How to Smooth Entropy?

Maciej Skorski(✉)

Cryptology and Data Security Group, University of Warsaw, Warsaw, Poland
maciej.skorski@mimuw.edu.pl

Abstract. Smooth entropy of X is defined as possibly biggest entropy of a distribution Y close to X. It has found many applications including privacy amplification, information reconciliation, quantum information theory and even constructing random number generators. However the basic question about the *optimal shape* for the distribution Y has not been answered yet. In this paper we solve this problem for Renyi entropies in *non-quantum settings*, giving a formal treatment to an approach suggested at TCC'05 and ASIACRYPT'05. The main difference is that we use a *threshold cut* instead of a *quantile cut* to rearrange probability masses of X. As an example of application, we derive tight lower bounds on the number of bits extractable from Shannon memoryless sources.

Keywords: Smooth Renyi entropy · Randomness extractors · Asymptotic equipartition property

1 Introduction

1.1 Entropy Smoothing

Security Based on Statistical Closeness. In most of cryptographic applications, probability distributions which are close enough in the variational (statistical) distance are considered indistinguishable. More informally, they have similar cryptographic "quality", when used as randomness sources (randomness extracting) [15] or secure keys (in the context of key derivation [1,6]).

Entropy Notions do not See Statistical Closeness. Unfortunately, standard entropy notions (including important min-entropy and collision entropy which are widely used as randomness measures in cryptography), are not robust with respect to small probability perturbations. Consider the AES cipher with a 256-bit key which is $\epsilon = 2^{-80}$-close to uniform. While such a key is considered secure nowadays, it may happen that it has no more than 81 bits of min-entropy (more precisely, fix $x \in \{0,1\}^{256}$ and consider the key X which is x_0 with probability $2^{-256} + 2^{-80}$ and uniform for $x \in \{0,1\}^{256} \setminus \{x_0\}$). This is a mismatch with respect to our intuitive understanding of min-entropy as a measure of how many almost random bits can be extracted.

M. Skorski—This work was partly supported by the WELCOME/2010-4/2 grant founded within the framework of the EU Innovative Economy Operational Programme.

R.M. Freivalds et al. (Eds.): SOFSEM 2016, LNCS 9587, pp. 392–403, 2016.
DOI: 10.1007/978-3-662-49192-8_32

Smooth Entropy takes Probability Perturbations into Account. To fix the issue described above, the concept of *smooth min-entropy* has been proposed [4,15]. Smooth entropy is defined as the maximal possible entropy within a certain distance to a given distribution. More precisely, for a given entropy notion $H(\cdot)$ (which is usually Renyi entropy, see Sect. 2 for its formal definition) we define the ϵ-smooth entropy of X as the value of the following optimization program

$$\begin{array}{ll} \text{maximize} & H(Y) \\ \text{s.t.} & \mathrm{SD}(X;Y) \leqslant \epsilon \end{array} \tag{1}$$

where SD() stands for statistical (variational) distance (see Sect. 2) for a formal definition). This definition is now well-suited for cryptographic applications, because does not depend anymore on negligible variations of the probability distribution. In particular, setting $\epsilon = 2^{-80}$ for our AES example we obtain the "correct" result of 256 bits of (smooth) entropy.

Importance of Smooth Entropy. Smooth Renyi entropy, formally introduced by Renner and Wolf in [15], found many applications including privacy amplification [13,15,19], information reconciliation [15] and quantum information theory [17,18]. The technique of perturbing a distribution to get more-entropy was actually known before. For example, entropy smoothing is implicitly used to prove the Asymptotic Equipartition Property [10] or more concretely in the construction of a pseudorandom generator from one-way functions [7–9]. However the simple question

Question 1. How does the shape of optimal Y depend on X?

has not been fully understood so far. In this paper we answer Question 1 by explicitly characterizing the shape of Y depending on X, and give some applications of the derived characterization.

1.2 Related Works and Our Contribution

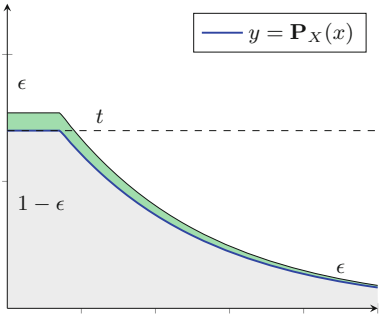
Related Works. The problem of finding the optimal shape for Y has been addressed in [13,15]. The authors argued intuitively that for min-entropy (which is a special case of Renyi entropy, particularly useful in randomness extraction) the optimal solution cuts down the biggest probabilities of X.

