Information-theoretic security proof for quantum-key-distribution protocols

Renato Renner,[1] Nicolas Gisin,[2] and Barbara Kraus[2]
[1]*Computer Science Department, ETH Zürich, CH-8092 Zürich, Switzerland*
[2]*Group of Applied Physics, University of Geneva, CH-1211 Genève 4, Switzerland*
(Received 1 March 2005; published 25 July 2005)

We present a technique for proving the security of quantum-key-distribution (QKD) protocols. It is based on direct information-theoretic arguments and thus also applies if no equivalent entanglement purification scheme can be found. Using this technique, we investigate a general class of QKD protocols with one-way classical post-processing. We show that, in order to analyze the full security of these protocols, it suffices to consider collective attacks. Indeed, we give new lower and upper bounds on the secret-key rate which only involve entropies of two-qubit density operators and which are thus easy to compute. As an illustration of our results, we analyze the Bennett-Brassard 1984, the six-state, and the Bennett 1992 protocols with one-way error correction and privacy amplification. Surprisingly, the performance of these protocols is increased if one of the parties adds noise to the measurement data before the error correction. In particular, this additional noise makes the protocols more robust against noise in the quantum channel.

DOI: 10.1103/PhysRevA.72.012332 PACS number(s): 03.67.Dd, 89.70.+c

I. INTRODUCTION

Classical key distribution schemes can only be secure under strong assumptions—e.g., that the computing power or the storage capacity of a potential adversary is limited. In contrast, quantum key distribution (QKD) allows for provable security under the sole assumption that the laws of physics are correct. This ultimate security is certainly one of the main reasons why so much theoretical and experimental effort is undertaken to investigate QKD protocols and, in particular, to make them practical [1–3].

One of the most challenging theoretical problems in the context of QKD is to determine sufficient and/or necessary conditions for the security of QKD protocols. This is exactly what we are concerned with in this paper. To be more precise, we investigate the security of a general class of QKD schemes which includes the most popular ones such as the Bennett-Brassard 1984 (BB84), the six-state, and the Bennett 1992 (B92) protocols [4–6]. Our results hold with respect to a model where two legitimate parties, traditionally called Alice and Bob, are connected by a quantum channel as well as an authentic, but otherwise fully insecure, classical channel.[1] We assume that Alice's source as well as Bob's detector is perfect, whereas an adversary (Eve) might have full control over the quantum channel.[2]

QKD protocols can usually be divided into a quantum and a classical part: In the quantum part, the transmitter (Alice) sends qubits (or more generally, some d-level physical systems) prepared in certain states to the receiver (Bob). The states of these qubits are encodings of bit values randomly chosen by Alice. Bob performs a measurement on the qubits to decode the bit values. For each of the bits, both the encoding and decoding are chosen at random from a certain set of operators. After the transmission step, Alice and Bob apply a *sifting* where they publicly compare the encoding and decoding operators they have used and keep only the bit pairs for which these operators match.

Once Alice and Bob have correlated bit strings, they proceed with the classical part of the protocol. In a first step, called *parameter estimation*, they compare the bit values for a randomly chosen sample of their strings, which gives an estimate of the quantum bit error rate (QBER)—i.e., the fraction of positions where Alice and Bob's strings differ. Note that the QBER is a direct measure for the secrecy of Alice and Bob's strings, since any eavesdropping strategy would, according to the laws of quantum mechanics (no-cloning theorem), perturb the correlations between them.[3] If the QBER is too high, Alice and Bob decide to abort the protocol. Otherwise, they apply a *classical (post-)processing* protocol to distill a secret key, using either one-way or two-way classical communication. One-way post-processing protocols usually consist of *error correction* and *privacy amplification*.[4] For the error correction, Alice sends certain information to Bob such that he can reconstruct Alice's string. Once Alice and Bob have identical strings, privacy amplification is used to compute a final key on which the adversary has virtually no information. We shall see that the performance of such one-way protocols can generally be increased if Alice additionally applies some *preprocessing* to her initial string before starting with the error correction.

Any realistic quantum channel is subject to noise. Consequently, even in the absence of an adversary Eve, the QBER

[1]If Alice and Bob initially share a short key, they can use a classical authentication scheme in order to implement an authentic channel.

[2]One possibility to deal with imperfections of the source or the detector is to include them in the model of the quantum channel (where dark counts might, e.g., be replaced by random bits). This, however, corresponds to a situation where Eve has partial control over these devices, which might be unreasonable

[3]For a fixed attack, the QBER might still take different values with certain probabilities. (Note that the average QBER is irrelevant in this context.)

[4]Error correction and privacy amplification might also be combined into one single protocol step.

 ©2005 The American Physical Society

is nonzero. On the other hand, Eve might in principle replace the real (noisy) quantum channel with an ideal noise-free channel and could thus tap mildly into the quantum communication such as to introduce precisely the original amount of noise. Hence, when proving the security of a protocol, one has to assume that all the noise is due to Eve. This raises the following question: What is the maximum QBER—i.e., the maximum tolerated channel noise—such that Alice and Bob can still generate a secure key? Clearly, the answer to this question depends on the amount of information that Eve might have gained by her attack.

Ideally, one does not want to impose any restriction on Eve's power. That is, any strategy allowed by the laws of physics has to be considered. On the other hand, the set of *all* possible attacks is usually difficult to handle. In order to cope with these conflicting objectives, three classes of attacks have been considered. The smallest class only contains the so-called *individual attacks*, where Eve is restricted to interacting with each of the signal systems sent by Alice separately. That is, for each of the signal systems, Eve attaches an auxiliary system and applies some fixed unitary operation. Finally, Eve measures each of these systems individually right after the sifting step—i.e., before Alice and Bob start with the classical processing. The class of *collective attacks* [7,8] is defined similarly, but the last requirement is dropped. That is, Eve might wait with her measurement until the very end of the protocol. In particular, the measurement she chooses might depend on the messages Alice and Bob exchange for error correction and privacy amplification. Moreover, she might measure all her auxiliary systems jointly. The security analysis of a protocol against collective attacks (see, e.g., [9]) can be seen as a step towards proving security in the most general case—i.e., against *coherent attacks*. The latter includes any attack allowed by the laws of quantum physics. In particular, Eve might let all the signal systems interact with one large auxiliary system, which she only measures at the very end of the protocol.

Many[5] of the previous security proofs of QKD protocols are based on the following observations [12–15].

(i) Instead of preparing a system in a certain state and then sending it to Bob, Alice can equivalently prepare an entangled state, send one of the qubits to Bob, and later measure her subsystem. In doing so, she effectively prepares Bob's system at a distance.

(ii) If the joint system of Alice and Bob is in a pure state, then it cannot be entangled with any third party; in particular it cannot be entangled with any of Eve's auxiliary systems. Hence, simple measurements provide Alice and Bob with data totally oblivious to Eve.

(iii) If furthermore the state shared by Alice and Bob is maximally entangled, then their measurement results are maximally correlated. Hence, if Alice and Bob performed some entanglement purification protocol [16,17], they would end up with the desired secret bits.

(iv) Since one is interested in the security of protocols implemented with nowadays technology, Alice and Bob's operations should not require the storage of quantum states; i.e., one does not want them to run a general entanglement distillation protocol. To overcome this problem, one uses the fact that certain entanglement distillation protocols are mathematically equivalent to quantum error correction codes. There exists a class of such codes, called CSS codes, which have the property that bit errors and phase errors can be corrected separately. Since the final key is classical, its value does not depend on the phase errors. Hence, Alice and Bob actually only have to correct the bit errors, which is a purely classical task.

This method for proving the security of QKD protocols is very elegant, but raises two different questions. First, is the detour via entanglement purification really necessary? Is it optimal? Or might other methods lead to better results? Second, must all cryptographers learn the intricate theory of entanglement? Is there an explanation of the results within the language of information theory? As we shall see, the theory of entanglement purification, as explained above, is not necessary and also too pessimistic (from Alice and Bob's point of view).

In fact, we present a technique for proving the security of QKD protocols which does *not* rely on entanglement purification. Instead, it is based on information-theoretic results on the security of privacy amplification [18,19]. These results were first applied in [20] to analyze the security of a generic QKD protocol similar to the one we are considering here[6] (see also [21] for a similar approach). Since secret key agreement might be possible even if the state describing Alice and Bob's joint system before error correction and privacy amplification does not allow for entanglement distillation, our method can lead to more optimistic results than any method based on entanglement purification.

In addition, we prove security with respect to a so-called *universally composable* security definition. The underlying idea is to characterize the security of a secret key by the maximum probability ε that it deviates from a perfect key which is uniformly distributed and independent of the adversary's information (see Sec. II B for a formal definition). This implies that the key can safely be used in *any* arbitrary context, except with some small probability ε. Remarkably, this is not the case for most of the known security definitions (cf. discussion in [19]).

One interesting example illustrating the strength of our technique is the BB84 protocol or the six-state protocol, where, in the classical processing step, Alice additionally adds some (large) amount of noise to her measurement data. We show that, surprisingly, this noise generally increases the rate at which Alice and Bob can generate secret key bits. However, the density operator describing Alice and Bob's system after the noise has been introduced is not entangled;

[5]This is not true for the first security proof of QKD against the most general attacks due to Mayers [10], which is based on different techniques. Also, the security proof of Biham *et al.* [11] uses different (information-theoretic) methods.

[6]The proof technique introduced in [20] applies to most of the known protocols with one-way error correction and privacy amplification (but without preprocessing). It is based on the result of [18] and the fact that the rank of a purification of Alice and Bob's joint system can be bounded.

i.e., the technique of entanglement purification cannot be applied in a straightforward way.

The paper is organized as follows: In Sec. II, we describe and analyze a generic QKD protocol using one-way classical post-processing. According to the discussion above, the protocol is subdivided into a quantum and a classical part. In Sec. II A, which is devoted to the quantum part, we review our result presented in [22]. It states that the density operator describing Alice and Bob's information after the quantum communication can be considered to be a symmetric (with respect to permutations of the qubit pairs) Bell-diagonal state. The classical part of the protocol is then studied in Sec. II B. Using some recent results of classical and quantum information theory [19,23], we analyze the performance of the classical post-processing. In Sec. III, we combine the main statements of Secs. II A and II B and derive an expression for the secret-key rate which only involves a minimization over a certain set of two-qubit states which correspond to collective attacks. In Sec. IV, we give an upper bound on the secret-key rate for any protocol with one-way classical post-processing, again involving only two-qubit density operators. Finally, in Sec. V, we apply our methods to the BB84, the six-state, and the B92 protocols. In addition, we show that the efficiency of each of these protocols can be increased if one of the parties adds noise to her measurement data.

II. GENERAL QKD PROTOCOL USING ONE-WAY COMMUNICATION

In this section, we describe a general class of QKD protocols employing one-way classical post-processing. This class contains the BB84, the six-state, and the B92 protocols [4–6], among many others. Each of these protocols consists of a quantum and a classical part: The quantum part includes the distribution and measurement of quantum information, and is determined by the operators Alice and Bob use for their encoding and decoding. Section II A is devoted to the analysis of this part. Generally speaking, we review our result proven in [22] which states that the density operator describing Alice and Bob's system after the distribution of quantum information can be assumed to be symmetric [cf. Eq. (1)]. Section II B deals with the classical part of the QKD protocol—i.e., parameter estimation and post-processing. We first give a description of a post-processing scheme and then derive an expression for the maximum length of the key that this scheme can generate, depending on the information that Alice and Bob share after the quantum part of the QKD protocol.

To simplify the presentation of our results, we assume that the physical systems which Alice sends to Bob are qubits. However, a generalization to higher dimensions is straightforward. Throughout the paper, we use the following notation: Vectors $(l_1,\ldots,l_n)$ are denoted by bold letters $\mathbf{l}$. We use capital letters as subscripts for density operators—e.g., σ_{AB}—to denote the subsystems they act on. A bold letter indicates that the corresponding subsystem is itself a product of many (identical) systems. Furthermore, for any state $|\Phi\rangle$, $P_{|\Phi\rangle}=|\Phi\rangle\langle\Phi|$ is the projector onto $|\Phi\rangle$.

A. Quantum part: Distribution of quantum information and measurement

The quantum part of a QKD protocol is specified by the encoding and decoding operations employed by Alice and Bob. For the following, we assume that Alice uses m different encodings, with index $j\in J:=\{1,\ldots,m\}$. For each $j\in J$, $|\phi_j^0\rangle$ and $|\phi_j^1\rangle$ denote the states used to encode the bit values 0 and 1, respectively.

In the first step of the protocol, Alice randomly chooses n bits $x_1,\ldots,x_n$ and sends n qubits prepared in the states $|\phi_{j_1}^{x_1}\rangle,\ldots,|\phi_{j_n}^{x_n}\rangle$ to Bob, for randomly chosen encodings $j_1,\ldots,j_n$. Upon receiving these states (which might have undergone some perturbation, possibly caused by an attack) Bob applies his measurements to obtain classical bits $(y_1,\ldots,y_n)$. Finally, Alice and Bob employ a sifting subprotocol, where they only keep the qubit pairs for which the encoding and measurement operations that they have applied are compatible.

As demonstrated in [22], this protocol can equivalently be described as a so-called entanglement-based scheme [24]. For this purpose, we define the encoding operators $A_j:=|0\rangle\langle(\phi_j^0)^*|+|1\rangle\langle(\phi_j^1)^*|$ and the decoding operators $B_j=|0\rangle\langle\phi_j^1|^{\perp}+|1\rangle\langle\phi_j^0|^{\perp}$, where $\{|0\rangle,\ |1\rangle\}$ is some orthonormal basis, in the following called the z basis. For $x=0$, 1 and $j\in J$, $|(\phi_j^x)^*\rangle$ denotes the complex conjugate of $|\phi_j^x\rangle$ in the z basis and $|\phi_j^x\rangle^{\perp}$ is some (not necessarily normalized) state orthogonal to $|\phi_j^x\rangle$.

For the entanglement-based scheme, Alice simply prepares n two-qubit systems in the state $A_{j_i}\otimes\mathbb{1}|\Phi^+\rangle$, where $|\Phi^+\rangle=1/\sqrt{2}(|0,0\rangle+|1,1\rangle)$, and sends the second qubits to Bob. Then, Bob randomly applies one of the operators B_j to each of the qubits he receives. We denote by $\tilde{\rho}_{\mathbf{AB}}^n$ the state describing the n qubit pairs shared by Alice and Bob after this step.[7] Finally, Alice and Bob measure their parts of $\tilde{\rho}_{\mathbf{AB}}^n$ and associate to the outcomes the bit values 0 or 1.

The description of a QKD protocol as an entanglement-based scheme is very convenient for the security analysis. In particular, instead of considering the quantum communication between Alice and Bob, it suffices to have a characterization of the quantum state $\tilde{\rho}_{\mathbf{AB}}^n$ held by Alice and Bob before they apply their measurements.

