Simple and Tight Bounds for Information Reconciliation and Privacy Amplification

Renato Renner[1] and Stefan Wolf[2]

[1] Computer Science Department, ETH Zürich, Switzerland.
`renner@inf.ethz.ch`.

[2] Département d'Informatique et R.O., Université de Montréal, QC, Canada.
`wolf@iro.umontreal.ca`.

Abstract. Shannon entropy is a useful and important measure in information processing, for instance, data compression or randomness extraction, under the assumption—which can typically safely be made in *communication theory*—that a certain random experiment is independently repeated many times. In *cryptography*, however, where a system's working has to be proven with respect to a malicious adversary, this assumption usually translates to a restriction on the latter's knowledge or behavior and is generally not satisfied. An example is quantum key agreement, where the adversary can attack each particle sent through the quantum channel differently or even carry out coherent attacks, combining a number of particles together. In information-theoretic key agreement, the central functionalities of *information reconciliation* and *privacy amplification* have, therefore, been extensively studied in the scenario of *general distributions*: Partial solutions have been given, but the obtained bounds are arbitrarily far from tight, and a full analysis appeared to be rather involved to do. We show that, actually, the general case is not more difficult than the scenario of independent repetitions—in fact, given our new point of view, even simpler. When one analyzes the possible efficiency of data compression and randomness extraction in the case of independent repetitions, then Shannon entropy H is the answer. We show that H can, in these two contexts, be generalized to two *very simple* quantities—H_0^ε and H_∞^ε, called *smooth Rényi entropies*—which are tight bounds for data compression (hence, information reconciliation) and randomness extraction (privacy amplification), respectively. It is shown that the two new quantities, and related notions, do not only extend Shannon entropy in the described contexts, but they also share central properties of the latter such as the chain rule as well as subadditivity and monotonicity.

Key words. Information-theoretic cryptography, entropy measures, data compression, randomness extraction, information reconciliation, privacy amplification, quantum key agreement.

1 Introduction, Motivation, and Main Results

1.1 Unconditional Cryptographic Security and Key Agreement

Unconditional cryptographic security does, in contrast to *computational* security, not depend on any assumption on an adversary's computing power nor on the hardness of computational problems. This type of security is, therefore, not threatened by potential progress in algorithm design or (classical and quantum) computer engineering. On the other hand, cryptographic functionalities such as encryption, authentication, and two- or multi-party computation can generally *not* be realized in an unconditionally secure way simply from scratch. It is, therefore, a natural question under what circumstances—as realistic as possible—they *can* be realized. In particular for encryption and authentication or, more specifically, *secret-key agreement*, this question has been studied extensively: In [23] and [9], unconditional secret key agreement is realized based on the existence of noisy channels between the legitimate partners and the adversary, whereas in [15], a scenario is introduced and studied where all parties have access to pieces of information (e.g., generated by repeated realizations of a certain random experiment). On the other hand, the possibility of information-theoretic key agreement has also been studied between parties connected not only by a classical, but also a quantum channel allowing for the transmission of quantum states [22, 1]. Here, the security can be shown under the condition that the laws of quantum physics are correct.

If, in a certain scenario, unconditional secret-key agreement is possible in principle, then it is a natural question what the maximum length of the generated secret key can be. To find the answer to this question has turned out to often reduce to analyzing two functionalities that form important building blocks of protocols for secret-key agreement (in any of the described settings), namely *information reconciliation* and *privacy amplification*.

Information reconciliation (see, for instance [4]) means that the legitimate partners generate identical shared strings from (possibly only weakly) correlated ones by noiseless and authenticated but public communication, hereby leaking to the adversary only a minimal amount of information about the original and, hence, the resulting string. The generated common but potentially highly compromised string must then be transformed into a virtually secret key by *privacy amplification*. On the technical level—but roughly speaking—, information reconciliation is error correction, whereas privacy amplification is hashing, e.g., by applying a universal hash function [13, 2] or an extractor [16] allowing for distilling a weakly random string's min-entropy H_∞. When these two functionalities are analyzed in a context where all pieces of information stem from many independent repetitions of the same random experiment, then the analysis shows that the amount of information to be exchanged in optimal information reconciliation is the conditional Shannon entropy of, say, one party Alice's information, given the other Bob's; on the other hand, privacy amplification, in the same independent-repetitions setting, allows for extracting a string the length of which equals the conditional Shannon entropy of the shared string given the adversary's information. Hence, as often in information theory, Shannon entropy turns out to be very useful in this asymptotic model. In a (classical or quantum) *cryptographic* context, however, the assumption of independent repetitions typically corresponds to a restriction on the adversary's behavior, and cannot realistically be made. It has been a common belief that in this case, the analysis of the described information-reconciliation and privacy-amplification protocols—and their combination—are quite involved and lead to rather complex (functional) bounds on the (operational) quantities such as the key length. It is the main goal of this paper to show that this is, actually, not the case.

1.2 Information Reconciliation and Privacy Amplification

Information reconciliation is *error correction*: Given that Alice and Bob hold random variables X and Y, respectively, Alice wants to send a minimal quantity of information C to Bob such that given Y and C, he can perfectly reconstruct X with high probability. (More generally, protocols for information reconciliation can use two-way communication. Such interactive protocols can be computationally much more efficient than one-way protocols, but do not reduce the minimal amount of information to be exchanged [4].) To determine the minimal amount of information to be sent from Alice to Bob such that the latter can reconstruct Alice's information with high probability reduces to the following data-compression problem.

Question 1. Given a distribution P_{XY} and $\varepsilon > 0$, what is the minimum length $H^{\varepsilon}_{\mathrm{enc}}(X|Y)$ of a binary string $C = e(X, R)$, computed from X and some additional independent randomness R, such that there exists an event Ω with probability at least $1-\varepsilon$ such that given Ω, X is uniquely determined by C, Y, and R?

Privacy amplification is *randomness extraction*: Given that Alice and Bob both know X and an adversary knows Y, Alice wants to send a message R to Bob such that from X and R, they can compute a (generally shorter) common string S about which the adversary, knowing Y and R but not X, has no information except with small probability. More specifically, privacy amplification deals with the following randomness-extraction problem.

Question 2. Given a distribution P_{XY} and $\varepsilon > 0$, what is the maximum length $H^{\varepsilon}_{\mathrm{ext}}(X|Y)$ of a binary string $S = f(X, R)$, where R is an additional random variable, such that there exists a uniformly distributed random variable U that is independent of (Y, R) together with an event Ω with probability at least $1-\varepsilon$ such that given Ω, we have $S = U$?

The problems of determining $H^{\varepsilon}_{\mathrm{enc}}(X|Y)$ and $H^{\varepsilon}_{\mathrm{ext}}(X|Y)$ have been studied by several authors. Note, first of all, that in the case where the distribution in question is of the form $P_{X^nY^n} = (P_{XY})^n$, corresponding to n independent repetitions of the random experiment P_{XY}, we have, for $\varepsilon > 0$,

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{H^{\varepsilon}_{\mathrm{enc}}(X^n|Y^n)}{n}=\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{H^{\varepsilon}_{\mathrm{ext}}(X^n|Y^n)}{n}=H(X|Y)\ .$$

Interestingly, the two—*a priori* very different—questions have the same answer in this case. We will show that in general, this is not true.

