Analysis of a rate-adaptive reconciliation protocol and the effect of leakage on the secret key rate

David Elkouss,[1,*] Jesus Martinez-Mateo,[2,†] and Vicente Martin[2,‡]

[1]*Departamento de Analisis Matemático and Instituto de Matemática Interdisciplinar, Universidad Complutense de Madrid, 28040 Madrid, Spain*

[2]*Facultad de Informática, Universidad Politécnica de Madrid (UPM) Campus de Montegancedo, 28660 Boadilla del Monte (Madrid), Spain*

(Received 22 February 2013; published 26 April 2013)

Quantum key distribution performs the trick of growing a secret key in two distant places connected by a quantum channel. The main reason is so that the legitimate users can bound the information gathered by the eavesdropper. In practical systems, whether because of finite resources or external conditions, the quantum channel is subject to fluctuations. A rate-adaptive information reconciliation protocol, which adapts to the changes in the communication channel, is then required to minimize the leakage of information in the classical postprocessing. We consider here the leakage of a rate-adaptive information reconciliation protocol. The length of the exchanged messages is larger than that of an optimal protocol; however, we prove that the min-entropy reduction is limited. The simulation results, both in the asymptotic and in the finite-length regime, show that this protocol allows to increase the amount of a distillable secret key.

DOI: 10.1103/PhysRevA.87.042334 PACS number(s): 03.67.Dd

I. INTRODUCTION

Shannon published his seminal "A Mathematical Theory of Communications" [1] in 1948 after eight years of intermittent work [2]. The paper meant the birth of communications and coding theory. Shannon not only established the frame under which communications systems could be studied and compared; he also proved their fundamental limits, i.e., the limiting rates for data compression and reliable transmission through noisy channels. This second result was especially surprising since there was no certainty that reliable transmission with a positive rate was even possible [3].

A year later, in 1949, Shannon's "Communication Theory of Secrecy Systems" [4] came to light. In the words of Gallager, "Shannon's cryptography work can be viewed as changing cryptography from an art to a science" [2]. Shannon successfully applied the tools developed in [1] to the problem of transmitting confidential messages through public channels. His main conclusion is that a message from a set of messages sent through a public channel can be obfuscated into a cypher text with the help of a secret key in such a way that the number of possible originating messages is the whole set of messages; that is, the cypher text leaks no information to a possible eavesdropper. The condition for this to happen is that the number of secret keys is equal to or greater than the number of messages. This condition only applies to eavesdroppers with unbounded resources; if we limit the storage or computing capability of the eavesdropper, secret communications are possible without fulfilling the condition. It is evident that computing power resources that today might be considered out of reach might become available in the near future. There is an implicit risk in assuming that an eavesdropper is limited in any way beyond the fundamental limits that physics impose upon her; therefore, the interest in establishing the scenarios in which some kind of security can be achieved without any assumption is self-evident.

The distribution of secret keys or SKD is a problem closely related to confidential communications. Two parties sharing a secret key can communicate privately through a channel in the conditions discussed in the previous paragraph. We can then study the problem of secret-key sharing as a way to achieve confidential communications. The main idea is that two distant parties can agree in a secret key if they have access to a shared source of randomness. The randomness source can take many incarnations, e.g., in the form of a source received from a trusted party or in the form of a noisy channel [5].

In most of the SKD scenarios the legitimate parties obtain instances of correlated sources, which means that they obtain similar but not identical strings. It is then assumed that there is an authentic though otherwise public channel available to all parties—including the eavesdropper. The legitimate parties can exchange additional information through this channel in order to reconcile their strings. They can do so by revealing some information about them, for instance the parities of carefully chosen positions. This process is known as information reconciliation [6]. It is not hard to see that the information exchanged through the public channel reduces the uncertainty that the eavesdropper has on the strings of the legitimate parties. Thus, a reduction in the leakage due to information reconciliation allows to increase the amount of the distillable secret key. A second step known as privacy amplification is then needed [7]. In the privacy amplification step the legitimate parties agree on a secret but shorter key of which the eavesdropper has a negligible amount of information.

These mathematical models can have a real (i.e., physical) correspondence. One such model is a physical fiber carrying single photons randomly polarized in one of two nonorthogonal bases [8]. Quantum key distribution (QKD) is probably the main practical application of SKD. In a QKD protocol [8–10], two legitimate parties, Alice and Bob, aim at sharing an information theoretic secret key, even in the presence of an eavesdropper Eve. In the quantum part of such a protocol, Alice and Bob exchange quantum signals, e.g., single photons, which carry classical information. For instance, Alice encodes a classical bit onto the polarization or the phase

*delkouss@mat.ucm.es
†jmartinez@fi.upm.es
‡vicente@fi.upm.es

 ©2013 American Physical Society

of a photon and sends this photon to Bob, who measures it. In any realistic implementation of a QKD protocol, the strings obtained after the exchange of the quantum signals suffer discrepancies mainly due to losses in the channel and noise in Bob's detectors but which are conservatively attributed to the action of an eavesdropper. Therefore, any QKD protocol must include the classical postprocessing steps described above in order to extract a secret key from the correlated strings.