Consider now a slight extension of the protocol where Alice and Bob randomly permute the positions of the measured bit pairs and, additionally, at each position, flip the values of both bits with probability one half. In the entanglement-based version of the protocol, these (purely classical) operations can equivalently be applied to the initial quantum state $\tilde{\rho}_{\mathbf{AB}}^n$ of Alice and Bob. For the following, we restrict our attention to the partial state $\tilde{\rho}_{\mathbf{AB}}^{n\mathrm{data}}$ containing only the n_{data} particle pairs which are later used for the computation of the final key (but not for parameter estimation) and

[7] Since we assume that the quantum channel is subject to noise (which might be controlled by the adversary), the state $\tilde{\rho}_{\mathbf{AB}}^n$ is generally a mixed state.

which are measured with respect to the z basis.[8] (To keep the notation simple, we write in the following n instead of n_{data}.) The common bit flip is then described by the quantum operation $\sigma_x \otimes \sigma_x$. Moreover, we can assume that Alice and Bob apply random phase flips $\sigma_z \otimes \sigma_z$ to their qubit pairs, since these do not change the distribution of the classical measurement outcomes. The resulting state $\rho_{\mathbf{AB}}^n$ of Alice and Bob is thus given by $\rho_{\mathbf{AB}}^n = \mathcal{D}_2^{\otimes n}[\mathcal{P}_n(\tilde{\rho}_{\mathbf{AB}}^n)]$ where the operator $\mathcal{P}_n$ denotes the completely positive map (CPM) which symmetrizes the state with respect to permutations of the n qubit pairs and where the CPM $\mathcal{D}_2$ describes the operation where both $\sigma_x \otimes \sigma_x$ and $\sigma_z \otimes \sigma_z$ are applied with probability $\frac{1}{2}$. This is equivalent to the random application of any of the operators $\mathbb{1} \otimes \mathbb{1}$, $\sigma_x \otimes \sigma_x$, $\sigma_y \otimes \sigma_y$, or $\sigma_z \otimes \sigma_z$; i.e., $\mathcal{D}_2$ can be interpreted as the action of a depolarizing channel transforming any two-qubit state to a Bell-diagonal state. Consequently, as shown in [22], $\rho_{\mathbf{AB}}^n$ has the simple form

$$\rho_{\mathbf{AB}}^n = \sum_{n_1,n_2,n_3,n_4}^{n} \mu_{n_1,n_2,n_3,n_4} \rho_{n_1,n_2,n_3,n_4}. \tag{1}$$

In this formula, the sum is taken over all $n_1, n_2, n_3, n_4 \in \mathbb{N}_0$ satisfying $n_1+n_2+n_3+n_4=n$ and μ_{n_1,n_2,n_3,n_4} are some (real-valued) non-negative coefficients. Moreover, ρ_{n_1,n_2,n_3,n_4} is the state of n qubit pairs defined by

$$\rho_{n_1,n_2,n_3,n_4} := \mathcal{P}_n(P_{|\Phi_1\rangle}^{\otimes n_1} \otimes P_{|\Phi_2\rangle}^{\otimes n_2} \otimes P_{|\Phi_3\rangle}^{\otimes n_3} \otimes P_{|\Phi_4\rangle}^{\otimes n_4}), \tag{2}$$

where $P_{|\Phi_1\rangle} := P_{|\Phi^+\rangle}$, $P_{|\Phi_2\rangle} := P_{|\Phi^-\rangle}$, $P_{|\Phi_3\rangle} := P_{|\Psi^+\rangle}$, and $P_{|\Phi_4\rangle} := P_{|\Psi^-\rangle}$ are projectors onto the Bell states $|\Phi^\pm\rangle = 1/\sqrt{2}|0,0\rangle \pm |1,1\rangle$ and $|\Psi^\pm\rangle = 1/\sqrt{2}|0,1\rangle \pm |1,0\rangle)$. Note that the state $\rho_{\mathbf{AB}}^n$ defined by Eq. (1) is, independently of the protocol, separable with respect to the different qubit pairs.

To prove the security of our protocol, we will assume that Eve holds a purification of $\rho_{\mathbf{AB}}^n$, which is the state describing Alice and Bob's joint system *after* they have applied the randomized permutation and bitflips. This is equivalent to saying that Eve knows *everything* that might be correlated (or entangled) with Alice and Bob's system.[9] In particular, the purification of $\rho_{\mathbf{AB}}^n$ includes any information (on Alice and Bob's qubit pair) that Eve might compute when learning the actual permutation and bit flips[10] applied by Alice and Bob.[11] It is explained in [22] that, if the encoding operators A_j are unitary, then this assumption is also tight; i.e., there actually exists an attack which provides Eve with this purification.

Formula (1) is already sufficient to prove our main results (see Sec. III). However, to simplify the analysis of certain protocols, it is often convenient to consider the additional symmetrization (see [22]) given by the CPM $\mathcal{D}_1$ defined by

$$\mathcal{D}_1(\rho) = 1/N \sum_j p_j A_j \otimes B_j(\rho) A_j^\dagger \otimes B_j^\dagger. \tag{3}$$

Here $p_j \geq 0$ determines the probability by which Alice and Bob decide (during the sifting phase) to keep their bits, if they have applied the operation $A_j \otimes B_j$, and N is used for the normalization. All classical data of Alice and Bob (including the bits used for parameter estimation) are then given by a measurement of the state $\mathcal{D}_2^{\otimes n}\{\mathcal{D}_1^{\otimes n}[\mathcal{P}_n(\tilde{\rho}_{AB}^n)]\}$ with respect to the z basis.

B. Classical part: Parameter estimation and classical post-processing

This section is devoted to the description and analysis of the classical part of the QKD protocol. We will use here techniques which partly have been developed in [20]. Assume that Alice and Bob already hold strings $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$, respectively, which they have obtained by measuring n-particle pairs $\rho_{\mathbf{AB}}^n$ distributed in the first part of the QKD protocol, as described in Sec. II A. Their goal is to generate a secure key pair $(\mathbf{S}_A, \mathbf{S}_B)$, using $\mathbf{X}$ and $\mathbf{Y}$.

The protocol we consider consists of two subprotocols, called *parameter estimation* and *classical (post-)processing*. The main purpose of the parameter estimation subprotocol is to estimate the amount of errors that have occurred during the distribution of the quantum information (see Sec. II A). To do this, Alice and Bob compare the measurement outcomes for some randomly chosen qubit pairs. If the quantum bit error rate QBER, they decide to abort the protocol.

In order to analyze a given QKD protocol, we need to characterize the initial states $\rho_{\mathbf{AB}}^n$ for which the protocol does not abort. Clearly, this characterization depends on the threshold QBER. Let Γ be the set of all two-qubit states σ_{AB} which correspond to a collective attack, meaning that there exists an operation of Eve such that $\rho_{\mathbf{AB}}^n = \sigma_{AB}^{\otimes n}$. The set Γ_{QBER} is then defined as the subset of Γ containing all states σ_{AB} for which the protocol does *not* abort (with probability almost 1). In other words, if $\sigma_{AB} \in \Gamma_{\text{QBER}}$, then the protocol is supposed to compute a secret key when starting with $\rho_{\mathbf{AB}}^n = \sigma_{AB}^{\otimes n}$. We will see in Sec. III that the characterization of the set Γ_{QBER} is sufficient to compute lower bounds on the secret-key rate.

After the parameter estimation, if the estimate for the QBER is below the threshold, Alice and Bob proceed with a classical subprotocol in order to turn their only partially secure strings $\mathbf{X}$ and $\mathbf{Y}$ into a highly secure key pair $(\mathbf{S}_A, \mathbf{S}_B)$. The protocol we consider is one way; i.e., only communication from Alice to Bob is needed. It consists of three steps.

(I) *Preprocessing:* Using her bit string $\mathbf{X}$, Alice computes two strings $\mathbf{U}$ and $\mathbf{V}$, according to some channels $\mathbf{U} \leftarrow \mathbf{X}$ and $\mathbf{V} \leftarrow \mathbf{U}$, defined by conditional probability distributions $P_{\mathbf{U}|\mathbf{X}}$ and $P_{\mathbf{V}|\mathbf{U}}$, respectively. She keeps $\mathbf{U}$ and sends $\mathbf{V}$ to Bob. (We will see that, for most protocols, the performance highly depends on a clever choice of $\mathbf{U}$, whereas the string $\mathbf{V}$ is usually not needed.)

(II) *Information reconciliation:* Alice sends error correction information $\mathbf{W}$ on $\mathbf{U}$ to Bob. Using $\mathbf{Y}$, $\mathbf{V}$, and $\mathbf{W}$, Bob computes a guess $\hat{\mathbf{U}}$ for $\mathbf{U}$.

[8] We will see in Sec. III that one can always assume that all these particle pairs are measured with respect to the same basis.

[9] Indeed, conditioned on any measurement of Eve's system, Alice and Bob's joint system is in a pure state.

[10] Note that Alice and Bob have to communicate over an insecure classical channel in order to agree on the common random permutation and bit flips.

[11] See [22] for a more formal statement and proof.

(III) *Privacy amplification:* Alice randomly chooses a function F from a family of two-universal hash functions[12] and sends a description of F to Bob. Then Alice and Bob compute their keys, $\mathbf{S}_A = F(\mathbf{U})$ and $\mathbf{S}_B = F(\hat{\mathbf{U}})$, respectively.

Before starting with the analysis of this protocol, let us introduce some notation. It is most convenient to describe the classical information of Alice and Bob as well as the quantum information of the adversary Eve by a tripartite density operator $\rho_{\mathbf{XY}E}$ of the form

$$\rho^n_{\mathbf{XY}E} = \sum_{\mathbf{x},\mathbf{y}} P_{\mathbf{XY}}(\mathbf{x},\mathbf{y}) P_{|\mathbf{x}\rangle} \otimes P_{|\mathbf{y}\rangle} \otimes \rho_E^{\mathbf{x},\mathbf{y}}, \tag{4}$$

where $\{|\mathbf{x}\rangle\}_{\mathbf{x}}$ and $\{|\mathbf{y}\rangle\}_{\mathbf{y}}$ are families of orthonormal vectors and where $\rho_E^{\mathbf{x},\mathbf{y}}$ is the quantum state of Eve given that Alice and Bob's random variables $\mathbf{X}$ and $\mathbf{Y}$ take the values $\mathbf{x}$ and $\mathbf{y}$, respectively. Similarly, the classical key pair $(\mathbf{S}_A, \mathbf{S}_B)$ together with the adversary's information $\rho_{E'}^{\mathbf{S}_A,\mathbf{S}_B}$ after the protocol execution is described by a quantum state $\rho_{\mathbf{S}_A\mathbf{S}_B E'}$.

To define the security of the final key pair $(\mathbf{S}_A, \mathbf{S}_B)$, we use the universally composable security definition introduced in [19]. The key pair $(\mathbf{S}_A, \mathbf{S}_B)$ is said to be ε *secure* (with respect to $\rho_{E'}$) if

$$\delta(\rho_{\mathbf{S}_A\mathbf{S}_B E'}, \rho_{\mathbf{SS}} \otimes \rho_{E'}) \leq \varepsilon, \tag{5}$$

where $\rho_{\mathbf{SS}} := \Sigma_{\mathbf{s}\in\mathcal{S}} 1/|\mathcal{S}| P_{|\mathbf{s}\rangle} \otimes P_{|\mathbf{s}\rangle}$ and where $\delta(\cdot,\cdot)$ denotes the trace distance. In other words, the state $\rho_{\mathbf{S}_A\mathbf{S}_B E'}$ describing the key of Alice and Bob together with the adversary's quantum system is ε close to a product state $\rho_{\mathbf{SS}} \otimes \rho_{E'}$ where the partial state $\rho_{\mathbf{SS}}$ describes a pair of identical and uniformly distributed keys. This is equivalent to saying that, with probability at least $\mathbf{1}-\varepsilon$, the keys $\mathbf{S}_A$ and $\mathbf{S}_B$ are *equal* to a *perfect key* $\mathbf{S}$ which is uniformly distributed and completely independent of the adversary's knowledge (cf. [19] for a proof). Hence, except with some small probability ε, Alice and Bob can safely use their key pair $(\mathbf{S}_A, \mathbf{S}_B)$ for any cryptographic task (e.g., for one-time-pad encryption) which is secure when using a perfect key $\mathbf{S}$.

The goal of the remaining part of this section is to derive an expression for the number $\ell^{(n)}$ of ε-secure key bits that can be generated from n qubit pairs by the above protocol, for an optimal choice of the protocol parameters. For this purpose, we first consider some fixed preprocessing, specified by the channels $\mathbf{U} \leftarrow \mathbf{X}$ and $\mathbf{V} \leftarrow \mathbf{U}$, for which we compute the maximum key length $\ell^n_{\mathbf{U}\leftarrow\mathbf{X},\mathbf{V}\leftarrow\mathbf{U}}$. The quantity $\ell^{(n)}$ is then obtained by optimizing over all choices of the preprocessing.

Our result is formulated in terms of an information-theoretic quantity, called *smooth Rényi entropy* [23] (see Appendix A for more details). Similarly to the Shannon entropy $H(X)$, the smooth Rényi entropy of a random variable X, denoted by $H^\varepsilon_\alpha(X)$, is a measure for the uncertainty about the value of X. We will also need an extension of this entropy measure to quantum states. Similarly to the von Neumann entropy $S(\rho)$, the smooth Rényi entropy $S^\varepsilon_\alpha(\rho)$ of a state ρ quantifies the amount of randomness contained in ρ.

The main ingredient needed for the following derivation is a recent result on the security of privacy amplification [19] (see lemma C.2). Generally speaking, it says that the length of the key that can be extracted from a string $\mathbf{U}$ held by both Alice and Bob is given by the uncertainty of the adversary about $\mathbf{U}$, measured in terms of smooth Rényi entropies. Applied to the last step of our protocol, we get

$$\ell^{(n)}_{\mathbf{U}\leftarrow\mathbf{X},\mathbf{V}\leftarrow\mathbf{U}} \approx S^\varepsilon_2(\rho^n_{\mathbf{UVWE}}) - S^\varepsilon_0(\rho^n_{\mathbf{VWE}}), \tag{6}$$

where ε depends on the desired security of the final key and where the approximation "$\approx$" means that equality holds up to some small additive term of the order $O(\ln(1/\varepsilon))$. In this formula, $\rho^n_{\mathbf{UVW}E}$ is the density operator describing the strings $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$, together with the adversary's knowledge—i.e.,

$$\rho^n_{\mathbf{UVWE}} = \sum_{\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{v},\mathbf{w}} P_{\mathbf{XYUVW}}(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{v},\mathbf{w}) P_{|u\rangle} \otimes P_{|v\rangle} \otimes P_{|w\rangle} \otimes \rho_E^{\mathbf{x},\mathbf{y}},$$

where $\{|\mathbf{u}\rangle\}_{\mathbf{u}}$, $\{|\mathbf{v}\rangle\}_{\mathbf{v}}$, and $\{|\mathbf{w}\rangle\}_{\mathbf{w}}$ are families of orthonormal vectors. Note that, since the channel connecting Alice and Bob might be arbitrarily insecure, the key must be secure even if the adversary knows $\mathbf{V}$ and $\mathbf{W}$.

In the next step, we will eliminate the dependence on $\mathbf{W}$ in Eq. (6). For this, we consider the amount m of (useful) information contained in $\mathbf{W}$. Since $\mathbf{W}$ is needed by Bob in order to guess $\mathbf{U}$, m depends on his uncertainty about $\mathbf{U}$. In fact, if an optimal error correction code is applied, then m is roughly equal to the entropy of $\mathbf{U}$ conditioned on Bob's information $\mathbf{Y}$ and $\mathbf{V}$. More precisely, using lemma C.3 described in Appendix C, we have $m \approx H^\varepsilon_0(\mathbf{U}|\mathbf{YV})$. Hence, when omitting $\mathbf{W}$ on the right-hand side of Eq. (6), the smooth Rényi entropies cannot decrease by more than m (see Appendix A for a summary of the properties of smooth Rényi entropy). We thus immediately obtain

$$\ell^{(n)}_{\mathbf{U}\leftarrow\mathbf{X},\mathbf{V}\leftarrow\mathbf{U}} \approx S^\varepsilon_2(\rho^n_{\mathbf{UV}E}) - S^\varepsilon_0(\rho^n_{\mathbf{V}E}) - H^\varepsilon_0(\mathbf{U}|\mathbf{VY}). \tag{7}$$

Since the channels $\mathbf{U} \leftarrow \mathbf{X}$ and $\mathbf{V} \leftarrow \mathbf{U}$ applied by Alice in the first step of the classical post-processing protocol are arbitrary, we can optimize over all choices of such channels. We thus conclude that the number $\ell^{(n)}$ of key bits that can be generated by the described protocol, for an optimal choice of all the parameters, is given by

$$\ell^{(n)} \approx \sup_{\substack{\mathbf{U}\leftarrow\mathbf{X} \\ \mathbf{V}\leftarrow\mathbf{U}}} S^\varepsilon_2(\rho^n_{\mathbf{UV}E}) - S^\varepsilon_0(\rho^n_{\mathbf{V}E}) - H^\varepsilon_0(\mathbf{U}|\mathbf{VY}). \tag{8}$$

In the following, we will often consider protocols where the strings $\mathbf{U}$ and $\mathbf{V}$ are computed bitwise from the string $\mathbf{X}$. The maximum length of the secret key that can be generated by such a protocol is then given by an expression similar to Eq. (8), but where the supremum is only taken over bitwise channels $U \leftarrow X$ and $V \leftarrow U$.

