The Chain Rule for HILL Pseudoentropy, Revisited

Krzysztof Pietrzak[1⋆], Maciej Skórski[2⋆⋆]

[1] IST Austria pietrzak@ist.ac.at
[2] University of Warsaw maciej.skorski@gmail.com

Abstract. Computational notions of entropy (a.k.a. pseudoentropy) have found many applications, including leakage-resilient cryptography, deterministic encryption or memory delegation. The most important tools to argue about pseudoentropy are chain rules, which quantify by how much (in terms of quantity and quality) the pseudoentropy of a given random variable X decreases when conditioned on some other variable Z (think for example of X as a secret key and Z as information leaked by a side-channel). In this paper we give a very simple and modular proof of the chain rule for HILL pseudoentropy, improving best known parameters. Our version allows for increasing the acceptable length of leakage in applications up to a constant factor compared to the best previous bounds. As a contribution of independent interest, we provide a comprehensive study of all known versions of the chain rule, comparing their worst-case strength and limitations.

1 Introduction

Min-entropy. Various notions of entropy are used to quantify the randomness in a random variable. The most important notion in cryptographic contexts is min-entropy, where a variable X (conditioned on Z) has min-entropy k if one cannot guess X (given Z) with probability better than 2^k.

Definition 1. *The* **min-entropy** *of a variable X is*

$$H_\infty(X) = -\log \max_x \Pr[X = x]$$

More generally, for a joint distribution (X, Z), the **average min-entropy** *of X conditioned on Z is [DRS04]*

$$\begin{aligned}\widetilde{\mathbf{H}}_\infty(X|Z) &= -\log \mathbb{E}_{z\leftarrow Z} \max_x \Pr[X = x|Z = z] \\ &= -\log \mathbb{E}_{z\leftarrow Z} 2^{-H_\infty(X|Z=z)} \ .\end{aligned}$$

⋆ Research supported by ERC starting grant (259668-PSPC)
⋆⋆ Research supported by the Ideas for Poland grant 2/2011 from the Foundation for Polish Science

Chain-Rules. Most entropy notions $H(.)$ satisfy a chain rule which roughly capture the fact that when additionally conditioning on a variable Z, the entropy can decrease by at most its length $|Z|$, i.e.,

$$H(X|Y,Z) \geq H(X|Y) - |Z| \tag{1}$$

In particular, average-case min-entropy satisfies such a rule [DRS04]

$$\widetilde{\mathbf{H}}_\infty(X|Y,Z) \;\geq\; \widetilde{\mathbf{H}}_\infty(X|Y) - H_0(Z) \;\geq\; \widetilde{\mathbf{H}}_\infty(X|Y) - |Z|\,, \tag{2}$$

where $H_0(Z) \leq |Z|$ denotes the logarithm of the support-size of Z.

Pseudoentropy. Information theoretic entropy notions refer to computationally unbounded parties, e.g., no algorithm can compress a distribution X (given Z) below its Shannon entropy $H(X|Z)$ and no algorithm can guess X (given Z) better than with probability $2^{-\widetilde{\mathbf{H}}_\infty(X|Z)}$. Under computational assumptions, in particular in cryptographic settings, one often has to deal with distribution that appear to have high entropy only for computationally bounded parties.

The classical example is a pseudorandom distribution [BM84, Yao82], where $X \in \{0,1\}^n$ is said to be pseudorandom if it cannot be distinguished from the uniform distribution over $\{0,1\}^n$ by polynomial size distinguishers. In this case X appears to have n bits of Shannon and n bits of min-entropy. More generally, $X \in \{0,1\}^n$ has k bits of HILL entropy, if it cannot be distinguished from some distribution Y with k bits of min-entropy. Note that for $k = n$ HILL entropy is simply pseudorandomness, as the only distribution over $\{0,1\}^n$ with n bits of min-entropy is the uniform distribution. HILL entropy was introduced in [HILL99], the more general conditional notion below is from [HLR07].

Definition 2 ([HLR07]). *Let (X,Z) be a joint distribution of random variables. Then X has* **conditional HILL entropy** *k conditioned on Z, denoted by $\mathbf{H}^{\text{HILL}}_{\varepsilon,s}(X|Z) \geq k$, if there exists a joint distribution (Y,Z) such that $\widetilde{\mathbf{H}}_\infty(Y|Z) \geq k$, and $(X,Z) \sim_{\varepsilon,s} (Y,Z)$.*[3]

Computational notions of entropy find important applications in leakage-resilient cryptography [DP08b], deterministic encryption [FOR12], memory delegation [CKLR11], computational complexity [RTTV08a] and foundations of cryptography [HRV10].

Chain Rules for Computational Entropy. When considering chain rules as in as in eq.(1) for computational notions of entropy, one must not only specify by how much the *quantity* of the entropy decreases, but also its *quality*. For some computational entropy notions like Yao or unpredictability entropy, chain rules are very easy to prove, and have been folklore for a long time (for the short proofs cf.

[3] Let us stress that using the same letter Z for the 2nd term in (X,Z) and (Y,Z) means that we require that the marginal distribution Z of (X,Z) and (Y,Z) is the same.

Appendix A in [KPWW14]). For HILL entropy the situation is much more complicated. The first chain rules were found independently by [RTTV08b, DP08a], and several proofs for the chain rule for HILL entropy were given subsequently, often as a corollary of a more general result. The various proofs give different qualitative bounds and are summarised below.

Theorem 1 (Chain Rules for HILL Entropy). *For any joint distribution* $(X, Z) \in \{0,1\}^n \times \{0,1\}^m$ *we have that*

$$\mathbf{H}^{\mathrm{HILL}}_{\varepsilon',s'}(X|Z) \geqslant \mathbf{H}^{\mathrm{HILL}}_{\varepsilon,s}(X) - m - \Delta \tag{3}$$

where $\varepsilon' = \varepsilon \cdot p(2^\ell, \varepsilon^{-1})$ *and* $s' = s/q(2^\ell, \varepsilon^{-1})$*, for some polynomial functions* $p(.)$ *and* $q(.)$ *as summarised in Table 1 (*$\Delta = 0$ *except for [DP08b], where* $\Delta = 2\log(1/\epsilon)$.

Reference	Technique	$s' =$	$\epsilon' =$	Meaningful range
(a) [DP08b]	Worst-Case Metric Entropy	$\Omega\left(s \cdot 2^{2m}\epsilon^2\right)$	$O(\sqrt{2^m\epsilon})$	$s > 2^{-2m}\epsilon^{-2},\ 2^{-m} > \epsilon$
(b) [RTTV08b]	Dense Model Theorem	$\Omega\left(s \cdot \mathrm{poly}(\epsilon, \min_z(\Pr[Z=z]))\right)$	$O(2^m\epsilon)$	$s > \max_z \frac{1}{\Pr[Z=z]^2} \cdot \epsilon^{-2},\ 2^{-m} > \epsilon$
(c) [FOR12]	Worst-Case Metric Entropy	$\Omega\left(s \cdot 2^{2m}\epsilon^2\right)$	$O(2^m\epsilon)$	$s > 2^{2m}\epsilon^{-2},\ 2^{-m} > \epsilon$
(d) [JP14]	Simulating Auxiliary Inputs	$\Omega\left(s \cdot \frac{\epsilon^2}{2^{3m}} - 2^m\right)$	$O(\epsilon)$	$s > 2^{4m}\epsilon^{-2} + 2^{3m}\epsilon^{-2}$
(e) [VZ13]	Simulating Auxiliary Inputs	$\Omega\left(s \cdot \frac{\epsilon^2}{2^m} - \frac{1}{\epsilon^2} - 2^m\right)$	$O(\epsilon)$	$s > 2^m\epsilon^{-4} + 2^{2m}\epsilon^{-2} + 2^m\epsilon^{-2}$
(f) **This paper** using [GW10]	Relaxed HILL Entropy	$\Omega\left(s \cdot \frac{\epsilon^2}{2^m} - 2^m\right)$	$O(\epsilon)$	$s > 2^{2m}\epsilon^{-2} + 2^m\epsilon^{-2}$
(g) **This paper**	Average Metric Entropy	$\Omega\left(s \cdot \frac{\epsilon^2}{2^m} - 2^m\epsilon^2\right)$	$O(\epsilon)$	$s > 2^m\epsilon^{-2} + 2^{2m}$

Table 1: Qualitative bounds on chain rules for HILL entropy. For simplicity, smaller order terms $\log(1/\epsilon)$, n, m are hidden under the big-O notation.