The channel connecting Alice and Bob in a real system may substantially vary over time. The motivation of this work is to analyze the *sp* protocol [11], an information-reconciliation protocol that adapts to these channel variations. We previously showed that in a classical repetition scenario (i.e., with classical attackers and independent, identically distributed sources) its reconciliation efficiency is only limited by the quality of the error-correcting code used to implement the protocol [12]. We consider here the leakage of the *sp* protocol with a quantum eavesdropper, both in the asymptotic and in the finite-length regime, and its impact on the amount of the distillable secret key.

II. PRELIMINARIES AND NOTATION

Let X be a discrete random variable taking values in the finite alphabet $\mathcal{X}$. The Shannon entropy [1], min-entropy, and max-entropy [13] of X are respectively defined by

$$H(X) = -\sum_{x\in\mathcal{X}} p_X(x)\log_2 p_X(x), \tag{1}$$

$$H_\infty(X) = \min_{x\in\mathcal{X}}[-\log_2 p_X(x)], \tag{2}$$

$$H_0(X) = \log_2 |x \in \mathcal{X} : p_X(x) > 0|, \tag{3}$$

where $|\cdot|$ stands for the cardinality of a set. It holds that $H_\infty(X) \leqslant H(X) \leqslant H_0(X)$, and the equality occurs when the outcomes in X are given by a uniform distribution.

Now let X and Y be two jointly distributed discrete random variables taking values on alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively. The conditional entropy, min-entropy, and max-entropy of X given Y are defined by

$$H(X|Y) = \sum_{y\in\mathcal{Y}} H(X|y), \tag{4}$$

$$H_\infty(X|Y) = \min_{y\in\mathcal{Y}} H_\infty(X|y), \tag{5}$$

$$H_0(X|Y) = \max_{y\in\mathcal{Y}} H_0(X|y), \tag{6}$$

where the entropy of a random variable given an event is the entropy of the induced random variable.

Let the state of a finite-dimensional quantum system be represented by a trace one, positive semidefinite operator on a (finite-dimensional) Hilbert space $\mathcal{H}$. We denote by $\mathcal{P}(\mathcal{H})$ the set of all states acting on $\mathcal{H}$.

Let us give some basic definitions about the quantum counterparts of these classical information measures. The equivalent of the entropy of a random variable is the von Neumann entropy of a state ρ_X [14]. It is defined as

$$H(X)_{\rho_X} = -\mathrm{tr}(\rho_X \log_2 \rho_X), \tag{7}$$

where tr denotes the trace operation and we indicate with a subscript the state on which the entropy is computed. Henceforth, it is explicitly written whenever it helps to clarify a statement.

Let $\rho_{XY} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_Y)$ be a bipartite quantum state. The conditional quantum min-entropy of ρ_{XY} given $\mathcal{H}_Y$ is defined as

$$H_\infty(X|Y) = \sup_{\sigma_Y}(-\log_2 \min\{\lambda|\lambda\,\mathrm{id}_X \otimes \sigma_Y \geqslant \rho_{XY}\}), \tag{8}$$

where $\lambda > 0$.

If $\mathcal{H}_Y$ is one-dimensional,

$$H_\infty(X|Y) = H_\infty(X) = -\log_2 \lambda_{\max}(\rho_X), \tag{9}$$

where $\lambda_{\max}(\rho_X)$ outputs the maximum eigenvalue of ρ_X.

We finally consider the smooth generalization of the conditional min-entropy introduced in [15]. Let $\{\rho,\sigma\} \in \mathcal{P}(\mathcal{H})$; the trace distance between ρ and σ is given by

$$\tfrac{1}{2}||\rho - \sigma||_1 = \mathrm{tr}(|\rho - \sigma|). \tag{10}$$

The smooth entropy was first defined as an optimization over all states ε-close in terms of the trace distance. The smooth entropies have been redefined in terms of other measures such as the purified distance and verify additional properties [16,17] but for the present study it suffices to consider the original definition.

Let $\rho_{XY} \in \mathcal{P}(\mathcal{H}_X \otimes \mathcal{H}_Y)$ and $\varepsilon \geqslant 0$. The smooth version of Eq. (8) is given by

$$H^\varepsilon_\infty(X|Y)_{\rho_{XY}} = \sup_{\hat{\rho}_{XY}} H_\infty(X|Y)_{\hat{\rho}_{XY}}, \tag{11}$$

where the supreme is found over all $\hat{\rho}_{XY}$ such that $\frac{1}{2}||\rho_{XY} - \hat{\rho}_{XY}||_1 \leqslant \varepsilon$.