[12] For a definition and constructions of two-universal hash functions, see, e.g., [25] or [26].

III. LOWER BOUND ON THE SECRET-KEY RATE

The goal of this section is to derive a lower bound for the secret-key rate which only involves two-qubit states and which is thus easy to compute. For this purpose, we use the general expression (8) of Sec. II B for the number of key bits that can be generated from a given state, together with the fact that, after symmetrization, any state of Alice and Bob has the simple form (1).

Let us start with a description of our main result. Consider the QKD protocol described in Sec. II, where we assume that Alice uses bitwise channels $U \leftarrow X$ and $V \leftarrow U$ to compute $\mathbf{U}=(U_1,\ldots,U_n)$ and $\mathbf{V}=(V_1,\ldots,V_n)$, respectively, from her data $\mathbf{X}=(X_1,\ldots,X_n)$. Let Γ_{QBER} be the set of two-qubit density operators σ_{AB} defined in Sec. II B; i.e., the protocol aborts (with high probability) whenever it starts with a product state $(\sigma_{AB})^{\otimes n}$ for any $\sigma_{AB} \notin \Gamma_{\mathrm{QBER}}$. We show that, for an optimal choice of parameters, the protocol of the previous section generates secret-key bits at rate $r := \lim_{n\to\infty}(\ell^{(n)}/n)$ where

$$r \geqslant \sup_{\substack{U\leftarrow X \\ V\leftarrow U}} \inf_{\sigma_{AB}\in\Gamma_{\mathrm{QBER}}} \left(S(U|VE) - H(U|YV)\right). \tag{9}$$

In this formula, $S(U|VE)$ denotes the von Neumann entropy of U conditioned on V and Eve's initial information—i.e., $S(U|VE) := S(\sigma_{UVE}) - S(\sigma_{VE})$. The state σ_{UVE} is obtained from σ_{AB} by taking a purification σ_{ABE} of the Bell diagonal state $\sigma_{AB}^{\mathrm{diag}} := \mathcal{D}_2(\sigma_{AB})$,[13] and applying the measurement of Alice followed by the classical channels $U \leftarrow X$ and $V \leftarrow U$. Similarly, Y is the outcome of Bob's measurement applied to the second subsystem of σ_{ABE}.

As Eq. (9) involves a minimization over the set Γ_{QBER} of two-qubit states, our lower bound on the secret-key rate only depends on the set of possible collective attacks. On the other hand, the security we prove holds against any arbitrary coherent attack. Note also that the statement extends to the situation where Alice—instead of applying a bitwise preprocessing on each of the n bits—uses some operation involving larger blocks—say, of length m. In this case, one has to consider all attacks $U^{\otimes r}$ where the adversary applies the same operation U on each of the $r=n/m$ blocks.

A crucial task when computing explicit values for Eq. (9) is to characterize the set Γ_{QBER}, This set is determined by the conditions under which the protocol aborts. In Sec. V, we will demonstrate how formula (9) is computed for concrete QKD schemes such as the BB84 or six-state protocol. It turns out that, in these examples, the maximum is taken if $V \leftarrow U$ is the trivial channel where V is independent of U; i.e., the random variable V can be omitted.

One method to further reduce the number of parameters is to consider the set $\mathcal{D}_2(\mathcal{D}_1(\Gamma_{\mathrm{QBER}}))$, which only contains normalized two-qubit density operators of the form

$$\rho^1[\boldsymbol{\lambda}] = \lambda_1 P_{|\Phi^+\rangle} + \lambda_2 P_{|\Phi^-\rangle} + \lambda_3 P_{|\Psi^+\rangle} + \lambda_4 P_{|\Psi^-\rangle}, \tag{10}$$

i.e., Eq. (1) for $n=1$. As mentioned in Sec. II A (see [22] for details), the state shared by Alice and Bob is—independently of the considered protocol—measured with respect to the z basis. Hence, we obtain for the QBER Q, computed as an average over the different encodings, $Q=\lambda_3+\lambda_4$. Apart from that, the state must be normalized, which implies that, for any given value of Q, there are at most two free parameters λ_2 and λ_3—i.e., $\lambda_1 = 1-Q-\lambda_2$ and $\lambda_4 = Q-\lambda_3$.

To prove Eq. (9), we will make use of a known result [20] on the relation between the statistics obtained when applying two different measurements $\mathcal{E}$ and $\mathcal{F}$ on the individual subsystems of a symmetric n-partite state ρ^n (cf. lemma C.1 in Appendix C). Let $\mathbf{Z}=(Z_1,\ldots,Z_k)$ be the outcomes when applying $\mathcal{E}$ to each of the first k subsystems of ρ^n, for $k \leqslant n$, and let $Q_{\mathbf{Z}}$ be the frequency distribution of the symbols in the string $\mathbf{Z}$, i.e., for any possible measurement outcome z,

$$Q_{\mathbf{Z}}(z) := \frac{|\{i: Z_i = z\}|}{k}.$$

Similarly, let $Q_{\mathbf{Z}'}$ be the frequency distribution of the outcomes $\mathbf{Z}'=(Z'_1,\ldots,Z'_{k'})$ of $\mathcal{F}$ applied to k' of the remaining $n-k$ subsystems of ρ^n. Lemma C.1 implies that, if k and k' are large enough, then, with probability almost 1, there exists a density operator σ on one subsystem which is compatible with both of these statistics. Formally, this means that $Q_{\mathbf{Z}} \approx P_{\mathcal{E}}[\sigma]$ and $Q_{\mathbf{Z}} \approx P_{\mathcal{F}}[\sigma]$, where $P_{\mathcal{E}}[\sigma]$ and $P_{\mathcal{F}}[\sigma]$ denote the probability distributions of the outcomes when measuring σ with respect to $\mathcal{E}$ and $\mathcal{F}$, respectively. Moreover, the state σ is contained in a certain set $\mathcal{B}$ which, roughly speaking, contains all density operators which correspond to the state of one single subsystem of ρ^n, conditioned on any measurement on the remaining subsystems.

We are now ready to prove expression (9) for the secret-key rate. As in Sec. II A, we consider an extension of the protocol where, before invoking the classical part of the QKD protocol, Alice and Bob symmetrize their strings $\mathbf{X}$ and $\mathbf{Y}$. More concretely, they both apply the *same* randomly chosen permutation on their strings. Clearly, this is equivalent to a protocol where Alice and Bob first permute and then measure their bits (see Sec. II A). The state $\rho_{\mathbf{AB}}^n$ of Alice and Bob's system before the measurement is then symmetric. We can thus assume without loss of generality that the first n_{pe} qubit pairs are used for the parameter estimation, while the actual key is generated from the measurement outcomes obtained from the next n_{data} pairs.

Consider now some fixed protocol where the preprocessing is defined by the channels $U \leftarrow X$ and $V \leftarrow U$. We show that this protocol is secure as long as the rate at which the key is generated is not larger than

$$r_{U\leftarrow X, V\leftarrow U} = \inf_{\sigma_{AB}\in\Gamma_{\mathrm{QBER}}} \left(S(U|VE) - H(U|YV)\right). \tag{11}$$

In other words, $r_{U\leftarrow X, V\leftarrow U}$ is the rate that can be achieved if the channels $U \leftarrow X$ and $V \leftarrow U$ are used for the preprocessing. The assertion (9) then follows by optimizing over all channels for the preprocessing.

[13] This means that $\sigma_{AB}^{\mathrm{diag}}$ has the same diagonal entries as σ_{AB} with respect to the Bell basis.

The proof of Eq. (11) is subdivided into two parts. In the first part, we show that the parameter estimation works correctly; i.e., if the adversary introduces too much noise, then the protocol aborts. The second part of the proof is concerned with the security of the classical post-processing step; that is, if the noise is below a certain level, then the final key is secure.

For this analysis, we need to consider the joint state $\rho_{\mathbf{AB}}^{n_{\mathrm{pe}}+n_{\mathrm{data}}}$ of the n_{pe} qubit pairs used for parameter estimation and the n_{data} pairs used for the classical post-processing. Additionally, in order to simplify the presentation of the proof, we assume that there is a small number $n_{\mathrm{aux}} := n - n_{\mathrm{pe}} - n_{\mathrm{data}} > 0$ of auxiliary qubit pairs exchanged by Alice and Bob which are not used in the classical part of the protocol.[14] In order to analyze the structure of the state $\rho_{\mathbf{AB}}^{n_{\mathrm{pe}}+n_{\mathrm{data}}}$, we consider an imaginary measurement $\mathcal{E}_{\mathrm{Bell}}$ with respect to the Bell basis applied to each of these n_{aux} auxiliary positions of $\rho_{\mathbf{AB}}^{n}$. We then prove the security of our QKD protocol conditioned on the statistics $Q_{\mathbf{W}}$ of the outcomes $\mathbf{W}=(W_1,\ldots,W_{n_{\mathrm{aux}}})$ of this imaginary measurement.[15] We show that the protocol is secure for any value of $Q_{\mathbf{W}}$, which implies that the protocol is secure in general (with probability almost 1).

Formally, let $P_{\mathcal{E}_{\mathrm{Bell}}}[\Gamma_{\mathrm{QBER}}]$ be the set of probability distributions obtained by measuring the states $\sigma_{AB} \in \Gamma_{\mathrm{QBER}}$ with respect to the Bell basis. We prove the following two statements.

(i) If $Q_{\mathbf{W}} \notin P_{\mathcal{E}_{\mathrm{Bell}}}[\Gamma_{\mathrm{QBER}}]$, then the protocol aborts after the parameter estimation; i.e., no key is generated.

(ii) If $Q_{\mathbf{W}} \in P_{\mathcal{E}_{\mathrm{Bell}}}[\Gamma_{\mathrm{QBER}}]$, then the key generated by the classical post-processing is secure.

To prove statement (i), let $\mathcal{F}$ be the measurement that Alice and Bob apply to each of the n_{pe} qubit pairs used for parameter estimation and let Q_{pe} be the frequency distribution of the measurement outcomes of $\mathcal{F}$. Since the state $\rho_{\mathbf{AB}}^{n}$ is symmetric, we can apply lemma C.1 described above, where $\mathcal{B}$ is defined by the set Γ of all two-qubit states characterizing the collective attacks of Eve (cf. Sec. II B). Consequently, there exists a state $\sigma_{AB} \in \Gamma$ (of a *single* qubit pair) which is compatible with both the statistics Q_{pe} and $Q_{\mathbf{W}}$—i.e., $P_{\mathcal{F}}[\sigma_{AB}] \approx Q_{\mathrm{pe}}$ and $P_{\mathcal{E}_{\mathrm{Bell}}}[\sigma_{AB}] \approx Q_{\mathbf{W}}$. Assume now that $Q_{\mathbf{W}} \notin P_{\mathcal{E}_{\mathrm{Bell}}}[\Gamma_{\mathrm{QBER}}]$. Because of $P_{\mathcal{E}_{\mathrm{Bell}}}[\sigma_{AB}] \approx Q_{\mathbf{W}}$, this implies that $\sigma_{AB} \notin \Gamma_{\mathrm{QBER}}$. Hence, by the definition of the set Γ_{QBER}, the protocol aborts.

We proceed with the proof of statement (ii). For any frequency distribution Q, let $\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$ be the state of the n_{data} qubit pairs used for generating the final key, conditioned on the event that the statistics of the measurement outcomes of the n_{aux} auxiliary pairs is equal to Q. Assume now that Alice and Bob measure their data bits according to one fixed basis,[16] called the z basis, and, additionally, apply common random bit flips. Then, according to the discussion in Sec. II A, it is sufficient to consider states of the form (1). In particular, the conditional state $\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$ can be written as

$$\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}} = \sum_{n_1,n_2,n_3,n_4} \mu_{n_1,n_2,n_3,n_4} \rho_{n_1,n_2,n_3,n_4}, \tag{12}$$

where ρ_{n_1,n_2,n_3,n_4} is defined by Eq. (2). Hence, if we applied the Bell measurement $\mathcal{E}_{\mathrm{Bell}}$ to each of the n_{data} subsystems, then, for any 4-tuple (n_1,n_2,n_3,n_4), with probability μ_{n_1,n_2,n_3,n_4}, the resulting frequency distribution Q_{data} would be equal to $Q_{n_1,n_2,n_3,n_4} := (n_1/n, n_2/n, n_3/n, n_4/n)$. On the other hand, because of the permutation symmetry of the state $\rho_{\mathbf{AB}}^{n}$, we have $Q_{\mathrm{data}} \approx Q_{\mathbf{W}}$ with probability almost one 1.[17] Hence, the coefficients μ_{n_1,n_2,n_3,n_4} can only be non-negligible if Q_{n_1,n_2,n_3,n_4} is close to $Q_{\mathbf{W}}=Q$; that is, we can restrict the sum in Eq. (12) to values (n_1,n_2,n_3,n_4) such that $Q_{n_1,n_2,n_3,n_4} \approx Q$.

Consider now the product state $(\sigma_{AB})^{\otimes n_{\mathrm{data}}}$, where $\sigma_{AB} := \rho^1[Q]$ is the two-qubit state depending on Q as defined by Eq. (10). Since the state $(\sigma_{AB})^{\otimes n_{\mathrm{data}}}$ is symmetric, we can also write it in the form (12), with some coefficients μ'_{n_1,n_2,n_3,n_4}. Again, these coefficients can only be non-negligible if Q_{n_1,n_2,n_3,n_4} is close to Q. Hence, the states $\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$ and $(\sigma_{AB})^{\otimes n_{\mathrm{data}}}$ have the same structure (12) where the coefficients μ_{n_1,n_2,n_3,n_4} and μ'_{n_1,n_2,n_3,n_4} are negligible except for $Q_{n_1,n_2,n_3,n_4} \approx Q$. Indeed, it is a consequence of the results presented in Appendix A 3 that the smooth Rényi entropies of the states derived from $\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$ are roughly equal to the corresponding entropies of the states derived from $(\sigma_{AB})^{\otimes n_{\mathrm{data}}}$. To make this a bit more precise, let $\rho_{\mathbf{UV}E|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$ be the state obtained when applying the measurement of Alice followed by the channels $U \leftarrow X$ and $V \leftarrow U$ to each of the subsystems of a purification of $\rho_{\mathbf{AB}|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}$. Then, lemma A.4 implies that

$$S_2^{\varepsilon}(\rho_{\mathbf{UV}E|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}) \gtrapprox n_{\mathrm{data}} S(\sigma_{UVE})$$

and

$$S_0^{\varepsilon}(\rho_{\mathbf{V}E|Q_{\mathbf{W}}=Q}^{n_{\mathrm{data}}}) \lessapprox n_{\mathrm{data}} S(\sigma_{VE}),$$

where σ_{UVE} is the state obtained from $\sigma_{AB} := \rho^1[Q]$, as described after Eq. (9).

[14] If this is not the case, one can always change the protocol such that some of the data bits are discarded, without reducing its rate.

[15] Note that the outcomes $\mathbf{W}$ of this imaginary Bell measurement on the auxiliary qubit pairs are only needed for the security analysis. In the actual protocol, Alice and Bob do never have to perform any measurement operation on the auxiliary qubit pairs.