As shown in the table, every chain rule losses a factor exponential in m in quality (either in the size s or in the advantage ϵ) and also a factor $\mathrm{poly}(\epsilon)$. The second loss is the reason for poor security bounds in applications, for example in security proofs for leakage resilient stream ciphers (cf. [Pie09] and related papers), but seems unavoidable given the current state of the art. The choice of whether we lose 2^m in size or advantage depends on an application, as we will see later.

All the chain rules in Table 1 can be slightly generalized. Namely, one can opt for a larger s' at the prize of a larger ϵ'. This is possible because the common part of all the corresponding proofs is an approximation argument (typically by the Chernoff Bound). The most general statements can be found in Table 2 below. Table 1 is recovered from Table 2 by setting the free parameter δ to be of the same order as the other additive term in ϵ' (ϵ or $2^m\epsilon$ in the table), in order to get the smallest possible ϵ', while keeping the number of parameters small. We stress that in later discussions, including applications and our results described in Section 1.2, we refer to the general bounds from Table 2.

New chain rule. We prove the following results

Reference	Technique	$s' =$	$\epsilon' =$
(a) [DP08b]	Worst-Case Metric Entropy	$\Omega\left(s \cdot 2^{2m}\delta^2\right)$	$O(\sqrt{2^m\epsilon} + \delta)$
(b) [RTTV08b]	Dense Model Theorem	$\Omega\left(s \cdot \frac{\delta^2}{\max_z(\Pr[Z=z])^2}\right)$	$O(2^m\epsilon + \delta)$
(c) [FOR12]	Worst-Case Metric Entropy	$\Omega\left(s \cdot \delta^2\right)$	$O(2^m\epsilon + \delta)$
(d) [JP14]	Simulating Auxiliary Inputs	$\Omega\left(s \cdot \frac{\delta^2}{2^{3m}} - 2^m\right)$	$O(\epsilon + \delta)$
(e) [VZ13]	Simulating Auxiliary Inputs	$\Omega\left(s \cdot \frac{\delta^2}{2^m} - \frac{1}{\delta^2} - 2^m\right)$	$O(\epsilon + \delta)$
(f) [GW10]	Relaxed HILL Entropy	$\Omega\left(s \cdot \frac{\delta^2}{2^m} - 2^m\right)$	$O(\epsilon + \delta)$
(g) **This paper**	Average Metric Entropy	$\Omega\left(s \cdot \frac{\delta^2}{2^m} - 2^m\delta^2\right)$	$O(\epsilon + \delta)$

Table 2: Qualitative bounds on chain rules for HILL entropy, in the most general form with the free parameter δ.

Theorem 2 (Chain rule for metric entropy with loss in size). *Let $X \in \{0,1\}^n$ and $Z \in \{0,1\}^m$ be correlated random variables. Then for any (ϵ, s) we have*

$$\mathbf{H}^{\mathrm{Metric,det},[0,1]}_{\epsilon',s'}(X|Z) \geqslant \mathbf{H}^{\mathrm{Metric,det},[0,1]}_{\epsilon,s}(X) - m \tag{4}$$

where $s' = s/2^m - 2^m$ and $\epsilon' = \epsilon$.

Corollary 1. *Let $X \in \{0,1\}^n$ and $Z \in \{0,1\}^m$ be correlated random variables. Then for any (ϵ, s) we have*

$$\mathbf{H}^{\mathrm{HILL}}_{\epsilon',s'}(X|Z) \geqslant \mathbf{H}^{\mathrm{HILL}}_{\epsilon,s}(X) - m \tag{5}$$

where $s' = \Omega\left(\frac{s}{2^m} \cdot \frac{\delta^2}{n+1-k} - 2^m \cdot \frac{\delta^2}{n+1-k}\right)$, $\epsilon' = \epsilon + \delta$, δ is arbitrary and $k = \mathbf{H}^{\mathrm{HILL}}_{\epsilon,s}(X)$ (actually $k = \mathbf{H}^{\mathrm{Metric,det},[0,1]}_{(\epsilon,s)}(X|Z)$ is enough).

The proofs can be found in Section 3. Our new chain rule (g) loses a leakage-dependent factor in s instead in ϵ, and can be viewed as complementary with respect to (c) which loses it only in ϵ. Later we will see that there are settings where both chain rules gives equivalent security (basically when ϵ can be chosen sufficiently small), but for other cases our chain rule might be preferable (when we start with moderate values of ϵ and aim for relatively small ϵ'). We will discuss these applications with practically meaningful numerical examples in Section 1.2.

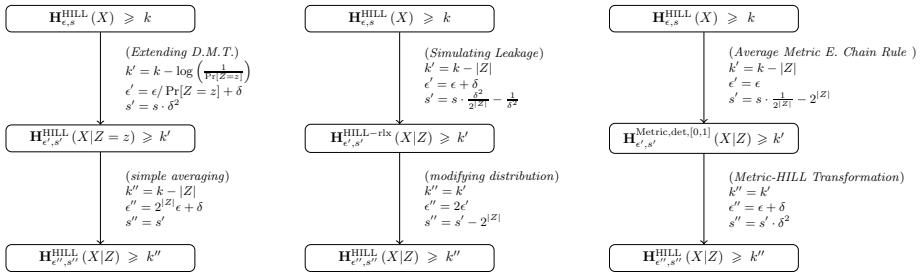
1.1 Proofs Techniques for Chain Rules

Basically, all the previously known chain rules have been obtained by one of the two following ways:

(a) bounding pseudoentropy for every leakage value separately

(b) using so called relaxed pseudoentropy

The first technique, so called decomposable entropy [FR12], can be viewed as an extension of the dense model theorem which is equvialent when the entropy amount is full (this equivalence holds up to a constant factor as demonstrated in [Sko15a]); this approach yields always an exponential (in m) loss for ϵ. The second way is to use the so called "relaxed" pseudoentropy, which offer an exponential (in m) loss for s, but no loss in ϵ. In this paper we come up with a different approach, namely we first prove a variant of a chain rule for average metric entropy which loses only in s and then use known transformations to convert it back to HILL entropy. As shown in Table 1 our approach yields best possible known loss in s compared to the known chain rules which do not decrease ϵ.



(a) Proofs based on bounding the pseudoentropy for every leakage outcome separately, which is an extension of the Dense Model Theorem: chain rules (a),(b) and (c).