III. INFORMATION RECONCILIATION

A. Impact of information reconciliation on the secret key length

One common assumption in a SKD protocol is that all the parties have access to the outcomes of an independent identically distributed experiment repeated many times. If this assumption holds, the parties can safely regard an average behavior as the law of large numbers guarantees that the joint outcome will be typical with high probability. However, assuming a repetition scenario might be unrealistic in some situations, in these cases key distillation can be considered for a finite number of outcomes of a joint experiment. This second, more restrictive, scenario is sometimes referred as finite-key distillation. Both the repetition [10] and the finite-key [18–20] scenarios have been addressed in QKD.

The secrecy of a key K can be measured in terms of its closeness to a perfect one which is uniformly random and decoupled from the eavesdropper's system Z. A key K is considered ε-secure if [21]

$$\tfrac{1}{2}||\rho_{KZ} - \tau_K \otimes \rho_Z||_1 \leqslant \varepsilon. \tag{12}$$

The communications on the public channel might be one way or two ways. We have chosen to restrict the channel to one-way communications since our focus is on practical protocols with reduced distillation complexity, network requirements, etc. However, it should be noted that two-way communications can be used to distill a key in scenarios where one-way secret key distillation is not possible [5] and,

in general, the amount of a distillable secret key with two-way communications can be strictly higher than with one-way communications [22,23].

In the repetition scenario and aided by one-way classical communications, the maximum rate at which a key can be extracted with ε approaching zero as the number of repetitions goes to infinity is given by [24]

$$K = H(X|Z) - H(X|Y), \tag{13}$$

where X and Y are classical systems available to the legitimate parties Alice and Bob and Z is a quantum system at the eavesdropper's site. The first term on the right-hand side of Eq. (13) amounts to the randomness that can be extracted which is independent of Z while the second term can be regarded as the information that Alice and Bob should exchange to reconcile X and Y.

Equation (13) is valid only in the asymptotic case. However, a real system has only access to finite resources, which means that not only do Alice and Bob have bounded computational power but also they have to distill a secret key from a finite number of experiments. Thus, in general there is no convergence toward an ideal key and security has to be considered for an acceptable security threshold ε.

Let us assume that Alice and Bob exchange N signals out of which they use m for estimating their correlations and $t \leqslant N - m$ for key distillation. If the correlations do not verify some conditions, Alice and Bob abort the protocol; $\varepsilon_{\mathrm{PE}}$ represents the probability that the parameter estimation procedure fails.

Given some reconciliation protocol, C stands for the set of all possible reconciliation messages and $\varepsilon_{\mathrm{EC}}$ represents the maximum probability that the estimate at Bob's site does not coincide with Alice's string.

Let $\varepsilon_{\mathrm{PA}}$ represent the failure probability in the privacy amplification procedure, and $\bar{\varepsilon}$ is a smoothing parameter; then the rate at which the legitimate parties can distill an ε-secure key is bounded by [25]

$$K^{\varepsilon} \leqslant \frac{1}{N}\left(H_{\infty}^{\bar{\varepsilon}}(X^t|Z^N C) - 2\log_2 \frac{1}{\varepsilon_{\mathrm{PA}}}\right), \tag{14}$$

where $\varepsilon = n_{\mathrm{PE}}\varepsilon_{\mathrm{PE}} + \varepsilon_{\mathrm{EC}} + \varepsilon_{\mathrm{PA}} + \bar{\varepsilon}$, and n_{PE} is the number of estimated parameters.

The smooth min-entropy in Eq. (14) can be evaluated to measure the net impact of information reconciliation [25]:

$$H_{\infty}^{\bar{\varepsilon}}(X^t|Z^N C) \geqslant H_{\infty}^{\bar{\varepsilon}}(X^t|Z^N) - \mathrm{leak}, \tag{15}$$

where leak is a purely classical term that tracks the amount of information correlated with X^t revealed during reconciliation. It is given by [15]

$$\mathrm{leak} = H_0(C) - H_{\infty}(C|X^t). \tag{16}$$

The main effect of an imperfect reconciliation is a reduction of the secret key rate, which in turn, in terms of the figures of merit of a QKD protocol, limits the distance range over which secret keys can be distilled [10,26].

B. Fundamental limits of information reconciliation

Let Alice and Bob be two parties holding x and y, two n-length strings that are respectively n outcomes of two jointly distributed random variables X and Y. A one-way reconciliation protocol on the strings x and y is a protocol that produces the strings s_x and s_y from x and y, respectively, after exchanging the message $c(x)$ through the public channel.

