conditioned on the sifting procedure passing satisfies:*

$$\rho_{A''B'S^\Phi C^\Sigma C^\Omega E|F^{\mathrm{si}}=\checkmark} = \rho_{A'BC^\Sigma C^\Omega E|F^{\mathrm{si}}=\checkmark} \otimes \rho_{S^\Phi}, \tag{136}$$

*where $\rho_{S^\Phi} = \frac{1}{2^m} \sum_{\phi\in\{0,1\}^m} |\phi\rangle\langle\phi|_\Phi$.*

*Proof.* Since the measurement map $\mathcal{M}_{B\to B\Omega}$ only acts on register $B$, independently of the value $\Phi_B$, we have

$$\mathcal{M}_{B\to B\Omega}(\rho_{A'BE} \otimes \rho_{S^{\Phi_A}} \otimes \rho_{S^{\Phi_B}}) = \rho_{A'B\Omega E} \otimes \rho_{S^{\Phi_A}} \otimes \rho_{S^{\Phi_B}}, \tag{137}$$

where the state $\rho_{A'BE\Omega} = \mathcal{M}_{B\to B\Omega}(\rho_{A'BE})$ is a classical-quantum state:

$$\rho_{A'BE\Omega} = \sum_{\omega\in 2^{[M]}} |\omega\rangle\langle\omega|_\Omega \otimes (\mathrm{id}_{A'} \otimes M_B^\omega \otimes \mathrm{id}_E)\rho_{A'BE}(\mathrm{id}_{A'} \otimes (M_B^\omega)^\dagger \otimes \mathrm{id}_E). \tag{138}$$

It is straightforward to check that the classical map 'sift' has the following property: for all strings $\phi_A, \phi_B, \theta \in \{0,1\}^M$ and any subset $\omega \subseteq [M]$, if the sifting succeeds, then

$$\mathrm{sift}(\phi_A + \theta, \phi_B + \theta, \omega) = \mathrm{sift}(\phi_A, \phi_B, \omega). \tag{139}$$

Indeed, this is true since the map 'sift' only examines whether Alice and Bob's measurement bases coincide or not, and not their actual value. In particular, if $\Phi_A$ and $\Phi_B$ are uniformly distributed, then $\Phi$, the restriction of $\Phi_A$ to the subset $\Sigma$ returned by the sifting map when it succeeds, will also be uniformly distributed over the set of strings of length $m$.

Finally, the discarding map $\mathcal{E}_{\mathrm{di}}$ examines register $S^{\Phi_A}$ and puts its content, restricted to the subset determined by the sifting map, into register $S^\Phi$, and traces over all the systems that do not belong to that subset. The above property of the sifting map ensures that the value of $\Phi$ does not depend on $\Omega$.

This establishes that whenever the sifting test passes, the output state takes a tensor product form: $\rho_{A''B'S^\Phi C^\Sigma C^\Omega E|F^{\mathrm{si}}=\checkmark} = \rho_{A''B'C^\Sigma C^\Omega E|F^{\mathrm{si}}=\checkmark} \otimes \rho_{S^\Phi}$. $\qquad\square$

This lemma shows in particular that it is legitimate to consider $S^\Phi$ as a uniform seed, and not as a transcript, hence its notation $S^\Phi$ instead of $C^\Phi$.

## 9.3 Security: reduction to the entanglement-based protocol

We are now ready to prove Theorem 13.

*Proof.* It is sufficient to consider the case where the sifting procedure succeeds, since otherwise the protocol aborts and its output is trivially secret. For this reason, let us define a slight variant $\texttt{qkd\_modified}_{m,\text{pe,ec,pa}}$ of the entanglement-based protocol of Part I which differs by taking an additional input register $F^{\text{si}} \in \{\checkmark, \varnothing\}$. The variant starts by examining the content of this registers, and either aborts if the flag is set to $\varnothing$, or proceeds with the protocol $\texttt{qkd\_eb}_{m,\text{pe,ec,pa}}$ if the flag is set to $\checkmark$. A second difference between $\texttt{qkd\_eb}_{m,\text{pe,ec,pa}}$ and $\texttt{qkd\_modified}_{m,\text{pe,ec,pa}}$ is that the randomness for the measurement basis choice is explicitly given as an input. In particular, for any state $\rho_{A'BE}$, it holds that:

$$\texttt{qkd\_modified}_{m,\text{pe,ec,pa}}(\rho_{A'BE} \otimes \rho_{S^{\Phi}} \otimes |\checkmark\rangle\langle\checkmark|_{F^{\text{si}}}) = \texttt{qkd\_eb}_{m,\text{pe,ec,pa}}(\rho_{A'BE}). \tag{140}$$

From it definition, it is immediate that if $\texttt{qkd\_eb}_{m,\text{pe,ec,pa}}$ is $\varepsilon$-secure, then so is $\texttt{qkd\_modified}_{m,\text{pe,ec,pa}}$. Indeed, the only quantitative difference between the two protocol is that the latter one is less robust since it will not output nontrivial keys as soon as the sifting flag is set to $\varnothing$.