[16] If the data bits are measured with respect to different bases, the argument must be repeated for each basis. This is, however, usually not needed. In fact, for an optical performance of the protocol, one of the encodings should be chosen with probability almost 1 whereas the other encodings should only be chosen with some small probability [27]. (The bit pairs resulting from the latter are then only used for parameter estimation.) This reduces the number of qubit pairs lost in the sifting step.

[17] One might use the lemma C.1 (with $\mathcal{E}=\mathcal{F}=\mathcal{E}_{\mathrm{Bell}}$) to get a quantitative statement.

Using these identities, it follows from Eq. (7) that the final key generated by the protocol of the previous section, for fixed channels $U \leftarrow X$ and $V \leftarrow U$, is secure as long as its length is not larger than

$$\ell_{U\leftarrow X,V\leftarrow U}[\sigma_{AB}] \approx n_{\text{data}}(S(\sigma_{UVE}) - S(\sigma_{VE}) - H(U|VY)),$$

for $\sigma_{AB}=\rho^1[Q]$. In other words, $\ell_{U\leftarrow X,V\leftarrow U}[\sigma_{AB}]$ is the length of a secure key that can be extracted when applying the protocol to a state of the form $\rho^{n_{\text{data}}}_{\mathbf{AB}|Q_{\mathbf{W}}=Q}$.

Since the final key must be secure for all possible initial states for which the protocol does not abort, we have to take the minimum of this quantity over the states $\sigma_{AB}=\rho^1[Q]$, for any $Q \in P_{\mathcal{E}_{\text{Bell}}}[\Gamma_{\text{QBER}}]$. Since, according to Eq. (10), $\rho^1[Q]$ is diagonal, the minimum ranges over all diagonal states $\sigma^{\text{diag}}_{AB}$ whose diagonal elements correspond to $Q \in P_{\mathcal{E}_{\text{Bell}}}(\Gamma_{\text{QBER}})$. This is equivalent to saying that the diagonal elements of $\sigma^{\text{diag}}_{AB}$ are equal to the diagonal entries of a density operator $\sigma_{AB} \in \Gamma_{\text{QBER}}$; i.e., the number ℓ of key bits generated by the protocol is given by

$$\ell_{U\leftarrow X,V\leftarrow U} := \inf_{\sigma_{AB}\in\Gamma_{\text{QBER}}} \ell_{U\leftarrow X,V\leftarrow U}[\sigma^{\text{diag}}_{AB}],$$

where $\sigma^{\text{diag}}_{AB} := \mathcal{D}_2(\sigma_{AB})$. This concludes the proof of Eq. (11) and thus also Eq. (9).

IV. UPPER BOUND ON THE SECRET-KEY RATE

As demonstrated in Sec. III, the rate of a QKD protocol is lower bounded by an expression which only involves von Neumann entropies of states of single-qubit pairs [cf. Eq. (9)]. In the following, we show that, roughly speaking, the right-hand side of Eq. (9) is also an upper bound on the rate if the supremum is taken over all *quantum* channels (instead of only classical channels) $U \leftarrow X$ and $V \leftarrow X$.

Clearly, in order to prove upper bounds, it is sufficient to consider collective attacks. We thus assume that the overall state $\rho^n_{\mathbf{AB}E}$ of Alice's, Bob's, and Eve's quantum system has product form—i.e., $\rho^n_{ABE}=\sigma^{\otimes n}_{ABE}$—for some tripartite state σ_{ABE}. Hence, before starting with the classical processing, the situation is fully specified by the n-fold product state $\sigma^{\otimes n}_{XYE}$, where σ_{XYE} is the state obtained when applying Alice's and Bob's measurements to σ_{ABE}. Similarly to Eq. (4), σ_{XYE} can be written as

$$\sigma_{XYE} = \sum_{x,y} P_{XY}(x,y) P_{|x\rangle} \otimes P_{|y\rangle} \otimes \sigma^{x,y}_E.$$

We show that the rate $r(\sigma_{XYE})$ at which secret-key bits can be generated from this situation, using only a public communication channel from Alice and Bob, is upper bounded by

$$r(\sigma_{XYE}) \leqslant \sup_{\substack{\sigma_U\leftarrow X\\ \sigma_V\leftarrow X}} (S(U|VE) - S(U|YV)). \tag{13}$$

In this formula, the supremum is taken over all density operators σ^x_U and σ^x_V depending on x. The density operators occurring in the entropies are then given by the appropriate traces of

$$\sigma_{UVYE} := \sum_{x,y} P_{XY}(x,y) \sigma^x_U \otimes \sigma^x_V \otimes P_{|y\rangle} \otimes \sigma^x_E. \tag{14}$$

A similar upper bound for the key rate follows from a result of Devetak and Winter [28]. In contrast to Eq. (13), their formula involves an additional limes over the number n of product states, whereas the supremum only involves classical channels $U \leftarrow X$ and $V \leftarrow U$.

Because of the optimization over the density operators σ^x_U and σ^x_V, expression (13) is generally hard to evaluate. To simplify this computation, it is convenient to consider measurements of Eve, resulting in classical values Z. In this case, the bound corresponds to a known result due to Csiszár and Körner [29],

$$r(X,Y,Z) = \sup_{\substack{U\leftarrow X\\ V\leftarrow U}} (H(U|VZ) - H(U|YV)). \tag{15}$$

The proof of the upper bound (13) is subdivided into two parts: First, in Sec. IV A, we give general conditions on a measure M such that $M(\sigma_{XYE})$ is an upper bound on the rate $r_{\sigma_{XYE}}$. Second, in Sec. IV B, we show that the measure M defined by the right-hand side of Eq. (13) satisfies these conditions.

A. General properties of upper bounds

Let M be a real-valued function on the set of tripartite density operators. We show that $M(\sigma_{XYE})$ is an upper bound on the rate $r_{\sigma_{XYE}}$ if the following conditions are satisfied [here, we also write $M(X;Y;E)$ instead of $M(\sigma_{XYE})$; moreover, if a random variable X' is computed from X, we write $X' \leftarrow X$]:

(i) $M(\sigma^{\otimes n}_{XYE}) \leqslant nM(\sigma_{XYE})$, for any $n \in \mathbb{N}$.
(ii) $M(X';Y;E) \leqslant M(X;Y;E)$ for $X' \leftarrow X$.
(iii) $M(X;Y';E) \leqslant M(X;Y;E)$ for $Y' \leftarrow Y$.
(iv) $M(XC;YC;EC) \leqslant M(X;Y;E)$ for $C \leftarrow X$.
(v) There exists a function α with $\lim_{\varepsilon\to 0} \alpha(\varepsilon)=0$ such that, for any state $\rho_{\mathbf{S}_A\mathbf{S}_B E}$ describing an ε-secure key pair of length ℓ [cf. Eq. (5)],

$$M(\rho_{\mathbf{S}_A\mathbf{S}_B E}) \geqslant [1-\alpha(\varepsilon)]\ell.$$

Consider an arbitrary secret-key agreement protocol and assume that the protocol starts with n copies of the state σ_{XYE}. Let $\rho^n_{\mathbf{S}_A\mathbf{S}_B E'}$ be the overall state of Alice's and Bob's key $\mathbf{S}_A$ and $\mathbf{S}_B$, respectively, together with the adversary's information E' after the protocol execution. Then, using properties (i)–(iv), we find

$$nM(\sigma_{XYE}) \geqslant M(\sigma^{\otimes n}_{XYE}) \geqslant M(\rho^n_{\mathbf{S}_A\mathbf{S}_B E'}). \tag{16}$$

For any $n \in \mathbb{N}$, the resulting state must be $\varepsilon(n)$ close to a state describing a secret key of length $\ell(n)$, for $\varepsilon(n)\to 0$ as n approaches infinity. Hence, from Eq. (16) and property (v),

$$M(\sigma_{XYE}) \geqslant \lim_{n\to\infty} \frac{\ell(n)}{n} = r(\sigma_{XYE}),$$

which concludes the proof.

B. Concrete expression for the upper bound

Let M be the measure defined by the right-hand side of Eq. (13); i.e., for any tripartite density operator σ_{XYE}, $M(\sigma_{XYE}) := M(X;Y;E)$ is given by

$$M(X;Y;E) := \sup_{\sigma_U^x, \sigma_V^x} (S(U|VE) - S(U|VY)).$$

The goal of this section is to show that this measure satisfies the conditions of Sec. IV A, which implies that $M(\sigma_{XYE})$ is an upper bound on the secret-key rate $r(\sigma_{XYE})$.

Let us start with property (i). It suffices to show that, for any state $\sigma_{XYEX'Y'E'} := \sigma_{XYE} \otimes \sigma_{X'Y'E'}$,

$$M(XX';YY';EE') \leqslant M(X;Y;E) + M(X';Y';E'),$$

i.e.,

$$\sup_{(\tilde{U},\tilde{V})\leftarrow(X,X')} S(\tilde{U}|\tilde{V}EE') - S(\tilde{U}|\tilde{V}YY') \leqslant \sup_{(U,V)\leftarrow X} S(U|VE) - S(U|VY) + \sup_{(U',V')\leftarrow X'} S(U'|V'E') - S(U'|V'Y'),$$

where $(U,V)\leftarrow X$ [and likewise $(U',V')\leftarrow X'$ and $(\tilde{U},\tilde{V}) \leftarrow (X,X')$] means that the density operators σ_U^x and σ_V^x used for the definition of σ_{UVE} and σ_{UVY} [cf. Eq. (14)] are computed from the classical random variable X. The left-hand side of this expression can be upper bounded by

$$\sup_{(\tilde{U},\tilde{V})\leftarrow(X,X')} S(\tilde{U}|\tilde{V}EE') - S(\tilde{U}|\tilde{V}YE') + \sup_{(\tilde{U},\tilde{V})\leftarrow(X,X')} S(\tilde{U}|\tilde{V}YE') - S(\tilde{U}|\tilde{V}YY').$$

It thus remains to be shown that for any $(\tilde{U},\tilde{V})\leftarrow(X,X')$ there exists $(U,V)\leftarrow X$ such that

$$S(\tilde{U}|\tilde{V}EE') - S(\tilde{U}|\tilde{V}YE') \leqslant S(U|VE) - S(U|VY) \quad (17)$$

and, similarly, for any $(\tilde{U},\tilde{V})\leftarrow(X,X')$ there exists $(U',V') \leftarrow X'$ such that

$$S(\tilde{U}|\tilde{V}YE') - S(\tilde{U}|\tilde{V}YY') \leqslant S(U'|V'E') - S(U'|V'Y'). \quad (18)$$

Inequality (17) follows from the observation that $(\tilde{U},\tilde{V},E') \leftarrow X \leftarrow (Y,E)$ is a Markov chain;[18] that is, we can set $U := \tilde{U}$ and $V := (\tilde{V},E')$, in which case the left-hand side and right-hand side of Eq. (17) become identical. Inequality (18) follows similarly from the fact that $(\tilde{U},\tilde{V},Y)\leftarrow X' \leftarrow (Y',E')$ is a Markov chain; i.e., we can set $U' := \tilde{U}$ and $V' := (\tilde{V},Y)$ to obtain equality.

[18] Let σ_{ABZ} be a tripartite quantum state of the form $\sigma_{ABZ} = \Sigma_z P_Z(z) \sigma_{AB}^z \otimes P_{|z\rangle}$, where $\{|z\rangle\}$ is a family of orthonormal vectors. We say that $A \leftarrow Z \leftarrow B$ is a Markov chain if $\sigma_{ABZ} = \Sigma_z P_Z(z) \sigma_A^z \otimes \sigma_B^z \otimes P_{|z\rangle}$; i.e., the state in the subsystem A is fully determined by the classical value z.

To prove property (ii), that is, for any $X' \leftarrow X$,

$$\sup_{(U',V')\leftarrow X'} S(U'|V'E) - S(U'|V'Y) \leqslant \sup_{(U,V)\leftarrow X} S(U|VE) - S(U|VY),$$

it suffices to show that if $(U',V') \leftarrow X' \leftarrow (X,Y,E)$ is a Markov chain, then $(U',V') \leftarrow X \leftarrow (Y,E)$ is a Markov chain. This is true since $X' \leftarrow X \leftarrow (Y,E)$ is a Markov chain.

For property (iii), we need to show that, for any $Y' \leftarrow Y$,

$$\sup_{(U,V)\leftarrow X} S(U|VE) - S(U|VY') \leqslant \sup_{(U,V)\leftarrow X} S(U|VE) - S(U|VY).$$

This is, however, a direct consequence of the strong subadditivity, implying that

$$S(U|VY') \geqslant S(U|VY'Y) = S(U|VY),$$

where the equality is a consequence of the fact that $Y' \leftarrow Y \leftarrow (U,V)$ is a Markov chain.

To prove property (iv), i.e., for $C \leftarrow X$,

$$\sup_{(U',V')\leftarrow(X,C)} S(U'|V'EC) - S(U'|V'YC) \leqslant \sup_{(U,V)\leftarrow X} S(U|VE) - S(U|VY),$$

note that $(U',V',C) \leftarrow X \leftarrow (Y,E)$ is a Markov chain. We can thus set $U := U'$ and $V := (V',C)$, in which case the left-hand side and right-hand side of the above expression become equal.

It remains to be shown that property (v) holds. Let $\sigma_U^x := P_{|x\rangle}$ and let σ_V^x be an arbitrary state independent of x. Then, from lemma B.2,

$$M(\mathbf{S}_A;\mathbf{S}_B;E) \geqslant S(\mathbf{S}_A|E) - S(\mathbf{S}_A|\mathbf{S}_B) \geqslant S(\mathbf{S}_A) - \sqrt{2\varepsilon}\ell - 1/e - S(\mathbf{S}_A|\mathbf{S}_B),$$

where $M(\mathbf{S}_A;\mathbf{S}_B;E) := M(\rho_{\mathbf{S}_A\mathbf{S}_B E})$. The assertion then follows from the fact that

$$I(\mathbf{S}_A;\mathbf{S}_B) \geqslant [1 - \varepsilon - 2h(\varepsilon)]\ell.$$

V. EXAMPLES: THE SIX-STATE, BB84, AND B92 PROTOCOLS

To compute expression (9) for the secret-key rate, we have to optimize over the choices of the channels $U \leftarrow X$ and $V \leftarrow U$ used for the classical processing. Clearly, every choice of these channels gives a lower bound on the rate. Surprisingly, for the QKD protocols considered below, a good choice is to define U as a noisy version of X, while V is set to a constant; i.e., it can be discarded. For the protocol, this means that, before doing error correction, Alice should simply add some noise to her measurement data. Intuitively, this puts Bob into a better position than Eve, since the effect of this noise on the correlation between Alice and Eve is worse than on those between Alice and Bob.

A. Six-state protocol

The six-state protocol [5] uses three different encodings, defined by the z basis $\{|0\rangle_z, |1\rangle_z\}$, the x basis $\{|0\rangle_x, |1\rangle_x\}$

$:=\{1/\sqrt{2}(|0\rangle_z \pm |1\rangle_z)\}$, and the y basis $\{|0\rangle_y, |1\rangle_y\}$ $:=\{1/\sqrt{2}(|0\rangle_z \pm i|1\rangle_z)\}$. Alice and Bob measure the QBER for each of these encodings. This gives three conditions on the diagonal entries $\lambda_1, \ldots, \lambda_4$ (with respect to the Bell basis) of the states σ_{AB} contained in the set Γ_{QBER} over which we have to minimize [see Eq. (9)]. In particular, if the QBER equals Q for all encodings, we get $\lambda_3+\lambda_4=Q$, $\lambda_2+\lambda_4=Q$, and $\lambda_2+\lambda_3=Q$. Together with the normalization, we immediately find $\lambda_1=1-\frac{3}{2}Q$ and $\lambda_2=\lambda_3=\lambda_4=\frac{1}{2}Q$.