(b) Proofs of going through relaxed pseudoentropy. The first step is either by a direct argument (chain rule (f)) or by leakage simulating techniques (chain rules (d) and (e)).

(c) **this paper**: A proof going through average metric entropy directly (g).

Fig. 1: Chain rules classified by used proof techniques.

1.2 Qualitative Comparison

Table 1 summarizes the known and our new bonds for the HILL entropy chain rule. In the first three bounds (a) to (c) the advantage $\epsilon' = 2^m\epsilon$ degrades exponentially in m (with the best result achieved by (c)), whereas in the bounds (d) to (g) we one have a degradation in the circuit size s', but the distinguishing advantage ϵ' stays the same (up to some small constant hidden on the big-Oh notation, which we'll ignore for the rest of this section).

The degradation in circuit size for all bounds (d) to (g) is of the form $s' = s/\alpha - \beta$, so the circuit size degrades by a factor α and an additive term β. The best factor $\alpha = 2^m/\epsilon^2$ is achieved by the bounds (e) to (g), and the best additive loss is achieved by (g). Below we give a numerical example showing that for some practical settings of parameters this really matters, and (g) gives meaningful security guarantees whereas (d),(e) and (f) don't. Comparing (g)

with (c) is less straight forward, because (c) has a degradation in the advantage whereas (g) does not. To get a meaningful comparison, we consider first settings where we can assume that the running time to advantage ratio s/ϵ is fixed, and then discuss the case when no such a simple tradeoff exists.

Fixed time-success ratio (application for weak PRFs). For concreteness, we assume that $X = (x_1, F(K, x_1), \ldots, (x_\ell, F(K, x_\ell))$ consists of input-output pairs of a weak PRF $F(.,.)$ with key K, and we want to know how good the HILL entropy of X is given some m bits of leakage about K. This is the setting in which the chain rule is e.g. used in the security proof of the leakage-resilient stream-cipher from [Pie09]. For example think of $F(.,.)$ as AES256, and assume its security as a weak PRF satisfies $s/\epsilon \approx 2^{256}$, which is the case if bruce force key-search is the best attack.[4] Under this assumption, the degradation in circuit size in [FOR12] and our new bounds (g)) are identical as illustrated with a concrete numerical example in Table 3.

Chain Rule	Before leakage		Leakage	After leakage			
	ϵ	s	m	$s' \approx$		$\epsilon' \approx$	
(e) [VZ13]	2^{-55}	2^{201}	46	$s \cdot \frac{\epsilon^2}{2^m} - \frac{1}{\epsilon^2} - 2^m$	< 0	ϵ	2^{-55}
(d) [JP14]				$s \cdot \frac{\epsilon^2}{2^{3m}} - 2^m$	< 0		
(f) [GW10]				$s \cdot \frac{\epsilon^2}{2^m} - 2^m$	< 0		
(g) **this paper**				$s \cdot \frac{\epsilon^2}{2^m} - 2^m\epsilon^2$	2^{45}		
(c) [FOR12]	2^{-101}	2^{155}		$s \cdot 2^{2m}\epsilon^2$	2^{45}	$2^m\epsilon$	

Table 3: Numerical example for the degradation is circuit size for the bound (c),(d),(e) and (f) from Table 1. We assume a distribution which is (ϵ, s) pseudorandom where for any s the ϵ is such that $s/\epsilon = 2^{256}$ (we can for example conjecture that the security of AES256 as a weak PRF satisfies this). Then we chose s such that we get $\epsilon' = 2^{-55}$ after $m = 46$ bits of leakage. Only the (g) and (c) bound give a non-zero bound for s' in this case, i.e. $s' \approx 2^{45}$ for both.

More generally, assume that we have a weak PRF that has k bits of security, i.e., it is (s, ϵ) secure for any $s/\epsilon = 2^k$ (see Definition 3). Then, after leaking m bits it is (s', ϵ') secure for any $s'/\epsilon' = 2^t$, where t satisfies the conditions from Table 4 Let us stress that the equivalence of our bounds and the ones from [FOR12] only holds in the setting where s/ϵ is basically constant. This is a

[4] We consider the security of AES256 as a weak PRF, and not a standard PRF, because of non-uniform attacks which show that no PRF with a k bit key can have $s/\epsilon \approx 2^k$ security [DTT09], at least unless we additionally require $\epsilon \gg 2^{-k/2}$.

reasonable assumption for secret-key primitives, but certainly not for most other settings like public-key crypto.[5]

Chain Rule	Technique	Security after leakage	Analysis
(e) [VZ13]	Simulating auxiliary inputs	$t = \frac{k}{5} - \frac{m}{5}$	Appendix A.1
(d) [JP14]	Simulating auxiliary inputs	$t = \frac{k}{3} - \frac{4m}{3}$	Appendix A.2
(f) [GW10]	Relaxed HILL Entropy	$t = \frac{k}{3} - \frac{2m}{3}$	Appendix A.3
(c) [FOR12]	Dense Model Theorem	$t = \frac{k}{3} - \frac{m}{3}$	Appendix A.4
(g) **This work**	Average Metric Entropy	$t = \frac{k}{3} - \frac{m}{3}$	Appendix A.5

Table 4: Consider an (s, ϵ) secure weak PRF where $s/\epsilon = 2^k$ (for any choice of of s), then after m bits of leakage on the key, the PRF is $s'/\epsilon' = 2^t$ secure, where depending on the chain rule used, t can take the values as indicated in the table.

No fixed time-sucess ratio (application for PRGs with weak seeds). To be more concrete, consider the problem of generating pseudorandomness from *weak seeds*. Let $\mathsf{PRG} : \{0,1\}^n \to \{0,1\}^{2n}$ be a length-doubling PRG with known parameters $(\epsilon_{\mathsf{PRG}}, s_{\mathsf{PRG}})$. Suppose that we have a "weak" source X with min-entropy only $n - d$. The output of PRG on X is not guaranteed to be pseudorandom, and in fact it is not secure [DY13]. One way to overcome this problem is so called "expand-extract" approach [DY13]. Namely, we simply take an extractor $\mathsf{Ext} : \{0,1\}^{2n} \to \{0,1\}^n$ and output $\mathsf{Ext}(\mathsf{PRG}(X))$. The proof that the output is pseudorandom goes by chain rules. In the second approach the roles of the extractor atincrseed and key X are swaped and then some facts about so called square-friendly applications are used [DY13]. Since the approach based on the square-friendly properties or on chain rules derived by extending the dense model theorem yield a loss in ϵ, for settings where d is relatively big (more specifically when $d > \log \epsilon_{\mathsf{PRG}}^{-1}$) one prefers the use of a chain rule with loss in s. Below in Table 5 we provide corresponding bounds and a numerical example, where only our chain rule guarantees meaningful security.

2 Preliminaries

Security definitions. Given a cryptographic primitive, we consider the probability ϵ of breaking its security (defined as winning the corresponding security game) by an adversary with running time (circuit size) s. The following definition is the standard way to define the security level.

[5] Consider e.g. RSA, here given our current understanding of the hardness of factoring, ϵ goes from basically 0 to 1 as the running time s reaches the time required to run the best factoring algorithms. In any case, it's not reasonable to assume that s/ϵ is almost constant over the entire range of s.