Our goal is therefore to show that in that case, for any input channel $\mathcal{N}_{A \to BE}$, there exists a state $\rho_{A''B'E'F^{\text{si}}}$ where $A''$ and $B'$ consist of $m$ systems such that

$$\texttt{qkd\_pm}_{M,m,\text{pe,ec,pa}}(\mathcal{N}_{A \to BE}) = \texttt{qkd\_modified}_{m,\text{pe,ec,pa}}(\rho_{A''B'E'F^{\text{si}}} \otimes \rho_{S^{\Phi}}). \tag{141}$$

Let us therefore consider the application of the prepare-and-measure protocol $\texttt{qkd\_pm}_{M,m,\text{pe,ec,pa}}$ to an arbitrary quantum channel $\mathcal{N}_{A \to BE}$. According to the description of the protocol, the classical-quantum state shared by Alice, Bob and Eve after the distribution step is given by some $\rho_{RBES^{\Phi_A}S^{\Phi_B}}$. The assumption made on Alice's preparation shows, as stated in Corollary 15, show that there exist a state $\tau_{AA'}$ and a measurement map $\mathcal{M}_{A' \to R|S^{\Phi_A}}$ such that

$$\rho_{RBES^{\Phi_A}S^{\Phi_B}} = \mathcal{N}_{A \to BE}(\rho_{RAS^{\Phi_A}S^{\Phi_B}}) \tag{142}$$

$$= \left(\mathcal{N}_{A \to BE} \circ \mathcal{M}_{A' \to R|S^{\Phi_A}}\right)(\tau_{AA'} \otimes \rho_{S^{\Phi_A}} \otimes \rho_{S^{\Phi_B}}) \tag{143}$$

$$= \mathcal{M}_{A' \to R|S^{\Phi_A}}(\rho_{A'BE} \otimes \rho_{\Phi_A} \otimes \rho_{\Phi_B}), \tag{144}$$

where we defined $\rho_{A'BE} = \mathcal{N}_{A \to BE}(\tau_{AA'})$. The last equality follows from the fact that the maps $\mathcal{N}_{A \to BE}$ and $\mathcal{M}_{A' \to R|S^{\Phi_A}}$ trivially commute since they act on distinct systems. After applying the measurement map $\mathcal{M}_{B \to B\Omega}$ promised by Lemma 16, followed by the sifting and discard maps, we obtain

$$\rho_{R'B'ES^{\Phi}S^{\Phi_B}C^{\Sigma}C^{\Omega}F^{\text{si}}} = \mathcal{E}_{\text{di}} \circ \mathcal{E}_{\text{si}} \circ \mathcal{M}_{B \to B\Omega}(\rho_{RBES^{\Phi_A}S^{\Phi_B}}) \tag{145}$$

$$= \mathcal{E}_{\text{di}} \circ \mathcal{E}_{\text{si}} \circ \mathcal{M}_{B \to B\Omega} \circ \mathcal{M}_{A' \to R|S^{\Phi_A}}(\rho_{A'BES^{\Phi_A}S^{\Phi_B}}), \tag{146}$$

where $R'$ and $B'$ are now restricted to the $m$ indices corresponding to the set $\Sigma$ provided by the sifting map. Indeed, recall that the discard map $\mathcal{E}_{\text{di}}$ replaces the $M$-system registers $R$ and $B$ by $m$-system registers $R'$ and $B'$ obtained by tracing over the systems not corresponding to $\Sigma$.

Since the measurement map $\mathcal{M}_{A' \to R|S^{\Phi_A}}$ of Alice's system $A'$ commutes with $\mathcal{E}_{\text{si}} \circ \mathcal{M}_{B \to B\Omega}$, we obtain:

$$\rho_{R'B'ES^{\Phi}S^{\Phi_B}C^{\Sigma}C^{\Omega}F^{\text{si}}} = \mathcal{E}_{\text{di}} \circ \mathcal{M}_{A' \to R|S^{\Phi_A}} \circ \mathcal{E}_{\text{si}} \circ \mathcal{M}_{B \to B\Omega}(\rho_{A'BES^{\Phi_A}S^{\Phi_B}}), \tag{147}$$

In the mathematical description of the protocol given in Section 7.3, the discard map was applied to registers $R, B$, but since it commutes with measurement maps on either Alice's or Bob's system, the map can be just as well applied to registers $A'$ and $U$, with outputs denoted $A''$ and $U'$, respectively. We deduce that we can replace the map $\mathcal{E}_{\text{di}} \circ \mathcal{M}_{A' \to R|S^{\Phi_A}}$ by $\mathcal{M}_{A'' \to R'|S^{\Phi}} \circ \mathcal{E}_{\text{di}}$ in (147). This yields:

$$\rho_{R'B'ES^{\Phi}S^{\Phi_B}C^{\Sigma}C^{\Omega}F^{\text{si}}} = \mathcal{M}_{A'' \to R'|S^{\Phi}} \circ \mathcal{E}_{\text{di}} \circ \mathcal{E}_{\text{si}} \circ \mathcal{M}_{B \to B\Omega}(\rho_{A'BES^{\Phi_A}S^{\Phi_B}}) \tag{148}$$

$$= \mathcal{M}_{A'' \to R'|S^{\Phi}}(\rho_{A''B'C^{\Sigma}C^{\Omega}S^{\Phi}S^{\Phi_B}EF^{\text{si}}}). \tag{149}$$

Lemma 17 now shows that

$$\rho_{R'B'ES^{\Phi}S^{\Phi_B}C^{\Sigma}C^{\Omega}|F^{\text{si}}=\checkmark} = \mathcal{M}_{A'' \to R'|S^{\Phi}}\left(\rho_{A''B'EC^{\Sigma}C^{\Omega}|F^{\text{si}}=\checkmark}\right) \otimes \rho_{S^{\Phi}}. \tag{150}$$

Let us finally collect $E' = ES^{\Phi_B}C^{\Sigma}C^{\Omega}$. If the sifting test passes, then

$$\texttt{qkd\_modified}_{m,\text{pe,ec,pa}}\left(\rho_{A''B'S^{\Phi}E'|F^{\text{si}}=\checkmark} \otimes \rho_{S^{\Phi}} \otimes |\checkmark\rangle\langle\checkmark|_{F^{\text{si}}}\right) = \texttt{qkd\_eb}_{m,\text{pe,ec,pa}}(\rho_{A''B'E'}), \tag{151}$$

which concludes the proof. $\qquad\square$

# 10 Conclusion

We provide a self-contained security proof of QKD detailing all the steps of the protocol and explicitly spelling out all the required assumptions for the security proof to go through. For simplicity, we focussed on a variant of the entanglement-based BBM92 protocol as well as the BB84 protocol and showed that practical secret key rates can be achieved, even for moderately large block size. These results, however, come at the price of several assumptions which are sometimes challenging to enforce in practice. This should not come as a surprise since many simplified implementations are known to be vulnerable to quantum hacking, illustrating that there exist necessary trade-offs between ease of implementation and security guarantees.

