

2011-01-07

# Computational entropy and information leakage (MA Thesis)

---

Fuller, Benjamin; Reyzin, Leonid. "Computational Entropy and Information Leakage (MA Thesis)", Technical Report BUCS-TR-2011-001, Computer Science Department, Boston University, January 7, 2011. [Available from: <http://hdl.handle.net/2144/11358>]

<https://hdl.handle.net/2144/11358>

*"Downloaded from OpenBU. Boston University's institutional repository."*

# Computational Entropy and Information Leakage

Benjamin Fuller

Leonid Reyzin

Boston University  
{bfuller,reyzin}@bu.edu

January 10, 2011

## Abstract

We investigate how information leakage reduces computational entropy of a random variable  $X$ . Recall that HILL and metric computational entropy are parameterized by quality (how distinguishable is  $X$  from a variable  $Z$  that has true entropy) and quantity (how much true entropy is there in  $Z$ ).

We prove an intuitively natural result: conditioning on an event of probability  $p$  reduces the quality of metric entropy by a factor of  $p$  and the quantity of metric entropy by  $\log_2 1/p$  (note that this means that the reduction in quantity and quality is the same, because the quantity of entropy is measured on logarithmic scale). Our result improves previous bounds of Dziembowski and Pietrzak (FOCS 2008), where the loss in the *quantity* of entropy was related to its original *quality*. The use of metric entropy simplifies the analogous the result of Reingold et. al. (FOCS 2008) for HILL entropy.

Further, we simplify dealing with information leakage by investigating *conditional* metric entropy. We show that, conditioned on leakage of  $\lambda$  bits, metric entropy gets reduced by a factor  $2^\lambda$  in quality and  $\lambda$  in quantity. Our formulation allow us to formulate a “chain rule” for leakage on computational entropy. We show that conditioning on  $\lambda$  bits of leakage reduces *conditional* metric entropy by  $\lambda$  bits. This is the same loss as leaking from *unconditional* metric entropy. This result makes it easy to measure entropy even after several rounds of information leakage.

# 1 Introduction

Suppose you have a pseudorandom generator that, during the computation, leaks some function of the seed to an adversary. How pseudorandom are the resulting outputs?

More generally, suppose you have a distribution that has *computational* entropy. Suppose some correlated information leaks to an adversary. How much computational entropy is left?

These questions come up naturally in the context of leakage-resilient cryptography. The question of pseudoentropy with a leaked “seed” has been addressed before primarily in two works. Dziembowski and Pietrzak posed the question —about pseudorandom generators— in their construction of a leakage-resilient stream cipher [DP08]. Reingold et. al. [RTTV08] consider the general case of pseudoentropy of a random variable after a particular leakage in the context of computational versions of the dense model theorem [GT08].

We consider both the leakage of a particular value and of a random variable. We provide a simple answer to both questions (Lemma 3.5, Theorem 3.2). Theorem 3.2 is particularly elegant:

If  $\lambda$  bits of information are leaked, then the amount of computational entropy decreases by at most  $\lambda$ .

Naturally, the answer becomes so simple only once the correct notion of entropy is in place. Our result holds for average-case **Metric\*** entropy (defined in [BSW03, DP08]). In case this notion of entropy seems esoteric, we point out that it is convertible (with a small loss) to average-case **HILL** entropy [HLR07] using the techniques of [BSW03], which can be used with randomness extractors to get pseudorandom bits [DORS08, HLR07].

When speaking about **HILL** entropy and its variants, one has to keep in mind that what matters is not only the number of bits of entropy, but also its quality. Namely, **HILL** entropy of a variable  $X$  is defined as the amount of entropy in a distribution  $Z$  that is indistinguishable from  $X$  (**Metric\*** entropy is defined similarly; the differences are discussed in Section 2). Indistinguishability is parameterized by the maximum size of the distinguishing circuit  $D$  and the maximum quality of its distinguishing—i.e.,  $\epsilon = |\mathbb{E}[D(X)] - \mathbb{E}[D(Z)]|$ . In our results, both the amount of entropy and its quality decrease: that is,  $\epsilon$  increases by a factor of  $2^\lambda$ . We note that because entropy is measured on a logarithmic scale (min-entropy is simply the negative logarithm of maximum probability), this loss in the quality and the quantity is actually the same.

Average-case entropy works well in situations in which not all leakage is equally informative. For instance, in case the leakage is equal to the Hamming weight of a uniformly distributed string, sometimes the entropy of the string gets reduced to nothing (if the value of the leakage is 0 or the length of the string), but most of the time it stays high. For the information-theoretic case, it is known that deterministic leakage of  $\lambda$  bits reduces the average entropy by at most  $\lambda$  [DORS08, Lemma 2.2(b)] (the reduction is less for randomized leakage). Thus, our result matches the information-theoretic case for deterministic leakage. For randomized leakage, our statement can be somewhat improved (Theorem 3.4.3).

If a worst-case, rather than an average-case guarantee is needed, we also provide a statement of the type “with probability at least  $1 - \delta$  over all possible leakage, entropy loss due to leakage is at most  $\lambda - \log 1/\delta$ ” (Lemma 3.3). Statements of this type are used for computational entropy in [DP08, FKPR10]. If one is interested in the entropy lost due to a specific leakage value, rather than

over a distribution of leakage values, we provide an answer, as well (Lemma 3.5): if the leakage has probably  $p$ , then the amount of entropy decreases by  $\log 1/p$  and the quality decreases by a factor of  $p$  (i.e.,  $\epsilon$  becomes  $\epsilon/p$ ). Reingold et. al. [RTTV08] provide a similar formulation for HILL entropy. The use of metric entropy allows for a tighter reduction than [RTTV08] and allows us to eliminate the loss in circuit size that occurs in the reduction of [RTTV08].

We also provide a chain rule: namely, our result for average-case **Metric\*** entropy holds even if the original distribution has only average-case **Metric\*** entropy. Thus, in case of multiple leakages, our result can be applied multiple times. The price for the conversion from **Metric\*** to HILL entropy needs to be paid only once. The chain rule highlights one of the advantages of average-case entropy: if one tried to use the worst-case statement “with probability at least  $1 - \delta$ , entropy is reduced by at most  $\lambda + \log 1/\delta$ ” over several instances of leakage, then total entropy loss bound would be greater and the probability that it is satisfied would be lower, because the  $\delta$ s would add up.

Our result can be used to improve the parameters of the leakage-resilient stream cipher of [DP08] and leakage-resilient signature scheme of [FKPR10].

## 2 Entropy and Extraction

We begin by clarifying previous definitions of entropy and introducing a few natural definitions for conditional entropy.

### 2.1 Preliminary Notation

Let  $x \in X$  denote an element  $x$  in the support of  $X$ . Let  $x \leftarrow X$  be the process of a sampling  $x$  from the distribution  $X$ . Let  $U_n$  represent the random variable with the uniform distribution over  $\{0, 1\}^n$ . Let  $\delta(X, Y)$  be the statistical distance between random variables  $X, Y$  drawn from a set  $\chi$ , defined as  $\delta(X, Y) = \frac{1}{2} \sum_{x \in \chi} |\Pr(X = x) - \Pr(Y = x)|$ . Let  $\mathcal{D}_s$  be the set of all probabilistic circuits of size  $s$  with binary output  $\{0, 1\}$ . Following the notation of [DP08], let  $\mathcal{D}_s^*$  be the set of all deterministic circuits of size  $s$  with output in  $[0, 1]$ . We say  $s \approx s'$  if the two sizes  $s, s'$  differ by a small additive constant. Given a circuit  $D$ , define the computational distance  $\delta^D$  between  $X$  and  $Y$  as  $\delta^D(X, Y) = |\mathbb{E}[D(X)] - \mathbb{E}[D(Y)]|$ . We denote the size of a circuit  $D$  as  $|D|$ . For a probability distribution  $X$ , let  $|X|$  denote the size of the support of  $X$ , that is  $|X| = |\{x \mid \Pr[X = x] > 0\}|$ . All logarithms without a base are considered base 2, that is,  $\log x = \log_2 x$ .

### 2.2 Unconditional Entropy

We begin with the standard notion of min-entropy and proceed to computational notions.

**Definition 1.** A distribution  $X$  has **min-entropy** at least  $k$ , denoted  $H_\infty(X) \geq k$  if

$$\forall x \in X, \Pr[X = x] \leq 2^{-k}.$$

Computational min-entropy has two additional parameters: distinguisher size  $s$  and quality  $\epsilon$ . Larger  $s$  and smaller  $\epsilon$  mean “better” entropy.

**Definition 2.** ([HILL99]) A distribution  $X$  has **HILL entropy** at least  $k$ , denoted  $H_{\epsilon, s}^{\text{HILL}}(X) \geq k$  if there exists a distribution  $Y$  where  $H_\infty(Y) \geq k$ , such that  $\forall D \in \mathcal{D}_s, \delta^D(X, Y) \leq \epsilon$ .