Technique	Real security for deficiency d		Comments	Numerical example ($n = 256$)				
	ϵ'	s'		ϵ_{PRG}	s_{PRG}	d	ϵ'	s'
square-security [DY13]	$\sqrt{2^d \epsilon_{\mathsf{PRG}}}$	$\Omega\left(s_{\mathsf{PRG}}\right)$		2^{-40}	2^{176}	50	1	2^{88}
chain rule (c)	$2^d \epsilon_{\mathsf{PRG}} + 2^{-\frac{n-d}{2}} + \delta$	$\Omega\left(s_{\mathsf{PRG}} \cdot \delta^2\right)$	δ arbitrary				1	2^{176}
chain rule (f) and (e)	$\epsilon_{\mathsf{PRG}} + 2^{-\frac{n-d}{2}} + \delta$	$\Omega\left(s_{\mathsf{PRG}} \cdot \frac{\delta^2}{2^d}\right)$	δ arbitrary $s_{\mathsf{PRG}} > 2^{2d}\delta^{-2}$				2^{-39}	< 0
chain rule (g)	$\epsilon_{\mathsf{PRG}} + 2^{-\frac{n-d}{2}} + \delta$	$\Omega\left(s_{\mathsf{PRG}} \cdot \frac{\delta^2}{2^d}\right)$	δ arbitrary $s_{\mathsf{PRG}} > 2^{2d}$				2^{-39}	2^{46}

Table 5: Security of a PRG fed with weak seeds, by "expand-extract-reseed" technique. We start with a 256-bit PRG output with security parameters $(\epsilon_{\mathsf{PRG}}, s_{\mathsf{PRG}}) = (2^{-40}, 2^{176})$, chosen to exclude best known non-uniform attacks [DTT09] which are of complexity $s > 2^n \epsilon^2$. We aim for $\epsilon' \approx 2^{-39}$.

Definition 3 (Security of cryptographic primitives, [Lub96]). *We say that a cryptographic primitive has λ bits of security (alternatively: it is 2^λ-secure) if every adversary has time-advantage ratio at least 2^λ.*

We note that for indistinguishability applications, that is when winning the security game is equivalent to distinguishing a given object from the ideal object (like PRFs, PRGs), the advantage is defined as the difference of the winning probability and $\frac{1}{2}$ which corresponds to chances that a random guess succeeds, whereas for unpredictability applications (like one-way functions) the advantage is simply equal the winning probability. In this paper we will consider indistinguishability applications only.

Some technical entropy definitions. We consider several classes of distinguishers. With $\mathcal{D}_s^{\mathsf{rand},\{0,1\}}$ we denote the class of randomized circuits of size at most s with boolean output (this is the standard non-uniform class of distinguishers considered in cryptographic definitions). The class $\mathcal{D}_s^{\mathsf{rand},[0,1]}$ is defined analogously, but with real valued output in $[0,1]$. $\mathcal{D}_s^{\mathrm{det},\{0,1\}}, \mathcal{D}_s^{\mathrm{det},[0,1]}$ are defined the corresponding classes for *deterministic* circuits. With $\delta^D(X,Y) = |\,\mathbb{E}_X[D(X)] - \mathbb{E}_Y[D(Y)]$ we denote D's advantage in distinguishing X and Y.

Definition 4 (Metric pseudoentropy [BSW03,FR12]). *A random variable X has* real deterministic Metric entropy *at least k if*

$$\mathbf{H}_{\epsilon,s}^{\mathrm{Metric},\mathcal{D}_s^{\mathrm{det},[0,1]}}(X) \geqslant k \iff \forall D \in \mathcal{D}_s^{\mathrm{det},[0,1]} \exists Y_D \,,\, \mathbf{H}_\infty(Y_D) = k \,:\, \delta^D(X, Y_D) \leqslant \epsilon$$

Relaxed versions of HILL and Metric entropy. A weaker notion of conditional HILL entropy allows the conditional part to be replaced by some computationally indistinguishable variable

Definition 5 (Relaxed HILL pseudoentropy [GW11,Rey11]). *For a joint distribution (X, Z), we say that X has* relaxed HILL entropy *k conditioned on*

Z if

$$\begin{aligned} & \mathbf{H}^{\mathrm{HILL-rlx}}_{\epsilon,s}(X|Z) \geqslant k \\ \iff & \exists (Y,Z'), \widetilde{\mathbf{H}}_{\infty}(Y|Z') = k, \forall D \in \mathcal{D}^{\mathsf{rand},\{0,1\}}_{s} \ , \ : \ \delta^{D}((X,Z),(Y,Z')) \leqslant \epsilon \end{aligned}$$

The above notion of *relaxed* HILL satisfies a chain rule whereas the chain rule for the standard definition of conditional HILL entropy is known to be false [**?**]. One can analogously define relaxed variants of metric entropy, we won't give these as they will not be required in this paper. The relaxed variant of HILL entropy is also useful because one can convert relaxed entropy into standard HILL entropy, losing in s an additive term exponential in the length of the conditional part.

Lemma 1 (HILL-rlx⟶HILL, [JP14]). *For any X and correlated Z of length m, we have $\mathbf{H}^{\mathrm{HILL}}_{\epsilon,s'}(X|Z) \geqslant \mathbf{H}^{\mathrm{HILL-rlx}}_{\epsilon,s}(X|Z)$ where $s' = s - 2^m$.*

Pseudoentropy against different distinguisher classes. For randomized distinguishers, it's irrelevant if the output is boolean or real values, as we can replace any $D \in \mathcal{D}^{\mathsf{rand},[0,1]}_{s}$ with a $D' \in \mathcal{D}^{\mathsf{rand},\{0,1\}}$ s.t. $\mathbb{E}[D'(X)] = \mathbb{E}[D(X)]$ by setting (for any x) $\Pr[D'(x) = 1] = \mathbb{E}[D(x)]$. For HILL entropy (as well as for its relaxed version), it also doesn't matter if we consider randomized or deterministic distinguishers in Definition 2, as we always can "fix" the randomness to an optimal value. This is no longer true for metric entropy,[6] and thus the distincition betwenn metric and metric star entropy is crucial.

3 Main Result

We start with the following recently proven characterization of the distribution maximizing expectations under min-entropy constraints (Section 3.1). Based on this auxiliary result, in Section 3.2 and Section 3.3 we prove our chain rules stated in Theorem 2 and Corollary 1.

3.1 An auxiliary result on constrained optimization

Lemma 2 (Optimizing expectations under entropy constraints [SGP15, Sko15b]). *Given $\mathrm{D} : \{0,1\}^{n+m} \times \{0,1\}^{m} \to [0,1]$ consider the following optimization problem*

$$\begin{aligned} & \max_{Y|Z} \mathbb{E}\mathrm{D}(Y,Z) \\ & \text{s.t. } \widetilde{\mathbf{H}}_{\infty}(Y|Z) \geqslant k \end{aligned} \tag{6}$$

The distribution $Y|Z = Y^*|Z$ satisfying $\widetilde{\mathbf{H}}_{\infty}(Y^*|Z) = k$ is optimal for (6) if and only if there exist real numbers $t(z)$ and a number $\lambda \geqslant 0$ such that for every z

[6] It might be hard to find a high min-entropy distribution Y that fools a randomised distinguisher D, but this task can become easy once D's randomness is fixed.