In order to evaluate the entropies occurring in expression (9), we need to consider a purification $|\psi\rangle_{ABE}$ of the diagonalization $\mathcal{D}_2(\sigma_{AB})$ of σ_{AB}—i.e.,

$$|\psi\rangle_{ABE} := \sum_{i=1}^{4} \sqrt{\lambda_i}|\Phi_i\rangle_{AB} \otimes |\nu_i\rangle_E,$$

where $|\Phi_1\rangle_{AB}, \ldots, |\Phi_4\rangle_{AB}$ denote the Bell states in Alice and Bob's joint system (with respect to the z basis[19]) and where $|\nu_1\rangle_E, \ldots, |\nu_4\rangle_E$ are some mutually orthogonal states in Eve's system. It is easy to verify that, if Alice and Bob apply their measurements (with respect to the z basis), resulting in outcomes x and y, respectively, the state of Eve's system is given by $|\theta^{x,y}\rangle$, where

$$|\theta^{0,0}\rangle = \frac{1}{\sqrt{2}}(\sqrt{\lambda_1}|\nu_1\rangle_E + \sqrt{\lambda_2}|\nu_2\rangle_E),$$

$$|\theta^{1,1}\rangle = \frac{1}{\sqrt{2}}(\sqrt{\lambda_1}|\nu_1\rangle_E - \sqrt{\lambda_2}|\nu_2\rangle_E),$$

$$|\theta^{0,1}\rangle = \frac{1}{\sqrt{2}}(\sqrt{\lambda_3}|\nu_3\rangle_E + \sqrt{\lambda_4}|\nu_4\rangle_E),$$

$$|\theta^{1,0}\rangle = \frac{1}{\sqrt{2}}(\sqrt{\lambda_3}|\nu_3\rangle_E - \sqrt{\lambda_4}|\nu_4\rangle_E).$$

In particular, the density operators σ_E^0 and σ_E^1 describing Eve's system, if Alice has the value 0 or 1, respectively, are given by $\sigma_E^0=\frac{1}{2}P_{|\theta^{0,0}\rangle}+\frac{1}{2}P_{|\theta^{0,1}\rangle}$ and $\sigma_E^1=\frac{1}{2}P_{|\theta^{1,0}\rangle}+\frac{1}{2}P_{|\theta^{1,1}\rangle}$. We can write these states with respect to the basis $\{|\nu_0\rangle_E, \ldots, |\nu_3\rangle_E\}$,

$$\sigma_E^x = \begin{pmatrix} \lambda_1 & \pm\sqrt{\lambda_1\lambda_2} & 0 & 0 \\ \pm\sqrt{\lambda_1\lambda_2} & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & \pm\sqrt{\lambda_3\lambda_4} \\ 0 & 0 & \pm\sqrt{\lambda_3\lambda_4} & \lambda_4 \end{pmatrix},$$

where $\pm$ is a plus sign if $x=0$ and a minus sign if $x=1$.

As mentioned above, we define U as a noisy version of X, with bit-flip probability q—i.e., $P_{U|X=0}(1)=P_{U|X=1}(0)=q$. Moreover, V is set to a constant, which means that it can simply be omitted. Using the fact that $S(UE)=H(U)+S(E|U)$ and, similarly, $H(UY)=H(U)+H(Y|U)$, the entropy difference on the right-hand side in the supremum of Eq. (9) is given by

$$S(U|E) - H(U|Y) = S(E|U) - S(E) - (H(Y|U) - H(Y)),$$

with

$$S(E|U) = \frac{1}{2}S((1-q)\sigma_E^0 + q\sigma_E^1) + \frac{1}{2}S(q\sigma_E^0 + (1-q)\sigma_E^1),$$

$$S(E) = S\left(\frac{1}{2}\sigma_E^0 + \frac{1}{2}\sigma_E^1\right).$$

Furthermore, $H(Y)=1$ and

$$H(Y|U) = h[q(1-Q) + (1-q)Q],$$

where h is the binary entropy function.

These expressions can easily be evaluated numerically. For an optimal choice of the parameter q, we get a positive secret-key rate if $Q \leqslant 0.141$. Without the preprocessing, we obtain the known bound $Q \leqslant 0.126$ [13] (see Fig. 1). Remarkably, this bound has already been improved to $Q \leqslant 0.127$ [13] using degenerate quantum codes, which can be interpreted as a certain type of pre-processing.

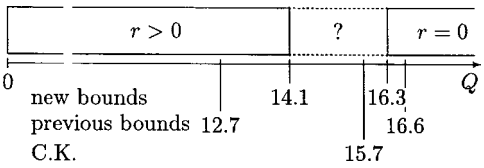

FIG. 1. Lower and upper bounds on the maximally tolerable QBER Q in percent for the six-state protocol. The last line (C.K.) indicates the QBER such that $I(X;Y)=I(X;Z)=I(Y;Z)$ where X and Y is Alice's and Bob's classical information, respectively, and where Z is the classical information that Eve can gain in an individual attack.

Another method to obtain conditions on the set Γ_{QBER} in Eq. (9) is to use some additional symmetrization. For this, we consider the operator $\mathcal{D}_1$ as defined by Eq. (3) with $A_1=V_x$, $A_2=V_y$, $A_3=V_z$ and $B_1=V_x$, $B_2=V_y^\dagger$, $B_3=V_z$, where V_x, V_y, and V_z denote the unitary operators transforming the z basis into the x, y, and z bases, respectively. This implies that $\mathcal{D}_2(\mathcal{D}_1(\sigma_{AB}))=\lambda_1 P_{|\Phi^+\rangle}+\lambda_2 P_{|\Phi^-\rangle}+\lambda_3 P_{|\Psi^+\rangle}+\lambda_4 P_{|\Psi^-\rangle}$, where $\lambda_3+\lambda_4=2\lambda_2$. As explained in [22], we can, instead of $\mathcal{D}_2$, apply another symmetrization operation $\mathcal{D}_2'(\rho)$, e.g.,

$$\mathcal{D}_2'(\rho) = \sum_l O_l' \otimes O_l' \rho (O_l')^\dagger \otimes (O_l')^\dagger,$$

where $O_l' \in \{UV : U \in \{\mathbb{1}, \sigma_z, \text{diag}(-i,1), \text{diag}(i,1)\}$ and $V \in \{\mathbb{1}, \sigma_x\}\}$. Apart from depolarizing any state to a Bell-diagonal state, this map also equalizes the coefficients λ_3 and λ_4 in Eq. (10). This implies that $\mathcal{D}_2'(\mathcal{D}_1(\Gamma_{\text{QBER}}) = \{(1-3Q/2)P_{|\Phi^+\rangle}+Q/2(P_{|\Phi^-\rangle}+P_{|\Psi^+\rangle}+P_{|\Psi^-\rangle})\}$. Thus, using this method, we find right away all the necessary conditions on the set Γ_{QBER}.

Finally, we can use Eq. (15) to compute an upper bound on the secret-key rate of the one-way six-state protocol. Let again $|\theta^{0,0}\rangle$ and $|\theta^{1,1}\rangle$ be the states of Eve conditioned on the

[19] We assume here that the encoding with respect to the z basis is chosen with probability almost 1 (see also the discussion in Sec. III and [27]) such that the number of bit pairs discarded in the sifting step is negligible.

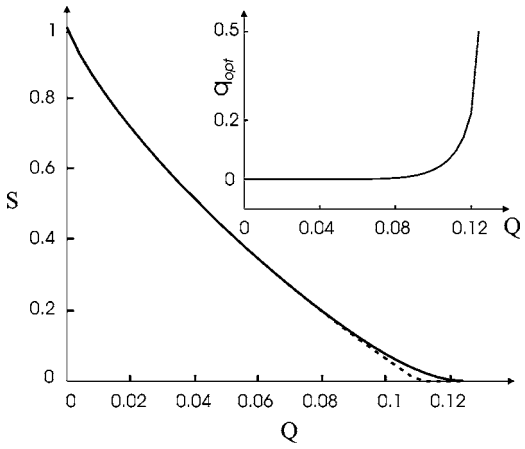

FIG. 2. Lower bound on the secret-key rate of the BB84 protocol as a function of the QBER Q. The dashed line represents the known result [4], whereas the solid line shows our new lower bound. The inset shows the optimal value q_{opt} for the probability by which Alice has to flip her bits in the preprocessing phase.

event that Alice and Bob have the values (0, 0) and (1, 1), respectively. If the adversary applies a von Neumann measurement with respect to projectors along $(1/\sqrt{2})(|\theta^{0,0}\rangle + |\theta^{1,1}\rangle)$ and $(1/\sqrt{2})(|\theta^{0,0}\rangle - |\theta^{1,1}\rangle)$, resulting in Z, we get $r(X,Y,Z)=0$ whenever $Q \geqslant 0.163$.

B. BB84 protocol

The BB84 protocol [4] is very similar to the six-state protocol, but uses only two of the three bases for the encoding. Hence, one only gets two conditions on the diagonal entries $\lambda_1,\ldots,\lambda_4$ (with respect to the Bell basis) of the density operator σ_{AB}: namely, $\lambda_3+\lambda_4=Q$ and $\lambda_2+\lambda_4=Q$. Hence, the set Γ_{QBER} contains all states with diagonal entries $\lambda_1=1-2Q+\lambda_4$ and $\lambda_2=\lambda_3=Q-\lambda_4$, for any $\lambda_4 \in [0,Q]$.

The evaluation of Eq. (9) now follows the same lines as described above for the six-state protocol. A straightforward calculation shows that, independently of the amount of noise added in the preprocessing, expression (9) takes its minimum for $\lambda_4=Q^2$. When optimizing over the preprocessing (i.e., the amount of noise introduced by Alice) we get a positive rate if $Q \leqslant 0.124$ (see Fig. 2). Note that, without the preprocessing, we obtain $Q \leqslant 0.110$, which is exactly the bound due to Shor and Preskill [12]. Computing the upper bound (15) reproduces the known result saying that the (one-way) secret-key rate is zero if $Q \geqslant 0.146$ [30].

C. B92 protocol

In contrast to the BB84 and six-state protocols, Alice uses two nonorthogonal states [6] $|\varphi^0\rangle=\alpha|0\rangle+\beta|1\rangle$ and $|\varphi^1\rangle=\alpha|0\rangle-\beta|1\rangle$ to encode her bit values 0 and 1, respectively, where α and β are (without loss of generality) real coefficients with $\alpha^2+\beta^2=1$. Bob randomly applies a measurement with respect to the basis $\{|\varphi^0\rangle,|\varphi^0\rangle^\perp\}$ or $\{|\varphi^1\rangle,|\varphi^1\rangle^\perp\}$, where $|\varphi^x\rangle^\perp$ denotes the normalized vector orthogonal to $|\varphi^x\rangle$, for $x=0,1$. He then assigns the bit values 0 and 1 to the measurement outcomes $|\varphi^1\rangle^\perp$ and $|\varphi^0\rangle^\perp$, respectively. In the sifting step, Alice and Bob discard all bit pairs where Bob measured $|\varphi_0\rangle$ or $|\varphi_1\rangle$.

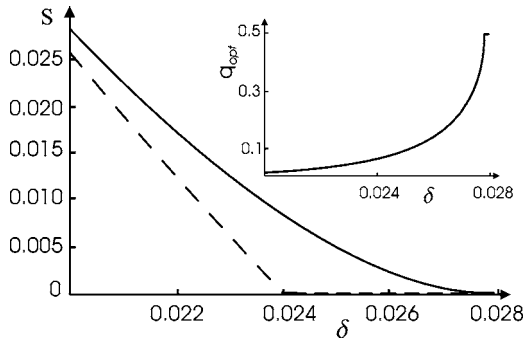

FIG. 3. Lower bound on the secret-key rate of the B92 protocol, for $\alpha=0.38$ (see text for an explanation of the parameter δ). The dashed line represents the known result without preprocessing [20], whereas the solid line is our new lower bound on the rate when Alice additionally adds noise q_{opt} to her measurement data.

In order to evaluate expression (9), we will rely on some of the calculations presented in [20]. Note that, in contrast to the BB84 or six-state protocol, the sifting only depends on the measurement outcomes of Bob. Therefore, we consider the operation $\mathcal{D}_1$ [see Eq. (3)] defined by $\mathcal{D}_1(\bar{\sigma}_{AB}) := A \otimes B \bar{\sigma}_{AB} A^\dagger \otimes B^\dagger$, where $A := |0\rangle\langle\varphi^0| + |1\rangle\langle\varphi^1|$ and $B := |0\rangle\langle\varphi^1|^\perp + |1\rangle\langle\varphi^0|^\perp$. We then need to minimize over the set Γ_{QBER} containing all states σ_{AB} which are compatible with the QBER and, in addition, can result from $\mathcal{D}_1$ applied to any two-qubit density operator $\bar{\sigma}_{AB}$ which corresponds to a collective attack of Eve—i.e., $\bar{\sigma}_{AB}=\mathrm{tr}_E(\mathbb{1}_A \otimes U_{BE} P_{|\Phi^+\rangle_{AB}} \otimes |0\rangle_E \mathbb{1}_A \otimes U_{BE}^\dagger)$ for some unitary operation U_{BE}. In [20], explicit conditions on the diagonal entries (with respect to the Bell basis) of these states have been computed. In particular, the first two diagonal entries are $\lambda_1=(1-Q)(1+s)/2$ and $\lambda_2=(1-Q)(1-s)/2$ where s is the scalar product between the states of the adversary, conditioned on the event that Alice and Bob have the values (0, 0) and (1, 1), respectively. This characterization is already sufficient to obtain reasonable lower bounds on the rate (9).

Similarly to the previous examples, adding noise on Alice's side turns out to be useful. The results of our computations are summarized in Fig. 3, parametrized by the noise δ of a corresponding depolarizing channel $\rho \mapsto (1-2\delta)\rho+\delta\mathbb{1}$.[20] The rate is positive as long as $\delta \leqslant 0.0278$ (compared to $\delta \lessapprox 0.0240$ without noise [15,20]). Within the region shown in the figure, the relation between the parameter δ and the QBER is $Q \approx 2\delta$.[21]

VI. CONCLUSIONS AND OPEN PROBLEMS

We have analyzed a general class of QKD protocols with one-way classical post-processing, thereby using a technique

[20]For any given value of the QBER, the value δ is defined as the parameter of a depolarizing channel $\rho \mapsto (1-2\delta)\rho+\delta\mathbb{1}$ which produces the same QBER when employing the protocol.

[21]In general we have $Q=\delta/(\gamma^2(1-2\delta)+2\delta)$, where $\gamma^2=4\alpha^2(1-\alpha^2)$.

which is not based on entanglement purification. We have shown that, in order to guarantee security against the most general attacks, it is sufficient to consider collective attacks. Moreover, we have derived a new general lower bound on the secret-key rate [formula (9)] which is very similar to the well-known expression for the classical one-way secret-key rate due to Csiszár and Körner [29]. While the latter applies if the information of the adversary is purely classical (i.e., if she is restricted to individual attacks), expression (9) can be seen as a quantum version of it.

In order to evaluate Eq. (9), one only needs to optimize over a certain set of two-qubit density operators, which is characterized by the possible collective attacks on the specific protocol. We have illustrated this for some of the most popular QKD schemes: namely the BB84, the six-state, and the B92-protocols, with one-way classical post-processing, say, from Alice to Bob. Surprisingly, our results imply that the performance of these protocols can be increased if Alice introduces noise to her measurement data. In particular, we get new lower bounds on the maximum tolerated channel noise which are between 10% and 15% larger than the previously known ones.

While our method allows one to exactly analyze the security of a general class of QKD protocols with one-way post-processing, it is still an open problem to identify the protocols which achieve the maximum rate. In particular, we do not know whether a bitwise preprocessing is optimal or whether it might be more advantageous for Alice and Bob to process larger blocks. Note, however, that the upper bound (13) on the secret-key rate of one-way protocols essentially has the same form as the lower bound (9), but involves a maximization over certain quantum states instead of only classical random variables. The question of whether bitwise preprocessing is optimal thus reduces to the problem of proving that these two expressions are equal.