Switching the quantifiers of  $Y$  and  $D$  gives us the following, weaker notion.

**Definition 3.** ([BSW03]) A distribution  $X$  has **Metric entropy** at least  $k$ , denoted  $H_{\epsilon,s}^{\text{Metric}}(X) \geq k$  if  $\forall D \in \mathcal{D}_s$  there exists a distribution  $Y$  with  $H_\infty(Y) \geq k$  and  $\delta^D(X, Y) \leq \epsilon$ .

Drawing  $D$  from  $\mathcal{D}_s^*$  instead of  $\mathcal{D}_s$  in the above two definitions gives us notions of “HILL-star” entropy  $H_{\epsilon,s}^{\text{HILL}^*}$  and “metric-star” entropy  $H_{\epsilon,s}^{\text{Metric}^*}$ , respectively (this notation was introduced in [DP08]).  $H^{\text{HILL}}$  and  $H^{\text{HILL}^*}$  are essentially equivalent, as shown in the following lemma (discovered jointly with the authors of [DP08]), whose proof in Appendix A.

**Lemma 2.1.**  $H_{\epsilon,s'}^{\text{HILL}}(X) \geq k \Leftrightarrow H_{\epsilon,s}^{\text{HILL}^*}(X) \geq k$ , for  $s' \approx s$ .

Unfortunately, only one direction of the equivalence holds for metric entropy (the proof is also in Appendix A).

**Lemma 2.2.**  $H_{\epsilon,s'}^{\text{Metric}}(X) \geq k \Rightarrow H_{\epsilon,s}^{\text{Metric}^*}(X) \geq k$ , for  $s' \approx s$ .

It is immediate that  $H_{\epsilon,s}^{\text{HILL}}(X) \geq k \Rightarrow H_{\epsilon,s}^{\text{Metric}}(X) \geq k$  and  $H_{\epsilon,s}^{\text{HILL}^*}(X) \geq k \Rightarrow H_{\epsilon,s}^{\text{Metric}^*}(X) \geq k$ . For the opposite direction, the implication is known to hold only with a loss in quality and circuit size, as proven by Barak, Shaltiel, and Wigderson [BSW03, Theorem 5.2]<sup>1</sup>.

**Theorem 2.3.** ([BSW03]) Let  $X$  be a discrete distribution over a finite set  $\chi$ . For every  $\epsilon, \epsilon_{\text{HILL}} > 0, \epsilon' \geq \epsilon + \epsilon_{\text{HILL}}, k$ , and  $s$ , if  $H_{\epsilon,s}^{\text{Metric}^*}(X) \geq k$  then  $H_{\epsilon',s_{\text{HILL}}}^{\text{HILL}^*}(X) \geq k$  where  $s_{\text{HILL}} = \Omega(\epsilon_{\text{HILL}}^2 s / \log |\chi|)$ .

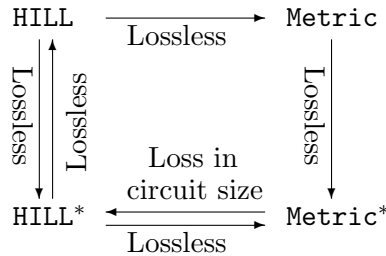


Figure 1: Known state of equivalence for HILL and Metric Entropy. All of these equivalences carry over to the conditional case.

### 2.3 Randomness Extractors

Originally defined for information-theoretic, rather than computational entropy, an *extractor* takes a distribution  $X$  of min-entropy  $k$ , and with the help of a uniform string called the seed, “extracts” the randomness contained in  $X$  and outputs a string of length  $m$  that is *almost uniform* even given the seed.

**Definition 4** ([NZ93]). Let  $\chi$  be a finite set. A polynomial-time computable deterministic function  $\text{ext} : \chi \times \{0, 1\}^d \rightarrow \{0, 1\}^m \times \{0, 1\}^d$  is a strong  $(k, \epsilon)$ -extractor if the last  $d$  outputs of bits of  $\text{ext}$

<sup>1</sup>The theorem statement in [BSW03] is incorrect, because it confuses Metric and Metric\* entropies. However, the proofs seems correct with respect to Metric\*. We generalize the theorem slightly to allow for distributions over a generic set  $\chi$  rather than just  $\{0, 1\}^n$ . Reingold et. al. [RTTV08, Theorem 1.3] contains a similar conversion but it is tightly coupled with their proof.

are equal to the last  $d$  input bits (these bits are called *seed*), and  $\delta(\text{ext}(X, U_d), U_m \times U_d) \leq \epsilon$  for every distribution  $X$  on  $\chi$  with  $H_\infty(X) \geq k$ . The number of extracted bits is  $m$ , and the entropy loss is  $k - m$ .

It turns out that extractors can be applied to distributions with computational entropy to obtain pseudorandom, rather than random, outputs: that is, outputs that are computationally indistinguishable from, rather than statistically close to, uniformly random strings. This fact is well-known for HILL entropy. However, we have not seen it proven for Metric entropy and, although the proof is quite straightforward, we provide it here for completeness. (Since HILL entropy implies Metric entropy, this proof also works for HILL entropy.)

**Theorem 2.4.** *Let  $\text{ext} : \chi \times \{0, 1\}^d \rightarrow \{0, 1\}^m \times \{0, 1\}^d$  be a  $(k, \epsilon_{\text{ext}})$ -extractor, computable by circuits of size  $s_{\text{ext}}$ . Let  $X$  be a distribution over  $\chi$  with  $H_{\epsilon_{\text{metric}}, s}^{\text{metric}}(X) \geq k$ . Then  $\forall D \in \mathcal{D}_{s'}$ , where  $s' \approx s_{\text{metric}} - s_{\text{ext}}$ ,*

$$\delta^D(\text{ext}(X, U_d), U_m \times U_d) \leq \epsilon_{\text{ext}} + \epsilon_{\text{metric}}.$$

*Proof.* We proceed by contradiction. Suppose not, that is,  $\exists D \in \mathcal{D}_{s'}$  such that

$$\delta^D(\text{ext}(X, U_d), U_m \times U_d) > \epsilon_{\text{ext}} + \epsilon_{\text{metric}}.$$

We use  $D$  to construct a distinguisher  $D'$  to distinguish  $X$  from all distributions  $Y$  where  $H_\infty(Y) \geq k$ , violating the metric-entropy of  $X$ . We define  $D'$  as follows: upon receiving input  $\alpha \in \chi$ ,  $D'$  samples  $\text{seed} \leftarrow U_d$ , runs  $\beta \leftarrow \text{ext}(\alpha, \text{seed})$  and then runs  $D(\beta, \text{seed})$  on the result. Note that  $D' \in \mathcal{D}_s$  where  $s \approx s' + s_{\text{ext}} = s_{\text{metric}}$ . Thus we have the following  $\forall Y$ , where  $H_\infty(Y) \geq k$ :

$$\begin{aligned} \delta^{D'}(X, Y) &= \delta^D(\text{ext}(X, U_d), \text{ext}(Y, U_d)) \\ &\geq \delta^D(\text{ext}(X, U_d), U_m \times U_d) - \delta^D((\text{ext}(Y, U_d), U_m \times U_d)) \\ &> \epsilon_{\text{ext}} + \epsilon_{\text{metric}} - \epsilon_{\text{ext}} = \epsilon_{\text{metric}} \end{aligned}$$

Thus  $D'$  is able to distinguish  $X$  from all  $Y$  with sufficient min-entropy. This is a contradiction.  $\square$

Unfortunately, the theorem does not extend to Metric\* entropy, because the distinguisher  $D'$  we construct in this proof is randomized. The only way to extract from Metric\* entropy that we know of is to convert Metric\* entropy to HILL\* entropy using Theorem 2.3 (which incurs some loss) and then use Theorem 2.4.

## 2.4 Conditional Entropy and Extraction

Conditional entropy measures the entropy that remains in a distribution after some information about the distribution is leaked. There are many different possible definitions. We follow the definition of “average min-entropy” [DORS08, Section 2.4]; the reasons for that particular choice of definition are detailed there. Because min-entropy is the negative logarithm of the highest probability, average min-entropy is defined the negative logarithm of the average, over the condition, of highest probabilities.

**Definition 5** ([DORS08]). Let  $(X, Y)$  be a pair of random variables. The **average min-entropy** of  $X$  conditioned on  $Y$  is defined as

$$\tilde{H}_\infty(X|Y) \stackrel{\text{def}}{=} -\log[\mathbb{E}_Y(2^{-H_\infty(X|Y)})] = -\log \sum_{y \in Y} \Pr[Y = y] 2^{-H_\infty(X|Y=y)}$$

The definition of average min-entropy has been extended to the computational case by Hsiao, Lu, Reyzin [HLR07].