Unfortunately, the assumption that the distribution has product form is generally unrealistic in a cryptographic context: In quantum key agreement, for instance, it corresponds to the assumption that the adversary attacks every particle individually, independently, and in exactly the same way. But what if she does not?

It is fair to say that the problem of optimizing privacy amplification and "distribution uniformizing" has been studied intensively in the general case and considered to be quite involved (see, for instance, [5], [6], [7], and references therein). It is our goal to show that this belief is, both for information reconciliation and privacy amplification, in fact unjustified.

An example of a previous result is that $H^{\varepsilon}_{\mathrm{ext}}(X|Y)$ is bounded from below by the minimum, over all $y \in \mathcal{Y}$, of the so-called *collision entropies* or *Rényi entropies of order* 2, $H_2(X|Y = y)$ (see below for a precise definition) [2]. However, this bound is not tight: For instance, the adversary can be given additional knowledge that increases the H_2-entropy from her viewpoint. In fact, such "spoiling-knowledge" arguments do not only show that the H_2-bound is arbitrarily far from tight, but also that the quantity H_2 has some very counter-intuitive properties that make it hard to handle.

We define two quantities that can be computed very easily and that represent tight bounds on $H^{\varepsilon}_{\mathrm{enc}}$ and $H^{\varepsilon}_{\mathrm{ext}}$, respectively. In a nutshell, we show that the general case is as easy as the special independent-repetitions scenario—or even easier when being looked at it in the right way. We also observe that, in general, the answers to Questions 1 and 2 above are not at all equal.

1.3 Two New Quantities: Conditional Smooth Rényi Entropies and Their Significance

For a distribution P_{XY} and $\varepsilon > 0$, let[3]

$$H^{\varepsilon}_0(X|Y) := \min_{\Omega}\max_{y}\log\left|\{x:\ P_{X\Omega|Y=y}(x)>0\}\right| \tag{1}$$

$$H^{\varepsilon}_{\infty}(X|Y) := \max_{\Omega}\min_{y}\min_{x}\bigl(-\log P_{X\Omega|Y=y}(x)\bigr)\ , \tag{2}$$

where the first minimum/maximum ranges over all events Ω with probability $\Pr[\Omega] \geq 1 - \varepsilon$.

First, we observe that these quantities are defined with respect to P_{XY} in a *very simple way* and are very easy to compute. Indeed, the involved optimization problems can easily be solved by eliminating the smallest probabilities and

[3] All logarithms in this paper are binary. $P_{X\Omega}(x)$ is the probability that Ω occurs and X takes the value x.

by cutting down the largest probabilities, respectively. On the other hand, they provide the answers to Questions 1 and 2 (Section 3).

Answer to Question 1. For $\varepsilon_1 + \varepsilon_2 = \varepsilon$, we have

$$H_0^{\varepsilon}(X|Y) \leq H_{\mathrm{enc}}^{\varepsilon}(X|Y) \leq H_0^{\varepsilon_1}(X|Y) + \log(1/\varepsilon_2) \ .$$

Answer to Question 2. For $\varepsilon_1 + \varepsilon_2 = \varepsilon$, we have

$$H_{\infty}^{\varepsilon_1}(X|Y) - 2\log(1/\varepsilon_2) \leq H_{\mathrm{ext}}^{\varepsilon}(X|Y) \leq H_{\infty}^{\varepsilon}(X|Y) \ .$$

We can say that—modulo a small error term—these results provide simple *functional* representations of the important and natural *operationally* defined quantities $H_{\mathrm{enc}}^{\varepsilon}$ and $H_{\mathrm{ext}}^{\varepsilon}$. In a way, H_0^{ε} (i.e., $H_{\mathrm{enc}}^{\varepsilon}$) and H_{∞}^{ε} ($H_{\mathrm{ext}}^{\varepsilon}$) are two natural generalizations of Shannon entropy to a cryptographic setting with an adversary potentially not following any rules. In particular, both H_0^{ε} and H_{∞}^{ε} fall back to Shannon entropy if the distribution is of the form $(P_{XY})^n$ for large n (Section 2.3). An example of an application of our results is the possibility of analyzing quantum key-agreement protocols or classical protocols based on correlated information. For instance, our results allow for deriving a simple tight bound on the efficiency of key agreement by one-way communication[4] (Section 3.3).

H_0^{ε} and H_{∞}^{ε} are special cases of *smooth Rényi entropies*. In Section 2.1 we give the general definition of conditional and unconditional smooth Rényi entropies of any order α, and in Section 2.2 we show that, roughly speaking, H_{α}^{ε} is, for any α ($\neq 1$), equal to either H_0^{ε} (if $\alpha < 1$) or H_{∞}^{ε} ($\alpha > 1$) up to an additive constant. *Unconditional* smooth Rényi entropy has been introduced in [19], applied in [18], and is, implicitly, widely used in the randomness-extraction literature (see, e.g., [21]). We will show, however, that the *conditional* quantities, introduced in this paper, are the ones that prove particularly useful in the context of cryptography.

If we have concluded that H_0^{ε} and H_{∞}^{ε} generalize Shannon entropy, then this is, in addition, true because they have similar properties (Section 2.4). We summarize the most important ones in a table. (Let $\varepsilon, \varepsilon', \varepsilon_1$, and ε_2 be nonnegative constants. The approximation "$\lessapprox$" holds up to $\log(1/(\varepsilon - \varepsilon_1 - \varepsilon_2))$.)

[4] Our results thus also apply to *fuzzy extractors* [10] which are technically the same as one-way secret-key agreement schemes (where the *generation* and the *reproduction procedures* correspond to the algorithms of Alice and Bob, respectively).

	Shannon entropy H	New entropies H_0^ε and H_∞^ε
chain rule (Lemmas 4 and 5)	$H(X\|Y) = H(XY) - H(Y)$	$H_0^{\varepsilon+\varepsilon'}(XY) - H_0^{\varepsilon'}(Y) \leq H_0^\varepsilon(X\|Y) \lessapprox H_0^{\varepsilon_1}(XY) - H_\infty^{\varepsilon_2}(Y)$ $H_\infty^{\varepsilon_1}(XY) - H_0^{\varepsilon_2}(Y) \lessapprox H_\infty^\varepsilon(X\|Y) \leq H_\infty^{\varepsilon+\varepsilon'}(XY) - H_\infty^{\varepsilon'}(Y)$
sub-additivity (Lemma 6)	$H(XY) \leq H(X) + H(Y)$	$H_0^{\varepsilon+\varepsilon'}(XY) \leq H_0^\varepsilon(X) + H_0^{\varepsilon'}(Y)$ $H_\infty^\varepsilon(XY) \leq H_\infty^{\varepsilon+\varepsilon'}(X) + H_0^{\varepsilon'}(Y)$
monotonicity (Lemma 7)	$H(X) \leq H(XY)$	$H_0^\varepsilon(X) \leq H_0^\varepsilon(XY)$ $H_\infty^\varepsilon(X) \leq H_\infty^\varepsilon(XY)$

Hence, all important properties of Shannon entropy also hold for the new quantities generalizing it. In contrast, note that the important chain rule, for instance, does *not* hold for the original, "non-smooth" Rényi entropies H_0, H_2, and H_∞. In fact, this drawback is one of the reasons for the somewhat limited applicability of these quantities.