ACKNOWLEDGMENTS

We would like to thank Robert König and Valerio Scarani for many helpful comments. This work has been supported by the Swiss NCCR, "Quantum Photonics," and the European IST project SECOQC.

APPENDIX A: SMOOTH RÉNYI ENTROPY

1. Basic properties

Smooth Rényi entropy has been introduced in [23] in order to characterize fundamental properties of classical random variables. For instance, the ε-smooth Rényi entropy of order 0 of a random variable X conditioned on Y, denoted $H_0^\varepsilon(X|Y)$, measures the minimum length of an encoding C of X such that X can be reconstructed from C and Y, except with probability roughly ε. Similarly, the ε-smooth Rényi entropy of order 2, denoted $H_2^\varepsilon(X|Y)$, quantifies the amount of uniform randomness independent of Y that can be extracted from X (with probability roughly $1-\varepsilon$).

The formal definition of smooth Rényi entropy H_α^ε (of order α, for $\alpha \in [0,\infty]$) looks very similar to the definition of (conventional) Rényi entropy H_α [31]. Indeed, the ε-smooth Rényi entropy $H_\alpha^\varepsilon(X)$ of a random variable X with distribution P_X can be seen as the minimum (if $\alpha<1$) or maximum (if $\alpha>1$) Rényi entropy $H_\alpha(Q)$ of any probability distribution Q which is ε close to P_X. Here, the distance between P_X and Q is measured with respect to the *statistical distance* (also called *variational distance*) $\bar{\delta}(\cdot,\cdot)$, which is the classical analog of the trace distance $\delta(\cdot,\cdot)$.[22]

Definition A.1. The ε-smooth Rényi entropy (of order $\alpha \in [0,\infty]$) of a probability distribution P is[23]

$$H_\alpha^\varepsilon(P) := \frac{1}{1-\alpha} \inf_Q \log_2\left(\sum_{z\in\mathcal{Z}} Q(z)^\alpha\right),$$

where the infimum ranges over all probability distributions Q such that $\bar{\delta}(P,Q)\leq\varepsilon$. For a random variable X with probability distribution P_X, we also write $H_\alpha^\varepsilon(X)$ instead of $H_\alpha^\varepsilon(P_X)$.

In particular, for $\varepsilon=0$, the smooth Rényi entropy is equal to the conventional Rényi entropy. Similarly to the above definition, the conditional Rényi entropy $H_\alpha^\varepsilon(X|Y)$ is defined by taking the maximum (if $\alpha<1$) or minimum (if $\alpha>1$) value of the smooth Rényi entropy of the probability distributions $P_{X|Y=y}$, for any possible value of y.

In [19], the notion of smooth Rényi entropy has been generalized to quantum states. For a density operator ρ, we denote by $S_\alpha^\varepsilon(\rho)$ the ε-smooth Rényi entropy of order α of ρ. Similar to the von Neumann entropy, $S_\alpha^\varepsilon(\rho)$ is defined as the (classical) smooth Rényi entropy $H_\alpha^\varepsilon(P)$ of the probability distribution P defined by the eigenvalues of ρ. We also write $S_\alpha^\varepsilon(UV)$ instead of $S_\alpha^\varepsilon(\rho_{UV})$ and, similarly, $S_\alpha^\varepsilon(U)$ instead of $S_\alpha^\varepsilon(\rho_U)$, where ρ_U is the partial state $\rho_U := \mathrm{tr}_V(\rho_{UV})$.

We start reviewing some basic properties of smooth Rényi entropy of quantum states. The proofs can be found in [23,19]. Most of these properties are very analogous to the properties of the von Neumann entropy $S(\cdot)$. For instance, if ρ_{UV} is a state on $\mathcal{H}_U\otimes\mathcal{H}_V$, then the difference between $S_\alpha^\varepsilon(UV)$ and $S_\alpha^\varepsilon(U)$ is bounded by the entropy of V, which corresponds to the well-known fact that $S(U)-S(V)\leq S(UV)\leq S(U)+S(V)$: For $\alpha=2$, we have

$$S_2^\varepsilon(UV) \leq S_2^{\varepsilon+\varepsilon'}(U) + S_0^{\varepsilon'}(V), \tag{A1}$$

$$S_2^{\varepsilon+\varepsilon'}(UV) \geq S_2^\varepsilon(U) - S_0^{\varepsilon'}(V), \tag{A2}$$

and, similarly, for $\alpha=0$,

$$S_0^{\varepsilon+\varepsilon'}(UV) \leq S_0^\varepsilon(U) + S_0^{\varepsilon'}(V), \tag{A3}$$

[22] Let σ and σ' be two density operators which are diagonal with respect to the same basis and let P and P' be the probability distributions defined by the eigenvalues of σ and σ', respectively. Then $\bar{\delta}(P,P')=\delta(\sigma,\sigma')$.

[23] If $\alpha=0$ or $\alpha=\infty$, $H_\alpha^\varepsilon(P)$ is defined by the continuous extension, $H_\alpha^\varepsilon(P) := \lim_{\beta\to\alpha} H_\beta^\varepsilon(P)$. For $\alpha=1$, set $H_1^\varepsilon(P) := H(P)$.

$$S_0^{\varepsilon}(UV) \geq S_0^{\varepsilon+\varepsilon'}(U) - S_0^{\varepsilon'}(V). \tag{A4}$$

Consider now a bipartite state ρ_{UZ} on $\mathcal{H}_U \otimes \mathcal{H}_Z$ where the second part is purely classical, i.e.,

$$\rho_{UZ} = \sum_z P_Z(z) \rho_U^z \otimes P_{|z\rangle},$$

for some probability distribution P_Z and a family of orthonormal vectors $\{|z\rangle\}_z$ on $\mathcal{H}_Z$. Then, the smooth Rényi entropy cannot increase when conditioning on Z, that is,

$$S_\alpha^{\varepsilon}(U|Z) \leq S_\alpha^{\varepsilon}(U), \tag{A5}$$

for $\alpha=0$ and $\alpha=2$. The following inequalities can be interpreted as extensions of the chain rule $S(UZ)=S(U|Z)+S(Z)$ to smooth Rényi entropy:

$$S_2^{\varepsilon}(U|Z) \leq S_2^{\varepsilon+\varepsilon'}(UZ) - H_2^{\varepsilon'}(Z), \tag{A6}$$

$$S_2^{\varepsilon+\varepsilon'+\varepsilon''}(U|Z) > S_2^{\varepsilon'}(UZ) - H_0^{\varepsilon''}(Z) - 2\log_2(1/\varepsilon), \tag{A7}$$

$$S_0^{\varepsilon}(U|Z) \geq S_0^{\varepsilon+\varepsilon'}(UZ) - H_0^{\varepsilon'}(Z), \tag{A8}$$

$$S_0^{\varepsilon+\varepsilon'+\varepsilon''}(U|Z) < S_0^{\varepsilon'}(UZ) - H_2^{\varepsilon''}(Z) + 2\log_2(1/\varepsilon). \tag{A9}$$

More generally, let ρ_{UZV} be a density operator on $\mathcal{H}_U \otimes \mathcal{H}_Z \otimes \mathcal{H}_V$ such that the states on $\mathcal{H}_U$ and $\mathcal{H}_V$ only depend on the classical subsystem $\mathcal{H}_Z$; i.e., there exist density operators ρ_U^z and ρ_V^z on $\mathcal{H}_U$ and $\mathcal{H}_V$, respectively, such that

$$\rho_{UVZ} = \sum_{z\in\mathcal{Z}} P_Z(z) \rho_U^z \otimes \rho_V^z \otimes P_{|z\rangle},$$

where P_Z is a probability distribution and $\{|z\rangle\}_{z\in\mathcal{Z}}$ a family of orthonormal vectors on $\mathcal{H}_Z$. Then

$$S_2^{\varepsilon+\varepsilon'}(UVZ) \geq S_2^{\varepsilon}(U|Z) + S_2^{\varepsilon'}(VZ), \tag{A10}$$

$$S_0^{\varepsilon+\varepsilon'}(UVZ) \leq S_0^{\varepsilon}(U|Z) + S_0^{\varepsilon'}(VZ). \tag{A11}$$

The following identities are useful to determine the conditional smooth Rényi entropy $S_\alpha^{\varepsilon}(U|Z)$ if the smooth Rényi entropy $S_\alpha^{\varepsilon}(U|Z=z)$, conditioned on certain values z, is known. For any $z\in\mathcal{Z}$, let $\varepsilon_z := \varepsilon \cdot P_Z(z)$. Then

$$S_2^{\varepsilon_z}(U|Z) \leq S_2^{\varepsilon}(U|Z=z), \tag{A12}$$

$$S_0^{\varepsilon_z}(U|Z) \geq S_0^{\varepsilon}(U|Z=z). \tag{A13}$$

Additionally, for any set $\bar{\mathcal{Z}} \subset \mathcal{Z}$ such that $\Pr_z[z\in\bar{\mathcal{Z}}] \geq 1-\varepsilon$,

$$S_2^{\varepsilon+\varepsilon'}(U|Z) \geq \min_{z\in\bar{\mathcal{Z}}} S_2^{\varepsilon'}(U|Z=z), \tag{A14}$$

$$S_0^{\varepsilon+\varepsilon'}(U|Z) \leq \max_{z\in\bar{\mathcal{Z}}} S_0^{\varepsilon'}(U|Z=z). \tag{A15}$$

Similarly to the von Neumann entropy, the smooth Rényi entropy can only increase when applying a unital quantum operation $\mathcal{E}$,[24] that is,

$$S_\alpha^{\varepsilon}(\mathcal{E}(\rho_U)) \geq S_\alpha^{\varepsilon}(\rho_U) \tag{A16}$$

for any $\alpha\in\mathbb{R}^+$ and $\varepsilon\in\mathbb{R}^+$.

The smooth Rényi entropies of order α are related for different values of α. In particular, we have

$$S_2^{\varepsilon}(U) \lessapprox S_0^{\varepsilon}(U), \tag{A17}$$

where the approximation holds up to $O(\varepsilon)$. Finally, the smooth Rényi entropy of an n-fold product state $\rho^{\otimes n}$ approaches the von Neumann entropy. Formally, for any $\alpha \in \mathbb{R}^+$ and $\varepsilon\in\mathbb{R}^+$,

$$|S_\alpha^{\varepsilon}(\rho^{\otimes n}) - nS(\rho)| \leq O(\ln(1/\varepsilon)). \tag{A18}$$

2. Smooth Rényi entropy and measurements

Let $\mathcal{E}$ be a measurement defined by a family of operators $\{E_z\}_{z\in\mathcal{Z}}$. Let $\rho_{\tilde{U}} := \mathcal{E}(\rho_U) = \Sigma_z E_z \rho_U E_z^\dagger$ be the state of the quantum system after applying $\mathcal{E}$ to a density operator ρ_U, and let Z be the classical measurement outcome—i.e., $P_Z(z) := \mathrm{tr}(E_z \rho_U E_z^\dagger)$, for $z\in\mathcal{Z}$. We have seen in the previous section [see (A16)] that the entropy $S_\alpha^{\varepsilon}(\tilde{U})$ of $\rho_{\tilde{U}}$ can only be larger than the entropy $S_\alpha^{\varepsilon}(U)$ of ρ_U if $\mathcal{E}$ is unital. The following lemma states that the maximum increase of the smooth Rényi entropy when applying $\mathcal{E}$ is bounded by the entropy $H_0^{\varepsilon}(Z)$ of the classical measurement outcome Z.

Lemma A.2. Let $\rho_{\tilde{U}}$ be the state obtained when applying the trace-preserving measurement $\mathcal{E}$ to ρ_U and let Z be the classical outcome. Then, for $\varepsilon, \varepsilon' \in \mathbb{R}^+$,

$$S_2^{\varepsilon}(\tilde{U}) \leq S_2^{\varepsilon+\varepsilon'}(U) + H_0^{\varepsilon'}(Z), \tag{A19}$$

$$S_0^{\varepsilon+\varepsilon'}(\tilde{U}) \leq S_0^{\varepsilon}(U) + H_0^{\varepsilon'}(Z). \tag{A20}$$

Proof. Let T be the linear operation from $\mathcal{H}_U$ to $\mathcal{H}_{\tilde{U}} \otimes \mathcal{H}_Z$ defined by

$$T: |\varphi\rangle \mapsto \sum_{z\in\mathcal{Z}} (E_z|\varphi\rangle) \otimes |z\rangle,$$

for any $|\varphi\rangle \in \mathcal{H}_U$, where $\{|z\rangle\}_z$ is a family of orthonormal vectors in $\mathcal{H}_Z$. Let $\rho'_{\tilde{U}Z} := T\rho_U T^\dagger$. It is easy to verify that $\rho_{\tilde{U}} = \mathrm{tr}_Z(\rho'_{\tilde{U}Z})$, and that the eigenvalues of ρ'_Z correspond to the probabilities $P_Z(z)$. Hence, since the smooth Rényi entropy of quantum states is defined by the classical smooth Rényi entropy of its eigenvalues, we have $S_\alpha^{\varepsilon'}(\rho'_Z) = H_\alpha^{\varepsilon'}(Z)$. Moreover, because $\mathcal{E}$ is trace preserving—i.e., $\Sigma_{z\in\mathcal{Z}} E_z^\dagger E_z = \mathbb{1}_U$—we have $T^\dagger T = \mathbb{1}_U$. Consequently, $\rho'_{\tilde{U}Z}$ has the same eigenvalues as ρ_U—i.e., $S_\alpha^{\varepsilon}(\rho'_{\tilde{U}Z}) = S_\alpha^{\varepsilon}(\rho_U)$. Hence, using Eq. (A2), we find

[24] A quantum operation $\mathcal{E}$ is *unital* if $\mathcal{E}$ is trace preserving and if the fully mixed state is a fixed point of $\mathcal{E}$. Formally, if $\rho \mapsto \Sigma_z E_z \rho E_z^\dagger$ is the operator-sum representation of $\mathcal{E}$, then $\Sigma_z E_z^\dagger E_z = E_z E_z^\dagger = \mathbb{1}$.

$$S_2^{\varepsilon}(\rho_{\tilde{U}}) = S_2^{\varepsilon}(\mathrm{tr}_Z(\rho'_{\tilde{U}Z})) \leqslant S_2^{\varepsilon+\varepsilon'}(\rho'_{\tilde{U}Z}) + S_0^{\varepsilon'}(\rho'_Z) = S_2^{\varepsilon+\varepsilon'}(\rho_U) + H_0^{\varepsilon'}(Z),$$

which concludes the proof of (A19). Inequality (A20) follows by the same argument, where (A2) is replaced by (A4). □

A similar relation holds between the smooth Rényi entropy $S_\alpha^\varepsilon(U)$ of the original quantum state ρ_U and the entropy $S_\alpha^\varepsilon(\tilde{U}|Z)$ of the state $\rho_{\tilde{U}}$ after the measurement, conditioned on the classical outcome Z. Lemma A.3 below states that the difference between these entropies is roughly bounded by the entropy $H_0^\varepsilon(Z)$ of Z.