**Definition 6.** ([HLR07]) Let  $(X, Y)$  be a pair of random variables.  $X$  has *conditional HILL entropy* at least  $k$  conditioned on  $Y$ , denoted  $H_{\epsilon, s}^{\text{HILL}}(X|Y) \geq k$  if there exists a collection of distributions  $Z_y$  for each  $y \in Y$ , giving rise to a joint distribution  $(Z, Y)$ , such that  $\tilde{H}_\infty(Z|Y) \geq k$  and  $\forall D \in \mathcal{D}_s, \delta^D((X, Y), (Z, Y)) \leq \epsilon$ .

Again, we can switch the quantifiers of  $Z$  and  $D$  to obtain the definition of conditional metric entropy.

**Definition 7.** Let  $(X, Y)$  be a pair of random variables.  $X$  has *conditional metric entropy* at least  $k$  conditioned on  $Y$ , denoted by  $H_{\epsilon, s}^{\text{Metric}}(X|Y) \geq k$ , if  $\forall D \in \mathcal{D}_s$  there exists a collection of distributions  $Z_y$  for each  $y \in Y$ , giving rise to a joint distribution  $(Z, Y)$ , such that  $\tilde{H}_\infty(Z|Y) \geq k$  and  $\delta^D((X, Y), (Z, Y)) \leq \epsilon$ .

Conditional HILL\* and conditional Metric\* can be defined similarly, replacing  $\mathcal{D}$  with  $\mathcal{D}^*$ . The same relations among the four notions of conditional entropy hold as in the unconditional case (see Lemma B.1 for an example).

Similarly to information-theoretic average-case entropy [DORS08, Lemma 2.2a], conditional Metric\* entropy implies, with some confidence, a smaller amount of lower-quality Metric\* entropy, as shown in Appendix B.

Average-case extractors, defined in [DORS08, Section 2.5], are extractors extended to work with average-case, rather than unconditional, min-entropy. It is also shown there that every extractor can be converted to an average-case extractor with some loss, and that some extractors are already average-case extractors without any loss.

**Definition 8.** Let  $\chi_1, \chi_2$  be finite sets. An extractor  $\text{ext}$  is a  $(k, \epsilon)$ -average-case extractor if for all pairs of random variables  $X, Y$  over  $\chi_1, \chi_2$  such that  $\tilde{H}_\infty(X|Y) \geq k$ , we have  $\delta((\text{ext}(X, U_d), Y), U_m \times U_d \times Y) \leq \epsilon$ .

Similar to extractors in the case of unconditional entropy, average-case extractors can be used on distributions that have Metric\* (and therefore also on distributions that have HILL or HILL\*) conditional entropy to produce pseudorandom, rather than random outputs. The proof is similar to [HLR07, Lemma 5].

### 3 Main Results: Computational Entropy after Leakage

We first present our main results. Proofs are presented in Section 3.4. As a starting point, consider Lemma 3 of [DP08], modified slightly to separate the quality of entropy parameter  $\epsilon_1$  from the confidence parameter  $\epsilon_2$  (both are called  $\epsilon$  in [DP08]):

**Lemma 3.1** ([DP08, Lemma 3]). *Let  $\text{prg} : \{0, 1\}^n \rightarrow \{0, 1\}^\nu$  and  $f : \{0, 1\}^n \rightarrow \{0, 1\}^\lambda$  (where  $1 \leq \lambda < n < \nu$ ) be any functions. If  $\text{prg}$  is a  $(\epsilon_{\text{prg}}, s)$ -secure pseudorandom-generator, then for any  $\epsilon_1, \epsilon_2, \Delta > 0$  satisfying  $\epsilon_{\text{prg}} \leq \epsilon_1 \epsilon_2 / 2^\lambda - 2^{-\Delta}$ , we have with  $X \sim U_n$ ,*

$$\Pr_{y=f(X)} [H_{\epsilon_1, s'}^{\text{Metric}^*}(\text{prg}(X)|f(X) = y) \geq \nu - \Delta] \geq 1 - \epsilon_2 \quad (1)$$

where  $s' \approx s$ .

Our results improve the parameters and simplify the exposition. Our main theorem, proven in Section 3.4.2, is as follows:

**Theorem 3.2.** *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon|Y|,s'}^{\text{Metric}^*}(X|Y) \geq H_{\epsilon,s}^{\text{Metric}^*}(X) - \log |Y|$$

where  $s' \approx s$ .

Intuitively, this theorem says that the quality and quantity of entropy reduce by the number of leakage values. It seems unlikely that the bounds can be much improved. The loss in the amount of entropy is necessary when, for example,  $X$  is the output of a pseudorandom generator and the leakage consists of  $\log |Y|$  bits of  $X$ . The loss in the quality of entropy seems necessary when the leakage consists of  $\log |Y|$  bits of the seed used to generate  $X$ . (However, we do not know how to show that both losses are necessary simultaneously.)

The theorem holds even if the metric entropy in  $X$  is also conditional (see Theorem 3.6 for the formal statement), and thus can be used in cases of repeated leakage as a chain rule. This theorem is more general than Lemma 3.1, because it applies to any discrete random variables with sufficient entropy, rather than just the output of a pseudorandom generator<sup>2</sup>. Because of the average-case formulation, it is also simpler. In addition, the average-case formulation allows one to apply average-case extractors (such as universal hash functions) without the additional loss of  $\epsilon_2$  (after the conversion to HILL entropy, see Corollary 3.7) and handles cases of repeated leakage better (because one does not have to account for  $\epsilon_2$  multiple times).

Simplicity and generality aside, this result is *quantitatively* better. To make the quantitative comparison, we present the following alternative formulation of our result, in the style of [DP08, Lemma 3]:

**Lemma 3.3.** *Let  $X, Y$  be discrete random variables with  $|Y| \leq 2^\lambda$  and  $H_{\epsilon_{ent},s}^{\text{Metric}^*}(X) \geq \nu$ , then for any  $\epsilon_1, \epsilon_2, \Delta > 0$  satisfying  $\epsilon_{ent} \leq \epsilon_1 \epsilon_2 / 2^\lambda$  and  $2^{-\Delta} \leq \epsilon_2 / 2^\lambda$ ,*

$$\Pr_{y \in Y} [H_{\epsilon_1, s'}^{\text{Metric}^*}(X|Y = y) \geq \nu - \Delta] \geq 1 - \epsilon_2$$

where  $s' \approx s$ .

To compare the bounds, observe that we have removed  $\epsilon_1$  from  $2^{-\Delta}$ , because the constraint  $\epsilon_{prg} \leq \epsilon_1 \epsilon_2 / 2^\lambda - 2^{-\Delta}$  in particular implies that  $\epsilon_{prg} \leq \epsilon_1 \epsilon_2 / 2^\lambda$  and  $\epsilon_1 \epsilon_2 / 2^\lambda \geq 2^{-\Delta}$ .

### 3.1 Structure of the Proof

We begin by presenting Theorem 1.3 of [RTTV08], restated in our language, which provides a similar result for HILL entropy.

**Lemma 3.4** ([RTTV08, Theorem 1.3]). *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon', s'}^{\text{HILL}}(X|Y = y) \geq H_{\epsilon, s}^{\text{HILL}}(X) - \log 1/P_y \tag{2}$$

where  $P_y = \Pr[Y = y]$ ,  $\epsilon' = \Omega(\epsilon/P_y)$ , and  $s' = s/\text{poly}(P_y/\epsilon, \log 1/P_y)$

---

<sup>2</sup>The output of a pseudorandom generator has full HILL entropy and thus full **Metric**<sup>\*</sup> entropy.

The works of [DP08][RTTV08] both utilize the proof technique presented in Figure 2(quality, quantity parameters are removed for clarity). In our lemma, we focus on the second conversion showing that

$$H^{\text{Metric}}(X) \geq \nu \Rightarrow H^{\text{Metric}}(X|Y = y) \geq \nu - \Delta.$$

The use of  $\text{Metric}^*$  entropy still captures the interesting aspects of [DP08] and [RTTV08] and

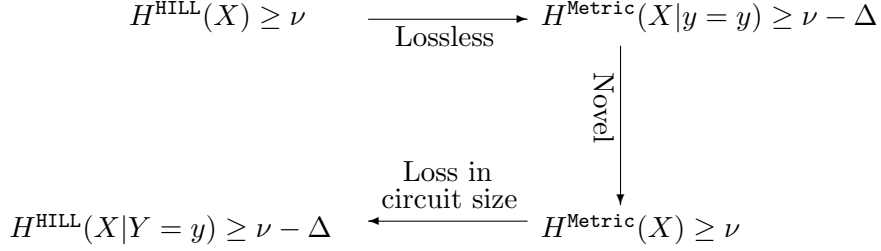


Figure 2: Structure of proof in [DP08], [RTTV08]

allows us to provide a tight reduction and will allow the proof of a “chain rule” (Theorem 3.6). This is because the chain rule only uses the second step multiple times, converting back to HILL entropy only once.

We now state the main technical lemma, a given leakage value decreases the  $\text{Metric}^*$  entropy quality and quantity proportionally to its probability:

**Lemma 3.5.** *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon/P_y, s'}^{\text{Metric}^*}(X|Y = y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - \log 1/P_y \quad (3)$$

where  $P_y = \Pr[Y = y]$  and  $s' \approx s$ .