(a) $\sum_x \max(\mathrm{D}(x,z) - t(z), 0) = \lambda$
(b) *If* $0 < \mathbf{P}_{Y^*|Z=z}(x) < \max_{x'} \mathbf{P}_{Y^*|Z=z}(x')$ *then* $\mathrm{D}(x,z) = t(z)$.
(c) *If* $\mathbf{P}_{Y^*|Z=z}(x) = 0$ *then* $\mathrm{D}(x,z) \leqslant t(z)$
(d) *If* $\mathbf{P}_{Y^*|Z=z}(x) = \max_{x'} \mathbf{P}_{Y^*|Z=z}(x')$ *then* $\mathrm{D}(x,z) \geqslant t(z)$

Remark 1. The characterization can be illustrated in an easy and elegant way. First, it says that the area under the graph of $\mathrm{D}(x,z)$ and above the threshold $t(z)$ is the same, no matter what z is (see Figure 2). Second, for every z the

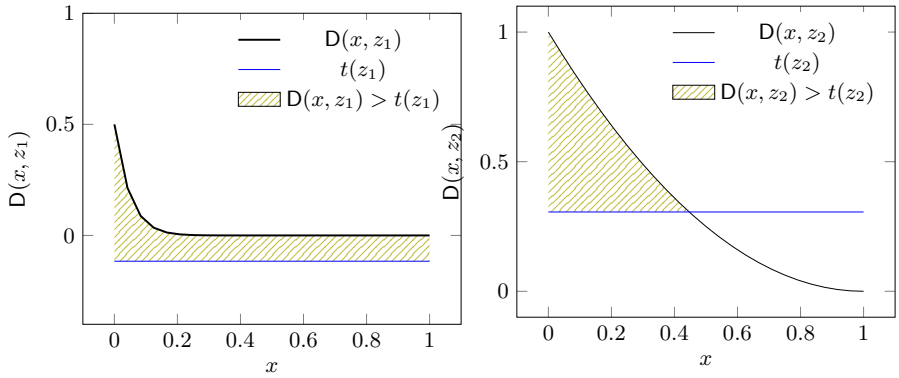


Fig. 2: For every z, the (green) area under $\mathrm{D}(\cdot,z)$ and above $t(z)$ equals λ

distribution $Y^*|Z=z$ is flat over the set $\{x:\ \mathrm{D}(x,z) > t(z)\}$ and vanishes for x satisfying $\mathrm{D}(x,z) < t(z)$, see Figure 3.

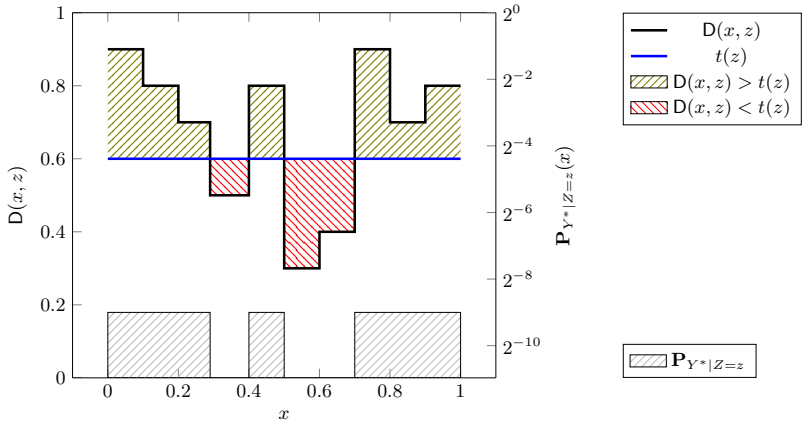


Fig. 3: Relation between distinguisher $\mathrm{D}(x,z)$, threshold $t(z)$ and distribution $Y^*|Z=z$.

Proof (Proof sketch of Lemma 2). Consider the following linear optimization program

$$\begin{aligned}
\underset{P_{x,z},a_z}{\text{maximize}} \quad & \sum_{x,z} \mathrm{D}(x,z)P(x,z) \\
\text{subject to} \quad & -P_{x,z} \leqslant 0,\ (x,z) \in \{0,1\}^n \times \{0,1\}^m \\
& \sum_x P_{x,z} - \mathbf{P}_Z(z) = 0,\ z \in \{0,1\}^m \\
& P_{x,z} - a_z \leqslant 0,\ z \in \{0,1\}^m \\
& \sum_z a_z - 2^{-k} \leqslant 0
\end{aligned} \tag{7}$$

This problem is equivalent to (6) if we define $\mathbf{P}_{Y,Z}(x,z) = P(x,z)$ and replace the condition $\sum_z \max_x \mathbf{P}_{Y,Z}(x,z) \leqslant 2^{-k}$, which is equivalent to $\widetilde{H}_\infty(Y|Z) \geqslant k$, by the existence of numbers $a_z \geqslant \max_x \mathbf{P}_{Y,Z}(x,z)$ such that $\sum_z a_z \leqslant 2^{-k}$. The solutions of (7) can be characterized as follows:

Claim 1. *The numbers $(P_{x,z})_{x,z}, (a_z)_z$ are optimal for (7) if and only if there exist numbers $\lambda^1(x,z) \geqslant 0$, $\lambda^2(z) \in \mathbb{R}$, $\lambda^3(x,z) \geqslant 0$, $\lambda^4 \geqslant 0$ such that*

(a) $\mathrm{D}(x,z) = -\lambda^1(x,z) + \lambda^2(z) + \lambda^3(x,z)$ *and* $0 = -\sum_x \lambda^3(x,z) + \lambda^4$
(b) We have $\lambda^1(x,z) = 0$ *if* $P_{x,z} > 0$, $\lambda^3(x,z) = 0$ *if* $P_{x,z} < a_z$, $\lambda^4 = 0$ *if* $\sum_z a_z < 2^{-k}$.

Proof (of Claim). This is a straightforward application of KKT conditions. □
□

It remains to apply and simplify the last characterization. Let $(P^*_{x,z})_{x,z}, (a^*_z)_z$ be optimal for (7), where $P^*(x,z) = \mathbf{P}_{Y^*,Z}(x,z)$, and $\lambda^1(x,z), \lambda^2(z), \lambda^3(x,z), \lambda^4(x)$ be corresponding multipliers given by the last claim. Define $t(z) = \lambda^2(z)$ and $\lambda = \lambda^4$. Observe that for every z we have $a^*_z \geqslant \max_x \mathbf{P}(x,z) \geqslant 2^{-n}\mathbf{P}_Z(z) > 0$ and thus for every (x,z) we have

$$\lambda^1(x,z) \cdot \lambda^3(x,z) = 0 \tag{8}$$

If $P^*(x,z) = 0$ then $P^*(x,z) < a^*(z)$ and $\lambda^3(x,z) = 0$, hence $\mathrm{D}(x,z) \leqslant t(z)$ which proves (c). If $P^*(x,z) = \max_{x'} P^*(x,z)$ then $P^*(x,z) < 0$ and $\lambda^1(x,z) = 0$ which proves (d). Finally observe that (8) implies

$$\max(\mathrm{D}(x,z) - t(z), 0) = \max(-\lambda^1(x,z) + \lambda^3(x,z), 0) = \lambda^3(x,z)$$

Hence, the assumption $\sum_x \lambda^3(x,z) = \lambda^4 = \lambda$ proves (a). Suppose now that the characterization given in the Lemma is satisfied. Define $P^*(x,z) = \mathbf{P}_{Y,Z}(x,z)$ and $a_z = \max_z \mathbf{P}_{Y^*,Z}(x,z)$, let $\lambda^3(x,z) = \max(\mathrm{D}(x,z) - t(z), 0)$, $\lambda^1(x,z) = \max(t(z) - \mathrm{D}(x,z), 0)$ and $\lambda^4 = \lambda$. We will show that these numbers satisfy the conditions described in the last claim. By definition we have $-\lambda^1(x,z) + \lambda^2(z) + \lambda^3(x,z) = \mathrm{D}(x,z)$, by the assumptions we get $\sum_x \lambda^3(x,z) = \lambda = \lambda^4$.