The proofs of the above properties of the new, more general, quantities are—just as are their definitions—in fact simpler than the corresponding proofs for Shannon entropy; they only apply counting arguments (instead of, for instance, the concavity of the logarithm function and Jensen's inequality). Since, on the other hand, Shannon entropy is simply a special case of the new quantities (for many independent repetitions), we obtain simpler proofs of the corresponding properties of Shannon entropy for free.

Note that although we state that all smooth Rényi entropies come down to either H_0^ε or H_∞^ε, we give *general* definitions and statements on H_α^ε for any α. This can be convenient in contexts in which the entropies have a natural significance, such as H_2 in connection with two-universal hashing [2].

2 Smooth Rényi Entropy: Definition and Properties

2.1 Definition

We start by briefly reviewing the notion of *smooth Rényi entropy* [19] and then generalize it to *conditional* smooth Rényi entropy.

Let X be a random variable on $\mathcal{X}$ with probability distribution P_X. We denote by $\mathcal{B}^\varepsilon(P_X)$ the set of non-negative functions Q_X with domain $\mathcal{X}$ such that $Q_X(x) \leq P_X(x)$, for any $x \in \mathcal{X}$, and $\sum_{x\in\mathcal{X}} Q_X(x) \geq 1 - \varepsilon$. The *$\varepsilon$-smooth*

Rényi entropy of order α, for $\alpha \in (0,1) \cup (1,\infty)$ and $\varepsilon \geq 0$, is defined by[5]

$$H_\alpha^\varepsilon(X) := \frac{1}{1-\alpha} \log r_\alpha^\varepsilon(X) ,$$

where

$$r_\alpha^\varepsilon(X) := \inf_{Q_X \in \mathcal{B}^\varepsilon(P_X)} \sum_{x \in \mathcal{X}} Q_X(x)^\alpha .$$

For $\alpha = 0$ and $\alpha = \infty$, smooth Rényi entropy is defined by the limit values, i.e., $H_0^\varepsilon(X) := \lim_{\alpha \to 0} H_\alpha^\varepsilon(X)$ and $H_\infty^\varepsilon(X) := \lim_{\alpha \to \infty} H_\alpha^\varepsilon(X)$.

It follows directly from the definition that, for $\alpha < 1$,

$$\varepsilon \geq \varepsilon' \longleftrightarrow H_\alpha^\varepsilon(X) \leq H_\alpha^{\varepsilon'}(X)$$

holds and, similarly, for $\alpha > 1$,

$$\varepsilon \geq \varepsilon' \longleftrightarrow H_\alpha^\varepsilon(X) \geq H_\alpha^{\varepsilon'}(X) .$$

Moreover, for $\varepsilon = 0$, smooth Rényi entropy $H_\alpha^0(X)$ is equal to "conventional" Rényi entropy $H_\alpha(X)$ [20]. Similarly to conditional Shannon entropy, we define a *conditional* version of smooth Rényi entropy.

Definition 1. *Let X and Y be random variables with range $\mathcal{X}$ and $\mathcal{Y}$, respectively, and joint probability distribution P_{XY}. The* conditional ε-smooth Rényi entropy of order α of X given Y, *for $\alpha \in (0,1) \cup (1,\infty)$ and $\varepsilon \geq 0$, is defined by*

$$H_\alpha^\varepsilon(X|Y) := \frac{1}{1-\alpha} \log r_\alpha^\varepsilon(X|Y)$$

where

$$r_\alpha^\varepsilon(X|Y) := \inf_{Q_{XY} \in \mathcal{B}^\varepsilon(P_{XY})} \max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} Q_{X|Y=y}(x)^\alpha ,$$

and where $Q_{X|Y=y}(x) := Q_{XY}(x,y)/P_Y(y)$, for any $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ (with the convention $Q_{X|Y=y}(x) = 0$ if $P_Y(y) = 0$).[6] *For $\alpha = 0$ and $\alpha = \infty$, we define $H_0^\varepsilon(X|Y) := \lim_{\alpha \to 0} H_\alpha^\varepsilon(X|Y)$ and $H_\infty^\varepsilon(X|Y) := \lim_{\alpha \to \infty} H_\alpha^\varepsilon(X|Y)$.*

For $\alpha = 0$ and $\alpha = \infty$, Definition 1 reduces to (1) and (2), respectively. Note that the *infimum* is in fact a *minimum* which is obtained by cutting away the smallest probabilities or cutting down the largest, respectively.

[5] The definition given here slightly differs from the original definition in [19]. However, it turns out that this version is more appropriate for our generalization to conditional smooth Rényi entropy (Definition 1).

[6] Since $\sum_x Q_{XY}(x,y)$ is generally smaller than $P_Y(y)$, the distribution $Q_{X|Y=y}(\cdot) := Q_{XY}(\cdot,y)/P_Y(y)$ is not necessarily normalized.

2.2 Basic Properties

We will now derive some basic properties of smooth Rényi entropy. In particular, we show that the smooth Rényi entropies can be split into two classes: It turns out that for any value $\alpha < 1$, $H_\alpha^\varepsilon(X|Y)$ is, up to an additive constant, equal to $H_0^\varepsilon(X|Y)$. Similarly, $H_\alpha^\varepsilon(X|Y)$, for $\alpha > 1$, is essentially $H_\infty^\varepsilon(X|Y)$.

For this, we need a generalization, to the smooth case, of the fact that

$$\alpha \leq \beta \longleftrightarrow H_\alpha(X) \geq H_\beta(X) \tag{3}$$

holds for any $\alpha, \beta \in [0, \infty]$.

Lemma 1. *Let X and Y be random variables. Then, for $\varepsilon \geq 0$ and for $\alpha \leq \beta < 1$ or $1 < \alpha \leq \beta$,*

$$H_\alpha^\varepsilon(X|Y) \ \geq \ H_\beta^\varepsilon(X|Y) \ .$$

Proof. For any probability distribution Q on $\mathcal{X}$, the right hand side of (3) can be rewritten as

$$\sqrt[1-\alpha]{\sum_{x\in\mathcal{X}} Q(x)^\alpha} \geq \sqrt[1-\beta]{\sum_{x\in\mathcal{X}} Q(x)^\beta} \ . \tag{4}$$

It is easy to verify that this inequality also holds for any (not necessarily normalized) nonnegative function Q with $\sum_{x\in\mathcal{X}} Q(x) \leq 1$.