We believe that there is room for improvement for these trade-offs and that further collaboration between theory and experiment will be essential for achieving this objective. In this context, it is crucial to model the protocols as thoroughly as possible in order to understand what level of security can be obtained, and under which assumptions. Finally, we would like to encourage more research into derving tight finite resource trade-offs for quantum key distribution protocols beyond BB84 and BBM92. This will likely require techniques beyond the entropic uncertainty relation presented here. An example of interest is the 6-state protocol [45] where entropic uncertainty relations do not currently yield optimal secret key rates (even asymptotically) and approaches based on a more complete tomography of the channel lead to large penalities for finite keys.

*Note added.* After completion of this work a novel and intriguing proof technique (based on Rényi entropy accumulation) has been proposed by Dupuis et al. [46]. This technique does not yet seem to yield tradeoffs between security and protocol parameters that match those found in [13] and [14], on which we improve upon here. However, the technique is more versatile and in particular allows to show security of device-independent protocols as demonstrated by Arnon-Friedman et al. [47]. Device-independent protocols have the advantage that fewer assumptions on the physical devices used in the protocol are required, but are beyond the scope of this work.

# References

[1] C.H. Bennett and G. Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. In *Proc. IEEE International Conference on Computers, Systems and Signal Processing 1984*, volume 1, pages 175–179, Bangalore, 1984.

[2] A.K. Ekert. Quantum Cryptography Based on Bell's Theorem. *Physical Review Letters*, 67(6):661–663, 1991. DOI: 10.1103/PhysRevLett.67.661.

[3] C. Bennett, G. Brassard, and N. Mermin. Quantum Cryptography Without Bell's Theorem. *Physical Review Letters*, 68(5):557–559, 1992. DOI: 10.1103/PhysRevLett.68.557.

[4] H.-K. Lo and H.F. Chau. Unconditional Security of Quantum Key Distribution over Arbitrarily Long Distances. *Science*, 283(5410):2050–2056, 1999. DOI: 10.1126/science.283.5410.2050.

[5] P.W. Shor and J. Preskill. Simple Proof of Security of the BB84 Quantum Key Distribution Protocol. *Physical Review Letters*, 85(2):441–444, 2000. DOI: 10.1103/PhysRevLett.85.441.

[6] D. Mayers. Unconditional Security in Quantum Cryptography. *Journal of the ACM*, 48(3):351–406, 2001. DOI: 10.1145/382780.382781.

[7] M. Koashi. Unconditional Security of Quantum Key Distribution and the Uncertainty Principle. *Journal of Physics: Conference Series*, 36(1):98–102, 2006. DOI: 10.1088/1742-6596/36/1/016.

[8] H. Maassen and J. Uffink. Generalized Entropic Uncertainty Relations. *Physical Review Letters*, 60(12): 1103–1106, 1988. DOI: 10.1103/PhysRevLett.60.1103.

[9] W. Heisenberg. Über den Anschaulichen Inhalt der Quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, mar 1927.

[10] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005. URL http://arxiv.org/abs/quant-ph/0512258.

[11] L.C. Comandar, M. Lucamarini, B. Fröhlich, J.F. Dynes, A.W. Sharpe, S.W.-B. Tam, Z.L. Yuan, R.V. Penty, and A.J. Shields. Quantum key distribution without detector vulnerabilities using optically seeded lasers. *Nature Photonics*, 10(5):312–315, 2016. DOI: 10.1038/nphoton.2016.50.

[12] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti. Experimental demonstration of long-distance continuous-variable quantum key distribution. *Nature Photonics*, 7(5):378–381, 2013. DOI: 10.1038/nphoton.2013.63.

[13] M. Tomamichel, C.C.W. Lim, N. Gisin, and R. Renner. Tight Finite-Key Analysis for Quantum Cryptography. *Nature Communications*, 3:634, 2012. DOI: 10.1038/ncomms1631.

[14] M. Hayashi and T. Tsurumaru. Concise and Tight Security Analysis of the Bennett-Brassard 1984 Protocol with Finite Key Lengths. *New Journal of Physics*, 14(9):093014, 2012. DOI: 10.1088/1367-2630/14/9/093014.

[15] V. Scarani and R. Renner. Quantum Cryptography with Finite Resources: Unconditional Security Bound for Discrete-Variable Protocols with One-Way Postprocessing. *Physical Review Letters*, 100(20), 2008. DOI: 10.1103/PhysRevLett.100.200501.

[16] R. Renner. Symmetry of Large Physical Systems Implies Independence of Subsystems. *Nature Physics*, 3 (9):645–649, 2007. DOI: 10.1038/nphys684.

[17] M. Christandl, R. König, and R. Renner. Postselection Technique for Quantum Channels with Applications to Quantum Cryptography. *Physical Review Letters*, 102(2), 2009. DOI: 10.1103/PhysRevLett.102.020504.