Our Contribution. We show that this characterization is not true, and the problem is more subtle: the optimal solution uses a threshold not a quantile cut (see Fig. 1).

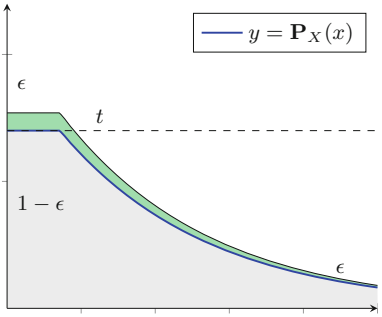
The precise answer to Question 1 is given in Theorem 1. We provide an intuitive explanation, as a three-step algorithm, in Fig. 2.

Theorem 1 (Optimal Renyi entropy smoothing). *Let $\alpha > 1$ be fixed, X be an arbitrary distribution over a finite set and $\epsilon \in (0,1)$. Let $t \in (0,1)$ be such that*

$$\sum_x \max\left(\mathbf{P}_X(x) - t, 0\right) = \epsilon, \tag{2}$$



(a) **Quantile Cut** - a folklore solution (TCC'05,ASIACRYPT'05), far from optimal.



(b) **Threshold Cut** - our idea, nearly optimal.

Fig. 1. Our result - the optimal shape for entropy smoothing

and Y be distributed according to

$$\mathbf{P}_Y(x) = \frac{\min\left(t, \mathbf{P}_X(x)\right)}{1-\epsilon}. \tag{3}$$

Then Y is nearly optimal, that is we have

$$\mathrm{SD}(X;Y) \leqslant \epsilon \tag{4}$$

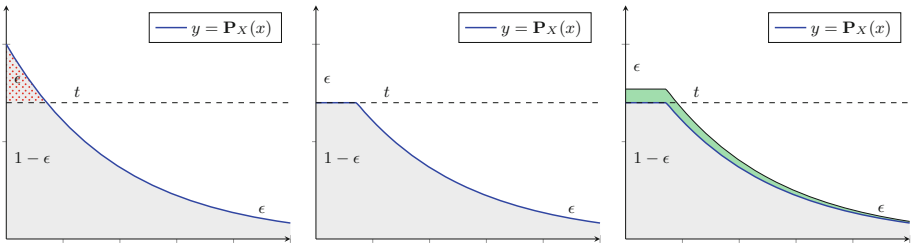
and

$$\mathbf{H}_\alpha(Y) \leqslant \mathbf{H}^\epsilon_\alpha(X) \leqslant \mathbf{H}_\alpha(Y) + \frac{\alpha}{\alpha-1}\log\left(\frac{1}{1-\epsilon}\right) \tag{5}$$

Corollary 1 (Tightness of Theorem 1). *Note that for fixed $\alpha > 1$ we have $\frac{\alpha}{\alpha-1}\log\left(\frac{1}{1-\epsilon}\right) = O(\epsilon)$. Thus, our solution differs from the ideal one by only a negligible additive constant in the entropy amount, which is good enough for almost all applications.*

1.3 Tight No-Go Results for Extracting from Stateless Shannon Sources

A *stateless source* (called also *memoryless*) is a source which produces consecutive samples independently. While this is a restriction, it is often assumed by practitioners working on random number generators (cf. [2,3,5,11]) and argued to be reasonable under some circumstances (so called *restart mode* which enforces fresh samples, see [3,5]). An important result is obtained from a more general fact called Asymptotic Equipartition Property (AEP). Namely, for a stateless source the min-entropy rate (min-entropy per sample) is close to its Shannon entropy per bit. The convergence holds *in probability*, for large number of samples.



(a) **Find the treshold** (b) **Cut the mass above** (c) **Rescale**

Fig. 2. Our result - details of optimal entropy smoothing

A variant of the AEP: The min entropy per bit in a sequence $X_1, \ldots, X_n$ of i.id. samples from X *converges*, when $n \to \infty$, to the Shannon entropy of X. More precisely

$$\frac{-\log \mathbf{P}_{X_1,\ldots,X_n}(\cdot)}{n} \xrightarrow{\text{in probability}} H(X), \tag{6}$$

where the probability is taken over $X_1, \ldots, X_n$.