As mentioned above, the infimum in the definition of r_α^ε is actually a minimum. Hence, there exists $Q_{XY} \in \mathcal{B}^\varepsilon(P_{XY})$ such that for any $y \in \mathcal{Y}$,

$$r_\alpha^\varepsilon(X|Y) \geq \sum_{x\in\mathcal{X}} Q_{X|Y=y}(x)^\alpha$$

holds. When this is combined with (4), we find

$$\sqrt[1-\alpha]{r_\alpha^\varepsilon(X|Y)} \geq \sqrt[1-\alpha]{\sum_{x\in\mathcal{X}} Q_{X|Y=y}(x)^\alpha} \geq \sqrt[1-\beta]{\sum_{x\in\mathcal{X}} Q_{X|Y=y}(x)^\beta} \ .$$

Because this holds for any $y \in \mathcal{Y}$, we conclude

$$\sqrt[1-\alpha]{r_\alpha^\varepsilon(X|Y)} \geq \sqrt[1-\beta]{r_\beta^\varepsilon(X|Y)} \ .$$

The assertion now follows from the definition of smooth Rényi entropy. □

Lemma 2 is, in some sense, the converse of Lemma 1. Since it is a straightforward generalization of a statement of [19][7], we omit the proof here.

Lemma 2. *Let X and Y be random variables. Then, for $\varepsilon \geq 0$, $\varepsilon' \geq 0$, and $\alpha < 1$, we have*

$$H_0^{\varepsilon+\varepsilon'}(X|Y) \leq H_\alpha^\varepsilon(X|Y) + \frac{\log(1/\varepsilon')}{1-\alpha}$$

and for $\alpha > 1$,

$$H_\infty^{\varepsilon+\varepsilon'}(X|Y) \geq H_\alpha^\varepsilon(X|Y) - \frac{\log(1/\varepsilon')}{\alpha - 1} \ .$$

[7] The result of [19] corresponds to the special case where Y is a constant.

When Lemmas 1 and 2 are combined, we obtain the following characterization of smooth Rényi entropy $H_\alpha^\varepsilon(X|Y)$, for $\alpha < 1$, in terms of smooth Rényi entropy of order 0:

$$H_0^{\varepsilon+\varepsilon'}(X|Y) - \frac{\log(1/\varepsilon')}{1-\alpha} \le H_\alpha^\varepsilon(X|Y) \le H_0^\varepsilon(X|Y) \ .$$

Similarly, for $\alpha > 1$,

$$H_\infty^{\varepsilon+\varepsilon'}(X|Y) + \frac{\log(1/\varepsilon')}{\alpha-1} \ge H_\alpha^\varepsilon(X|Y) \ge H_\infty^\varepsilon(X|Y) \ .$$

If $\varepsilon = 0$, this leads to an approximation of the (conventional) Rényi entropy H_α, of any order α, in terms of the smooth Rényi entropies H_0^ε and H_∞^ε. For example, the collision entropy $H_2(X)$ cannot be larger than $H_\infty^\varepsilon(X) + \log(1/\varepsilon)$ (whereas $H_2(X) \approx 2H_\infty(X)$, for certain probability distributions P_X).

2.3 Smooth Rényi Entropy as a Generalization of Shannon Entropy

Interestingly, one obtains as an immediate consequence of the asymptotic equipartition property (AEP) (cf. [8]) that, for many independent realizations of a random experiment, smooth Rényi entropy is asymptotically equal to Shannon entropy. (Note that the same is not true at all for the usual Rényi entropies.)

Lemma 3. *Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be n independent pairs of random variables distributed according to P_{XY}. Then we have, for any $\alpha \neq 1$,*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{H_\alpha^\varepsilon(X^n|Y^n)}{n} = H(X|Y) \ ,$$

where $H(X|Y)$ is the conditional Shannon entropy.

For a proof as well as a more detailed (non-asymptotic) version of this statement, we refer to [12].

2.4 Shannon-like Properties of Smooth Rényi Entropy

Smooth Rényi entropy shares basic properties with Shannon entropy—this is in contrast to the usual Rényi entropies, which do not have these properties. Therefore, the smooth versions are much more natural and useful quantities in many contexts, as we will see.

Chain Rule We first prove a property corresponding to the chain rule $H(X|Y) = H(XY) - H(Y)$ of Shannon entropy. More precisely, Lemmas 4 and 5 below are two different inequalities, which, combined, give a chain rule for smooth Rényi entropies of any order α.

Lemma 4. *Let X and Y be random variables and let $\varepsilon \geq 0$, $\varepsilon' \geq 0$, $\varepsilon'' \geq 0$. Then, for $\alpha < 1 < \beta$, we have*

$$H_\alpha^{\varepsilon+\varepsilon'+\varepsilon''}(X|Y) < H_\alpha^{\varepsilon'}(XY) - H_\beta^{\varepsilon''}(Y) + \frac{\beta-\alpha}{(1-\alpha)(\beta-1)} \log(1/\varepsilon) \ ,$$

and, similarly, for $\alpha > 1 > \beta$,

$$H_\alpha^{\varepsilon+\varepsilon'+\varepsilon''}(X|Y) > H_\alpha^{\varepsilon'}(XY) - H_\beta^{\varepsilon''}(Y) - \frac{\alpha-\beta}{(\alpha-1)(1-\beta)} \log(1/\varepsilon) \ .$$

Proof. It is easy to verify that the assertion can be rewritten as

$$\log r_\alpha^{\varepsilon+\varepsilon'+\varepsilon''}(X|Y) < \log r_\alpha^{\varepsilon'}(XY) + \frac{1-\alpha}{\beta-1} \log r_\beta^{\varepsilon''}(Y) + \frac{\beta-\alpha}{\beta-1} \log(1/\varepsilon) \ . \quad (5)$$

By the definition of $r_\alpha^{\varepsilon'}(XY)$ there exists an event Ω_1 with probability $\Pr[\Omega_1] = 1-\varepsilon'$ such that $r_\alpha^{\varepsilon'}(XY) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY\Omega_1}(x,y)^\alpha$. Similarly, one can find an event Ω_2 such that $\Pr[\Omega_2] = 1-\varepsilon''$ and $r_\beta^{\varepsilon''}(Y) = \sum_{y\in\mathcal{Y}} P_{Y\Omega_2}(y)^\beta$. Hence, the event $\Omega := \Omega_1 \cap \Omega_2$ has probability $\Pr[\Omega] \geq 1-\varepsilon'-\varepsilon''$ and satisfies

$$\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY\Omega}(x,y)^\alpha \leq r_\alpha^{\varepsilon'}(XY)$$

as well as

$$\sum_{y\in\mathcal{Y}} P_{Y\Omega}(y)^\beta \leq r_\beta^{\varepsilon''}(Y) \ .$$

For any $y \in \mathcal{Y}$, let $\bar{r}_y := \sum_{x\in\mathcal{X}} P_{X\Omega|Y=y}(x)^\alpha$. Since inequality (5) is independent of the labeling of the values in $\mathcal{Y}$, we can assume without loss of generality that these are natural numbers, $\mathcal{Y} = \{1,\ldots,n\}$, for $n := |\mathcal{Y}|$, and that the values $\bar{r}_y$ are arranged in increasing order, $\bar{r}_y > \bar{r}_{y'} \longrightarrow y > y'$. Let $\bar{y} \in \mathcal{Y}$ be the minimum value such that $\Pr[Y > \bar{y}, \Omega] \leq \varepsilon$ holds. In particular,

$$\Pr[Y \geq \bar{y}, \Omega] = \Pr[Y > \bar{y}-1, \Omega] > \varepsilon \ . \quad (6)$$