[18] L. Sheridan, T.P. Le, and V. Scarani. Finite-Key Security Against Coherent Attacks in Quantum Key Distribution. *New Journal of Physics*, 12:123019, 2010.

[19] C. Pfister, N. Lütkenhaus, S. Wehner, and P.J. Coles. Sifting Attacks in Finite-Size Quantum Key Distribution. *New Journal of Physics*, 18(5):053001, 2016. DOI: 10.1088/1367-2630/18/5/053001.

[20] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner. A Monogamy-of-Entanglement Game with Applications to Device-Independent Quantum Cryptography. *New Journal of Physics*, 15(10):103002, 2013. DOI: 10.1088/1367-2630/15/10/103002.

[21] M. Tomamichel and R. Renner. Uncertainty Relation for Smooth Entropies. *Physical Review Letters*, 106 (11):110506, 2011. DOI: 10.1103/PhysRevLett.106.110506.

[22] M. Tomamichel. *Quantum Information Processing with Finite Resources — Mathematical Foundations*, volume 5 of *SpringerBriefs in Mathematical Physics*. Springer International Publishing, 2016. ISBN 978-3-319-21890-8. DOI: 10.1007/978-3-319-21891-5.

[23] C.W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, NY, 1976.

[24] M. Tomamichel, R. Colbeck, and R. Renner. Duality Between Smooth Min- and Max-Entropies. *IEEE Transactions on Information Theory*, 56(9):4674–4681, 2010. DOI: 10.1109/TIT.2010.2054130.

[25] J.L. Carter and M.N. Wegman. Universal Classes of Hash Functions. *Journal of Computer and System Sciences*, 18(2):143–154, 1979. DOI: 10.1016/0022-0000(79)90044-8.

[26] M.N. Wegman and J.L. Carter. New Hash Functions and their Use in Authentication and Set Equality. *Journal of Computer and System Sciences*, 22(3):265–279, 1981. DOI: 10.1016/0022-0000(81)90033-7.

[27] A. Rényi. On Measures of Information and Entropy. In *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–561, Berkeley, California, USA, 1961. University of California Press.

[28] R. König, R. Renner, and C. Schaffner. The Operational Meaning of Min- and Max-Entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, 2009. DOI: 10.1109/TIT.2009.2025545.

[29] S. Winkler, M. Tomamichel, S. Hengl, and R. Renner. Impossibility of Growing Quantum Bit Commitments. *Physical Review Letters*, 107(9):090502, 2011. ISSN 0031-9007. DOI: 10.1103/PhysRevLett.107.090502.

[30] H.-K. Lo, H.F. Chau, and M. Ardehali. Efficient Quantum Key Distribution Scheme and a Proof of Its Unconditional Security. *Journal of Cryptology*, 18(2):133–165, 2004 DOI: 10.1007/s00145-004-0142-y

[31] D. Frauchiger, R. Renner, and M. Troyer. True randomness from realistic quantum devices, 2013. URL http://arxiv.org/abs/1311.4547.

[32] C. Portmann and R. Renner. Cryptographic Security of Quantum Key Distribution, 2014. URL http://arxiv.org/abs/1409.3525.

[33] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov. Hacking Commercial Quantum Cryptography Systems by Tailored Bright Illumination. *Nature Photonics*, 4(10):686–689, 2010. DOI: 10.1038/nphoton.2010.214.

[34] M. Tomamichel and Esther Hänggi. The Link Between Entropic Uncertainty and Nonlocality. *Journal of Physics A: Mathematical and Theoretical*, 46(5):055301, 2013. DOI: 10.1088/1751-8113/46/5/055301.

[35] C.C.W. Lim, C. Portmann, M. Tomamichel, R. Renner, and Nicolas Gisin. Device-Independent Quantum Key Distribution with Local Bell Test. *Physical Review X*, 3(3):031006, 2013. DOI: 10.1103/PhysRevX.3.031006.

[36] I. Devetak and A. Winter. Distillation of Secret Key and Entanglement from Quantum States. *Proceedings of the Royal Society A*, 461(2053):207–235, 2005. DOI: 10.1098/rspa.2004.1372.

[37] D. Elkouss, A. Leverrier, R. Alleaume, and J.J. Boutros. Efficient Reconciliation Protocol for Discrete-Variable Quantum Key Distribution. In *Proc. IEEE ISIT 2009*, pages 1879–1883, 2009. DOI: 10.1109/ISIT.2009.5205475.

[38] M. Tomamichel, J. Martinez-Mateo, C. Pacher, and D. Elkouss. Fundamental Finite Key Limits for Information Reconciliation in Quantum Key Distribution, 2014. URL http://arxiv.org/abs/1401.5194.

[39] M. Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory.* PhD thesis, ETH Zurich, 2012. URL http://arxiv.org/abs/1203.2142.

[40] R.J. Serfling. Probability Inequalities for the Sum in Sampling without Replacement. *Annals of Statistics*, 2(1):39–48, 1974.

[41] J.H. van Lint. *Introduction to Coding Theory.* Graduate Texts in Mathematics. Springer, third edition, 1999.

[42] H.-K. Lo, X. Ma, and K. Chen. Decoy State Quantum Key Distribution. *Physical Review Letters*, 94(23), 2005. DOI: 10.1103/PhysRevLett.94.230504.

[43] J. Hasegawa, M. Hayashi, T. Hiroshima, and A. Tomita. Security analysis of decoy state quantum key distribution incorporating finite statistics, 2007. URL http://arxiv.org/abs/0707.3541.

[44] C.C.W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden. Concise security bounds for practical decoy-state quantum key distribution. *Physical Review A*, 89(2):022307, 2014. DOI: 10.1103/PhysRevA.89.022307.