Thus, the AEP is a bridge connecting the heuristic use of Shannon entropy as a measure of extractable randomness (practice) and the provable security (randomness extractors theory). The best known quantitative form of Eq. (6) appears in [9].

Lemma 1 (Asymptotic Equipartition Property [9]). *Let $X_1, \ldots, X_n$ be i.i.d. samples from a distribution X of Shannon entropy k. Then the sequence $(X_1, \ldots, X_n)$ is ϵ-close to a distribution of min entropy $kn - O\left(\sqrt{kn \log(1/\epsilon)}\right)$.*

Corollary 2. *In particular, one can extract $kn - O\left(\sqrt{kn \log(1/\epsilon)}\right) - 2\log(1/\epsilon)$ bits which are ϵ-close to uniform (e.g. using independent hash functions [8] as an extractor).*

Based on Theorem 1 we reprove the following result which matches the bound in [9]. Our result can be understood as the lower bound on the convergence speed in the Asymptotic Equipartition Property given in Lemma 1.

Theorem 2 (An upper bound on the extraction rate [16]). *Let $X_1, X_2, \ldots$ be of i.i.d. random variables, each of Shannon entropy k. Then from the sequence $(X_1, X_2, \ldots, X_n)$ no extractor can get more than*

$$N = kn - \Theta(\sqrt{kn \log(1/\epsilon)}) \tag{7}$$

bits which are ϵ-close (in the variation distance) to uniform (the constant under $\Theta(\cdot)$ depends on the source).

Remark 1 (The bound is tight for most settings). Since from N bits of min-entropy we can extract at least $N - 2\log(1/\epsilon)$ bits ϵ-close to uniform, and since in most cases $\log(1/\epsilon) = o(kn)$

From Theorem 2 we conclude that the error in Eq. (6) is significant and has to be taken into account no matter what the extractor is. It is worth of noting that our separation between Shannon entropy and extractable entropy holds *in the most favorable case*, when the bits are independent.

Corollary 3 (A significant error in the heuristic estimate). *In the above setting, the gap between the Shannon entropy and the number of extractable bits ϵ-close to uniform equals at least $\Theta(kn - \sqrt{\log(1/\epsilon)})$. In particular, for the recommended security level ($\epsilon = 2^{-80}$) we obtain the loss of $kn - N \approx \sqrt{80kn}$ bits, no matter what an extractor we use.*

1.4 Organization

Notions we use, as well as some auxiliary technical facts, are explained in Sect. 2. We prove our main result, that is Theorem 1, in Sect. 3. The proof of Theorem 2 appears in Sect. 4.

2 Preliminaries

2.1 Basic Definitions

The most popular way of measuring how two distributions are close is the statistical distance.

Definition 1 (Statistical Distance). *The statistical (or total variation) distance of two distributions X, Y over the same finite set is defined as*

$$\mathrm{SD}\left(X;Y\right) = \sum_x |\Pr[X = x] - \Pr[Y = x]| \tag{8}$$

We also say that X and Y are ϵ-close.

Below we recall the definition of Renyi entropy of order α. The logarithms are taken at base 2.

Definition 2 (Renyi Entropy). *The Renyi entropy of order α of a distribution X equals $\mathbf{H}_\alpha(X) = \frac{1}{1-\alpha} \log\left(\sum_x \Pr[X = x]^\alpha\right)$.*

Choosing $\alpha \to 1$ and $\alpha \to \infty$ we recover two important notions: Shannon entropy and min entropy.

Definition 3 (Shannon Entropy). *The Shannon Entropy of a distribution X equals $H(X) = -\sum_x \Pr[X = x] \log \Pr[X = x]$.*

Definition 4 (Min Entropy). *The min entropy of a distribution X equals* $H_\infty(X) = -\max_x \log \Pr[X = x]$.

Smooth Renyi Entropy is defined as the value of the program (1).