Let Ω' be the event that $Y \leq \bar{y}$ holds, i.e., we have $\Pr[\overline{\Omega'}, \Omega] \leq \varepsilon$ and, consequently,

$$\Pr[\Omega', \Omega] = 1 - \Pr[\overline{\Omega}] - \Pr[\overline{\Omega'}, \Omega] \geq 1 - \varepsilon - \varepsilon' - \varepsilon'' \ .$$

Hence, since $P_{X\Omega\Omega'|Y=y}(x) = 0$ holds for any $x \in \mathcal{X}$ and $y > \bar{y}$, we have

$$r_\alpha^{\varepsilon+\varepsilon'+\varepsilon''}(X|Y) \leq \max_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} P_{X\Omega\Omega'|Y=y}(x)^\alpha \leq \max_{y\leq\bar{y}} \bar{r}_y \leq \bar{r}_{\bar{y}} \ .$$

Therefore, it remains to be proven that

$$\log \bar{r}_{\bar{y}} < \log\biggl(\sum_{x\in\mathcal{X},y\in\mathcal{Y}} P_{XY\Omega}(x,y)^\alpha\biggr) + \frac{1-\alpha}{\beta-1} \log\biggl(\sum_{y\in\mathcal{Y}} P_{Y\Omega}(y)^\beta\biggr) - \frac{\beta-\alpha}{\beta-1} \log\varepsilon \ . \quad (7)$$

Let $s := \sum_{y=\bar{y}}^{n} P_Y(y)^\alpha$. Then,

$$\bar{r}_{\bar{y}} \cdot s = \sum_{y=\bar{y}}^{n} \bar{r}_{\bar{y}} P_Y(y)^\alpha \leq \sum_{y=\bar{y}}^{n} \bar{r}_y P_Y(y)^\alpha \leq \sum_{y=1}^{n} \bar{r}_y P_Y(y)^\alpha \ , \tag{8}$$

where the first inequality follows from the fact that $\bar{r}_y \geq \bar{r}_{\bar{y}}$ holds for all $y \geq \bar{y}$. When the definition of $\bar{r}_y$ is inserted into inequality (8), we get

$$\bar{r}_{\bar{y}} \leq \frac{1}{s} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{XY\Omega}(x,y)^\alpha$$

i.e.,

$$\log \bar{r}_{\bar{y}} \leq \log\left(\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{XY\Omega}(x,y)^\alpha \right) - \log s \ . \tag{9}$$

In order to find a bound on s, let $p_y := P_{Y\Omega}(y)$, $p := \frac{\beta-\alpha}{\beta-1}$, $q := \frac{\beta-\alpha}{1-\alpha}$, and $\gamma := \frac{\alpha(\beta-1)}{\beta-\alpha}$, i.e., $\gamma p = \alpha$ and $(1-\gamma)q = \beta$. We then have $\frac{1}{p} + \frac{1}{q} = 1$ and can apply *Hölder's inequality*, yielding

$$\begin{aligned}
\sqrt[p]{s} \cdot \sqrt[q]{\sum_{y \in \mathcal{Y}} P_{Y\Omega}(y)^\beta} &\geq \sqrt[p]{\sum_{y=\bar{y}}^{n} (p_y)^\alpha} \cdot \sqrt[q]{\sum_{y=\bar{y}}^{n} (p_y)^\beta} \\
&= \sqrt[p]{\sum_{y=\bar{y}}^{n} \big((p_y)^\gamma\big)^p} \cdot \sqrt[q]{\sum_{y=\bar{y}}^{n} \big((p_y)^{1-\gamma}\big)^q} \\
&\geq \sum_{y=\bar{y}}^{n} (p_y)^\gamma (p_y)^{1-\gamma} = \sum_{y=\bar{y}}^{n} p_y = \Pr[Y \geq \bar{y}, \Omega] > \varepsilon \ .
\end{aligned}$$

Hence,

$$\log s > p \log \varepsilon - \frac{p}{q} \log\left(\sum_{y \in \mathcal{Y}} P_{Y\Omega}(y)^\beta \right) \ .$$

Combining this with (9) implies (7) and, thus, concludes the proof. □

Lemma 5. *Let X and Y be random variables and let $\varepsilon \geq 0$, $\varepsilon' \geq 0$. Then, for any $\alpha < 1$, we have*

$$H_\alpha^{\varepsilon+\varepsilon'}(XY) \leq H_\alpha^{\varepsilon}(X|Y) + H_\alpha^{\varepsilon'}(Y) \ ,$$

and, similarly, for $\alpha > 1$,

$$H_\alpha^{\varepsilon+\varepsilon'}(XY) \geq H_\alpha^{\varepsilon}(X|Y) + H_\alpha^{\varepsilon'}(Y) \ .$$

Proof. Let Ω be an event with $\Pr[\Omega] \geq 1 - \varepsilon$ such that

$$\max_y \sum_{x \in \mathcal{X}} P_{X\Omega|Y=y}(x)^\alpha \leq r_\alpha^\varepsilon(X|Y) \ .$$

Similarly, let Ω' be an event with $\Pr[\Omega'] \geq 1 - \varepsilon'$ such that $\Omega' \leftrightarrow Y \leftrightarrow (X, \Omega)$ is a Markov chain and

$$\sum_{y \in \mathcal{Y}} P_{Y\Omega'}(y)^\alpha \leq r_\alpha^\varepsilon(Y) \ .$$

Since $\Pr[\Omega, \Omega'] \geq 1 - \varepsilon - \varepsilon'$ holds, we have

$$r_\alpha^{\varepsilon+\varepsilon'}(XY) \leq \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{XY\Omega\Omega'}(x, y)^\alpha \ .$$

The assertion thus follows from

$$\begin{aligned} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{XY\Omega\Omega'}(x, y)^\alpha &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{Y\Omega'}(y)^\alpha P_{X\Omega|Y=y}(x)^\alpha \\ &\leq \left(\sum_{y \in \mathcal{Y}} P_{Y\Omega'}(y)^\alpha \right) \left(\max_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} P_{X\Omega|Y=y}(x)^\alpha \right) \ . \end{aligned}$$

□

It is easy to see that the statements of Lemma 4 and Lemma 5 still hold if all entropies are conditioned on an additional random variable Z. For example, the statement of Lemma 5 then reads, for $\alpha < 1$,

$$H_\alpha^{\varepsilon+\varepsilon'}(XY|Z) - H_\alpha^{\varepsilon'}(Y|Z) \leq H_\alpha^\varepsilon(X|YZ) \tag{10}$$

and for $\alpha > 1$,

$$H_\alpha^{\varepsilon+\varepsilon'}(XY|Z) - H_\alpha^{\varepsilon'}(Y|Z) \geq H_\alpha^\varepsilon(X|YZ) \ . \tag{11}$$

Sub-Additivity The Shannon entropy $H(XY)$ of a pair of random variables X and Y cannot be larger than the sum $H(X) + H(Y)$. The following statement generalizes this sub-additivity property to smooth Rényi entropy. The proof of this statement is straightforward and, in fact, very similar to the (simple) proof of Lemma 5.