[45] D. Bruss. Optimal Eavesdropping in Quantum Cryptography with Six States. *Physical Review Letters*, 81 (14):3018–3021, 1998. DOI: 10.1103/PhysRevLett.81.3018.

[46] F. Dupuis, O. Fawzi, and R. Renner. Entropy accumulation, 2016. URL http://arxiv.org/abs/1607.01796.

[47] R. Arnon-Friedman, R. Renner, and T. Vidick. Simple and tight device-independent security proofs, 2016. URL http://arxiv.org/abs/1607.01797.

[48] R. Bhatia. *Matrix Analysis.* Graduate Texts in Mathematics. Springer, 1997. ISBN 0-387-94846-5.

# A  Proof of entropic uncertainty relation in Proposition 4

Let us restate the desired inequality for the convenience of the reader.

**Proposition 4 (restated).** Let $\tau_{APRS} \in \mathcal{S}_\bullet(APRS)$ be an arbitrary sub-normalized state with $P$ a classical register, and set $t := \mathrm{tr}\{\tau_{APRS}\}$. Furthermore, let $\varepsilon \in [0, \sqrt{t})$ and let $q$ be a bijective function on $P$ that is a symmetry of $\tau_{APRS}$ in the sense that $\tau_{ARS,P=p} = \tau_{ARS,P=q(p)}$ for all $p \in P$. Then, we have

$$H_{\min}^\varepsilon(X|PR)_\sigma + H_{\max}^\varepsilon(X|PS)_\sigma \geq \log \frac{1}{c_q}, \qquad \text{where} \tag{152}$$

where $c_q = \max_{p \in P} \max_{x,z \in X} \left\| F_A^{q(p),x} \left(F_A^{p,z}\right)^\dagger \right\|_\infty^2$. Here, $\sigma_{XPRS} = \mathcal{M}_{A \to X|P}(\tau_{APRS})$ for the map

$$\mathcal{M}_{A \to X|P}[\,\cdot\,] = \mathrm{tr}_A \left\{ \sum_{p \in P} \sum_{x \in X} |x\rangle_X \left( |p\rangle\langle p|_P \otimes F_A^{p,x} \right) \cdot \left( |p\rangle\langle p|_P \otimes F_A^{p,x} \right)^\dagger \langle x|_X \right\}. \tag{153}$$

and any set (indexed by $p \in P$) of generalized measurements $\{F_A^{p,x}\}_{x \in X}$.

*Proof.* The condition on $\varepsilon$ ensures that all smooth entropies are well-defined. We first introduce the Stinespring dilation isometry of the measurement map $\mathcal{M}_{A \to X|P}$. This is the isometry $V : A \to AXX'|P$ given by

$$V := \sum_{p \in P} \sum_{x \in X} |x\rangle_X \otimes |x\rangle_{X'} \otimes |p\rangle\langle p|_P \otimes F_A^{p,x}. \tag{154}$$

Now note that the measured state $\sigma_{XPRS}$ in (153) has a natural purification in

$$\sigma_{PP'AXX'RSD} = V\left(\tau_{PP'ARSD}\right)V^\dagger, \quad \text{where} \tag{155}$$

$$|\tau_{PP'ARSD}\rangle = \sum_{p \in P} \sqrt{\Pr[P=p]_\tau} \, |p\rangle_P \otimes |p\rangle_{P'} \otimes |\tau_{ARSD|P=p}\rangle, \tag{156}$$

where $P'$ is isomorphic to $P$ and $|\tau_{ARSD|P=p}\rangle$ is any purification of $\tau_{ARS|P=p}$ on a sufficiently large auxiliary system $D$. (The choice $|D| = |A||R||S|$ ensures that all purifications can be accommodated.) This now allows us to rephrase our target inequality. Using the duality relation for smooth min- and max-entropy in (17) together with the fact that $H_{\min}^\varepsilon(X|PR)_\sigma = H_{\min}^\varepsilon(X|P'R)_\sigma$ since $P'$ is a copy of $P$, we find that (152) is equivalent to

$$H_{\min}^\varepsilon(X|PR)_\sigma \geq H_{\min}^\varepsilon(X|PAX'RD)_\sigma + \log \frac{1}{c_q} \,. \tag{157}$$

Moreover, the data-processing inequality for the smooth min-entropy in (18) applied for the map $\mathrm{tr}_D$ yields $H^\varepsilon_{\min}(X|PAX'RD)_\sigma \leq H^\varepsilon_{\min}(X|PAX'R)_\sigma$ and thus it in fact suffices to show that[13]

$$H^\varepsilon_{\min}(X|PR)_\sigma \geq H^\varepsilon_{\min}(X|PAX'R)_\sigma + \log \frac{1}{c_q}. \tag{158}$$

The remainder of the proof will be concerned with showing the inequality in (158).

For this purpose, let us consider the following unitary rotation:

$$Q_P = \sum_{p \in P} \left| q^{-1}(p) \middle\rangle\!\middle\langle p \right|_P. \tag{159}$$

that exchanges $p$ with its conjugate, $q(p)$. It clearly acts as a permutation when acting on the classical register $P$ and furthermore we have $Q_P(\tau_{APRS})Q_P^\dagger = \tau_{APRS}$ due to the symmetry condition that we imposed on $q$ and $\tau_{PARS}$ in the statement of the proposition. Based on this we define the isometry

$$\bar{V} := Q_P V Q_P^\dagger = \sum_{p \in P} \sum_{x \in X} |x\rangle_X \otimes |x\rangle_{X'} \otimes |p\rangle\!\langle p|_P \otimes F_A^{q(p),x}, \tag{160}$$

that corresponds to a measurement in the basis determined by $q(p)$ instead of $p$. We find

$$\bar{V}V^\dagger \sigma_{AXX'PRS} V\bar{V}^\dagger = Q_P V Q_P^\dagger(\tau_{APRS})Q_P V^\dagger Q_P^\dagger = Q_P \sigma_{AXX'PRS} Q_P^\dagger, \tag{161}$$

which shows that the trace non-increasing map $\bar{V}V^\dagger(\cdot)V\bar{V}^\dagger$ coherently undoes the measurement in the basis determined by $p$ and then instead measures in the basis determined by $q(p)$.