This proves part (a). Now we verify the conditions in (b). Note that $\mathrm{D}(x,z) < t(z)$ is possible only if $\mathbf{P}_{Y^*|Z=z}(x) = 0$ and $\mathrm{D}(x,z) > t(z)$ is possible only if $\mathbf{P}_{Y^*|Z=z}(x) = \max_{x'} \mathbf{P}_{Y^*|Z=z}(x')$. Therefore, if $\mathbf{P}_{Y,Z}(x,z) > 0$ then we must have $\mathrm{D}(x,z) \geqslant t(z)$ which means that $\lambda^1(x,z) = 0$. Similarly if $\mathbf{P}_{Y,Z}(x,z) < \max_z \mathbf{P}_{Y^*,Z}(x,z)$ then $\mathrm{D}(x,z) \leqslant t(z)$ and $\lambda^3(x,z) = 0$. Finally, since we assume $\widetilde{H}_\infty(Y^*|Z) = k$ we have $\sum_z a_z = 2^{-k}$ and thus there is no additional restrictions on λ^4. □

3.2 New chain rule for Metric entropy

We start by sketching the idea of the proof. Assuming contrarily, we have a function D of complexity D' which distinguishes between (X,Z) and all distributions (Y,Z) such that $\widetilde{\mathbf{H}}_\infty(Y|Z) \geqslant k-m$. By Lemma 2 we can replace D by a distinguisher D' which is regular conditioned on the second argument, that is $\mathbb{E}\,\mathrm{D}(U,z) = \text{const}$ independently on z. This is the key trick in our proof.

Proof (of Theorem 2). Suppose not. There exists real-valued D of size s' such that

$$\mathbb{E}\,\mathrm{D}(X,Z) - \mathbb{E}\,\mathrm{D}(Y,Z) \geqslant \epsilon, \quad \forall Y : \ \mathbf{H}_\infty(Y|Z) \geqslant k-m. \tag{9}$$

The distribution Y^* which minimizes the left-hand side is optimal to the program in (6) (where k is replaced by $k-m$). We start by showing that we can actually assume that D has a very strong property, namely is regular.

Claim (Regular distinguisher). There exists D' of complexity $\mathrm{size}(\mathrm{D})+2^m$ which satisfies Equation (9) in place of D, and is regular, that is $\sum_x \mathrm{D}(x,z) = \lambda$ for some λ and every z.

Proof (Proof of Claim). Let $t(z)$ and λ be as in Lemma 2. Define $\mathrm{D}'(x,z) = \max(\mathrm{D}(x,z) - t(z), 0)$. It is easy to see that Y^* is optimal also when D is replaced by D'. Moreover, we have $\mathbb{E}\,\mathrm{D}'(X,Z) \geqslant \mathbb{E}\,\mathrm{D}(X,Z) - \lambda$ and $\mathbb{E}\,\mathrm{D}'(Y^*,Z) = \mathbb{E}\,\mathrm{D}'(Y^*,Z) - \lambda$ and thus $\mathbb{E}\,\mathrm{D}'(X,Z) - \mathbb{E}\,\mathrm{D}'(Y^*,Z) \geqslant \epsilon$. Therefore,

$$\mathbb{E}\,\mathrm{D}'(X,Z) - \mathbb{E}\,\mathrm{D}'(Y,Z) \geqslant \epsilon, \quad \forall Y : \ \mathbf{H}_\infty(Y|Z) \geqslant k-m \tag{10}$$

note that we have

$$\sum_x \mathrm{D}'(x,z) = \lambda, \quad \forall z \tag{11}$$

which finishes the proof. □

Having transformed our distinguisher into a more convenient form we define

$$\mathrm{D}''(x,z) = \max_z \mathrm{D}'(x,z). \tag{12}$$

Claim. We have $\mathbb{E}\,\mathrm{D}''(X) \geqslant \mathbb{E}\,\mathrm{D}'(X,Z)$.

Proof. This follows by the definition of D''. □

Claim. For every Y such that $\mathbf{H}_\infty(Y) \geqslant k$ we have $\mathbb{E}\,\mathrm{D}''(Y) \leqslant \mathbb{E}\,\mathrm{D}'(Y^*, Z)$

Proof. We get

$$\begin{aligned}\mathbb{E}\,\mathrm{D}''(Y) &\leqslant 2^{-k}\sum_x \max_z \mathrm{D}'(x,z)\\ &\leqslant 2^{-k}\sum_{x,z}\mathrm{D}'(x,z)\\ &= 2^{-k+m}\cdot\lambda = \mathbb{E}\,\mathrm{D}'(Y^*,Z)\end{aligned} \tag{13}$$

where in the last line we have used the fact that D' is regular (see Equation (11)) and that $HminAv(Y^*|Z) = k - m$ □

Combining the last two claims we get $\mathbb{E}\,\mathrm{D}''(X) - \mathbb{E}\,\mathrm{D}''(Y) \geqslant \epsilon$ for all Y of min-entropy k. It remains to observe that the complexity of D'' equals $s = (s' + 2^m)\cdot 2^m$. □

3.3 The chain rule for HILL entropy

Corollary 1 follows from Theorem 2 by the following result being a tight version of the transformation originally due to [BSW03]

Theorem 3 (Metric⟶HILL entropy, [Sko15a]). *For any n-bit random variable X and a correlated random variable Z we have*

$$\mathbf{H}^{\mathrm{HILL}}_{(s',\epsilon')}(X|Z) \geqslant \mathbf{H}^{\mathrm{Metric},\mathcal{D}^{\mathrm{det},[0,1]}}_{(s,\epsilon)}(X|Z)$$

where $\delta \in (0,1)$ is an arbitrary parameter, $s' = \Omega\left(s\cdot\delta^2/(\Delta+1)\right)$, $\epsilon' = \epsilon + \delta$ and $\Delta = n - k$ is the entropy deficiency.

References

BM84. Manuel Blum and Silvio Micali, *How to generate cryptographically strong sequences of pseudorandom bits*, no. 4, 850–864.

BSW03. Boaz Barak, Ronen Shaltiel, and Avi Wigderson, *Computational analogues of entropy.*, RANDOM-APPROX, Lecture Notes in Computer Science, vol. 2764, Springer, 2003, pp. 200–215.

CKLR11. Kai-Min Chung, Yael Tauman Kalai, Feng-Hao Liu, and Ran Raz, *Memory delegation*, Cryptology ePrint Archive, Report 2011/273, 2011, http://eprint.iacr.org/.

DP08a. Stefan Dziembowski and Krzysztof Pietrzak, *Leakage-resilient cryptography*, 2008, pp. 293–302.

DP08b. ———, *Leakage-resilient cryptography in the standard model*, IACR Cryptology ePrint Archive **2008** (2008), 240.

DRS04. Yevgeniy Dodis, Leonid Reyzin, and Adam Smith, *Fuzzy extractors: How to generate strong keys from biometrics and other noisy data*, 2004, pp. 523–540.