Lemma 6. *Let X and Y be random variables and let $\varepsilon \geq 0$. Then, for any $\alpha < 1$,*

$$H_\alpha^{\varepsilon+\varepsilon'}(XY) \leq H_\alpha^\varepsilon(X) + H_0^{\varepsilon'}(Y)$$

holds. Similarly, for $\alpha > 1$, we have

$$H_\alpha^\varepsilon(XY) \leq H_\alpha^{\varepsilon+\varepsilon'}(X) + H_0^{\varepsilon'}(Y) \ .$$

Monotonicity The uncertainty on a pair of random variables X and Y cannot be smaller than the uncertainty on X alone. This is formalized by the following lemma. The proof is again similar to Lemma 5.

Lemma 7. *Let X and Y be random variables and let $\varepsilon \geq 0$. Then, for $\alpha \neq 1$, we have*

$$H_\alpha^\varepsilon(X) \leq H_\alpha^\varepsilon(XY) \ .$$

In particular, the smooth Rényi entropy does not increase when a function is applied:

$$H_\alpha^\varepsilon(f(X)) \leq H_\alpha^\varepsilon(X) \ . \tag{12}$$

Independence, Conditional Independence, and Markov Chains Conditioning on independent randomness cannot have any effect on the entropy.

Lemma 8. *Let X and Y be independent random variables and let $\varepsilon \geq 0$, $\varepsilon' \geq 0$. Then, for any $\alpha \neq 1$, we have*

$$H_\alpha^\varepsilon(X|Y) = H_\alpha^\varepsilon(X) \ .$$

This statement can be generalized to random variables X, Y, and Z such that $X \leftrightarrow Z \leftrightarrow Y$ is a Markov chain:

$$H_\alpha^\varepsilon(X|YZ) = H_\alpha^\varepsilon(X|Z) \ .$$

When this is combined with inequalities (10) and (11), we obtain, for $\alpha < 1$,

$$H_\alpha^{\varepsilon+\varepsilon'}(XY|Z) \leq H_\alpha^\varepsilon(X|Z) + H_\alpha^{\varepsilon'}(Y|Z)$$

and, for $\alpha > 1$,

$$H_\alpha^{\varepsilon+\varepsilon'}(XY|Z) \geq H_\alpha^\varepsilon(X|Z) + H_\alpha^{\varepsilon'}(Y|Z) \ .$$

3 Smooth Rényi Entropy in Cryptography

3.1 Randomness Extraction and Privacy Amplification

The problem of extracting uniform randomness from a non-uniform source has first been studied in [3, 13], and later been defined explicitly in [16]. Today, randomness extraction is a well-known and widely-used concept in theoretical computer science and, in particular, cryptography. A (strong) extractor is a function f which takes as input a random variable X and some additional uniformly distributed randomness R and is such that if X satisfies a certain entropy condition, the output $S := f(X, R)$ is almost independent of R and uniformly distributed.

For two random variables Z and W with joint distribution P_{ZW}, we define the *distance from uniform* by $d(Z|W) := \frac{1}{2}\delta(P_{ZW}, P_U \times P_W)$ where P_U is the uniform distribution on the range of Z and where $\delta(\cdot,\cdot)$ denotes the statistical distance.[8]

Definition 2. *A* strong $(\tau, \kappa, \varepsilon)$-extractor *on a set $\mathcal{X}$ is a function with domain $\mathcal{X} \times \mathcal{R}$ (for a set $\mathcal{R}$) and range $\mathcal{U}$ of size $|\mathcal{U}| = 2^\tau$ such that, for any random variable X on $\mathcal{X}$ satisfying $H_\infty(X) \geq \kappa$ and R uniformly distributed over $\mathcal{R}$, $d(f(X,R)|R) \leq \varepsilon$ holds.*

The following result has originally been proven in [13] based on two-universal hashing (where the randomness R is used to select a function from a two-universal[9] class of functions.). Later, similar statements have been shown in [2] and [11].[10]

[8] The *statistical distance* between two probability distributions P and Q is defined by $\delta(P,Q) := \frac{1}{2}\sum_v |P(v) - Q(v)|$.

[9] A *two-universal class of functions* from $\mathcal{Z}$ to $\mathcal{W}$ is a family $\mathcal{F}$ of functions $f : \mathcal{Z} \mapsto \mathcal{W}$ such that for any $z \neq z'$ and for f chosen at random from $\mathcal{F}$, $\Pr[f(z) = f(z')] \leq \frac{1}{|\mathcal{W}|}$.

[10] For a simple proof of Lemma 9, see, e.g., [14], p. 20.

Lemma 9 (Leftover hash lemma). *For any $\kappa > \tau$, there exists a strong $(\tau, \kappa, 2^{-(\kappa-\tau)/2})$-extractor.*

The following measure is closely related to *smooth entropy* as defined in [7] and [5]. For a distribution P_{XY}, it quantifies the amount of uniform randomness, conditioned on Y, which can be extracted from X.

Definition 3. *Let X and Y be random variables and let $\varepsilon \geq 0$. The ε-*extractable randomness *of X* conditioned on Y *is*

$$H_{\text{ext}}^{\varepsilon}(X|Y) := \max_{\mathcal{U}:\, \exists f \in \Gamma_{XY}^{\varepsilon}(\mathcal{X}\to\mathcal{U})} \log|\mathcal{U}| \ ,$$

where $\Gamma_{XY}^{\varepsilon}(\mathcal{X} \to \mathcal{U})$ denotes the set of functions f from $\mathcal{X} \times \mathcal{R}$ (for some set $\mathcal{R}$) to $\mathcal{U}$ such that $d(f(X,R)|YR) \leq \varepsilon$ holds, for R independent of (X,Y) and uniformly distributed on $\mathcal{R}$.

As mentioned in the introduction, smooth Rényi entropy equals the amount of extractable uniform randomness, up to some small additive constant. Here, the lower bound follows directly from the leftover hash lemma and the definition of H_∞^ε. The upper bound, on the other hand, is a special case of the bound on one-way key agreement derived in Section 3.3.