Definition 5 (Smooth Renyi Entropy, [4]). *The ϵ-smooth Renyi entropy of order α of a distribution X equals $H^\epsilon_\alpha(X) = \max_Y H_\alpha(Y)$ where the maximum is taken over Y satisfying the constraint $\mathrm{SD}(X;Y) \leqslant \epsilon$.*

Definition 6 (Extractable Entropy, [14]). *We say that X has k extractable bits within distance ϵ, denoted $H^\epsilon_{\mathrm{ext}}(X) \geqslant k$, if for some randomized function Ext we have $\mathrm{SD}\left(\mathrm{Ext}(X,S);U_k,S\right) \leqslant \epsilon$, where U_k is a uniform k-bit string and S is an independent uniform string.*

2.2 Technical Facts

We will need the following simple fact on convex functions

Proposition 1. *Let f be a strictly convex differentiable real-valued function and $x < y$. Then for any $\delta > 0$ we have*

$$f(x) - f(x-\delta) \leqslant f(y) - f(y-\delta).$$

Our proof uses the following characterization of "extractable" distributions.

Theorem 3 (An Upper Bound on Extractable Entropy, [14]). *If $H^\epsilon_{\mathrm{ext}}(X) \geqslant k$ then X is ϵ-close to Y such that $H_\infty(Y) \geqslant k$.*

Another important fact we use is the sharp bound on binomial tails.

Theorem 4 (Tight Binomial Tails [12]). *Let $B(n,p)$ be a sum of independent Bernoulli trials with success probability p. Then for $\gamma \leqslant \frac{3}{4}q$ we have*

$$\Pr\left[B(n,p) \geqslant pn + \gamma n\right] = Q\left(\sqrt{\frac{n\gamma^2}{pq}}\right) \cdot \psi\left(p,q,n,\gamma\right) \tag{9}$$

with the error term satisfies

$$\psi\left(p,q,n,\gamma\right) = \exp\left(\frac{n\gamma^2}{2pq} - n\mathrm{KL}\left(p+\gamma \parallel p\right) + \frac{1}{2}\log\left(\frac{p+\gamma}{p}\cdot\frac{q}{q-\gamma}\right) + O_{p,q}\left(n^{-\frac{1}{2}}\right)\right) \tag{10}$$

where $\mathrm{KL}\left(a \parallel b\right) = a\log(a/b) + (1-a)\log((1-a)/(1-b))$ is the Kullback-Leibler divergence, and Q is the complement of the cumulative distribution function of the standard normal distribution.

3 Proof of Theorem 1

We start by rewriting Eq. (1) in the following way

$$\begin{array}{ll} \text{minimize} & \left(\sum_x (\mu(x)+\epsilon(x))^\alpha\right)^{\frac{1}{\alpha-1}} \\ \text{s.t.} & \begin{cases} \sum_x \epsilon(x) = 0 \\ \sum_x |\epsilon(x)| = 2\epsilon \\ \forall x \quad 0 \leqslant \mu(x)+\epsilon(x) \leqslant 1 \end{cases} \end{array} \tag{11}$$

where for the sake of clarity we replace $\mathbf{P}_X$ by μ_X.

Claim. Let $\epsilon(x)$ be optimal for Eq. (11). Define $S^+ = \{x : \ \epsilon(x) < 0\}$. Then $\mu(x)+\epsilon(x) = \mu(y)+\epsilon(y)$ for all $x, y \in S^+$. We will show the optimality by a mass-shifting argument.

Proof (of Claim). Suppose that

$$\epsilon(x_1), \epsilon(x_2) < 0 \text{ and } \mu(x_1)+\epsilon(x_1) > \mu(x_2)+\epsilon(x_2) > 0 \tag{12}$$

for two different points x_1, x_2 (the statement is trivially true when there is only one point). Take a number δ such that

$$0 < \delta < \min\left(-\epsilon(x_2), \frac{(\mu(x_1)+\epsilon(x_1)) - (\mu(x_2)+\epsilon(x_2))}{2}\right) \tag{13}$$

and modify μ by shifting the mass from x_2 to x_1 in the following way

$$\epsilon'(x) = \begin{cases} \epsilon(x), & x \notin \{x_1, x_2\} \\ \epsilon(x)-\delta, & x = x_1 \\ \epsilon(x)+\delta, & x = x_2 \end{cases}$$