The lemma is quite intuitive: the more surprising a leakage value is, the more it decreases the entropy. Its proof proceeds by contradiction: assuming that a distinguisher  $D$  exists for  $X|Y = y$ , we build a distinguisher  $D'$  for  $X$ . The structure of the proof is similar to the structure of the proof of [DP08], but the new  $D'$  is more complicated, as is the analysis of its performance. Here is the outline of the proof (see Section 3.4.1 for the details). Let  $\nu = H_{\epsilon, s}^{\text{Metric}^*}(X)$ .

1. Suppose  $D$  distinguishes  $X|Y = y$  from any distribution  $Z$  of min-entropy  $\nu - \Delta$ . Show that either for all such  $Z$ ,  $\mathbb{E}[D(Z)]$  is lower than  $\mathbb{E}[D(X|Y = y)]$ , or for all such  $Z$ ,  $\mathbb{E}[D(Z)]$  is higher than  $\mathbb{E}[D(X|Y = y)]$ . Assume the former without loss of generality. This initial step allows us to remove absolute values and to find a high-entropy distribution  $Z^+$  on which  $\mathbb{E}[D(Z)]$  is the highest.
2. Show that  $Z^+$  must be essentially flat, and let  $\alpha$  be  $\min_{z \in Z^+} D(z)$ .
3. Create a new distinguisher  $D'$  that rescales the output of  $D$  by stretching the interval  $[\alpha, 1]$  to  $[0, 1]$ . In other words, if the output  $x$  of  $D$  is less than  $\alpha$ , then  $D'$  outputs 0; else it outputs  $\frac{x-\alpha}{1-\alpha}$ . (In [DP08],  $D'$  chose a higher cut-off point and, instead of rescaling, simply output

0 if  $x$  was less than the cut-off point and 1 otherwise. It thus lost in performance if that particular cut-off point was not a good one. Rescaling allows us to avoid hitting a single bad cut-off point; cutting off at  $\alpha$  improves the performance of  $D'(X)$ .)

4. Show an upper bound on  $\mathbb{E}[D'(W)]$  for any  $W$  of min-entropy  $\nu$ , by using the fact that outside of  $Z^+$ ,  $D'$  outputs 0.
5. Show a lower bound on  $\mathbb{E}[D'(X)]$ , by first calculating  $\mathbb{E}[D'(X|Y = y)]$ .
6. Show a significant gap exists between these two values, regardless of  $\alpha$ .

Theorem 3.2 follows in a straightforward way from Lemma 3.5. In fact, the lemma allows us to prove a stronger version—a chain rule.

**Theorem 3.6.** *Let  $X, Y_1, Y_2$  be discrete random variables. Then*

$$H_{\epsilon|Y_2, s'}^{\text{Metric}^*}(X|Y_1, Y_2) \geq H_{\epsilon, s}^{\text{Metric}^*}(X|Y_1) - \log |Y_2|$$

where  $s' \approx s$ .

The proof of the theorem, presented in Section 3.4.2, first translates the conditional **Metric**<sup>\*</sup> entropy of  $X|Y_1$  to **Metric**<sup>\*</sup> entropy for each distribution  $X|Y_1 = y_1$ . We then apply Lemma 3.5 and obtain conditional **Metric**<sup>\*</sup> entropy by averaging over all points in  $Y_1, Y_2$ . Note that  $Y_1, Y_2$  do not need to be independent. (Indeed, this makes sense: if two leakage functions are correlated, then they are providing the adversary with less information.)

This combined with Lemma B.1 allows us to state a HILL-entropy version, as well.

**Corollary 3.7.** *Let  $X, Y_1, Y_2$  be discrete random variables and let  $\epsilon_{\text{HILL}} > 0$ . Then*

$$H_{\epsilon|Y_2 + \epsilon_{\text{HILL}}, s_{\text{HILL}}}^{\text{HILL}^*}(X|Y_1, Y_2) \geq H_{\epsilon, s}^{\text{Metric}^*}(X|Y_1) - \log |Y_2|$$

where  $s_{\text{HILL}} = \Omega\left(\frac{\epsilon_{\text{HILL}}^2}{\log |X|}\right)$ .

### 3.2 HILL $\Rightarrow$ HILL lemma

To facilitate comparison with [RTTV08] we present a “HILL-to-HILL” version of Lemma 3.5.

**Lemma 3.8.** *Let  $X$  be a discrete random variable over  $\chi$  and let  $Y$  be a discrete random variable. Then,*

$$H_{\epsilon', s'}^{\text{HILL}}(X|Y = y) \geq H_{\epsilon, s}^{\text{HILL}}(X) - \log 1/P_y \tag{4}$$

where  $P_y = \Pr[Y = y]$ ,  $\epsilon' = \epsilon/P_y + \epsilon_{\text{HILL}}$ , and  $s' = \Omega(s\epsilon_{\text{HILL}}^2/\log |\chi|)$ .

The lemma follows by application of Lemma 3.5 and [BSW03, Theorem 5.2]. By setting  $\epsilon_{\text{HILL}} = \Omega(\epsilon/P_y)$  one obtains the following result:

**Lemma 3.9.** *Let  $X$  be a discrete random variable over  $\chi$  and let  $Y$  be a discrete random variable. Then,*

$$H_{\epsilon', s'}^{\text{HILL}}(X|Y = y) \geq H_{\epsilon, s}^{\text{HILL}}(X) - \log 1/P_y \tag{5}$$

where  $P_y = \Pr[Y = y]$ ,  $\epsilon' = \Omega(\epsilon/P_y)$ , and  $s' = \Omega(s(\epsilon/P_y)^2/\log |\chi|)$ .

Recall the result from [RTTV08]:

**Lemma 3.4.** *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon', s'}^{\text{HILL}}(X|Y = y) \geq H_{\epsilon, s}^{\text{HILL}}(X) - \log 1/P_y \quad (6)$$

where  $P_y = \Pr[Y = y]$ ,  $\epsilon' = \Omega(\epsilon/P_y)$ , and  $s' = s/\text{poly}(P_y/\epsilon, \log 1/P_y)$

Note all the parameters are the same, except the losses in circuit size. The exact comparison is difficult because the polynomial in [RTTV08] is not specified, and  $\log |\chi|$  may be bigger or smaller than  $\log 1/P_y$ . However, the current work has the added benefit of the chain rule (Theorem 3.6) before the conversion back to HILL entropy. In the case of repeated leakage, the gain of only paying the Metric to HILL conversion once should dominate the difference between the two results.

### 3.3 Improvement for randomized leakage

There are many meaningful situations where there is randomness inherent in  $Y$  that has nothing to do with  $X$ . In this case we can prove a stronger result than in Theorem 3.2. The result is:

**Theorem 3.10.** *Let  $X, Y$  be discrete random variables and let  $L = \frac{2^{-\tilde{H}_\infty(X|Y)}}{\min_{x \in X} \Pr[X=x]}$ . Then  $H_{\epsilon L, s'}^{\text{Metric}^*}(X|Y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - \log L$  where  $s' \approx s$ .*

Notice that this result is the same as Theorem 3.2, except  $|Y|$  is replaced with  $L$ . For a uniform  $X$ , this theorem provides an optimal bound:

$$H_\infty(X) - \tilde{H}_\infty(X|Y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - H_{\epsilon(H_\infty(X) - \tilde{H}_\infty(X|Y)), s'}^{\text{Metric}^*}(X|Y)$$

where  $s' \approx s$ . However, because of the  $\min_{x \in X} \Pr[X = x]$  the result breaks down for repeated leakage, as the leakage  $Y$  can make a particular event arbitrarily unlikely. The intuition behind the theorem is the  $E[D'(X)]$  can be measured more carefully; the proof is in Section 3.4.3. Lemma 3.5 can also be improved for the case of randomized leakage: the improved version replaces  $P_y$  with  $P_y/\max_x \Pr[Y = y|X = x]$ .

### 3.4 Proofs

In this section we present a detailed exposition of our results. Let  $P_y = \Pr[Y = y]$  for the random variable  $y \in Y$ ,  $Y$  should be clear from context. Let  $1_B$  define the indicator random variable that is 1 when  $B$  is true and 0 otherwise.

#### 3.4.1 Proof of Lemma 3.5

Recall the main technical lemma.

**Lemma 3.5.** *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon/P_y, s'}^{\text{Metric}^*}(X|Y = y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - \log 1/P_y \quad (7)$$

where  $P_y = \Pr[Y = y]$  and  $s' \approx s$ .