A reconciliation protocol is considered ε robust [6] if

$$\sum_{x\in\mathcal{X}^n, y\in\mathcal{Y}^n} p(x,y)p(s_x \neq s_y) \leqslant \varepsilon. \tag{17}$$

The efficiency of a reconciliation protocol can be measured using a quality parameter ξ^{ε} that compares the amount of disclosed information with the minimum theoretical disclosure:

$$\xi^{\varepsilon} = \frac{\mathrm{leak}}{nH(X|Y)}, \tag{18}$$

where the minimum $nH(X|Y)$ is known as the Slepian-Wolf bound; it delimits the minimum rate for reliably describing a source X to a distant party with access to side information Y [27].

It is well known the appropriateness of (linear) error correcting codes for the Slepian-Wolf problem [28]. In consequence, good error-correcting codes can be used for information reconciliation. Let $\mathcal{C}(n,k)$ be a linear code with coding rate $R_0 = k/n$; a message of length $n-k$ called the *syndrome* [29] can be used to reconcile two sources with conditional entropy $nH(X|Y)$. Even if $n-k$ is greater than the theoretical minimum, for finite lengths there is always nonzero error probability. We denote the rate of decoding errors or frame error rate (FER) by the parameter ε. Then, a reconciliation protocol based on sending the syndrome of a linear code is ε robust, and the reconciliation efficiency is given by

$$\xi_{\mathcal{C}}^{\varepsilon} = \frac{n-k}{nH(X|Y)} = \frac{1-R_0}{H(X|Y)}. \tag{19}$$

However, an acceptable FER in a communications protocol might not be sufficient in a security context. It is a common practice to divide the reconciliation process into two steps [18,30]. In the first one, a common string is produced, for instance using an error-correcting code as we just described. In the second one, Alice uniformly at random chooses a function f from a family of 2-universal hash functions [31] and computes a hash of her string, $f(s_x)$. Alice sends to Bob her choice f together with $f(s_x)$. Bob computes his own hash value $f(s_y)$ and the protocol aborts if $f(s_x) \neq f(s_y)$. Since the choice of the hash function is independent of X, only the length of the hash $\lceil -\log_2 \varepsilon_{\mathrm{EC}} \rceil$ for some $\varepsilon_{\mathrm{EC}} > 0$ is added to the leakage:

$$\mathrm{leak}_{\mathcal{C}}^{\varepsilon_{\mathrm{EC}}} = n(1-R_0) + \left\lceil \log_2 \frac{1}{\varepsilon_{\mathrm{EC}}} \right\rceil. \tag{20}$$

The joint reconciliation process is $\varepsilon_{\mathrm{EC}}$ robust where $\varepsilon_{\mathrm{EC}}$ can be chosen to be much smaller than the FER.

It is clear from Eq. (19) that the length of the conversation when using a code is fixed to $n-k$. That is, the amount of information does not adapt to the error rate in the channel. This is a perfect solution for the Slepian-Wolf problem since the correlations are fixed and known beforehand. However, in QKD it is common that the error rate varies from one execution

to the next. In consequence, an adaptation of the coding rate is needed in order to use linear codes for reconciliation.

IV. STUDY OF A RATE-ADAPTIVE PROTOCOL

In this section we study the efficiency and impact of a rate-adaptive protocol, which is in essence the *sp* protocol in [11] with an additional error-verification step.

A. Description of the rate-adaptive protocol

In the following we detail the steps of a rate-adaptive information reconciliation protocol.

Step 0: Preconditions. Alice and Bob agree on the following parameters: (i) a pool of shared mother codes of length n, constructed for different rates; (ii) d, the maximum number of symbols (bits) that will be used to adapt the coding rate; and (iii) the target $\varepsilon_{\mathrm{EC}}$ which characterizes the length of the hashes.

Step 1: Raw key exchange. Alice and Bob obtain two correlated strings x and y, respectively, of length $n-d$ and a precise estimate of the error rate p_e. If p_e is outside their target rates they abort the protocol. Otherwise, both parties select the appropriate code $\mathcal{C}$ and compute the adequate number of symbols (bits) s to reveal, with $s < d$, such that the coding rate is then adapted to p_e.

Step 2: Coding. Alice creates an extended string $\hat{x}$ of length n by concatenating x and x', a uniformly random string of length d. Alice sends to Bob the hash value $f(\hat{x})$, the syndrome of $\hat{x}$ on $\mathcal{C}$, and the values and positions of s symbols among the d symbols randomly generated.

Step 3: Decoding. Bob creates an extended string of length n by concatenating y and y', a uniformly random string of length d. Bob sets the values of the received s symbols to their correct value. Bob computes $\hat{y}$, his estimate of $\hat{x}$, and $f(\hat{y})$, his own hash value. If $f(\hat{y}) \neq f(\hat{x})$ they abort the protocol.

We would like to remark that in step 2 both the verification tag and the reconciliation message are jointly encoded and sent to Bob. There is no extra interactivity coming from error verification, still only one message is exchanged for reconciliation, and a second one from Bob to Alice is sent to notify the success or failure of the protocol.