Theorem 1. *Let X and Y be random variables and let $\varepsilon \geq 0$, $\varepsilon' \geq 0$. Then we have*

$$H_\infty^\varepsilon(X|Y) - 2\log(1/\varepsilon') \leq H_{\text{ext}}^{\varepsilon+\varepsilon'}(X|Y) \leq H_\infty^{\varepsilon+\varepsilon'}(X|Y) \ .$$

Using Lemma 2, we can, in particular, conclude that Rényi entropy of order α, for any $\alpha > 1$, is a lower bound on the number of uniform random bits that can be extracted, i.e.,

$$H_\alpha(X|Y) - \frac{\log(1/\varepsilon)}{\alpha - 1} - 2\log(1/\varepsilon') \leq H_{\text{ext}}^{\varepsilon+\varepsilon'}(X|Y) \ .$$

3.2 Data Compression, Error Correction, and Information Reconciliation

Another fundamental property of a probability distribution P is the minimum length of an encoding $C = E(X)$ of a random variable X with $P_X = P$ such that X can be retrieved from C with high probability. (A similar quantity can be defined for *a set* $\mathcal{P}$ of probability distributions.) As a motivating example, consider the following setting known as *information reconciliation* [4].[11] An entity (Alice) holds a value X which she wants to transmit to another (Bob), using τ bits of communication C. Clearly the minimum number τ of bits needed depends on the initial knowledge of Bob, which might be specified by some additional random variable Y (not necessarily known to Alice). From Bob's point of view,

[11] In certain cryptographic applications, (one-way) information reconciliation schemes are also called *secure sketches* [10] (where Bob's procedure is the *recovery function*).

the random variable X is thus initially distributed according to $P_{X|Y=y}$ for some $y \in \mathcal{Y}$. Consequently, in order to guarantee that Bob can reconstruct the value of X with high probability, the error correcting information C sent by Alice must be useful for most of the distributions $P_{X|Y=y}$.

For the following, note that any probabilistic encoding function E corresponds to a deterministic function e taking as input some additional randomness R, i.e., $E(X) = e(X, R)$.

Definition 4. *A $(\tau, \kappa, \varepsilon)$-*encoding *on a set $\mathcal{X}$ is a pair of functions (e, g) together with a random variable R with range $\mathcal{R}$ where e, the* encoding function*, is a mapping from $\mathcal{X} \times \mathcal{R}$ to $\mathcal{C}$, for some set $\mathcal{C}$ of size $|\mathcal{C}| = 2^\tau$, and g, the* decoding function*, is a mapping from $\mathcal{C} \times \mathcal{R}$ to $\mathcal{X}$ such that, for any random variable X with range $\mathcal{X}$ satisfying $H_0(X) \leq \kappa$, $\Pr[g(e(X, R), R) \neq X] \leq \varepsilon$ holds.*

The following result has originally been shown in the context of information reconciliation [4].

Lemma 10. *For any $\tau > \kappa$, there exists a $(\tau, \kappa, 2^{-(\tau-\kappa)})$-encoding.*

For a distribution P_{XY}, the measure defined below quantifies the minimum length of an encoding $C = e(X, R)$ of X such that X can be reconstructed from C, Y, and R (with high probability).

Definition 5. *Let X and Y be random variables and let $\varepsilon \geq 0$. The ε-*encoding length *of X given Y is*

$$H^\varepsilon_{\mathrm{enc}}(X|Y) := \min_{\mathcal{C}:\, \exists e \in \Lambda^\varepsilon_{XY}(\mathcal{X} \to \mathcal{C})} \log |\mathcal{C}|$$

where $\Lambda^\varepsilon_{XY}(\mathcal{X} \to \mathcal{C})$ denotes the set of function e from $\mathcal{X} \times \mathcal{R}$ (for some set $\mathcal{R}$) to $\mathcal{C}$ such that there exists a decoding function g from $\mathcal{Y} \times \mathcal{C} \times \mathcal{R}$ to $\mathcal{X}$ such that $\Pr[g(Y, e(X, R), R) \neq X] \leq \varepsilon$ holds, for R independent of (X, Y) and uniformly distributed on $\mathcal{R}$.

Similarly to the amount of extractable randomness, smooth Rényi entropy can also be used to characterize the minimum encoding length.

Theorem 2. *Let X and Y be random variables and let $\varepsilon \geq 0$, $\varepsilon' \geq 0$. Then we have*

$$H_0^{\varepsilon+\varepsilon'}(X|Y) \leq H_{\mathrm{enc}}^{\varepsilon+\varepsilon'}(X|Y) \leq H_0^\varepsilon(X|Y) + \log(1/\varepsilon') \ .$$

3.3 A Tight Bound for Key Agreement by One-Way Communication

As an application of Theorems 1 and 2, we prove tight bounds on the maximum length of a secret key that can be generated from partially secret and weakly correlated randomness by one-way communication.

Let X, Y, and Z be random variables. For $\varepsilon \geq 0$, define

$$M^\varepsilon(X;Y|Z) := \sup_{V \leftrightarrow U \leftrightarrow X \leftrightarrow (Y,Z)} H^\varepsilon_\infty(U|ZV) - H^\varepsilon_0(U|YV) \ . \tag{13}$$

Note that this is equivalent to[12]

$$M^{\varepsilon}(X;Y|Z) = \sup_{(U,V)\leftrightarrow X\leftrightarrow(Y,Z)} H_\infty^{\varepsilon}(U|ZV) - H_0^{\varepsilon}(U|YV) \ . \tag{14}$$

Consider now a setting where two parties, Alice and Bob, hold information X and Y, respectively, while the knowledge of an adversary Eve is given by Z. Additionally, they are connected by a public but authenticated one-way communication channel from Alice to Bob, and their goal is to generate an ε-secure key pair (S_A, S_B). Let $S^{\varepsilon}(X \to Y||Z)$ be the maximum length of an ε-secure key that can be generated in this situation. Here, ε-secure means that, except with probability ε, Alice and Bob's keys are equal to a perfect key which is uniformly distributed and independent of Eve's information. Note that, if $\Pr[S_A \neq S_B] \leq \varepsilon_1$ and $d(S_A|W) \leq \varepsilon_2$, where W summarizes Eve's knowledge after the protocol execution, then the pair (S_A, S_B) is ε-secure, for $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Theorem 3. *Let X, Y, and Z be random variables. Then, for $\varepsilon \geq 0$ and $\varepsilon' = \Theta(\varepsilon)$, we have*

$$M^{\varepsilon'}(X;Y|Z) - O(\log(1/\varepsilon')) \leq S^{\varepsilon}(X \to Y||Z) \leq M^{\varepsilon}(X;Y|Z) \ .$$

Proof. We first show that the measure $M^{\varepsilon'}(X;Y|Z)$ is a *lower bound* on the number of ε-secure bits that can be generated. To see this, consider the following simple three-step protocol.

1. *Pre-processing:* Alice computes U and V from X. She sends V to Bob and keeps U.
2. *Information reconciliation:* Alice sends error-correcting information to Bob. Bob uses this information together with Y and V to compute a guess $\hat{U}$ of U.
3. *Privacy amplification:* Alice chooses a hash function F and sends a description of F to Bob. Alice and Bob then compute $S_A := F(U)$ and $S_B := F(\hat{U})$, respectively.

It follows immediately from the analysis of information reconciliation and privacy amplification that the parameters of the protocol (i.e., the amount of error correcting information and the size of the final keys) can be chosen such that the final keys have length $M^{\varepsilon'}(X;Y|Z)$ and the key pair (S_A, S_B) is ε-secure.

On the other hand, it is easy to see that *any* measure $M^{\varepsilon}(X;Y|Z)$ is an *upper* bound on the amount of key bits that can be generated if the following conditions, which imply that the quantity cannot increase during the execution of any protocol, are satisfied:

[12] To see that the measure defined by (14) is not larger than the measure defined by (13), observe that the entropies on the right-hand side of (14) do not change when the random variable U is replaced by $U' := (U, V)$. This random variable U' then satisfies $V \leftrightarrow U' \leftrightarrow X \leftrightarrow (Y, Z)$.