Lemma A.3. Let $\rho_{\tilde{U}}$ be the state obtained when applying a von Neumann measurement $\mathcal{E}$ to a state ρ_U. Let $S_\alpha^\varepsilon(\tilde{U}|Z)$ be the entropy of $\rho_{\tilde{U}}$, conditioned on the classical outcome Z. Then, for $\varepsilon, \varepsilon', \varepsilon'' \in \mathbb{R}^+$,

$$S_2^{\varepsilon+\varepsilon'}(U) \geqslant S_2^{\varepsilon}(\tilde{U}|Z) - H_0^{\varepsilon'}(Z), \tag{A21}$$

$$S_2^{\varepsilon}(U) < S_2^{\varepsilon+\varepsilon'+\varepsilon''}(\tilde{U}|Z) + H_0^{\varepsilon'}(Z) + 2\log_2(1/\varepsilon'') \tag{A22}$$

and

$$S_0^{\varepsilon+\varepsilon'}(U) \leqslant S_0^{\varepsilon}(\tilde{U}|Z) + H_0^{\varepsilon'}(Z), \tag{A23}$$

$$S_0^{\varepsilon}(U) \geqslant S_0^{\varepsilon+\varepsilon'}(\tilde{U}|Z) - H_0^{\varepsilon'}(Z). \tag{A24}$$

Proof. Let E_z be the projectors defined by the measurement $\mathcal{E}$ and let $\rho_{\tilde{U}Z}$ be the state as defined in the proof of lemma A.2. Since, by assumption, the ranges of the operators E_z, for $z \in \mathcal{Z}$, are mutually orthogonal, the states $\rho_{\tilde{U}Z}$ and $\rho_{\tilde{U}}$ have the same eigenvalues and thus $S_\alpha^{\bar{\varepsilon}}(\tilde{U}Z) = S_\alpha^{\bar{\varepsilon}}(\tilde{U})$. Using this identity, (A21) follows from (A19) and (A5),

$$S_2^{\varepsilon+\varepsilon'}(U) \geqslant S_2^{\varepsilon}(\tilde{U}) - H_0^{\varepsilon'}(Z) \geqslant S_2^{\varepsilon}(\tilde{U}|Z) - H_0^{\varepsilon'}(Z).$$

Similarly, (A22) follows from (A16) and (A7),

$$S_2^{\varepsilon}(U) \leqslant S_2^{\varepsilon}(\tilde{U}) = S_2^{\varepsilon}(\tilde{U}Z) < S_2^{\varepsilon+\varepsilon'+\varepsilon''}(\tilde{U}|Z) + H_0^{\varepsilon'}(Z) + 2\log_2(1/\varepsilon'').$$

To prove (A23), we use (A16) and (A8),

$$S_0^{\varepsilon+\varepsilon'}(U) \leqslant S_0^{\varepsilon+\varepsilon'}(\tilde{U}) = S_0^{\varepsilon+\varepsilon'}(\tilde{U}Z) \leqslant S_0^{\varepsilon}(\tilde{U}|Z) + H_0^{\varepsilon'}(Z).$$

Finally, (A24) follows from (A20) and (A5),

$$S_0^{\varepsilon}(U) \geqslant S_0^{\varepsilon+\varepsilon'}(\tilde{U}) - H_0^{\varepsilon'}(Z) \geqslant S_0^{\varepsilon+\varepsilon'}(\tilde{U}|Z) - H_0^{\varepsilon'}(Z).$$

□

3. Smooth Rényi entropy of symmetric states

The goal of this section is to derive an expression for the smooth Rényi entropies of a symmetric state over n subsystems in terms of the von Neumann entropy of a corresponding state over only *one* subsystem.

Let $\sigma_1, \ldots, \sigma_d$ be density operators on $\mathcal{H}_U$ and let $\rho_{\mathbf{U}}^n$ be the symmetric state over $\mathcal{H}_U^{\otimes n}$ defined by

$$\rho_{\mathbf{U}}^n := \mathcal{P}_n\left(\sum_{\mathbf{n}\in\Gamma_d^n} \mu_{\mathbf{n}} \sigma_1^{\otimes n_1} \otimes \cdots \otimes \sigma_d^{\otimes n_d}\right), \tag{A25}$$

where, for any $\mathbf{n} \in \Gamma_d^n := \{(n_1, \ldots, n_d) : \Sigma_i n_i = n\}$, $\mu_{\mathbf{n}}$ are nonnegative coefficients such that $\Sigma_{\mathbf{n}} \mu_{\mathbf{n}} = 1$.

Similarly, for any d-tuple $\lambda = (\lambda_1, \ldots \lambda_d)$ over $\mathbb{R}^+$, let $\sigma_U[\lambda]$ be the density operator on $\mathcal{H}_U$ defined by

$$\sigma_U[\lambda] := \sum_i \lambda_i \sigma_i. \tag{A26}$$

Let $\mathcal{E}$ be a quantum operation from $\mathcal{H}_U$ to $\mathcal{H}_V$. The following lemma gives a relation between the smooth Rényi entropy of the symmetric state obtained by applying $\mathcal{E}$ to each of the subsystems of a purification of ρ_U^n and the von Neumann entropy of the state obtained by applying $\mathcal{E}$ to a purification of $\sigma_U[\lambda]$.

Lemma A.4. Let $\rho_{\mathbf{UW}}^n$ be a purification of the state $\rho_{\mathbf{U}}^n$ defined by (A25) with coefficients $\mu_{\mathbf{n}}$ and let $\rho_{\mathbf{VW}}^n := (\mathcal{E} \otimes \mathbb{1}_W)^{\otimes n}(\rho_{\mathbf{UW}}^n)$. Similarly, for any d-tuple λ, let $\sigma_{UW}[\lambda]$ be a purification of the state $\sigma_U[\lambda]$ defined by (A26) and let $\sigma_{VW}[\lambda] := (\mathcal{E} \otimes \mathbb{1}_W)(\sigma_{UW}[\lambda])$. Let $\bar{\Gamma}$ be a subset of Γ_d^n such that $\Sigma_{\mathbf{n}\in\bar{\Gamma}} \mu_{\mathbf{n}} \geqslant 1 - \varepsilon/2$. Then

$$S_2^{\varepsilon}(\rho_{\mathbf{VW}}^n) \gtrapprox n \min_{\lambda} S(\sigma_{\mathbf{VW}}[\lambda]),$$

$$S_0^{\varepsilon}(\rho_{\mathbf{VW}}^n) \lessapprox n \max_{\lambda} S(\sigma_{\mathbf{VW}}[\lambda]),$$

where the minimum and maximum are taken over all $\lambda = (\lambda_1, \ldots, \lambda_d)$ such that $n(\lambda_1, \ldots, \lambda_d) \in \bar{\Gamma}$, and where the approximation is up to $O(d\ln(n) + \ln(n/\varepsilon))$.

The proof of lemma A.4 is based on the fact that there exists a measurement on $\sigma_U[\lambda]^{\otimes n}$ such that the resulting state, conditioned on a certain measurement outcome, is equal to the state $\rho_{\mathbf{U}}^n$. The assertion then follows from the observation that this measurement does only change the entropies by a small constant.

We start with the proof of a restricted version of the statement, formulated as lemma A.5 below, which holds for states of the form (A25) where only one of the weights $\mu_{\mathbf{n}}$ is nonzero. Let $|\varphi_1\rangle, \ldots, |\varphi_d\rangle \in \mathcal{H}_U \otimes \mathcal{H}_W$ be purifications of the states $\sigma_1, \ldots, \sigma_d$, respectively, such that the partial traces $\mathrm{tr}_U(P_{|\varphi i\rangle})$ are mutually orthogonal. For $\mathbf{n} = (n_1, \ldots, n_d) \in \Gamma_d^n$, let

$$|\psi\rangle_{\mathbf{UW}}^{\mathbf{n}} := \frac{1}{\sqrt{|S_n|}} \sum_{\pi\in S_n} \pi(|\varphi_1\rangle^{\otimes n_1} \otimes \cdots \otimes |\varphi_d\rangle^{\otimes n_d}), \tag{A27}$$

where S_n denotes the set of all permutations π on n-tuples. Similarly, for $\lambda = (\lambda_1, \ldots, \lambda_d)$, let

$$|\varphi\rangle_{UW}^{\lambda} := \sum_{i=1}^{d} \sqrt{\lambda_i} |\varphi_i\rangle. \tag{A28}$$

Lemma A.5. Let $\rho_{\mathbf{UW}}^n[\mathbf{n}] := P_{|\psi\rangle_{\mathbf{UW}}^{\mathbf{n}}}$ be the pure state defined by (A27), for some fixed $\mathbf{n} = (n_1, \ldots, n_d) \in \Gamma_d^n$, and let $\rho_{\mathbf{VW}}^n[\mathbf{n}] := (\mathcal{E} \otimes \mathbb{1}_W)^{\otimes n}(\rho_{\mathbf{UW}}^n[\mathbf{n}])$. Moreover, for λ

$:=(n_1/n,\ldots,n_d/n)$, let $\sigma_{UW}[\lambda]:=P_{|\varphi\rangle^{\lambda}_{UW}}$ be the pure state defined by (A28) and let $\sigma_{VW}[\lambda]:=(\mathcal{E}\otimes\mathbb{1})(\sigma_{UW}[\lambda])$. Then, for $\alpha\in\{0,2\}$,

$$|S^{\varepsilon}_{\alpha}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])-nS(\sigma_{VW}[\lambda])|\leqslant O(\ln(n/\varepsilon)).$$

Proof. For any $i\in\{1,\ldots,d\}$, let P_i be the projector onto the support of $(\mathcal{E}\otimes\mathbb{1}_W)(P_{|\varphi_i\rangle})$, which, by the definition of the vectors $|\varphi_i\rangle$, are orthogonal for distinct i. Additionally, let $\mathcal{F}:\rho\mapsto F_0\rho F_0^{\dagger}+F_1\rho F_1^{\dagger}$ be the measurement on $\mathcal{H}_V^{\otimes n}$ defined by

$$F_0:=\sum_{\pi\in S_n}\pi(P_1^{\otimes n_1}\otimes\cdots\otimes P_d^{\otimes n_d})$$

and $F_1:=\mathbb{1}-F_0$. We first show that

$$\rho^{n}_{\mathbf{VW}}[\mathbf{n}]=\frac{1}{N}F_0(\sigma_{VW}[\lambda]^{\otimes n})F_0^{\dagger},\tag{A29}$$

where $N:=|S_n|\Pi_{i=1}^{d}\lambda_i^{n_i}$.

Let $(\mathcal{E}\otimes\mathbb{1}_W)(\rho)=\Sigma_{\alpha=1}^{m}\bar{E}_{\alpha}\rho\bar{E}_{\alpha}^{\dagger}$ be the operator-sum representation of $\mathcal{E}\otimes\mathbb{1}_W$. Moreover, for any $\bar{\alpha}:=(\alpha_1,\ldots,\alpha_n)$, let $\bar{E}_{\bar{\alpha}}:=\bar{E}_{\alpha_1}\otimes\cdots\otimes\bar{E}_{\alpha_n}$. The above equality can then be rewritten as

$$\sum_{\bar{\alpha}}\bar{E}_{\bar{\alpha}}(\rho^{n}_{\mathbf{UW}}[\mathbf{n}])\bar{E}_{\bar{\alpha}}^{\dagger}=\frac{1}{N}\sum_{\bar{\alpha}}F_0\bar{E}_{\bar{\alpha}}(\sigma_{UW}[\lambda]^{\otimes n})\bar{E}_{\bar{\alpha}}^{\dagger}F_0^{\dagger}.$$

It suffices to verify that equality holds for any term in the sum, i.e.,

$$\bar{E}_{\bar{\alpha}}|\psi\rangle^{n}_{\mathbf{UW}}=\frac{1}{\sqrt{N}}F_0\bar{E}_{\bar{\alpha}}|\varphi\rangle^{\otimes n}_{\mathbf{UW}},\tag{A30}$$

for any n-tuple $\bar{\alpha}=(\alpha_1,\ldots,\alpha_n)$ on $\{1,\ldots,m\}$. Because of the definition of the projectors P_i, we have $P_i\bar{E}_{\alpha}|\varphi_j\rangle=\bar{E}_{\alpha}|\varphi_j\rangle$, if $i=j$, and $P_i\bar{E}_{\alpha}|\varphi_j\rangle=0$ otherwise. Hence, for any $|\varphi_{i_1,\ldots,i_n}\rangle:=|\varphi_{i_1}\rangle\otimes\cdots\otimes|\varphi_{i_n}\rangle$,

$$F_0\bar{E}_{\bar{\alpha}}|\varphi_{i_1,\ldots,i_n}\rangle=\begin{cases}\bar{E}_{\bar{\alpha}}|\varphi_{i_1,\ldots,i_n}\rangle & \text{if } |\varphi_{i_1,\ldots,i_n}\rangle\in\Theta_{\mathbf{n}},\\ 0 & \text{otherwise,}\end{cases}.$$

where $\Theta_{\mathbf{n}}:=\{\pi(|\varphi_1\rangle^{n_1}\otimes\cdots\otimes|\varphi_d\rangle^{n_d}):\pi\in S_n\}$. This implies (A30) and thus (A29).

Let $\rho^{n}_{\tilde{\mathbf{V}}\tilde{\mathbf{W}}}$ be the state of the system after applying the measurement $\mathcal{F}$ to $\sigma_{VW}[\lambda]^{\otimes n}$, and let Z be the classical measurement outcome. In the following, we write $S^{\varepsilon}_{\alpha}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=0)$ to denote the entropy of $\rho^{n}_{\tilde{\mathbf{V}}\tilde{\mathbf{W}}}$ conditioned on $Z=0$. Then, according to (A29),

$$S^{\varepsilon}_{\alpha}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])=S^{\varepsilon}_{\alpha}(\tilde{\mathbf{V}}\tilde{W}|Z=0).\tag{A31}$$

Let $\varepsilon':=\frac{1}{2}P_Z(0)\varepsilon$ where $P_Z(0)=\mathrm{tr}[F_0(\sigma_{VW}^{\otimes n})F_0^{\dagger}]$. Using (A12) and (A22), we find

$$S^{\varepsilon}_{2}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=0)\geqslant S^{2\varepsilon'}_{2}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z)>S^{\varepsilon'}_{2}(\sigma_{VW}[\lambda]^{\otimes n})-1-2\log_2(1/\varepsilon').$$

Similarly, using (A13) and (A24),

$$S^{\varepsilon}_{0}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=0)\leqslant S^{2\varepsilon'}_{0}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}^{n}|Z)\leqslant S^{2\varepsilon'}_{0}(\sigma_{VW}[\lambda]^{\otimes n})+1.$$

Hence, because the smooth Rényi entropy of order 0 is larger than the smooth Rényi entropy of order 2 [cf. (A17)], we have

$$S^{\varepsilon'}_{2}(\sigma_{VW}^{\otimes n})S^{\varepsilon}_{2}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=0)\lessapprox S^{\varepsilon}_{0}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=0)\lessapprox S^{2\varepsilon'}_{0}(\sigma_{VW}^{\otimes n}),$$

where the approximation holds up to $O(\ln(1/\varepsilon'))$. Combining this with (A31), we conclude

$$S^{\varepsilon'}_{2}(\sigma_{VW}^{\otimes n})S^{\varepsilon}_{\alpha}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])S^{2\varepsilon'}_{0}(\sigma_{VW}^{\otimes n}).$$

The assertion then follows from the observation that $P_Z(0)\geqslant 1/n$, which implies $\varepsilon'\geqslant\varepsilon/2n$, and the fact that the smooth Rényi entropy of product states approaches the von Neumann entropy [see (A18)]. □

Proof of lemma A.4. It is easy to see that it suffices to prove the assertion for one specific purification of the states ρ^{n}_{U} and σ_U. Let thus $|\varphi_1\rangle,\ldots,|\varphi_d\rangle\in\mathcal{H}_U\otimes\mathcal{H}_W$ be the purifications of $\sigma_1,\ldots,\sigma_d$ defined above. Moreover, for any $\mathbf{n}\in\Gamma^{n}_{d}$, let $\rho^{n}_{UW}[\mathbf{n}]:=P_{|\psi\rangle^{\mathbf{n}}_{\mathbf{UW}}}$ be the state defined by (A27) and let $\rho^{n}_{\mathbf{UW}}:=P_{|\psi\rangle}$ where

$$|\psi\rangle:=\sum_{\mathbf{n}\in\Gamma}\sqrt{\mu_{\mathbf{n}}}|\psi\rangle^{\mathbf{n}}_{\mathbf{UV}}.$$

Similarly, for any $\lambda=(\lambda_1,\ldots,\lambda_d)$, let $\sigma_{UW}[\lambda]:=P_{|\varphi\rangle^{\lambda}_{UW}}$ be the state defined by (A28). It follows from these definitions that $\rho^{n}_{\mathbf{UW}}$ is a purification of $\rho^{n}_{\mathbf{U}}$ and, similarly, $\sigma_{UW}[\lambda]$ is a purification of $\rho_U[\lambda]$.