B. Leakage

The *sp* protocol creates an extended system X^tX' by adding d symbols (bits) with random values. The Slepian-Wolf bound implies that for successful reconciliation the length of the reconciliation message should be greater than

$$H(X^tX'|Y^t) = H(X^t|Y^t) + d, \tag{21}$$

which is trivially larger than $H(X^t|Y^t)$ if $d > 0$.

However, the appropriate comparison is in terms of the conditional smooth entropy on the reconciled system, since it is the magnitude that limits the distillable key after the reconciliation step. Lemma 1 shows that the smooth min-entropy decrease produced by the *sp* protocol on the extended system is equivalent to the decrease produced by an error-correcting code with rate R on the original system. This equivalent coding rate R is given by

$$R = \frac{k-s}{n-d}. \tag{22}$$

The dependence of R on d and s allows to understand how the protocol adapts the amount of information disclosed for reconciling errors. Since the value of d is fixed previous to the execution of the protocol, it is s, the number of symbols (bits) revealed to Bob on the public channel and the parameter available to Alice for modulating the coding rate. A higher value of s increases the information available to the decoder, allowing him to reconcile noisier strings, while a lower value of s allows him to reduce the leakage by increasing the coding rate. On the other hand, d sets the range of achievable rates, from $(k-d)/(n-d)$ to $k/(n-d)$. The extremal values correspond to the limiting cases of revealing the d symbols (bits) and revealing no information on the public channel.

Lemma 1. Let $\rho_{X^tZ^N}$ be a bipartite state and $\sigma_{X^tX'Z^NC}$ the extension resulting from the application of the *sp* protocol. Then the smooth min-entropy of the extended system X^tX' given Z^NC can be bounded by

$$H_\infty^\varepsilon(X^tX'|Z^NC)_\sigma \geqslant H_\infty^\varepsilon(X^t|Z)_\rho - t(1-R) - \left\lceil \log_2 \frac{1}{\varepsilon_{\mathrm{EC}}} \right\rceil.$$

Proof. The proof of Lemma 1 follows:

$$\begin{aligned} H_\infty^{\varepsilon+\varepsilon'}(X^tX'|Z^NC)_\sigma &\geqslant H_\infty^{\varepsilon+\varepsilon'}(X^tX'|Z^N)_\sigma - \text{leak} \\ &= H_\infty^{\varepsilon+\varepsilon'}(X^tX'|Z^NI)_\phi - \text{leak} \\ &\geqslant H_\infty^{\varepsilon}(X^t|Z^N)_\phi + H_\infty^{\varepsilon'}(X'|I)_\phi - \text{leak}. \end{aligned}$$

Let $\varepsilon' \geqslant 0$. The first inequality follows from Eq. (15) which bounds the impact of the conversation. We can trivially extend the state on $\sigma_{X^tX'Z^N}$ to $\phi_{X^tX'Z^NI} = \sigma_{X^tX'Z^N} \otimes \mathrm{id}_I$, where I is a one-dimensional system, without changing the value of the smooth min-entropy $[H_\infty^{\varepsilon+\varepsilon'}(X^tX'|Z^N)_\sigma = H_\infty^{\varepsilon+\varepsilon'}(X^tX'|Z^NI)_\phi]$; the first equality holds by this argument. We can apply Renner's superadditivity theorem in [15] for product states to obtain the second inequality. If we now consider just the second and third terms from this last relation, we obtain

$$\begin{aligned} &H_\infty^{\varepsilon'}(X'|I)_\phi - \text{leak} \\ &\quad = (s+p) - \left(s + n(1-R_0) + \left\lceil \log_2 \frac{1}{\varepsilon_{EC}} \right\rceil \right) \\ &\quad = -t(1-R) - \left\lceil \log_2 \frac{1}{\varepsilon_{EC}} \right\rceil. \end{aligned}$$

We can choose $\varepsilon' = 0$ and since I is one-dimensional $H_\infty(X'|I)_\phi$ reduces to $H_\infty(X')_\phi$. Furthermore, X' is classical and uniformly distributed, thus maximizing the min-entropy. The leakage is obtained by tracking the amount of information sent from Alice to Bob during the protocol and subtracting the part that is independent of X^tX'. We recover the desired result if we consider that $\phi_{X^tX'Z^NI}$ is also an extension of $\rho_{X^tZ^N}$, which means that $H_\infty^\varepsilon(X^t|Z^N)_\phi = H_\infty^\varepsilon(X^t|Z^N)_\rho$. ■

V. SIMULATION RESULTS

In this section we compare the tradeoffs between using the *sp* protocol, nonadapted error-correcting codes, and Cascade