that is shifting the mass from the biggest point to the smallest point. Note that from Eqs. (11) and (13) it follows that the constraints in (11) are satisfied with $\epsilon(x)$ replaced by $\epsilon'(x)$. Let Y be a random variable distributed according to $\mathbf{P}_Y(x) = \mu(x)+\epsilon(x)$ and let Y' be distributed as $\mathbf{P}_{Y'}(x) = \mu(x)+\epsilon'(x)$. Note that we have

$$\sum_x (\mu(x)+\epsilon'(x))^\alpha - \sum_x (\mu(x)+\epsilon(x))^\alpha = \\ ((\mathbf{P}_Y(x_2)+\delta)^\alpha - (\mathbf{P}_Y(x_2))^\alpha) - ((\mathbf{P}_Y(x_1))^\alpha - (\mathbf{P}_Y(x_1)-\delta)^\alpha)$$

Note that we have

$$\mathbf{P}_Y(x_2) < \mathbf{P}_Y(x_2)+\delta < \mathbf{P}_Y(x_1)-\delta < \mathbf{P}_Y(x_1)$$

by Eqs. (12) and (13). Now from Proposition 1 applied to $f(u) = u^\alpha$, $x = \mathbf{P}_Y(x_2) + \delta$, $y = \mathbf{P}_Y(x_2)$ and δ (here we also use the assumption $\alpha > 1$), it follows that

$$\sum_x (\mu(x) + \epsilon'(x))^\alpha - \sum_x (\mu(x) + \epsilon(x))^\alpha < 0$$

which means $\mathbf{H}_\alpha(Y') > \mathbf{H}_\alpha(Y)$. In other words, Y is not optimal. □

By the last claim it is clear that there is a number $t \in (0,1)$ such that the set $S^+ = \{x : \mathbf{P}_{Y^*}(x) < \mu(x)\}$ is contained in $\{x : \mu(x) \geqslant t\}$ and that $\mathbf{P}_{Y^*}(x) \geqslant t$ for $x \in S^+$. Therefore

$$\begin{aligned}\sum_x (\mathbf{P}_{Y^*}(x))^\alpha &\geqslant \#\{x : \mu(x) \geqslant t\} \cdot t^\alpha + \sum_{x:\ \mu(x)<t} (\mu(x))^\alpha \\ &= \sum_x \min(\mu(x), t)^\alpha \\ &\geqslant (1-\epsilon)^\alpha \sum_x (\mathbf{P}_Y(x))^\alpha \end{aligned} \tag{14}$$

which, since $\mathbf{H}^\epsilon_\alpha(X) = \mathbf{H}_\alpha(Y^*)$, proves the second inequality in Eq. (5). To prove the first inequality in Eq. (5) note that

$$\mathrm{SD}(X;Y) = \sum_{x:\ \mathbf{P}_X(x) > \mathbf{P}_Y(x)} (\mathbf{P}_X(x) - \mathbf{P}_Y(x)) = \sum_{x:\ \mu(x) > t/(1-\epsilon)} \left(\mu(x) - \frac{t}{1-\epsilon}\right).$$

Since $\frac{t}{1-\epsilon} > t$ we have

$$\mathrm{SD}(X;Y) = \sum_{x:\ \mu(x) > t/(1-\epsilon)} \left(\mu(x) - \frac{t}{1-\epsilon}\right) \leqslant \sum_{x:\ \mu(x) > t} (\mu(x) - t)$$

and therefore by Eq. (2) we obtain

$$\mathrm{SD}(X;Y) < \sum_{x:\ \mu(x) > t} (\mu(x) - t) = \epsilon.$$

which finishes the proof.

4 Proof of Theorem 2

4.1 Characterizing Extractable Entropy

We state the following fact with an explanation in Fig. 3.

Lemma 2 (Lower bound on the extractable entropy). *Let X be a distribution. Then for every distribution Y which is ϵ-close to X, we have $H_\infty(Y) \leqslant -\log t$ where t satisfies*

$$\sum_x \max(\mathbf{P}_X(x) - t, 0) = \epsilon. \tag{15}$$

The proof follows by Theorem 1.