1. $M^\varepsilon(X;Y|Z) \geq M^\varepsilon(X';Y|Z)$ for any X' computed from X.
2. $M^\varepsilon(X;Y|Z) \geq M^\varepsilon(X;Y'|Z)$ for any Y' computed from Y.
3. $M^\varepsilon(X;Y|Z) \geq M^\varepsilon(X;YC|ZC)$ for any C computed from X.
4. $M^\varepsilon(X;Y|Z) \leq M^\varepsilon(X;Y|Z')$ for any Z' computed from Z.
5. $M^\varepsilon(S_A;S_B\,|W) \geq n$ if the pair (S_A,S_B) is ε-secure with respect to an adversary knowing W.

The measure $M^\varepsilon(X;Y|Z)$ defined by (13) does in fact satisfy these properties. It is thus an upper bound on the length of an ε-secure key which can be generated by Alice and Bob.

Property 1 holds since any pair of random variables U and V that can be computed from X' can also be computed from X.

Property 2 follows from $H_0^\varepsilon(A|BC) \leq H_0^\varepsilon(A|B)$.

Property 3 holds since $M^\varepsilon(X;YC|ZC)$ can be written as the supremum over U and V' of $H_\infty^\varepsilon(U|ZV') - H_0^\varepsilon(U|YV')$, where V' is restricted to values of the form $V' = (V,C)$.

Property 4 follows from $H_\infty^\varepsilon(A|BC) \leq H_\infty^\varepsilon(A|B)$.

Property 5 follows from $M^\varepsilon(S_A;S_B|Z) \geq H_\infty^\varepsilon(S_A|Z) - H_0^\varepsilon(S_A|S_B)$, $H_\infty^\varepsilon(S_A|Z) \geq n$, and $H_0^\varepsilon(S_A|S_B) = 0$.

□

4 Concluding Remarks

We have analyzed data compression and randomness extraction in the cryptographic scenario where the assumption, usually made in classical information and communication theory, that the pieces of information stem from a large number of repetitions of a random experiment, has to be dropped. We have shown that Shannon entropy—the key quantity in independent-repetitions settings—then generalizes, depending on the context, to two different entropy measures H_0^ε and H_∞^ε. These new quantities, which are tight bounds on the optimal length of the compressed data and of the extracted random string, respectively, are very simple—in fact, simpler than Shannon information. Indeed, they can be computed from the distribution simply by leaving away the smallest probabilities or cutting down the largest ones, respectively. Moreover, the new quantities share all central properties of Shannon entropy.

An application of our results is the possibility of a simple yet general and tight analysis of protocols for quantum (see, e.g., [17]) and classical key agreement, where no assumption on an adversary's behavior has to be made. For instance, we give a simple tight bound for the possibility and efficiency of secret-key agreement by one-way communication.

It is conceivable that the new quantities have further applications in cryptography and in communication and information theory in general. We suggest as an open problem to find such contexts and applications.

Acknowledgment

The authors would like to thank Ueli Maurer for inspiring and helpful discussions.

References

1. C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, pp. 175–179. IEEE, 1984.
2. C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, Generalized privacy amplification, *IEEE Transactions on Information Theory*, Vol. 41, No. 6, 1915–1923, 1995.
3. C. H. Bennett, G. Brassard, and J. M. Robert, Privacy amplification by public discussion. *SIAM Journal on Computing*, Vol. 17, pp. 210–229, 1988.
4. G. Brassard and L. Salvail, Secret-key reconciliation by public discussion, *EUROCRYPT '93*, LNCS, Vol. 765, pp. 410–423. Springer-Verlag, 1994.
5. C. Cachin, Smooth entropy and Rényi entropy, *EUROCRYPT '97*, LNCS, Vol. 1233, pp. 193–208. Springer-Verlag, 1997.
6. C. Cachin, *Entropy Measures and Unconditional Security in Cryptography*, Ph. D. Thesis, ETH Zürich, Hartung-Gorre Verlag, Konstanz, 1997.
7. C. Cachin and U. Maurer, Smoothing probability distributions and smooth entropy, *Proceedings of International Symposium on Information Theory (ISIT) '97*. IEEE, 1997.
8. T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.
9. I. Csiszár and J. Körner, Broadcast channels with confidential messages, *IEEE Transactions on Information Theory*, Vol. 24, pp. 339–348, 1978.
10. Y. Dodis, L. Reyzin, and A. Smith, Fuzzy extractors: How to generate strong keys from biometrics and other noisy data, *EUROCRYPT 2004*, LNCS, Vol. 3027, pp. 523–540. Springer-Verlag, 2004.
11. J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, A pseudorandom generator from any one-way function, *SIAM Journal on Computing*, Vol. 28, No. 4, pp. 1364–1396, 1999.
12. T. Holenstein and R. Renner, On the smooth Rényi entropy of independently repeated random experiments. manuscript, 2005.
13. R. Impagliazzo, L. A. Levin, and M. Luby, Pseudo-random generation from one-way functions (extended abstract), *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing (STOC '89)*, pp. 12–24, 1989.
14. M. Luby and A. Wigderson, Pairwise independence and derandomization, Technical Report CSD-95-880, Computer Science Institute, Berkeley, CA, 1995. `http://citeseer.ist.psu.edu/luby95pairwise.html`.
15. U. M. Maurer, Secret key agreement by public discussion from common information, *IEEE Transactions on Information Theory*, Vol. 39, No. 3, pp. 733–742, 1993.
16. N. Nisan and D. Zuckerman, Randomness is linear in space, *Journal of Computer and System Sciences*, Vol. 52, pp. 43–52, 1996.
17. R. Renner, N. Gisin, and B. Kraus, Information-theoretic security proof for quantum-key-distribution protocols, *Physical Review A*, Vol. 72, 012332, 2005.
18. R. Renner and R. König, Universally composable privacy amplification against quantum adversaries, *Proc. of TCC 2005*, LNCS, Vol. 3378, pp. 407–425. Springer-Verlag, 2005.

19. R. Renner and S. Wolf, Smooth Rényi entropy and its properties, *Proceedings of International Symposium on Information Theory (ISIT) 2004*, p. 233. IEEE, 2004.
20. A. Rényi, On measures of entropy and information, *Proceedings of the 4th Berkeley Symp. on Math. Stat. and Prob.*, Vol. 1, pp. 547–561. Univ. of Calif. Press, 1961.
21. R. Shaltiel, Recent developments in explicit constructions of extractors, *Current trends in theoretical computer science. The Challenge of the New Century.*, Vol. 1, Algorithms and Complexity, 2002.
22. S. Wiesner, Conjugate coding, *SIGACT News*, Vol. 15, pp. 78–88, 1983.
23. A. D. Wyner, The wire-tap channel, *Bell System Technical Journal*, Vol. 54, No. 8, pp. 1355–1387, 1975.