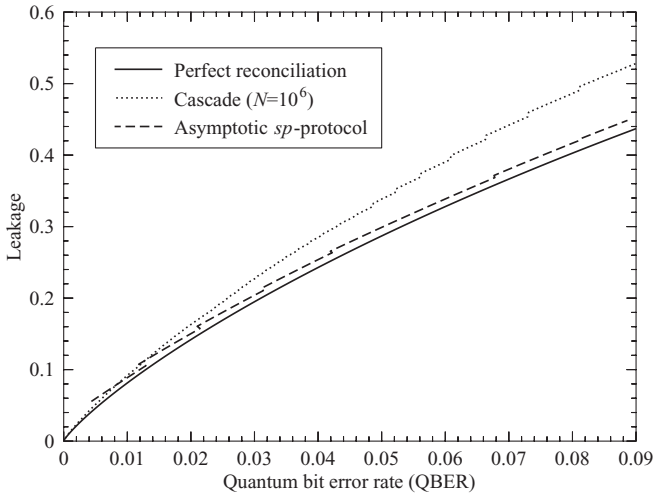

FIG. 1. The asymptotic leakage of the sp protocol, the leakage of Cascade, and the leakage for a perfect reconciliation procedure are compared as a function of the QBER.

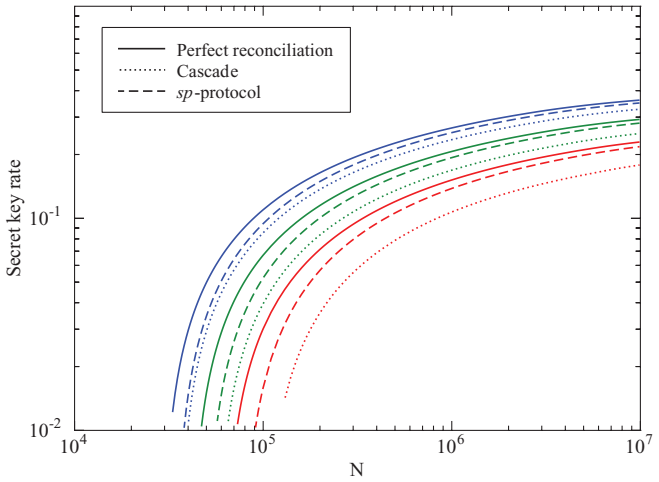

FIG. 2. (Color online) Secret key rate in the finite-key regime for a perfect reconciliation procedure, the sp protocol, and considering the efficiency of Cascade. Three different QBER values are considered (from left to right): 4% (blue), 5% (green), and 6% (red). Other parameters are $\varepsilon = 10^{-5}$ and $\varepsilon_{\mathrm{EC}} = 10^{-10}$.

(a well-known interactive protocol proposed in [6] and implemented in most QKD systems). First we present the difference of the reconciliation protocols in terms of asymptotic leakage and then we plug them into a QKD protocol and compare the distillable secret key with finite resources.

The strings are assumed to be binary and are modeled as the input and output of a binary symmetric channel (BSC). This is appropriate in the case of some QKD protocols [8,32,33] if errors on the quantum channel are symmetric and independent.

For convenience, we have implemented the rate-adaptive sp protocol with irregular binary low-density parity-check (LDPC) codes since there is a wealth of material and information available: a number of matrices, decoding algorithms, and communication standards have been proposed in the past years for these codes. However, nonbinary LDPC codes [34] or other code families [35] could probably be adapted to implement the sp protocol. We fixed the proportion of modulated symbols to $d/n = 5\%$.

Figure 1 shows the leakage rate ($\mathrm{leak}_C^{\varepsilon_{\mathrm{EC}}}/t$) as a function of the quantum bit error rate (QBER). An optimal protocol achieving the Slepian-Wolf bound (solid line) is compared to the asymptotic sp protocol computed using the theoretical analysis described in Appendix B (dashed line) and to Cascade (dotted line). Note that for Cascade, instead of upper bounding the leakage with the analytical estimate given in [6] which might be overly pessimistic, we used as the upper bound the leakage rate with large blocks of length 10^6 (see Appendix A for numerical justification).

Both Cascade and the sp protocol are close to optimal for small QBERs. However, over approximately 3% they begin to diverge and while the former closely follows the Slepian-Wolf bound the latter clearly has a higher leakage.

To analyze the impact of reconciliation on the achievable secret key rate, we have chosen the prepare and measure a version of the Bennett-Brassard 1984 (BB84) protocol and consider for simplicity, and in order to highlight the effect of reconciliation, an idealized scenario: we assume that Alice and Bob have access to single-photon sources and perfect detectors. Following [36] the secret key in this setting can be distilled at a rate

$$K^{\varepsilon} \leqslant \frac{t}{N}\{[1 - h(Q)] - \Delta(t) - \mathrm{leak}/t\}, \tag{23}$$

where h is the binary entropy function, Q is the estimated QBER that takes into account statistical fluctuations due to the finite-length case, and Δ is the smoothing parameter that allows us to lower bound the smooth min-entropy in Eq. (14) [25].