Now we have the tools at hand to prove the inequality in (158). By the definition of the smooth min-entropy, $H^\varepsilon_{\min}(X|PAX'R)_\sigma$, there exists a state $\omega_{PAX'R} \in \mathcal{S}(PAX'R)$ and a sub-normalized state $\tilde{\sigma}_{PAXX'R} \in \mathcal{S}_\bullet(PAXX'R)$ that is $\varepsilon$-close to $\sigma_{PAXX'R}$ in the sense that $P(\sigma_{PAXX'R}, \tilde{\sigma}_{PAXX'R}) \leq \varepsilon$ such that the following inequality holds:

$$\tilde{\sigma}_{PAXX'R} \leq 2^{-\lambda}\,\mathrm{id}_X \otimes \omega_{PAX'R} \quad \text{with} \quad \lambda := H^\varepsilon_{\min}(X|PAX'R)_\sigma. \tag{162}$$

Next we consider the CP trace non-increasing map

$$\mathcal{F}_{XX'A \to X|P}[\,\cdot\,] = \sum_{p \in P} \mathrm{tr}_{AX'} \left( Q_P^\dagger \bar{V}V^\dagger |p\rangle\!\langle p|_P \cdot |p\rangle\!\langle p|_P V\bar{V}^\dagger Q_P \right). \tag{163}$$

From (161) we learn that $\mathcal{F}\big[\sigma_{PAXX'R}\big] = \sigma_{PXR}$. Thus, using the fact that the purified distance contracts (3) when we apply $\mathcal{F}$, we find that the state $\hat{\sigma}_{PXR} := \mathcal{F}\big[\tilde{\sigma}_{PAXX'R}\big]$ satisfies

$$P(\hat{\sigma}_{PXR}, \sigma_{PXR}) \leq P(\tilde{\sigma}_{PAXX'R}, \sigma_{PAXX'R}) \leq \varepsilon. \tag{164}$$

Furthermore, applying $\mathcal{F}$ on both sides of (162) yields

$$\hat{\sigma}_{PXR} \leq 2^{-\lambda}\mathcal{F}\big[\mathrm{id}_X \otimes \omega_{PAX'R}\big] = 2^{-\lambda}\,\mathrm{tr}_{X'A} \left( Q_P^\dagger \bar{V}V^\dagger(\mathrm{id}_X \otimes \hat{\omega}_{PAX'R})V\bar{V}^\dagger Q_P \right), \tag{165}$$

where $\hat{\omega}_{PAX'R} = \sum_{p \in P} |p\rangle\!\langle p|_P \otimes \hat{\omega}_{AX'R}^p$ with $\hat{\omega}_{AX'R}^p = \langle p|\, \omega_{PAX'R}\, |p\rangle_P$.

Let us now simplify the right-hand side of this inequality, hoping to capture the incompatibility of the measurements in the basis $p$ versus the basis $q(p)$. First, we note that

$$Q_P^\dagger \bar{V}V^\dagger = \sum_{p \in P} \sum_{x,z \in X} |z\rangle\!\langle x|_X \otimes |z\rangle\!\langle x|_{X'} \otimes \left| q(p) \middle\rangle\!\middle\langle p \right|_P \otimes F_A^{q(p),z}\left(F_A^{p,x}\right)^\dagger. \tag{166}$$

---

[13]In its cryptographic application, another intuitive way to justify that we can remove the purifying system $D$ is that, without loss of generality, we may assume that the eavesdropper already holds a purification of the state shared by the honest parties and $D$ is thus trivial. Luckily this physical intuition is corroborated by a mathematical argument — the data-processing inequality.

and, hence, we can write

$$\text{tr}_{X'A}\left(Q_P^\dagger \bar{V}V^\dagger(\text{id}_X \otimes \hat{\omega}_{PAX'R})V\bar{V}^\dagger Q_P\right)$$

$$= \sum_{p\in P}\sum_{x,z\in X} |x\rangle\langle x|_X \otimes |q(p)\rangle\langle q(p)|_P \otimes \langle z|\,\text{tr}_A\left(F_A^{q(p),z}\big(F_A^{p,x}\big)^\dagger \hat{\omega}_{AX'R}^p F_A^{p,x}\big(F_A^{q(p),z}\big)^\dagger\right)|z\rangle_{X'} \tag{167}$$

$$\leq \sum_{p\in P}\sum_{x,z\in X} |x\rangle\langle x|_X \otimes |q(p)\rangle\langle q(p)|_P \otimes \left\|F_A^{p,x}\big(F_A^{q(p),z}\big)^\dagger F_A^{q(p),z}\big(F_A^{p,x}\big)^\dagger\right\|_\infty \langle z|\,\text{tr}_A\left(\hat{\omega}_{AX'R}^p\right)|z\rangle_{X'} \tag{168}$$

$$= \sum_{p\in P}\sum_{x,z\in X} |x\rangle\langle x|_X \otimes |q(p)\rangle\langle q(p)|_P \otimes \left\|F_A^{q(p),x}\big(F_A^{p,z}\big)^\dagger\right\|_\infty^2 \langle z|\hat{\omega}_{X'R}^p|z\rangle_{X'} \tag{169}$$

$$\leq \max_{p\in P}\max_{x,z\in X}\left\|F_A^{q(p),x}\big(F_A^{p,z}\big)^\dagger\right\|_\infty^2 \cdot \sum_{p\in P}\sum_{x,z\in X}|x\rangle\langle x|_X \otimes |p\rangle\langle p|_P \otimes \langle z|\hat{\omega}_{X'R}^p|z\rangle_{X'} \tag{170}$$

$$= c_q \cdot \sum_{p\in P}\text{id}_X \otimes |p\rangle\langle p|_P \otimes \hat{\omega}_R^p\,. \tag{171}$$