*Proof.* Assume  $H_{\epsilon, s}^{\text{Metric}^*}(X) \geq \nu$ . We denote  $\epsilon' = \epsilon/P_y$ . Let  $\chi$  be the outcome space of  $X$ . We assume for contradiction that

$$H_{\epsilon'/P_y, s'}^{\text{Metric}^*}(X|Y = y) \geq \nu - \log 1/P_y$$

does not hold. By definition of metric entropy there exists distinguisher  $D_y \in \mathcal{D}_s^*$  such that  $\forall Z$  with  $H_\infty(Z) \geq \nu - \log 1/P_y$  we have

$$|E[D_y(X)|Y = y] - E[D_y(Z)]| \geq \epsilon'.$$

Let  $Z^-$  and  $Z^+$  be distributions of min-entropy  $\nu - \log 1/P_y$  minimizing  $\mathbb{E}[D_y(Z^-)]$  and maximizing  $\mathbb{E}[D_y(Z^+)]$  respectively. Let  $\beta^- \stackrel{\text{def}}{=} \mathbb{E}[D_y(Z^-)]$ ,  $\beta^+ \stackrel{\text{def}}{=} \mathbb{E}[D_y(Z^+)]$  and  $\beta \stackrel{\text{def}}{=} \mathbb{E}[D_y(X)|Y = y]$ .

**Claim 3.11.** *Either  $\beta^- < \beta^+ + \epsilon' \leq \beta$  or  $\beta \leq \beta^- - \epsilon' < \beta^+$ .*

From Equation 8 and the fact that  $Z^+, Z^-$  have min-entropy at least  $\nu - \log 1/P_y$  it suffices to show that either  $\beta^- \leq \beta^+ \leq \beta$  or  $\beta \leq \beta^- \leq \beta^+$ . Suppose it does not hold. Then we can define a distribution  $Z$  as a convex combination of  $Z^+, Z^-$  with  $H_\infty(Z) \geq \nu - \log 1/P_y$  and  $E[D_y(Z)] = \beta$ . This is a contradiction of Equation 8.

For the rest of the proof we will assume that the first case  $\beta^- < \beta^+ + \epsilon' \leq \beta$  holds<sup>3</sup>. Let  $\alpha = \min_{z \in Z^+} D_y(z)$ . We define a distinguisher  $D'_y$  as follows

$$D'_y(z) = \begin{cases} 0 & D_y(z) \leq \alpha \\ \frac{D_y(z) - \alpha}{1 - \alpha} & \text{otherwise} \end{cases} \quad (8)$$

**Claim 3.12.** *For all  $z$  if  $\Pr[Z^+ = z] \neq 2^{-\nu + \log 1/P_y}$ , then  $D_y(z) \leq \alpha$  and therefore  $D'_y(z) = 0$ .*

*Proof.* It suffices to show there does not exist a  $z$  where  $\Pr[Z^+ = z] < 2^{-\nu + \log 1/P_y}$  and  $D_y(z) > \alpha$ . Suppose for contradiction that there exists an  $z$  where  $\Pr[Z^+ = z] < 2^{-\nu + \log 1/P_y}$  and  $D_y(z) > \alpha$ . Choose a  $w$ ,  $\Pr[Z^+ = w] > 0$  such that  $D_y(w) = \alpha$ . Then define  $gap = \min\{2^{-\nu + \log 1/P_y} - \Pr[Z^+ = z], \Pr[Z^+ = w]\} > 0$ . Thus, we define a new distribution  $Z'$  identical to  $Z^+$  at all other points and with

$$\Pr[Z' = z] = \Pr[Z^+ = z] + gap, \quad \Pr[Z' = w] = \Pr[Z^+ = w] - gap$$

Note that  $H(Z') \geq \nu - \log 1/P_y$ . Then we have  $\mathbb{E}[D_y(Z')] > \mathbb{E}[D_y(Z^+)]$  which is a contradiction.  $\square$

**Claim 3.13.** *For all  $W$  over  $\chi$  where  $H_\infty(W) \geq \nu$ ,  $E[D'_y(W)] \leq \frac{\beta^+ - \alpha}{1 - \alpha} 2^{-\log 1/P_y}$ .*

*Proof.* One has

$$\mathbb{E}[D'_y(W)] = \sum_{z \in W} \Pr[W = z] 1_{D_y(z) \geq \alpha} \frac{D_y(z) - \alpha}{1 - \alpha}$$

---

<sup>3</sup>The main differences are that we work with  $Z^-$ , define  $\alpha = \max_{z \in Z^-} D_y(z)$ , and define  $D'_y(z)$  to be 0 when  $D_y(z) > \alpha$  and to be  $(1 - \frac{D_y(z)}{\alpha})$  otherwise. We provide a full proof in Appendix D.

By Claim 3.12 we can write the sum as

$$\mathbb{E}[D'_y(W)] = \sum_{z \in Z^+} \Pr[W = z] \frac{D_y(z) - \alpha}{1 - \alpha}$$

We also know that  $\Pr[W = z] \leq 2^{-\nu}$  and that for all terms that contribute to the sum  $\Pr[Z^+ = w] = 2^{-\nu + \log 1/P_y}$ . Thus one has

$$\begin{aligned} \mathbb{E}[D'_y(W)] &= \sum_{z \in Z^+} \Pr[W = z] \frac{D_y(z) - \alpha}{1 - \alpha} \\ &\leq \sum_{z \in Z^+} \frac{1}{2^\nu} \frac{D_y(z) - \alpha}{1 - \alpha} \\ &= \sum_{z \in Z^+} \frac{1}{2^{\nu - \log 1/P_y}} 2^{-\log 1/P_y} \frac{D_y(z) - \alpha}{1 - \alpha} \\ &= \sum_{z \in Z^+} \Pr[Z^+ = z] P_y \frac{D_y(z) - \alpha}{1 - \alpha} \\ &= \frac{\mathbb{E}[D_y(Z^+)] - \alpha}{1 - \alpha} P_y \\ &= \frac{\beta^+ - \alpha}{1 - \alpha} P_y \end{aligned}$$

This completes the claim. □

**Claim 3.14.**  $\mathbb{E}[D'_y(X)] \geq \frac{\beta - \alpha}{1 - \alpha} P_y$

*Proof.* One computes

$$\begin{aligned} \mathbb{E}[D'_y(X)] &= \mathbb{E}[D'_y(X)|Y = y] \Pr[Y = y] + \mathbb{E}[D'_y(X)|Y \neq y] \Pr[Y \neq y] \\ &\geq \mathbb{E}[D'_y(X)|Y = y] \Pr[Y = y] \\ &= \sum_{x \in X} \Pr[X = x|Y = y] \frac{D_y(x) - \alpha}{1 - \alpha} P_y 1_{D_y(x) \geq \alpha} \\ &\geq \sum_{x \in X} \Pr[X = x|Y = y] \frac{D_y(x) - \alpha}{1 - \alpha} P_y \\ &\geq \frac{\mathbb{E}[D_y(X|Y = y)] - \alpha}{1 - \alpha} P_y \\ &\geq \frac{\beta - \alpha}{1 - \alpha} P_y \end{aligned}$$

□

By combining Claims 3.13 and 3.14 we have that for all  $W$  over  $\chi$  with  $H_\infty(W) \geq \nu$  we have that

$$\mathbb{E}[D'_y(X)] - \mathbb{E}[D'_y(W)] \geq \frac{\beta - \alpha}{1 - \alpha} P_y - \frac{\beta^+ - \alpha}{1 - \alpha} P_y = \frac{\beta - \beta^+}{1 - \alpha} P_y \quad (9)$$

Now recall that by Claim 3.11 that  $\beta^+ + \epsilon' < \beta$ . One has:

$$\begin{aligned}\mathbb{E}[D'_y(X)] - \mathbb{E}[D'_y(W)] &\geq \frac{\beta - \beta^+}{1 - \alpha} P_y \\ &\geq (\beta - \beta^+) P_y > \epsilon' P_y = \epsilon\end{aligned}$$

Thus, we have successfully distinguished the distribution  $X$  from all distributions  $W$  of sufficient min-entropy. This is a contradiction.  $\square$

### 3.4.2 Proof of Theorems 3.6 and 3.2

We recall and prove Theorem 3.6 and note that Theorem 3.2 is a special case.