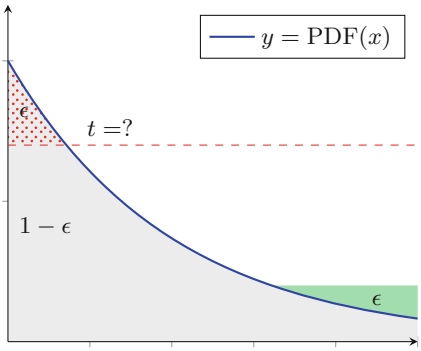


Fig. 3. Entropy Smoothing Problem. For a given probability density function, we want to cut a total mass of up to ϵ above a possibly highest threshold (in dotted red) and rearrange it (in green), to keep the upper bound smallest possible (Color figure online)

Without losing generality, we assume from now that $X \in \{0,1\}$ where $\Pr[X = 1] = p, q = 1 - p$. Define $X^n = (X_1, \ldots, X_n)$. For any $x \in \{0,1\}^n$ we have

$$\Pr[X^n = x] = p^{\|x\|} q^{n - \|x\|}. \tag{16}$$

According to the last lemma and Theorem 3, we have

$$H^{\epsilon}_{\mathrm{ext}}(X^n) \leqslant -\log t \tag{17}$$

where

$$\sum_x \max(\mathbf{P}_{X^n}(x) - t, 0) = \epsilon. \tag{18}$$

From now we assume that

$$t = p^{pn+\gamma n} q^{qn-\gamma n}. \tag{19}$$

4.2 Determining the Threshold t

The next key observation is that t is actually small and can be omitted. That is, we can simply cut the $(1-\epsilon)$-quantile. This is stated in the lemma below.

Lemma 3 (Replacing the threshold by the quantile). *Let $x_0 \in \{0,1\}^n$ be a point such that $\|x_0\| = pn + \gamma n$. Then we have*

$$\sum_{x:\ \|x\| \geqslant \|x_0\|} \max(\mathbf{P}_{X^n}(x) - \mathbf{P}_{X^n}(x_0)) \geqslant \frac{1}{2} \sum_{x:\ \|x\| \geqslant \|x_0\|} \mathbf{P}_{X^n}(x) \tag{20}$$

To prove the lemma, note that from Theorem 4 it follows that setting

$$\gamma' = \gamma + n^{-1} \log \left(\frac{p}{q} \right) \tag{21}$$

we obtain

$$\sum_{j \geqslant pn+\gamma' n} \binom{n}{j} \geqslant \frac{3}{4} \cdot \sum_{j \geqslant pn+\gamma n} \binom{n}{j} \tag{22}$$

when γ is sufficiently small comparing to p and q (formally this is justified by calculating the derivative with respect to γ and noticing that it is bigger by at most a factor of $1 + \frac{\gamma}{\sqrt{npq}}$). But we also have

$$p^j q^{n-j} \geqslant 2 \cdot p^{(p+\gamma)n} q^{(q-\gamma)n} \quad \text{for } j \geqslant \gamma' n \tag{23}$$

Therefore,

$$\begin{aligned} \sum_{j \geqslant pn+\gamma n} \binom{n}{j} p^j q^{n-j} &\geqslant \sum_{j \geqslant pn+\gamma' n} \binom{n}{j} p^j q^{n-j} \\ &\geqslant 2 \cdot p^{(p+\gamma)n} q^{(q-\gamma)n} \cdot \sum_{j \geqslant pn+\gamma' n} \binom{n}{j} \\ &\geqslant 2 \cdot \frac{3}{4} \cdot p^{(p+\gamma)n} q^{(q-\gamma)n} \cdot \sum_{j \geqslant pn+\gamma n} \binom{n}{j} \end{aligned} \tag{24}$$

which finishes the proof.

4.3 Putting This All Together

Now, by combining Lemmas 2 and 3 and the estimate $Q(x) \approx x^{-1} \exp(-x^2/2)$ for $x \gg 0$ we obtain

$$\epsilon \geqslant \exp \left(-n \mathrm{KL} \left(p + \gamma \parallel p \right) - \log \left(\frac{n\gamma^2}{2pq} \right) + O_{p,q}(1) \right) \tag{25}$$

which, because of the Taylor expansion $\mathrm{KL} \left(p + \gamma \parallel p \right) = \frac{\gamma^2}{2pq} + O_{p,q}(\gamma^3)$, gives us

$$\gamma \geqslant \Omega \left(\sqrt{\frac{\log(1/\epsilon)}{pqn}} \right) \tag{26}$$

Setting $\gamma = c \cdot \sqrt{\frac{\log(1/\epsilon)}{pqn}}$, with sufficiently big c, we obtain the claimed result.