To establish (168) and (169) we used the fact that $L^\dagger L \leq \|L^\dagger L\|_\infty\,\text{id} = \|L\|_\infty^2\,\text{id}$ for every linear operator $L$ by definition of the operator norm. The final equality (171) follows from the definition of $c_q$.

Combining this bound with (165) yields

$$\hat{\sigma}_{PXR} \leq 2^{-\lambda}c_q \cdot \text{id}_X \otimes \sum_{p\in P}|p\rangle\langle p|_P \otimes \hat{\omega}_R^p\,. \tag{172}$$

Since $\sum_{p\in P}\text{tr}(\hat{\omega}_R^p) = 1$ by construction and $P(\hat{\sigma}_{PXR},\sigma_{PXR}) \leq \varepsilon$ due to (164), the definition of the smooth entropy implies that

$$H_{\min}^\varepsilon(X|PR)_\sigma \geq \lambda - \log c_q = H_{\min}^\varepsilon(X|PAX'R)_\sigma - \log c_q\,, \tag{173}$$

concluding the proof. □

## B  Proof of Leftover Hashing Lemma in Proposition 9

Our proof of the leftover hashing lemma is based on the following result due to Renner [10, Corollary 5.5.2]:

**Lemma 18.** *Let $\sigma_{XD} \in \mathcal{S}_\bullet(XD)$ be a classical on $X$ and let $\mathcal{H}$ be a universal$_2$ family of hash functions from $\mathcal{X} = \{0,1\}^n$ to $\mathcal{K} = \{0,1\}^\ell$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{H}|}\sum_{h\in\mathcal{H}}|h\rangle\langle h|_{S^H}$ be fully mixed. Then,*

$$\left\|\omega_{KS^HD} - \chi_K \otimes \omega_{S^HD}\right\|_{\text{tr}} \leq \frac{1}{2}\sqrt{\text{tr}\{\sigma_{XD}\}}\,2^{-\frac{1}{2}(H_{\min}(X|D)_\sigma - \ell)} \tag{174}$$

*where $\chi_K = \frac{1}{2^\ell}\text{id}_K$ is the fully mixed state and $\omega_{KS^HD} = \text{tr}_X\left(\mathcal{E}_f(\sigma_{XD}\otimes\rho_{S^H})\right)$ for the function $f : (x,h) \mapsto h(x)$ that acts on the registers $X$ and $S^H$.*

We provide a short proof for the convenience of the reader (see [22, Section 7.3.2]).

*Proof.* First, by definition of the min-entropy, there exists a state $\tau_D \in \mathcal{S}(D)$ such that $\sigma_{XD} \leq 2^{-H_{\min}(X|D)_\sigma}\text{id}_X \otimes \tau_D$. Next, note that by definition of the trace distance, we have

$$\left\|\omega_{KS^HD} - \chi_K \otimes \omega_{S^HD}\right\|_{\text{tr}} = \sum_{h\in\mathcal{H}}\frac{1}{2|\mathcal{H}|}\left\|\omega_{KD|S^H=h} - \chi_K \otimes \omega_{D|S^H=h}\right\|_1\,, \tag{175}$$

where $\|\cdot\|_1$ denotes the Schatten 1-norm. Moreover, by construction of $\omega_{KS^HD}$ it is evident that $\omega_{D|S^H=h} = \omega_D = \sigma_D$ for all $h \in \mathcal{H}$. Due to Hölder's inequality for Schatten norms [48, Corollary IV.2.6], we have

$$\left\|\omega_{KD|S^H=h} - \chi_K \otimes \sigma_D\right\|_1^2 \leq \left\|\text{id}_K \otimes \tau_D^{\frac{1}{2}}\right\|_2^2\left\|\text{id}_K \otimes \tau_D^{-\frac{1}{2}}\big(\omega_{KD|S^H=h} - \chi_K \otimes \sigma_D\big)\right\|_2^2 \tag{176}$$

$$= 2^\ell\,\text{tr}\{\tau_D\}\,\text{tr}\left\{\text{id}_K \otimes \tau_D^{-1}\big(\omega_{KD|S^H=h} - \chi_K \otimes \sigma_D\big)^2\right\} \tag{177}$$

$$= 2^\ell\,\text{tr}\left\{\big(\text{id}_K \otimes \tau_D^{-1}\big)\omega_{KD|S^H=h}^2\right\} - \text{tr}\left\{\tau_D^{-1}\sigma_D^2\right\}\,, \tag{178}$$

where $\tau_D$ is inverted on its support. We took advantage of the fact that $\chi_K = \frac{1}{2^\ell} \mathrm{id}_K$ is proportional to the identity to simplify the above expression. Combining this with (175), Jensen's inequality thus ensures that

$$\left\| \omega_{KS^H D} - \chi_K \otimes \omega_{S^H D} \right\|_{\mathrm{tr}} \leq \frac{1}{2} \sqrt{\sum_{h \in \mathcal{H}} \frac{2^\ell}{|\mathcal{H}|} \mathrm{tr} \left\{ \left( \mathrm{id}_K \otimes \tau_D^{-1} \right) \omega_{KD|S^H=h}^2 \right\} - \mathrm{tr} \left\{ \tau_D^{-1} \sigma_D^2 \right\}}. \tag{179}$$