**Theorem 3.6.** *Let  $X, Y_1, Y_2$  be discrete random variables. Then  $H_{\epsilon|Y_2|s'}^{\text{Metric}^*}(X|Y_1, Y_2) \geq H_{\epsilon,s}^{\text{Metric}^*}(X|Y_1) - \log |Y_2|$  where  $s' \approx s$ .*

*Proof.* Assume  $H_{\epsilon,s}^{\text{Metric}^*}(X|Y_1) \geq \nu$ . By the definition of conditional metric entropy we know that there exists two functions  $\nu(\cdot)$  representing the metric entropy for each  $y_1$  and  $\epsilon(\cdot)$  representing the quality of distinguishing for each  $y_1$ . That is:

$$\epsilon, \nu : Y_1 \rightarrow \mathbb{R}^+ \cup \{0\}$$

subject to the following constraints where  $s' \approx s$ :

$$\begin{aligned}H_{\epsilon(y_1),s'}^{\text{Metric}^*}(X|Y_1 = y_1) &\geq \nu(y_1) \\ \sum_{y_1 \in Y_1} \Pr[Y_1 = y_1] \epsilon(y_1) &< \epsilon \\ H_{\epsilon,s}^{\text{Metric}^*}(X|Y_1) &= -\log \left( \mathbb{E}_{y_1 \in Y_1} [2^{-\nu(y_1)}] \right) \geq \nu\end{aligned}$$

Fix  $y_1 \in Y_1, y_2 \in Y_2$ . Let  $\epsilon(y_1, y_2) = \frac{\epsilon(y_1)}{\Pr[Y_2 = y_2 | Y_1 = y_1]}$  and let  $\Delta = -\log \Pr[Y_2 = y_2 | Y_1 = y_1]$  then by Lemma 3.5 we know that  $H_{\epsilon(y_1, y_2), s'}^{\text{Metric}^*}(X|Y_1 = y_1 \wedge Y_2 = y_2) \geq \nu(y_1) - \Delta$  where  $s' \approx s$ . Fix  $D \in \mathcal{D}_{s'}$ . Denote by  $Z_{y_1, y_2}$  a distribution with  $H_\infty(Z_{y_1, y_2}) \geq \nu(y_1) - \Delta$  and  $|\mathbb{E}[D(X|Y_1 = y_1, Y_2 = y_2, y_1, y_2)] - \mathbb{E}[D(Z_{y_1, y_2}, y_1, y_2)]| < \epsilon(y_1, y_2)$ . These  $Z_{y_1, y_2}$  give rise to a distribution  $Z$ . We calculate the performance of  $D$  on all of  $X, Z$

$$\begin{aligned}&|\mathbb{E}[D(X, Y_1, Y_2)] - \mathbb{E}[D(Z, Y_1, Y_2)]| \\ &= \sum_{y_1 \in Y_1} \sum_{y_2 \in Y_2} \Pr[Y_1 = y_1 \wedge Y_2 = y_2] |\mathbb{E}[D((X|Y_1 = y_1, Y_2 = y_2), y_1, y_2)] - \mathbb{E}[D(Z_{y_1, y_2}, y_1, y_2)]| \\ &< \sum_{y_1 \in Y_1} \sum_{y_2 \in Y_2} \Pr[Y_1 = y_1] \Pr[Y_2 = y_2 | Y_1 = y_1] \epsilon(y_1, y_2) \\ &= \sum_{y_1 \in Y_1} \Pr[Y_1 = y_1] \sum_{y_2 \in Y_2} \epsilon(y_1) \\ &= |Y_2| \sum_{y_1 \in Y_1} \Pr[Y_1 = y_1] \epsilon(y_1) < \epsilon |Y_2|\end{aligned}$$

It now suffices to show that  $Z$  has sufficient average min-entropy, we calculate:

$$\begin{aligned}
\tilde{H}_\infty(Z|Y_1, Y_2) &= -\log \left( \sum_{y_1 \in Y_1} \sum_{y_2 \in Y_2} \Pr[Y_1 = y_1 \wedge Y_2 = y_2] 2^{-H_\infty(Z_{y_1, y_2})} \right) \\
&\geq -\log \left( \sum_{y_1 \in Y_1} P_{y_1} \sum_{y_2 \in Y_2} \Pr[Y_2 = y_2 | Y_1 = y_1] 2^{-(\nu(y_1) - \Delta)} \right) \\
&= -\log \left( \sum_{y_1 \in Y_1} P_{y_1} \sum_{y_2 \in Y_2} \Pr[Y_2 = y_2 | Y_1 = y_1] \frac{2^{-(\nu(y_1))}}{\Pr[Y_2 = y_2 | Y_1 = y_1]} \right) \\
&= -\log |Y_2| - \log \left( \sum_{y_1 \in Y_1} P_{y_1} 2^{-\nu(y_1)} \right) \\
&\geq \nu - \log |Y_2|
\end{aligned}$$

This completes the proof. □

### 3.4.3 Proof of Theorem 3.10

The largest change is to the Claim 3.14, which gets replaced with the following.

**Claim 3.15.**  $\mathbb{E}[D'_y(X)] \geq \frac{\beta - \alpha}{1 - \alpha} \frac{\min_{x' \in X} \Pr[X = x']}{2^{-H_\infty(X|Y=y)}}$

*Proof.* One computes

$$\begin{aligned}
\mathbb{E}[D'_y(X)] &= \sum_x \Pr[X = x] D'_y(x) \\
&= \sum_x \Pr[X = x] D'_y(x) \frac{\max_{x' \in X} \Pr[X = x' | Y = y]}{\max_{x' \in X} \Pr[X = x' | Y = y]} \\
&= \frac{1}{2^{-H_\infty(X|Y=y)}} \sum_x \Pr[X = x] D'_y(x) \max_{x' \in X} \Pr[X = x] \Pr[Y = y] \\
&\geq \frac{1}{2^{-H_\infty(X|Y=y)}} \sum_x \min_{x' \in X} \Pr[X = x'] D'_y(x) \max_{x' \in X} \Pr[X = x | Y = y] \\
&\geq \frac{1}{2^{-H_\infty(X|Y=y)}} \sum_x \min_{x' \in X} \Pr[X = x'] D'_y(x) \Pr[X = x | Y = y] \\
&\geq \frac{\min_{x' \in X} \Pr[X = x']}{2^{-H_\infty(X|Y=y)}} \sum_x D'_y(x) \Pr[X = x | Y = y] \\
&= \frac{\min_{x' \in X} \Pr[X = x']}{2^{-H_\infty(X|Y=y)}} \frac{\beta - \alpha}{1 - \alpha}
\end{aligned}$$

□

This allows us to state a modified version of Lemma 3.5.

**Lemma 3.16.** *Let  $X, Y$  be discrete random variables. Then*

$$H_{\epsilon', s'}^{\text{Metric}^*}(X|Y = y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - (H_0(X) - H_\infty(X|Y = y)) \quad (10)$$

where  $\epsilon' = \frac{2^{-H_\infty(X|Y=y)}}{2^{-H_0(X)}}\epsilon$  and  $s' \approx s$ .

This allows us to state our modified theorem:

**Theorem 3.10.** *Let  $X, Y$  be discrete random variables and let  $L = \frac{2^{-\hat{H}_\infty(X|Y)}}{\min_{x \in X} \Pr[X=x]}$ . Then  $H_{\epsilon L, s'}^{\text{Metric}^*}(X|Y) \geq H_{\epsilon, s}^{\text{Metric}^*}(X) - \log L$  where  $s' \approx s$ .*

*Proof.* Assume  $H_{\epsilon, s}^{\text{Metric}^*}(X) \geq \nu$ . Then by Lemma 3.16 for each  $y \in Y$  we know that  $H_{\epsilon L_y, s'}^{\text{Metric}^*}(X|Y = y) \geq \nu - \log L_y$  where  $s' \approx s$  and  $L_y = \frac{2^{-H_\infty(X|Y=y)}}{\min_{x \in X} \Pr[X=x]}$ . Fix  $D \in \mathcal{D}_{s'}^*$ . Denote by  $Z_y$  a distribution with  $H_\infty(Z_y) \geq \nu - \log L_y$  and  $|\mathbb{E}[D(X|Y = y)] - \mathbb{E}[D(Z_y)]| < \epsilon L_y$ . These  $Z_y$  give rise to a distribution  $Z$ . We calculate the performance of  $D$  on all of  $X, Z$

$$\begin{aligned} |\mathbb{E}[D(X)] - \mathbb{E}[D(Z)]| &= \sum_{y \in Y} \Pr[Y = y] |\mathbb{E}[D((X|Y = y))] - \mathbb{E}[D(Z_y)]| \\ &< \sum_{y \in Y} \Pr[Y = y] \epsilon L_y \\ &= \frac{\epsilon}{\min_{x \in X} \Pr[X = x]} \sum_{y \in Y} \Pr[Y = y] 2^{-H_\infty(X|Y=y)} \\ &= \frac{\epsilon 2^{-\hat{H}_\infty(X|Y)}}{\min_{x \in X} \Pr[X = x]} = \epsilon L \end{aligned}$$

It now suffices to show that  $Z$  has sufficient average min-entropy, we calculate:

$$\begin{aligned} \tilde{H}_\infty(Z|Y) &= -\log \left( \sum_{y \in Y} \Pr[Y = y] 2^{-H_\infty(Z_y)} \right) \\ &\geq -\log \left( \sum_{y \in Y} P_y 2^{-(\nu - \log L_y)} \right) \\ &= -\log \left( \sum_{y \in Y} P_y 2^{-\nu} L_y \right) \\ &= -\log \left( \sum_{y \in Y} P_y 2^{-\nu} \frac{2^{-H_\infty(X|Y=y)}}{\min_{x \in X} \Pr[X = x]} \right) \\ &= \nu + \log \left( \min_{x \in X} \Pr[X = x] \right) - \log \left( \sum_{y \in Y} P_y 2^{-H_\infty(X|Y=y)} \right) \\ &= \nu + \log \left( \min_{x \in X} \Pr[X = x] \right) + \hat{H}_\infty(X|Y) \end{aligned}$$

This completes the proof. □

## Acknowledgements

The authors wish to thank Bhavana Kanukurthi and Peter Gacs for their insight in improving and simplifying our analysis. We would also like to thank Krzysztof Pietrzak and Sebastian Faust for helping to formulate the problem and important technical discussions.