For any $\mathbf{n}\in\Gamma^{n}_{d}$, let $\mathcal{H}_W^{\mathbf{n}}$ be the smallest subspace of $\mathcal{H}_W^{\otimes n}$ containing the support of the traces $\rho^{n}_{\mathbf{W}}[\mathbf{n}]=\mathrm{tr}_{\mathcal{H}_U^{\otimes n}}(\rho^{n}_{\mathbf{UW}}[\mathbf{n}])$. By the definition of the vectors $|\varphi_i\rangle$, the subspaces $\mathcal{H}_W^{\mathbf{n}}$ are orthogonal for distinct $\mathbf{n}\in\Gamma^{n}_{d}$. Hence, there exists a projective measurement $\mathcal{F}$ onto the subspaces $\mathcal{H}_U\otimes\mathcal{H}_W^{\mathbf{n}}$. Consider the state $\rho_{\tilde{\mathbf{V}}\tilde{\mathbf{W}}}$ obtained when applying $\mathcal{F}$ to $\rho^{n}_{\mathbf{VW}}$, and let Z be the classical outcome; i.e., Z takes values from the set Γ^{n}_{d}. The entropy $S^{\varepsilon}_{\alpha}(\tilde{V}^n\tilde{W}^n|Z=\mathbf{n})$ of the state $\rho^{n}_{\tilde{\mathbf{V}}\tilde{\mathbf{W}}}$ after the measurement, conditioned on $Z=\mathbf{n}$, is equal to the entropy of $\rho^{n}_{\mathbf{VW}}[\mathbf{n}]$ as defined by lemma A.5—i.e.,

$$S^{\varepsilon}_{\alpha}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=\mathbf{n})=S^{\varepsilon}_{\alpha}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}]).$$

Hence, from (A21) and (A14),

$$S^{\varepsilon}_{2}(\rho^{n}_{\mathbf{VW}})\geqslant S^{\varepsilon}_{2}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z)-H_0(Z)\geqslant\min_{\mathbf{n}\in\bar{\Gamma}}S^{\varepsilon/2}_{2}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z=\mathbf{n})-H_0(Z)=\min_{\mathbf{n}\in\bar{\Gamma}}S^{\varepsilon/2}_{2}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])-H_0(Z)$$

and, similarly, from (A23) and (A15),

$$S^{\varepsilon}_{0}(\rho^{n}_{\mathbf{VW}})\leqslant S^{\varepsilon}_{0}(\tilde{\mathbf{V}}\tilde{\mathbf{W}}|Z)+H_0(Z)\leqslant\max_{\mathbf{n}\in\bar{\Gamma}}S^{\varepsilon/2}_{0}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])+H_0(Z).$$

Finally, from lemma A.5,

$$|S^{\varepsilon/2}_{\alpha}(\rho^{n}_{\mathbf{VW}}[\mathbf{n}])-nS_{\alpha}(\sigma_{VW}[\lambda])|\leqslant O(\ln(2n/\varepsilon)),$$

where $\lambda=(n_1/n,\ldots,n_d/n)$. The assertion then follows from the observation that $H_0(Z)\leqslant\log_2(|\Gamma^{n}_{d}|)\leqslant d\log_2(n)$. □

APPENDIX B: ENTROPY OF ALMOST PRODUCT STATES

Let X be a classical random variable and let ρ_B^x be a quantum state depending on X. Clearly, if the states ρ_B^x are equal for all x, then the entropy of X does not change when conditioning on the quantum system—i.e., $S(X)=S(X|B)$. In this section, we show that, if the joint state describing X and ρ_B^x is close to a product state, then the entropy change of X when conditioning on the quantum system is still small (cf. lemma B.2).

We first need a lemma relating the trace distance of two density operators to the trace distance of purifications of them.

Lemma B.1. Let ρ and ρ' be density operators and let $|\psi\rangle$ be a purification of ρ. Then there exists a purification $|\psi'\rangle$ of ρ' such that

$$\delta(P_{|\psi\rangle},P_{|\psi'\rangle}) \leqslant \sqrt{2\delta(\rho,\rho')}.$$

Proof. Note that the fidelity F is related to the trace distance δ according to

$$1-F(\rho,\sigma) \leqslant \delta(\rho,\sigma) \leqslant \sqrt{1-F(\rho,\sigma)^2}.$$

Moreover, Uhlmann's theorem states that there exists a purification $|\psi'\rangle$ of ρ' such that

$$F(\rho,\rho') = F(P_{|\psi\rangle},P_{|\psi'\rangle}).$$

Hence,

$$\begin{aligned}\delta(P_{|\psi\rangle},P_{|\psi'\rangle}) &\leqslant \sqrt{1-F(P_{|\psi\rangle},P_{|\psi'\rangle})^2} = \sqrt{1-F(\rho,\rho')^2} \\ &\leqslant \sqrt{2[1-F(\rho,\rho')]} \leqslant \sqrt{2\delta(\rho,\rho')}.\end{aligned}$$

□

Lemma B.2. Let ρ_{XB} be a bipartite density operator of the form

$$\rho_{XB} = \sum_{x=1}^{d} \mu_x P_{|x\rangle} \otimes \rho_B^x,$$

where $\{|x\rangle\}_{x\in\{1,\ldots,d\}}$ is an orthonormal basis of the first subsystem. If

$$\delta(\rho_{XB},\rho_X\otimes\rho_B) \leqslant \varepsilon,$$

then

$$S(X|B) \geqslant S(X) - \sqrt{2\varepsilon}\log_2(d) - 1/e.$$

Proof. It is easy to see that the trace distance between ρ_{XB} and $\rho_X\otimes\rho_B$ can be written as

$$\delta(\rho_{XB},\rho_X\otimes\rho_B) = \sum_x \mu_x(\delta(\rho_B^x,\rho_B)).$$

Let ψ be a purification of ρ_B. According to lemma B.1, for all $x\in\{1,\ldots,d\}$, there exists a purification $|\psi_x\rangle$ of ρ_B^x such that

$$\delta(P_{|\psi_x\rangle},P_{|\psi\rangle}) \leqslant \sqrt{2\delta(\rho_B^x,\rho_B)}.$$

Hence, using Jensen's inequality,

$$\sum_x \mu_x(\delta(P_{|\psi_x\rangle},P_{|\psi\rangle})) \leqslant \sqrt{2\sum_x \mu_x(\delta(\rho_B^x,\rho_B))} \leqslant \sqrt{2\varepsilon}.$$

Let now $\rho_{XBB'}$ be the state defined by

$$\rho_{XBB'} := \sum_x \mu_x(P_{|x\rangle}\otimes P_{|\psi_x\rangle}).$$

Note that, by this definition, $\rho_{XB}=\mathrm{tr}_{B'}(\rho_{XBB'})$.

From the strong subadditivity, we have

$$S(X|B) \geqslant S(X|BB') = S(XBB') - S(BB') \geqslant S(X) - S(BB'),$$

where the last inequality holds since

$$S(BB'|X) = \sum_x \mu_x S(\rho_x^{BB'}) \geqslant 0.$$

Because the rank of $\rho_{BB'}$ is not larger than d, $S(BB')$ can be bounded using Fannes' inequality—i.e.,

$$S(\rho_{BB'}) \leqslant S(P_{|\psi\rangle}) + \delta(\rho_{BB'},P_{|\psi\rangle})\log_2(d) + 1/e. \quad \text{(B1)}$$

Since $\rho_{BB'}=\Sigma_x\mu_x(P_{|\psi_x\rangle})$, it follows from the convexity of the trace distance that

$$\delta(\rho_{BB'},P_{|\psi\rangle}) \leqslant \sum_x \mu_x(\delta(P_{|\psi_x\rangle},P_{|\psi\rangle})) \leqslant \sqrt{2\varepsilon}.$$

Inserting this into (B1) and observing that $S(P_{|\psi\rangle})=0$ concludes the proof. □

APPENDIX C: KNOWN RESULTS

Consider two different measurement operations $\mathcal{E}$ and $\mathcal{F}$ applied to the individual parts of a symmetric state ρ^n. Lemma C.1 gives a relation between the measurement statistics of $\mathcal{E}$ and $\mathcal{F}$ (see [20] for a proof). The distance between these statistics is measured with respect to the statistical distance $\bar{\delta}(\cdot,\cdot)$.

Lemma C.1. Let ρ^n be a symmetric quantum state on $\mathcal{H}^{\otimes n}$, and let $\mathcal{E}$ and $\mathcal{F}$ be POVM's on $\mathcal{H}$ with $|\mathcal{E}|$ and $|\mathcal{F}|$ POVM elements, respectively. Let $Q_{\mathbf{X}}$ and $Q_{\mathbf{Y}}$ be the frequency distribution of the outcomes when applying the measurements $\mathcal{E}^{\otimes k}$ and $\mathcal{F}^{\otimes n-k}$, respectively, to different subsystems of ρ^n. Finally, let $\mathcal{B}$ be any convex set of density operators such that, for any operator A on $n-1$ subsystems, the normalization of $\mathrm{tr}_{n-1}(\mathbb{1}\otimes A\rho^n\mathbb{1}\otimes A^\dagger)$ is contained in $\mathcal{B}$. Then, for any $\varepsilon>0$, with probability at least $1-2^{|\mathcal{E}|+|\mathcal{F}|}e^{-n\varepsilon^2/8}$, there exists a state $\sigma\in\mathcal{B}$ such that

$$\frac{k}{n}\bar{\delta}(Q_{\mathbf{X}},P_{\mathcal{E}}[\sigma]) + \frac{n-k}{n}\bar{\delta}(Q_{\mathbf{Y}},P_{\mathcal{F}}[\sigma]) \leqslant \varepsilon,$$

where $P_{\mathcal{E}}[\sigma]$ and $P_{\mathcal{F}}[\sigma]$ denote the probability distributions of the outcomes when measuring σ with respect to $\mathcal{E}$ and $\mathcal{F}$, respectively.

Lemma C.2 below provides an expression for the maximum length of a key S that can be generated from a string Z such that S is secure against an adversary holding a quantum state ρ_E^z depending on Z. The proof can be found in [19] (see also [18]). Note that lemma C.2 holds with respect to the universally composable security definition described in Sec. II B.

Lemma C.2. Let ρ_{ZE} be a density operator such that ρ_Z is classical—i.e., $\rho_{ZE}=\Sigma_z P_Z(z)P_{|z\rangle}\otimes\rho_E^z$, where $\{|z\rangle\}_z$ is a family of orthonormal vectors—and let $\varepsilon\in\mathbb{R}^+$. Let S be the key computed by applying a two-universal hash function F mapping the value of Z to a value in $\{0,1\}^\ell$. Then S is ε secure with respect to ρ_{EF} if

$$\ell \leqslant S_2^{\varepsilon'}(ZE) - S_0^{\varepsilon'}(E) - 2\log_2(1/\varepsilon),$$

where $\varepsilon'=(\varepsilon/8)^2$.

The following lemma on error correction is a direct consequence of lemma 4 from [32] (see also [23]). Roughly speaking, it states that a message of length $H_0^\varepsilon(X|Y)$ is sufficient to guess the value of X when only Y is known.

Lemma C.3. Let $\mathcal{X}$ and $\mathcal{Y}$ be sets, let $\varepsilon\in\mathbb{R}^+$, and let $m\in\mathbb{N}$. Then there exists a probabilistic encoding function $e:\mathcal{X}\times\mathcal{R}\to\mathcal{C}$, taking randomness with some distribution P_R such that the following holds: For all probability distributions P_{XY} on $\mathcal{X}\times\mathcal{Y}$ satisfying $H_0^{\varepsilon'}(X|Y)+\log_2(1/\varepsilon')\leqslant m$, for $\varepsilon'=\varepsilon/2$, there exists a decoding function $d:\mathcal{C}\times\mathcal{Y}\to\mathcal{X}$ such that the probability of a decoding error is smaller than ε, i.e.,

$$\Pr_{(x,y,r)\leftarrow P_{XY}\times P_R}[d(e(x,r),y)=x]\geqslant 1-\varepsilon,$$

and the encoding $C:=e(X,R)$ gives no more than m bits of information on X—i.e.,

$$H_0(C) - H_\infty(C|X) \leqslant m.$$

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[2] http://www.idquantique.com.
[3] http://www.magiqtech.com.
[4] C. H. Bennett and G. Brassard, *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Banglore, India, 1984*, (IEEE, New York, 1984), pp.175—179 [IBM Tech. Dist. Bull. 28, 3153(1985)].
[5] D. Bruss, Phys. Rev. Lett. **81**, 3018 (1998); H. Bechmann-Pasquinucci and N. Gisin, Phys. Rev. A **59**, 4238 (1999).
[6] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[7] E. Biham and T. Mor, Phys. Rev. Lett. **78**, 2256 (1997).
[8] E. Biham and T. Mor, Phys. Rev. Lett. **79**, 4034 (1997).
[9] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor, Algorithmica **34**, 372 (2002).
[10] D. Mayers, in *Advances in Cryptology—CRYPTO 1996*, Lecture Notes in Computer Science, Vol. 1109 (Springer, Berlin, 1996), pp. 343–357,.
[11] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury, in *Proceedings of the 32nd Annual ACM Symposium on the Theory of Computing* (ACM, New York, 2000), pp. 715–724.
[12] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).
[13] H.-K. Lo, QIC **1**, 2 (2001).
[14] D. Gottesman and H.-K. Lo, IEEE Trans. Inf. Theory **49**, 457 (2003).
[15] K. Tamaki, M. Koashi, and N. Imoto, Phys. Rev. Lett. **90**, 167904 (2003).
[16] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[17] H. Bechmann-Pasquinucci, B. Huttner, and N. Gisin, Phys. Lett. A **242**, 198 (1998).
[18] R. König, M. Maurer, and R. Renner, e-print quant-ph/0305154.
[19] R. Renner and R. König, in *Proceedings of the Second Theory of Cryptography Conference (TCC) 2005*, Lecture Notes in Computer Science, Vol. 3378 (Springer, Berlin, 2005), pp. 407–425.
[20] M. Christandl, R. Renner, and A. Ekert, e-print quant-ph/0402131.
[21] M. Ben-Or (unpublished).
[22] B. Kraus, N. Gisin, and R. Renner, e-print quant-ph/0410215.
[23] R. Renner and S. Wolf (unpublished).
[24] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[25] J. L. Carter and M. N. Wegman, J. Comput. Syst. Sci. **18**, 143 (1979).
[26] M. N. Wegman and J. L. Carter, J. Comput. Syst. Sci. **22**, 265 (1981).
[27] H.-K. Lo, H. F. Chau, and M. Ardehali, e-print quant-ph/0011056.
[28] I. Devetak and A. Winter, Proc. R. Soc. London, Ser. A **461**, 207 (2005).
[29] I. Csiszár and J. Körner, IEEE Trans. Inf. Theory **IT-24**, 339 (1978).
[30] C. A. Fuchs, N. Gisin, R. B. Griffiths, C. S. Niu, and A. Peres, Phys. Rev. A **56**, 1163 (1997).
[31] A. Rényi, in *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, 1961), Vol. 1, pp. 547–561.
[32] R. Renner and S. Wolf, in *Advances in Cryptology—EUROCRYPT 2004*, Lecture Notes in Computer Science, Vol. 3027 (Springer, Berlin, 2004), pp. 109–125.