DTT09. Anindya De, Luca Trevisan, and Madhur Tulsiani, *Non-uniform attacks against one-way functions and prgs*, Electronic Colloquium on Computational Complexity (ECCC) **16** (2009), 113.

DY13. Yevgeniy Dodis and Yu Yu, *Overcoming weak expectations*, Theory of Cryptography (Amit Sahai, ed.), Lecture Notes in Computer Science, vol. 7785, Springer Berlin Heidelberg, 2013, pp. 1–22 (English).

FOR12. Benjamin Fuller, Adam O'Neill, and Leonid Reyzin, *A unified approach to deterministic encryption: New constructions and a connection to computational entropy*, Cryptology ePrint Archive, Report 2012/005, 2012, http://eprint.iacr.org/.

FR12. Benjamin Fuller and Leonid Reyzin, *Computational entropy and information leakage*, Cryptology ePrint Archive, Report 2012/466, 2012, http://eprint.iacr.org/.

GW10. Craig Gentry and Daniel Wichs, *Separating succinct non-interactive arguments from all falsifiable assumptions*, Cryptology ePrint Archive, Report 2010/610, 2010, http://eprint.iacr.org/.

GW11. ———, *Separating succinct non-interactive arguments from all falsifiable assumptions*, 2011, pp. 99–108.

HILL99. Johan Hastad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby, *A pseudorandom generator from any one-way function*, SIAM J. Comput. **28** (1999), no. 4, 1364–1396.

HLR07. Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin, *Conditional computational entropy, or toward separating pseudoentropy from compressibility*, 2007, pp. 169–186.

HRV10. Iftach Haitner, Omer Reingold, and Salil Vadhan, *Efficiency improvements in constructing pseudorandom generators from one-way functions*, Proceedings of the 42nd ACM symposium on Theory of computing (New York, NY, USA), STOC '10, ACM, 2010, pp. 437–446.

JP14. Dimitar Jetchev and Krzysztof Pietrzak, *How to fake auxiliary input*, Theory of Cryptography - 11th Theory of Cryptography Conference, TCC 2014, San Diego, CA, USA, February 24-26, 2014. Proceedings (Yehuda Lindell, ed.), Lecture Notes in Computer Science, vol. 8349, Springer, 2014, pp. 566–590.

KPWW14. Stephan Krenn, Krzysztof Pietrzak, Akshay Wadia, and Daniel Wichs, *A counterexample to the chain rule for conditional HILL entropy*, IACR Cryptology ePrint Archive **2014** (2014), 678.

Lub96. Michael Luby, *Pseudorandomness and cryptographic applications*, Princeton computer science notes, Princeton University Press, 1996.

Pie09. Krzysztof Pietrzak, *A leakage-resilient mode of operation*, 2009, pp. 462–482.

Rey11. Leonid Reyzin, *Some notions of entropy for cryptography - (invited talk)*, 2011, pp. 138–142.

RTTV08a. Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil Vadhan, *Dense subsets of pseudorandom sets*, Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science (Washington, DC, USA), FOCS '08, IEEE Computer Society, 2008, pp. 76–85.

RTTV08b. Omer Reingold, Luca Trevisan, Madhur Tulsiani, and Salil P. Vadhan, *Dense subsets of pseudorandom sets*, 2008, pp. 76–85.

SGP15. Maciej Skorski, Alexander Golovnev, and Krzysztof Pietrzak, *Condensed unpredictability*, To appear in ICALP 2015, vol. 2015, 2015, p. 384.
Sko15a. Maciej Skorski, *Metric pseudoentropy: Characterizations, transformations and applications*, 2015.
Sko15b. ———, *Metric pseudoentropy: Characterizations, transformations and applications*, Information Theoretic Security - 8th International Conference, ICITS 2015, Lugano, Switzerland, May 2-5, 2015. Proceedings (Anja Lehmann and Stefan Wolf, eds.), Lecture Notes in Computer Science, vol. 9063, Springer, 2015, pp. 105–122.
VZ13. Salil Vadhan and ColinJia Zheng, *A uniform min-max theorem with applications in cryptography*, Advances in Cryptology CRYPTO 2013 (Ran Canetti and JuanA. Garay, eds.), Lecture Notes in Computer Science, vol. 8042, Springer Berlin Heidelberg, 2013, pp. 93–110 (English).
Yao82. Andrew Chi-Chih Yao, *Theory and applications of trapdoor functions (extended abstract)*, 1982, pp. 80–91.

A Time-Success Ratio Analysis

A.1 Chain rule given by Vadhan and Zheng

Theorem 4 (Time-success ratio for chain rule (e)). *Suppose that X has n bits of HILL entropy of quality (s,ϵ) for every $s/\epsilon \geqslant 2^k$. Then X conditioned on leakage of m bits has $n-m$ bits of HILL entropy of quality (s',ϵ') for every $s'/\epsilon' \geqslant 2^t$ where*

$$t = \frac{k}{5} - \frac{m}{5} \tag{14}$$

and this is the best possible bound guaranteed by chain rule (e).

Proof (Proof of Theorem 4). Suppose that we have $s' = s \cdot 2^{-m}\delta^2 - \delta^{-2} - 2^m$ and $\epsilon' = \epsilon + \delta$. We want to find the minimum value of the ratio $\frac{s'}{\epsilon'}$ under the assumption that ϵ, δ, s can be chosen in the possibly most plausible way. Therefore, we want to solve the following min-max problem

$$\begin{aligned} &\min_{\epsilon',s'} \quad \max_{s,\epsilon,\delta} \quad \frac{s'}{\epsilon'} \\ &\text{s.t.} \qquad\qquad \frac{s}{\epsilon} = 2^k,\ \epsilon + \delta = \epsilon',\ s' = s \cdot 2^{-m}\delta^2 - \delta^{-2} - 2^m \end{aligned} \tag{15}$$

First, we note that

$$s' = 2^{k-m}(\epsilon' - \delta)\delta^2 - \delta^{-2} - 2^m$$

Also, since $\delta < \epsilon'$, we need to assume $\epsilon' > 2^{-\frac{k-m}{5}}$ and $\epsilon' > 2^{-\frac{k-2m}{3}}$ to guarantee that $s' > 0$. Now, for $\delta = \Theta(\epsilon')$ we get

$$\frac{s'}{\epsilon'} = \Omega\left(2^{k-m}\epsilon'^2 - \epsilon'^{-3} - 2^m\epsilon'^{-1}\right) = \Omega\left(2^{\max\left(\frac{3}{5}\cdot(k-m),\frac{k+m}{3}\right)}\right) \tag{16}$$

provided that $\epsilon' \gg 2^{-\frac{k-m}{5}}$ and $\epsilon' \gg 2^{-\frac{k-2m}{3}}$. □

A.2 Chain rule given by Jetchev and Pietrzak

Theorem 5 (Time-success ratio for chain rule (d)). *Suppose that X has n bits of HILL entropy of quality (s, ϵ) for every $s/\epsilon \geqslant 2^k$. Then X conditioned on leakage of m bits has $n - m$ bits of HILL entropy of quality (s', ϵ') for every $s'/\epsilon' \geqslant 2^t$ where*

$$t = \frac{k}{3} - \frac{4m}{3} \tag{17}$$

and this is the best possible bound guaranteed by chain rule (d).