## References

- [BSW03] B. Barak, R. Shaltiel, and A. Wigderson. Computational Analogues of Entropy. In *11th International Conference on Random Structures and Algorithms*, pages 200–215, 2003.
- [DORS08] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. Technical Report. *Previous version appeared at Advances in Cryptology - EUROCRYPT*, 38:79–100, 2008.
- [DP08] S. Dziembowski and K. Pietrzak. Leakage-Resilient cryptography. In *IEEE 49th Annual IEEE Symposium on Foundations of Computer Science, 2008.*, pages 293–302, 2008.
- [FKPR10] S. Faust, E. Kiltz, K. Pietrzak, and G. Rothblum. Leakage-Resilient Signatures. *Advances in Cryptology-Theory of Cryptography*, pages 343–360, 2010.
- [GT08] Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics*, pages 481–547, 2008.
- [HILL99] J. Hastad, R. Impagliazzo, L.A. Levin, and M. Luby. A Pseudorandom Generator from Any One-Way Function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.
- [HLR07] C.Y. Hsiao, C.J. Lu, and L. Reyzin. Conditional Computational Entropy, or Toward Separating Pseudoentropy from Compressibility. *Advances in Cryptology-EUROCRYPT 2007*, pages 169–186, 2007.
- [NZ93] N. Nisan and D. Zuckerman. Randomness is Linear in Space. *Journal of Computer and System Sciences*, pages 43–52, 1993.
- [RTTV08] O. Reingold, L. Trevisan, M. Tulsiani, and S. Vadhan. Dense Subsets of Pseudorandom Sets. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 76–85. IEEE, 2008.

## A HILL $\Leftrightarrow$ HILL\* and Metric $\Rightarrow$ Metric\*

**Lemma 2.1.**  $H_{\epsilon, s'}^{\text{HILL}}(X) \geq k \Leftrightarrow H_{\epsilon, s}^{\text{HILL}^*}(X) \geq k$ , for  $s' \approx s$ .

*Proof.* First, suppose  $H_{\epsilon, s}^{\text{HILL}^*}(X) < k$ —that is,  $\forall Y$  such that  $H_\infty(Y) \geq k$ ,  $\exists D \in \mathcal{D}_s^*$  such that  $|\mathbb{E}[D(X)] - \mathbb{E}[D(Y)]| > \epsilon$ . Fix  $Y$  with  $H_\infty(Y) \geq k$ . Choose  $D \in \mathcal{D}_s^*$  such that  $|\mathbb{E}[D(X)] - \mathbb{E}[D(Y)]| > \epsilon$ . Then we construct a distinguisher  $D'(\cdot)$  as follows:

$$D'(x) = \begin{cases} 0 & \text{with probability } 1 - D(x) \\ 1 & \text{with probability } D(x) \end{cases}$$

It is clear that  $D' \in \mathcal{D}_{s'}$  for  $s'$  close to  $s$  ( $D'$  can be implemented by choosing a random number  $r \in [0, 1)$  of the same precision as the output of  $D$  and performing a single comparison: output 1 if  $r < D(x)$  and 0 otherwise). Note also that for all  $x$ ,  $\mathbb{E}[D'(x)] = D(x)$ , and therefore  $\forall X, \mathbb{E}[D'(X)] = \mathbb{E}[D(X)]$ , and thus

$$|\mathbb{E}[D'(X)] - \mathbb{E}[D'(Y)]| = |\mathbb{E}[D(X)] - \mathbb{E}[D(Y)]| > \epsilon,$$

which implies that  $H_{\epsilon, s'}^{\text{HILL}}(X) < k$ .

Now suppose  $H_{\epsilon, s'}^{\text{HILL}}(X) < k$ —that is,  $\forall Y$  such that  $H_\infty(Y) \geq k$ ,  $\exists D \in \mathcal{D}_s$  such that  $|\mathbb{E}[D(X; U)] - \mathbb{E}[D(Y; U)]| > \epsilon$ . Here  $U$  is the distribution of the randomness needed by the circuit  $D$ . Fix  $Y$  with  $H_\infty(Y) \geq k$ . Choose  $D \in \mathcal{D}_s$  such that  $|\mathbb{E}[D(X; U)] - \mathbb{E}[D(Y; U)]| > \epsilon$ . Thus we have that

$$\begin{aligned} \epsilon &< \left| \mathbb{E}_{u \in U}[D(X; u)] - \mathbb{E}_{u \in U}[D(Y; u)] \right| \\ &\leq \sum_{u \in U} \frac{1}{|U|} |\mathbb{E}[D(X; u)] - \mathbb{E}[D(Y; u)]| \end{aligned}$$

Thus, there exists a  $u \in U$  such that  $|\mathbb{E}[D(X; u)] - \mathbb{E}[D(Y; u)]| > \epsilon$ . We hardwire that  $u$  in place of the random gates of  $D$  to define  $D'(\cdot)$  which on input  $x$  outputs the result of  $D(x; u)$ . Clearly,  $D \in \mathcal{D}_s^*$  and  $|\mathbb{E}[D'(X)] - \mathbb{E}[D'(Y)]| = |\mathbb{E}[D(X; u)] - \mathbb{E}[D(Y; u)]| > \epsilon$ . This completes the proof.  $\square$

**Lemma 2.2.**  $H_{\epsilon, s'}^{\text{Metric}}(X) \geq k \Rightarrow H_{\epsilon, s}^{\text{Metric}^*}(X) \geq k$ , for  $s' \approx s$ .

*Proof.* Assume not. Then there exists  $D \in \mathcal{D}_s^*$  such that  $\forall Y, \forall Y$  with  $H_\infty(Y) \geq k$ , we have  $|E[D(X)] - E[D(Y)]| > \epsilon$ . Build  $D' \in \mathcal{D}_{s'}$  out of  $D$  the same way as in the corresponding case of Lemma 2.1.  $\square$

## B Conditional Metric\* implies conditional HILL\*

To demonstrate that the equivalences of Figure 2.2 hold in the conditional case we present the most complicated case. That is, we extend the proof of [BSW03] of equivalence of HILL\* and Metric\* to the conditional case. The other transformations are done similarly by breaking the conditional distribution into the component distributions  $X|Y = y$  performing the transformation there and translating back.

**Lemma B.1.** *Let  $X, Y$  be distributions with discrete outcome spaces with  $X$  taking values over  $\chi$ . Then  $\forall \epsilon, \epsilon_{\text{HILL}} > 0, k$  if  $H_{\epsilon, s}^{\text{Metric}^*}(X|Y) \geq k$  then  $H_{\epsilon + \epsilon_{\text{HILL}}, s_{\text{HILL}}}^{\text{HILL}^*}(X|Y) \geq k$  where  $s_{\text{HILL}} \approx \epsilon_{\text{HILL}}^2 s / \log |\chi|$ .*

*Proof.* By the definition of average min-entropy we know that there exists two functions,  $\nu(\cdot), \epsilon(\cdot)$  that represent the quantity of each conditional entropy ( $H(X|Y = y)$ ) and the quality of each conditional entropy respectively. More formally there exists functions:

$$\epsilon, \nu : Y \rightarrow \mathbb{R}^+ \cup \{0\}$$

subject to the following constraints where  $s' \approx s$ :

$$\begin{aligned} H_{\epsilon(y), s'}^{\text{Metric}^*}(X|Y = y) &\geq \nu(y) \\ \sum_{y \in Y} P_y \epsilon(y) &< \epsilon \\ H_{\epsilon, s}^{\text{Metric}^*}(X|Y) &= -\log \left( \mathbb{E}_{y \in Y} [2^{-\nu(y)}] \right) \geq k \end{aligned}$$

We make no claims about any of the individual values only their averaged values. Then by Theorem 2.3 we know that for each  $y$

$$H_{\epsilon(y)+\epsilon_{HILL},s_{HILL}}^{\text{HILL}^*}(X|Y=y) \geq \nu(y)$$

where  $s_{HILL} \approx \epsilon_{HILL}^2 s / \log |\mathcal{X}|$ . Thus we know there exists a collection of distributions  $Z_y$  with  $H_\infty(Z_y) \geq \nu(y)$  giving rise to the distribution  $Z$ . Let  $D \in \mathcal{D}_{s_{HILL}}^*$ , we calculate the performance of  $D$  on  $(Z, Y)$ :

$$\begin{aligned} |E[D(X, Y)] - E[D(Z, Y)]| &= \sum_{y \in Y} P_y |E[D(X, y)] - E[D(Z, y)]| \\ &< \sum_{y \in Y} P_y (\epsilon(y) + \epsilon_{HILL}) \\ &< \epsilon + \sum_{y \in Y} P_y \epsilon_{HILL} \\ &= \epsilon + \epsilon_{HILL} \leq \epsilon' \end{aligned}$$