Next, observe that $\omega_{KD|S^H=h} = \sum_{k \in \{0,1\}^\ell} \sum_{x \in \{0,1\}^n, h(x)=k} |k\rangle\langle k| \otimes \sigma_{D \wedge X=x}$ by construction. We then use the universal$_2$ property of $\mathcal{H}$ which implies that $\frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} 1\{h(x) = h(y)\} = 2^{-\ell}$ when $x \neq y$. This yields

$$\sum_{h \in \mathcal{H}} \frac{1}{|\mathcal{H}|} \mathrm{tr} \left\{ \mathrm{id}_K \otimes \tau_D^{-1} \omega_{KD|S^H=h}^2 \right\} = \sum_{x,y \in \{0,1\}^n} \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} 1\{h(x) = h(y)\} \mathrm{tr} \left\{ \tau_D^{-1} \sigma_{D \wedge X=x} \sigma_{D \wedge X=y} \right\} \tag{180}$$

$$= \sum_{x \in \{0,1\}^n} \mathrm{tr} \left\{ \tau_D^{-1} \sigma_{D \wedge X=x}^2 \right\} + \sum_{\substack{x,y \in \{0,1\}^n \\ x \neq y}} 2^{-\ell} \mathrm{tr} \left\{ \tau_D^{-1} \sigma_{D \wedge X=x} \sigma_{D \wedge X=y} \right\} \tag{181}$$

$$= \left( 1 - 2^{-\ell} \right) \mathrm{tr} \left\{ \left( \mathrm{id}_X \otimes \tau_D^{-1} \right) \sigma_{XD}^2 \right\} + 2^{-\ell} \mathrm{tr} \left\{ \tau_D^{-1} \sigma_D^2 \right\}. \tag{182}$$

Bounding $1 - 2^\ell \leq 1$ and plugging this into (179), we find

$$\left\| \omega_{KS^H D} - \chi_K \otimes \omega_{S^H D} \right\|_{\mathrm{tr}} \leq \frac{1}{2} \sqrt{2^\ell \mathrm{tr} \left\{ \left( \mathrm{id}_X \otimes \tau_D^{-1} \right) \sigma_{XD}^2 \right\}}. \tag{183}$$

Finally, due to the operator anti-monotonicity of the inverse and the definition of $\tau_D$, we have $\mathrm{id}_X \otimes \tau_D^{-1} \leq 2^{-H_{\min}(X|D)} \sigma_{XD}^{-1}$. Combined with (183) this yields the desired result. $\qquad \square$

Let us restate the desired inequality for the convenience of the reader.

**Proposition 9 (restated).** Let $\sigma_{XD} \in \mathcal{S}_\bullet(XD)$ be a classical on $X$ and let $\mathcal{H}$ be a universal$_2$ family of hash functions from $\mathcal{X} = \{0,1\}^n$ to $\mathcal{K} = \{0,1\}^\ell$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{H}|} \sum_{h \in \mathcal{H}} |h\rangle\langle h|_{S^H}$ be fully mixed. Then,

$$\left\| \omega_{KS^H D} - \chi_K \otimes \omega_{S^H D} \right\|_{\mathrm{tr}} \leq \frac{1}{2} 2^{-\frac{1}{2}(H_{\min}^\varepsilon(X|D)_\sigma - \ell)} + 2\varepsilon \tag{184}$$

where $\chi_K$ and $\omega_{KS^H D}$ are define as in Lemma 18.

*Proof.* Let $\tilde{\sigma}_{XD} \in \mathcal{S}_\bullet(XD)$ be a sub-normalized state such that $H_{\min}^\varepsilon(X|D)_\sigma = H_{\min}(X|D)_{\tilde\sigma}$ and $P(\tilde{\sigma}_{XD}, \sigma_{XD}) \leq \varepsilon$. Without loss of generality we can assume that $\tilde{\sigma}_{XD}$ is classical on $X$. Now, the Lemma 18 yields the inequality

$$\left\| \tilde\omega_{KS^H D} - \chi_K \otimes \tilde\omega_{S^H D} \right\|_{\mathrm{tr}} \leq \frac{1}{2} \sqrt{\mathrm{tr}\{\tilde\sigma_{XD}\}} \cdot 2^{-\frac{1}{2}(H_{\min}(X|D)_{\tilde\sigma} - \ell)} \leq \frac{1}{2} \cdot 2^{-\frac{1}{2}(H_{\min}^\varepsilon(X|D)_\sigma - \ell)}, \tag{185}$$

where we constructed $\tilde\omega_{KS^H D} = \mathrm{tr}_X \left( \mathcal{E}_f(\tilde\sigma_{XD} \otimes \rho_{S^H}) \right)$ and bounded $\mathrm{tr}\{\tilde\sigma_{XD}\} \leq 1$ to arrive at the second inequality. Using the monotonicity of the purified distance under CPTP maps we conclude that $P(\tilde\omega_{S^H D}, \omega_{S^H D}) \leq P(\tilde\omega_{KS^H D}, \omega_{KS^H D}) \leq \varepsilon$. Finally, exploiting the triangle inequality for the trace norm we find

$$\left\| \omega_{KS^H D} - \chi_K \otimes \omega_{S^H D} \right\|_{\mathrm{tr}} \leq 2\varepsilon + \left\| \tilde\omega_{KS^H D} - \chi_K \otimes \tilde\omega_{S^H D} \right\|_{\mathrm{tr}}. \tag{186}$$

$\qquad \square$