References

1. Barak, B., Dodis, Y., Krawczyk, H., Pereira, O., Pietrzak, K., Standaert, F.-X., Yu, Y.: Leftover hash lemma, revisited. Cryptology ePrint Archive, Report 2011/088 (2011). http://eprint.iacr.org/
2. Bouda, J., Krhovjak, J., Matyas, V., Svenda, P.: Towards true random number generation in mobile environments. In: Jøsang, A., Maseng, T., Knapskog, S.J. (eds.) NordSec 2009. LNCS, vol. 5838, pp. 179–189. Springer, Heidelberg (2009). http://dx.doi.org/10.1007/978-3-642-04766-4_13
3. Bucci, M., Luzzi, R.: Design of testable random bit generators. In: Rao, J.R., Sunar, B. (eds.) CHES 2005. LNCS, vol. 3659, pp. 147–156. Springer, Heidelberg (2005)
4. Cachin, C.: Smooth entropy and Rényi entropy. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 193–208. Springer, Heidelberg (1997)
5. Dichtl, M., Golić, J.D.: High-speed true random number generation with logic gates only. In: Paillier, P., Verbauwhede, I. (eds.) CHES 2007. LNCS, vol. 4727, pp. 45–62. Springer, Heidelberg (2007)
6. Dodis, Y., Pietrzak, K., Wichs, D.: Key derivation without entropy waste. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 93–110. Springer, Heidelberg (2014)
7. Hastad, J., Impagliazzo, R., Levin, L.A., Luby, M.: Pseudo-random generation from one-way functions. In: Proceedings of the 20th STOC, pp. 12–24 (1988)
8. Hastad, J., Impagliazzo, R., Levin, L.A., Luby, M.: A pseudorandom generator from any one-way function. SIAM J. Comput. **28**(4), 1364–1396 (1999). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.3930
9. Holenstein, T.: Pseudorandom generators from one-way functions: a simple construction for any hardness. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 443–461. Springer, Heidelberg (2006)
10. Holenstein, T., Renner, R.: On the randomness of independent experiments. IEEE Trans. Inf. Theory **57**(4), 1865–1871 (2011)
11. Lacharme, P., Röck, A., Strubel, V., Videau, M.: The linux pseudorandom number generator revisited. Cryptology ePrint Archive, Report 2012/251 (2012). http://eprint.iacr.org/
12. McKay, B.D.: On Littlewood's estimate for the binomial distribution. Adv. Appl. Probab. **21**(2), 475–478 (1989)
13. Renner, R.S., König, R.: Universally composable privacy amplification against quantum adversaries. In: Kilian, J. (ed.) TCC 2005. LNCS, vol. 3378, pp. 407–425. Springer, Heidelberg (2005). http://dx.doi.org/10.1007/978-3-540-30576-7_22
14. Renner, R., Wolf, S.: Smooth Renyi entropy and applications. In: ISIT 2004, Chicago, Illinois, USA, p. 232 (2004)
15. Renner, R.S., Wolf, S.: Simple and tight bounds for information reconciliation and privacy amplification. In: Roy, B. (ed.) ASIACRYPT 2005. LNCS, vol. 3788, pp. 199–216. Springer, Heidelberg (2005)
16. Skorski, M.: How much randomness can be extracted from memoryless shannon entropy sources. In: WISA 2015 (2015)
17. Schoenmakers, B., Tjoelker, J., Tuyls, P., Verbitskiy, E.: Smooth Renyi entropy of ergodic quantum information sources. In: 2007 IEEE International Symposium on Information Theory. ISIT 2007, pp. 256–260 (2007)

18. Tomamichel, M.: A framework for non-asymptotic quantum information theory. Ph.D. thesis, ETH Zurich (2012)
19. Watanabe, S., Hayashi, M.: Non-asymptotic analysis of privacy amplification via Renyi entropy and inf-spectral entropy. In: 2013 IEEE International Symposium on Information Theory Proceedings (ISIT), pp. 2715–2719 (2013)