Then note that

$$\begin{aligned} \tilde{H}_\infty(Z|Y) &\geq -\log \left( \mathbb{E}_{y \in Y} [2^{-H_\infty(Z|Y=y)}] \right) \\ &\geq -\log \left( \mathbb{E}_{y \in Y} [2^{\nu(y)}] \right) \\ &\geq k \end{aligned}$$

That is  $\tilde{H}_\infty(Z|Y) \geq k$ . Thus we know that  $H_{\epsilon+\epsilon_{HILL},s_{HILL}}^{\text{HILL}^*}(X|Y) \geq k$ . □

## C Proof of Lemma 3.3

**Lemma 3.3.** *Let  $X, Y$  be discrete random variables with  $|Y| \leq 2^\lambda$  and  $H_{\epsilon_{ent},s}^{\text{Metric}^*}(X) \geq \nu$ , then for any  $\epsilon_1, \epsilon_2, \Delta > 0$  satisfying  $\epsilon_{ent} \leq \epsilon_1 \epsilon_2 / 2^\lambda$  and  $2^{-\Delta} \leq \epsilon_2 / 2^\lambda$ ,*

$$\Pr_{y \in Y} [H_{\epsilon_1, s'}^{\text{Metric}^*}(X|Y=y) \geq \nu - \Delta] \leq 1 - \epsilon_2$$

where  $s' \approx s$ .

*Proof.* The proof is similar to the corresponding portion of the proof of [DP08, Lemma 3]. We denote  $P_y = \Pr[Y = y]$ . Assume for contradiction that the theorem does not hold. By the definition of metric entropy, there exists a set

$$S \text{ where } \Pr[y \in S] > \epsilon_2$$

and  $\forall y \in S, \exists D_y \in \mathcal{D}_s^*$  such that for all random variables  $Z$  with  $H_\infty(Z) \geq m - \Delta$  we have

$$|\mathbb{E}[D_y(Z)] - \mathbb{E}[D_y(X|Y=y)]| \geq \epsilon_1. \quad (11)$$

Then consider a  $y \in S$  such that

$$P_y > 2^{-\lambda} \epsilon_2.$$

Such a  $y$  exists, because otherwise  $\Pr[y \in S] \leq |Y|2^{-\lambda}\epsilon_2 = \epsilon_2$ . Then by Lemma 3.5 we know that

$$\begin{aligned} H_{\epsilon_{ent}/P_y, s'}^{\text{Metric}^*}(X|Y=y) &\geq \nu - \log 1/P_y, \text{ and therefore} \\ H_{2^\lambda\epsilon_{ent}/\epsilon_2, s'}^{\text{Metric}^*}(X|Y=y) &> \nu - \log 2^\lambda/\epsilon_2, \end{aligned}$$

where  $s' \approx s$ . Thus, there exists a  $Z$  with  $H_\infty(Z) \geq \nu - \Delta \geq \nu - \log \epsilon_2/2^\lambda$  for which

$$|\mathbb{E}[D_y(Z)] - \mathbb{E}[D_y(X|y=y)]| < 2^\lambda\epsilon_{ent}/\epsilon_2 \leq \epsilon_1.$$

This contradicts equation 11. □

## D Lemma 3.5 using other side of Claim 3.11

In this section we proceed to show that Lemma 3.5 is valid if the assumption is made that  $\beta \leq \beta^- - \epsilon < \beta^+$  in Claim 3.11. The other choice is made in text of the paper. We continue the proof directly from the definition of  $D'_y$ .

*Proof.* Assume that  $\beta \leq \beta^- - \epsilon < \beta^+$ . Let  $\alpha = \max_{z \in Z^-} D_y(z)$ . We define a distinguisher  $D'_y$  as follows

$$D'_y(z) = \begin{cases} 0 & D_y(z) \geq \alpha \\ 1 - \frac{D_y(z)}{\alpha} & \text{otherwise} \end{cases} \quad (12)$$

**Claim D.1.** *For all  $z$  if  $\Pr[Z^- = z] \neq 2^{-\nu+\log 1/P_y}$ , then  $D_y(z) \geq \alpha$  and therefore  $D'_y(z) = 0$ .*

*Proof.* It suffices to show there does not exist a  $z$  where  $\Pr[Z^- = z] < 2^{-\nu+\log 1/P_y}$  and  $D_y(z) < \alpha$ . Suppose for contradiction that there exists an  $z$  where  $\Pr[Z^- = z] < 2^{-\nu+\log 1/P_y}$  and  $D_y(z) < \alpha$ . Choose a  $w$ ,  $\Pr[Z^- = w] > 0$  such that  $D_y(w) = \alpha$ . Then define  $gap = \min\{2^{-\nu+\log 1/P_y} - \Pr[Z^- = z], \Pr[Z^- = w]\} > 0$ . Thus, we define a new distribution  $Z'$  identical to  $Z^-$  at all other points and with

$$\begin{aligned} \Pr[Z' = z] &= \Pr[Z^- = z] + gap \\ \Pr[Z' = w] &= \Pr[Z^- = w] - gap \end{aligned}$$

Note that  $H(Z') \geq \nu - \log 1/P_y$ . Then we have  $E[D_y(Z')] < E[D_y(Z^-)]$  which is a contradiction. □

**Claim D.2.** *For all  $W$  over  $\chi$  where  $H_\infty(W) \geq \nu$ ,  $E[D'_y(W)] \leq \left(1 - \frac{\beta^-}{\alpha}\right) P_y$ .*

*Proof.* One has

$$E[D'_y(W)] = \sum_{z \in W} \Pr[W = z] 1_{D_y(z) \leq \alpha} \left(1 - \frac{z}{\alpha}\right)$$

By Claim D.1 we can write the sum as

$$E[D'_y(W)] = \sum_{z \in Z^-} \Pr[W = z] \left(1 - \frac{z}{\alpha}\right)$$

We also know that  $\Pr[W = z] \leq 2^{-\nu}$  and that for all terms that contribute to the sum  $\Pr[Z^- = w] = 2^{-\nu + \log 1/P_y}$ . Thus one has

$$\begin{aligned}
E[D'_y(W)] &= \sum_{z \in Z^+} \Pr[W = z] \left(1 - \frac{z}{\alpha}\right) \\
&\leq \sum_{z \in Z^+} \frac{1}{2^\nu} \left(1 - \frac{z}{\alpha}\right) \\
&= \sum_{z \in Z^+} \frac{1}{2^{\nu - \log 1/P_y}} 2^{-\log 1/P_y} \left(1 - \frac{z}{\alpha}\right) \\
&= \sum_{z \in Z^+} \Pr[Z^+ = z] P_y \left(1 - \frac{z}{\alpha}\right) \\
&= \left(1 - \frac{E[D_y(Z^-)]}{\alpha}\right) P_y = \left(1 - \frac{\beta^-}{\alpha}\right) P_y
\end{aligned}$$

This completes the claim.  $\square$

**Claim D.3.**  $E[D'_y(X)] \geq \left(1 - \frac{\beta}{\alpha}\right) P_y$

*Proof.* One computes

$$\begin{aligned}
E[D'_y(X)] &= E[D'_y(X)|Y = y] \Pr[Y = y] + E[D'_y(X)|Y \neq y] \Pr[Y \neq y] \\
&\geq E[D'_y(X)|Y = y] \Pr[Y = y] \\
&= \sum_{x \in X} \Pr[X = x|Y = y] \left(1 - \frac{D_y(x)}{\alpha}\right) P_y \mathbf{1}_{D_y(x) \leq \alpha} \\
&\geq \sum_{x \in X} \Pr[X = x|Y = y] \left(1 - \frac{D_y(x)}{\alpha}\right) P_y \\
&\geq \left(1 - \frac{E[D_y(X|Y = y)]}{\alpha}\right) P_y \\
&\geq \left(1 - \frac{\beta}{\alpha}\right) P_y
\end{aligned}$$

$\square$

By combining Claims D.2 and D.3 we have that for all  $W$  over  $\chi$  with  $H_\infty(W) \geq \nu$  we have that

$$E[D'_y(X)|Y = y] - E[D'_y(W)] \geq \left(1 - \frac{\beta}{\alpha}\right) P_y - \left(1 - \frac{\beta^-}{\alpha}\right) P_y = \frac{\beta^- - \beta}{\alpha} P_y \quad (13)$$

Now recall that by Claim 3.11 that  $\beta^- - \epsilon' > \beta$ . One has:

$$\begin{aligned}
\mathbb{E}[D'_y(X)] - \mathbb{E}[D'_y(W)] &\geq \frac{\beta^- - \beta}{\alpha} P_y \\
&= (\beta^- - \beta) P_y > \epsilon' P_y = \epsilon
\end{aligned}$$

Thus, we have successfully distinguished the distribution  $X$  from all distributions  $W$  of sufficient min-entropy. This is a contradiction.  $\square$