Proof (Proof of Theorem 5). Suppose that we have $s' = s \cdot 2^{-3m}\delta^2 - 2^m$ and $\epsilon' = \epsilon + \delta$. We want to find the minimum value of the ratio $\frac{s'}{\epsilon'}$ under the assumption that ϵ, δ, s can be chosen in the possibly most plausible way. Therefore, we want to solve the following min-max problem

$$\begin{aligned} &\min_{\epsilon', s'} \quad \max_{s, \epsilon, \delta} \quad \frac{s'}{\epsilon'} \\ &\text{s.t.} \qquad\qquad \frac{s}{\epsilon} = 2^k,\ \epsilon + \delta = \epsilon',\ s' = s \cdot 2^{-3m}\delta^2 - 2^m \end{aligned} \tag{18}$$

First, we note that

$$s' = 2^{k-3m}(\epsilon' - \delta)\delta^2 - 2^m$$

Also, since $\delta < \epsilon'$, we need to assume $\epsilon' > 2^{-\frac{k-4m}{3}}$ to guarantee that $s' > 0$. Now, setting $\delta = \Theta(\epsilon')$ we have

$$\frac{s'}{\epsilon'} = \Omega\left(2^{k-m}\epsilon'^2\right) - 2^m\epsilon'^{-1} = \Omega\left(2^{\frac{k-2m}{3}}\right) \tag{19}$$

provided that $\epsilon' \gg 2^{-\frac{k-4m}{3}}$. □

A.3 Chain rule given by Gentry and Wichs

Theorem 6 (Time-success ratio for chain rule (f)). *Suppose that X has n bits of HILL entropy of quality (s, ϵ) for every $s/\epsilon \geqslant 2^k$. Then X conditioned on leakage of m bits has $n - m$ bits of HILL entropy of quality (s', ϵ') for every $s'/\epsilon' \geqslant 2^t$ where*

$$t = \frac{k}{3} - \frac{2m}{3} \tag{20}$$

and this is the best possible bound guaranteed by chain rule (f).

Proof (Proof of Theorem 6). Suppose that we have $s' = s \cdot 2^{-m}\delta^2 - 2^m$ and $\epsilon' = \epsilon + \delta$. We want to find the minimum value of the ratio $\frac{s'}{\epsilon'}$ under the assumption

that ϵ, δ, s can be chosen in the possibly most plausible way. Therefore, we want to solve the following min-max problem

$$\begin{aligned} &\min_{\epsilon', s'} \quad \max_{s,\epsilon,\delta} \quad \frac{s'}{\epsilon'} \\ &\text{s.t.} \qquad\qquad \frac{s}{\epsilon} = 2^k,\ \epsilon + \delta = \epsilon',\ s' = s \cdot 2^{-m}\delta^2 - 2^m \end{aligned} \tag{21}$$

First, we note that

$$s' = 2^{k-m}(\epsilon' - \delta)\delta^2 - 2^m$$

Also, since $\delta < \epsilon'$, we need to assume $\epsilon' > 2^{-\frac{k-2m}{3}}$ to guarantee that $s' > 0$. Now, setting $\delta = \Theta(\epsilon')$ we have

$$\frac{s'}{\epsilon'} = \Omega\left(2^{k-m}\epsilon'^2\right) - 2^m\epsilon'^{-1} = \Omega\left(2^{\frac{k+m}{3}}\right) \tag{22}$$

provided that $\epsilon' \gg 2^{-\frac{k-2m}{3}}$. □

A.4 Chain rule given by Fuller and Reyzin

Theorem 7 (Time-success ratio for chain rule (c)). *Suppose that X has n bits of HILL entropy of quality (s, ϵ) for every $s/\epsilon \geqslant 2^k$. Then X conditioned on leakage of m bits has $n - m$ bits of HILL entropy of quality (s', ϵ') for every $s'/\epsilon' \geqslant 2^t$ where*

$$t = \frac{k}{3} - \frac{m}{3} \tag{23}$$

and this is the best possible bound guaranteed by chain rule (c).

Proof (Proof of Theorem 7). Suppose that we have $s' = s \cdot \delta^2$ and $\epsilon' = 2^m\epsilon + \delta$. We want to find the minimum value of the ratio $\frac{s'}{\epsilon'}$ under the assumption that ϵ, δ, s can be chosen in the possibly most plausible way. Therefore, we want to solve the following min-max problem

$$\begin{aligned} &\min_{\epsilon', s'} \quad \max_{s,\epsilon,\delta} \quad \frac{s'}{\epsilon'} \\ &\text{s.t.} \qquad\qquad \frac{s}{\epsilon} = 2^k,\ 2^m\epsilon + \delta = \epsilon',\ s' = s \cdot \delta^2 \end{aligned} \tag{24}$$

First, we note that

$$s' = 2^{k-m}(\epsilon' - \delta)\delta^2$$

Also, since $\delta < \epsilon'$, we need to assume $\epsilon' > 2^{-\frac{k-m}{3}}$ to guarantee that $s' > 1$. Now, setting $\delta = \Theta(\epsilon')$ we have

$$\frac{s'}{\epsilon'} = \Omega\left(2^{k-m}\epsilon'^2\right) = \Omega\left(2^{\frac{k-m}{3}}\right), \tag{25}$$

provided that $\epsilon' > 2^{-\frac{k-m}{3}}$. □

A.5 Chain rule in this paper

Theorem 8 (Time-success ratio for chain rule (g)). *Suppose that X has n bits of HILL entropy of quality (s, ϵ) for every $s/\epsilon \geqslant 2^k$. Then X conditioned on leakage of m bits has $n - m$ bits of HILL entropy of quality (s', ϵ') for every $s'/\epsilon' \geqslant 2^t$ where*

$$t = \frac{k}{3} - \frac{m}{3} \tag{26}$$

and this is the best possible bound guaranteed by chain rule (g).

Proof (Proof of Theorem 8). Suppose that we have $s' = s \cdot 2^{-m}\delta^2 - 2^m\delta^2$ and $\epsilon' = \epsilon + \delta$. We want to find the minimum value of the ratio $\frac{s'}{\epsilon'}$ under the assumption that ϵ, δ, s can be chosen in the possibly most plausible way. Therefore, we want to solve the following min-max problem

$$\begin{array}{lll} \min\limits_{\epsilon', s'} & \max\limits_{s, \epsilon, \delta} & \dfrac{s'}{\epsilon'} \\ \text{s.t.} & & \dfrac{s}{\epsilon} = 2^k,\ \epsilon + \delta = \epsilon',\ s' = s \cdot 2^{-m}\delta^2 - 2^m\delta^2 \end{array} \tag{27}$$

First, we note that

$$s' = 2^{k-m}(\epsilon' - \delta)\delta^2 - 2^m\delta^2$$

Also, since $\delta < \epsilon'$, we need to assume $\epsilon' > 2^{-(k-2m)}$ and $\epsilon' > 2^{-\frac{k-m}{3}}$ to guarantee that $s' > 0$. Setting $\delta = \Theta(\epsilon')$ we obtain

$$\frac{s'}{\epsilon'} = \Omega\left(2^{k-m}\epsilon'^2\right) - 2^m\epsilon' = \Omega\left(2^{k-m}\epsilon'^2\right) \tag{28}$$

provided that $\epsilon' \gg 2^{-(k-2m)}$ and $\epsilon' > 2^{-\frac{k-m}{3}}$. If t is the security level, we must have $t < \min\left(k - 2m, \frac{k-m}{3}\right)$ and $k - m - 2t > t$. □