Figure 2 shows the secret key rate as a function of the number of exchanged signals (N). We compare in this figure the secret key rate for three different QBER values (4%, 5%, and 6%) using a perfect reconciliation protocol, Cascade, and the sp protocol. The security parameter ε is set to 10^{-5}, and $\varepsilon_{\mathrm{EC}} = 10^{-10}$, as suggested in [36].

The convergence of LDPC codes towards the asymptotic value is slower than that of Cascade (see Appendix A). In consequence the optimality of the distillable key with this implementation of the sp protocol increases with the length, shifting from close to Cascade for small lengths to close to the optimal value asymptotically. For low QBERs and small lengths, the slow convergence of LDPC codes together with the good efficiency of Cascade in this region makes both secret key rates very similar. For higher QBERs, even for small lengths the LDPC implementation of the sp protocol clearly outperforms Cascade.

VI. DISCUSSION

This paper analyzes some improvements in the classical postprocessing of QKD protocols. The key distillation process can be divided into two steps: information reconciliation and privacy amplification. Information reconciliation allows to establish a common string and in the privacy-amplification step a shorter but more secure key is created. Both steps are highly coupled: in essence every bit exchanged in the information-reconciliation step implies that one additional bit has to be removed from the final key in the privacy-amplification step.

The problem of correcting the discrepancies between the strings of the legitimate parties is also known as the problem of source coding with side information by the information theory community. Under this paradigm, the theoretical limits of information reconciliation are given by the Slepian-Wolf bound. Information reconciliation is, then, basically error correction.

We have adopted a pragmatic approach towards error correction and used modern coding techniques well suited for QKD purposes. In a real QKD scenario we have to deal with a broad range of error rates. Furthermore, the number of times the classical public communication channel is accessed should be limited. As opposed to the eavesdropper that should, for the sake of security, be assumed to have access to unbounded resources, the legitimate parties are equipped with a finite amount of resources.

The *sp* protocol, induced by a mother code of rate R_0, allows the legitimate parties to adapt the reconciliation step to varying conditions. However, it exchanges a message longer than the optimal one. We proved that the *sp* protocol is equivalent to the use of a code with an adapted rate R. The claim holds in the sense that the smooth min-entropy reduction of the former in an extended system is bounded by the reduction of the latter in the original system.

We implemented the *sp* protocol with irregular LDPC codes. The results obtained indicate that the *sp* protocol asymptotically behaves close to the theoretical limit. We claim no optimality in our implementation of the *sp* protocol and certainly it could be expected that other code families are better suited to short key lengths or to other kinds of correlations different from those modeled by a BSC. The analysis, however, applies to any linear error-correcting code. In consequence, it allows us to consider rate-adaptive information reconciliation as a specific code design problem. We believe that this protocol opens the doors to consider simpler and possibly better schemes for classical postprocessing in secret key distillation protocols.

ACKNOWLEDGMENTS

This work has been partially supported by the project Quantum Information Technologies in Madrid (QUITEMAD), Project No. P2009/ESP-1594, and the CHIST-ERA project Composing Quantum Channels, Project No. PRI-PIMCHI-2011-1071.

APPENDIX A: CASCADE SIMULATIONS

In order to estimate the asymptotic leakage of Cascade we simulated the protocol with strings of lengths 10^4, 10^5, and 10^6. The results in Table I show that with a string length of 10^6 the leakage rate has already converged.

TABLE I. Leakage rate of Cascade for strings of length 10^4, 10^5, and 10^6 as a function of the QBER.

QBER	10^4	10^5	10^6
0.01	0.0917	0.0914	0.0914
0.04	0.285	0.284	0.284
0.05	0.338	0.338	0.338
0.06	0.390	0.390	0.390

APPENDIX B: THEORETICAL ANALYSIS OF RATE-MODULATED CODES

Binary linear codes admit a bipartite graph representation in which symbols are linked with parity checks. An ensemble of irregular binary LDPC codes can be defined by the degree distributions on the edges of symbols and checks [37]. We can study the behavior of an ensemble under a message-passing algorithm by tracking the evolution of the message distributions. This recursive tracking is known as density evolution [37] and allows us to compute the asymptotic decoding threshold of a code family on a communications channel. In general, densities are updated following this recurrence relation:

$$p^{\ell+1}(x) = \rho(p_0(x) * \lambda(p^{\ell}(x))), \tag{B1}$$

where p^{ℓ} is the average probability on symbols on the decoding iteration ℓ if the code graph is treelike, $\lambda(x)$ and $\rho(x)$ are the symbol and check node degree polynomials, respectively, $p_0(x)$ is the initial message density, and $*$ stands for convolution.

In Sec. V, we focused our attention on the BSC. This channel is characterized by a single parameter: the crossover probability ε. That is, a bit is either noiselessly transmitted with probability $1-\varepsilon$ or flipped with probability ε. The channel is then modeled by the following initial density distribution:

$$p_0(x) = \varepsilon\Delta_{L(\varepsilon)}(x) + (1-\varepsilon)\Delta_{-L(\varepsilon)}(x), \tag{B2}$$

where $L(\varepsilon) = \log_2 \frac{\varepsilon}{1-\varepsilon}$ is a log-likelihood ratio, and $\Delta_t(x) = \delta(x-t)$ is the Dirac δ function displaced at position t.

Now, in the *sp* protocol, an n-length raw string is composed of $n-d$ bits sent through a noisy channel, in this case the above-described BSC, and d bits with randomly assigned values out of which s are revealed through the public and noiseless channel. Letting σ and π stand for the fraction of bits that are completely known and unknown to the decoder, respectively, we can compute the asymptotic behavior of the *sp* protocol with the following initial density:

$$p_0(x) = (1-\pi-\sigma)[\varepsilon\Delta_{L(\varepsilon)}(x) + (1-\varepsilon)\Delta_{-L(\varepsilon)}(x)] + \pi\Delta_0(x) + \sigma\Delta_\infty(x). \tag{B3}$$

[1] C. E. Shannon, Bell Labs Tech. J. **27**, 379 (1948).
[2] R. G. Gallager, IEEE Trans. Inf. Theory **47**, 2681 (2001).
[3] D. J. C. Mackay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, Cambridge, UK, 2003).
[4] C. E. Shannon, Bell Labs Tech. J. **28**, 656 (1949).
[5] U. Maurer, IEEE Trans. Inf. Theory **39**, 733 (1993).
[6] G. Brassard and L. Salvail, in *Advances in Cryptology—EUROCRYPT' 93, Lecture Notes in Computer Science* Vol. **765** (Springer, New York, 1994), pp. 410–423.

[7] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, IEEE Trans. Inf. Theory **41**, 1915 (1995).
[8] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (IEEE, New York, 1984), pp. 175–179.
[9] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[10] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).
[11] D. Elkouss, J. Martinez-Mateo, and V. Martin, Quantum Inform. Comput. **11**, 226 (2011).
[12] D. Elkouss, J. Martinez-Mateo, and V. Martin, in *Proceedings of 2010 International Symposium on Information Theory and its Applications (ISITA)* (IEEE, Piscataway, NJ, 2010), pp. 179–184.
[13] A. Rényi, in *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, 1960), pp. 547–561.
[14] J. Von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, New York, 1932).
[15] R. Renner, Int. J. Quantum Inf. **6**, 1 (2008).
[16] M. Tomamichel, R. Colbeck, and R. Renner, IEEE Trans. Inf. Theory **56**, 4674 (2010).
[17] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011).
[18] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nat. Commun. **3**, 634 (2012).
[19] M. Hayashi and T. Tsurumaru, New J. Phys. **14**, 093014 (2012).
[20] P. J. Salas, Quantum Inform. Comput. **13**, 861 (2013).
[21] R. König, R. Renner, A. Bariska, and U. Maurer, Phys. Rev. Lett. **98**, 140502 (2007).
[22] D. Gottesman and H. K. Lo, IEEE Trans. Inf. Theory **49**, 457 (2003).
[23] S. Watanabe, R. Matsumoto, T. Uyematsu, and Y. Kawano, Phys. Rev. A **76**, 032312 (2007).
[24] I. Devetak and A. Winter, Proc. R. Soc. A **461**, 207 (2005).
[25] V. Scarani and R. Renner, Phys. Rev. Lett. **100**, 200501 (2008).
[26] M. Peev *et al.*, New J. Phys. **11**, 075001 (2009).
[27] D. Slepian and J. Wolf, IEEE Trans. Inf. Theory **19**, 471 (1973).
[28] R. Zamir, S. Shamai, and U. Erez, IEEE Trans. Inf. Theory **48**, 1250 (2002).
[29] Let H be a parity check matrix of the code $\mathcal{C}$ and x a vector of length n; the syndrome of x is the vector $s(x) = H \cdot x$ of length $n - k$.
[30] Chi-Hang Fred Fung, X. Ma, and H. F. Chau, Phys. Rev. A **81**, 012318 (2010).
[31] M. N. Wegman and L. Carter, J. Comput. Syst. Sci. **22**, 265 (1981).
[32] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, J. Cryptology **5**, 3 (1992).
[33] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).
[34] K. Kasai, R. Matsumoto, and K. Sakaniwa, in *Proceedings of 2010 International Symposium on Information Theory and its Applications (ISITA)* (IEEE, Piscataway, NJ, 2010), pp. 922–927.
[35] P. Jouguet, S. Kunz-Jacques, and A. Leverrier, Phys. Rev. A **84**, 062317 (2011).
[36] R. Y. Q. Cai and V. Scarani, New J. Phys. **11**, 045024 (2009).
[37] T. Richardson and R. Urbanke, IEEE Trans. Inf. Theory **47**, 599 (2001).