Information Processing Letters 134 (2018) 62–66

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Quasi chain rule for min-entropy [☆]

Stefan Dziembowski [a], Tomasz Kazana [a,*], Maciej Zdanowicz [b,*]

[a] *Institute of Informatics, Department of Mathematics, Informatics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland*
[b] *Institute of Mathematics, Department of Mathematics, Informatics and Mechanics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland*

ARTICLE INFO

Article history:
Received 27 October 2016
Received in revised form 12 September 2017
Accepted 4 February 2018
Available online 5 February 2018
Communicated by L. Viganò

Keywords:
Min-entropy
Chain-rule
Cryptography
Information theory

ABSTRACT

The classical Shannon's entropy possesses a natural definition of conditional entropy and a useful chain rule whose application is ubiquitous in information theory. On the contrary, for the case of min-entropy both: the definition of conditional min-entropy and the formulation of chain rule are still subject of discussion. This paper goes along this line of research and proposes new candidate for chain rule for conditional min-entropy as defined in Dodis et al. paper [1]. We derive our *quasi chain rule* based on so-called *spoiling knowledge* idea.

1. Introduction

Since the beginning of the formal treatment of information theory most works have heavily relied on different flavours of the notion of *entropy*. Depending on the context, those were used to measure compressibility, unpredictability or uncertainty of outcomes of random processes. In his seminal work, Shannon applied the simplest compressibility notion of entropy, defined for a random variable X by the formula

$$\begin{aligned}\mathbf{H}(X) &\overset{\text{def}}{=} \mathbb{E}_x \log_2 \frac{1}{\Pr(X=x)}\\ &= \sum_x -\Pr(X=x)\cdot\log_2\Pr(X=x),\end{aligned}$$

to prove that in a perfectly secure symmetric key encryption scheme the length of the secret key is necessarily as large as the length of the message. The version of entropy which turned out to be most useful in the area of cryptography is the *min-entropy*, defined by the formula

$$\mathbf{H}_\infty(X) \overset{\text{def}}{=} -\log_2(\max_x \Pr(X=x)),$$

which quantitatively reflects the difficulty of guessing a random sample of a random variable. More precisely, the number $2^{-\mathbf{H}_\infty(X)}$ is the maximal probability of guessing the value of a random sample of X.

1.1. Conditional entropy

The present paper deals with the conditional counterparts of the aforementioned entropy notions. Shannon's compressibility entropy possesses a natural generalization to its conditional version $\mathbf{H}(X|Y)$, which satisfies the formula $\mathbf{H}((X,Y)) = \mathbf{H}(X|Y) + \mathbf{H}(Y)$ expressing an intuitive interpretation stating that the information contained in the pair (X,Y) consists of the information in Y extended by

[☆] This work was supported by Polish National Science Centre (NCN) SONATA GRANT UMO-2014/13/D/ST6/03252 and the WELCOME/2010-4/2 grant founded within the framework of the EU Innovative Economy (National Cohesion Strategy) Operational Programme.

* Corresponding authors.
E-mail addresses: std@mimuw.edu.pl (S. Dziembowski), tkazana@mimuw.edu.pl (T. Kazana), mez@mimuw.edu.pl (M. Zdanowicz).

https://doi.org/10.1016/j.ipl.2018.02.007
0020-0190/© 2018 Elsevier B.V. All rights reserved.

the conditional information in X given Y. Dodis et al. [1] provided an analogous notion for min-entropy. Namely, for two random variables X, Y the conditional min-entropy $\widetilde{\mathbf{H}}_\infty(X|Y)$ is given by the formula

$$\widetilde{\mathbf{H}}_\infty(X|Y) \overset{\text{def}}{=} -\log_2\left(\mathbb{E}_y 2^{-\mathbf{H}_\infty(X|Y=y)}\right).$$

This definition turns out to preserve the natural interpretation of min-entropy as the maximal probability of success in guessing X given Y, that is, for any algorithm $\mathcal{A}$ we have

$$\begin{aligned}\Pr(\mathcal{A}(Y)=X) &= \mathbb{E}_y \Pr\left(\mathcal{A}(y)=(X|\{Y=y\})\right)\\ &\le \mathbb{E}_y 2^{-\mathbf{H}_\infty(X|Y=y)} = 2^{-\widetilde{\mathbf{H}}_\infty(X|Y)}.\end{aligned}$$

Regrettably, the above definition possesses serious drawbacks illustrated by the following example.

Example 1.1. Let $X = (X_1, X_2) \in (\{0,1\}^n)^2$ be a random variable distributed uniformly over "the cross", that is, a set $\{0,1\}^n \times e \cup e \times \{0,1\}^n$ for some fixed $e \in \{0,1\}^n$. Note that, we have $\mathbf{H}_\infty((X_1,X_2)) = -\log_2 \frac{1}{2^{n+1}-1} \in [n, n+1]$ and $\mathbf{H}_\infty(X_i) = -\log_2 \frac{2^n}{2^{n+1}-1} < 1$ and therefore, the sum property $\mathbf{H}_\infty((X_1,X_2)) \le \mathbf{H}_\infty(X_1) + \mathbf{H}_\infty(X_2)$ does not hold without any further assumptions or conditions. Moreover, $\widetilde{\mathbf{H}}_\infty(X_2|X_1) < \mathbf{H}_\infty(X_2)$ and therefore $\widetilde{\mathbf{H}}_\infty(X_2|X_1) + \mathbf{H}_\infty(X_1) < 2$ which consequently means that the most natural chain rule does not hold either.

However, among many other things the authors of [1] prove the following result.

Lemma 1.2 *(Lemma 2.2 in [1]). Let X, Y, Z be random variables. Then*

(a) *For any $\delta > 0$, the conditional entropy $\mathbf{H}_\infty(X|Y=y)$ is at least $(\widetilde{\mathbf{H}}_\infty(X|Y) - \log_2(1/\delta))$ with probability at least $(1-\delta)$ over the random choice of $y \leftarrow Y$.*

(b) *If Y has at most 2^λ possible values, then*

$$\widetilde{\mathbf{H}}_\infty(X|(Y,Z)) \ge \widetilde{\mathbf{H}}_\infty((X,Y)|Z) - \lambda \ge \widetilde{\mathbf{H}}_\infty(X|Z) - \lambda.$$

In particular,

$$\widetilde{\mathbf{H}}_\infty(X|Y) \ge \mathbf{H}_\infty((X,Y)) - \lambda \ge \mathbf{H}_\infty(X) - \lambda.$$

The item Lemma 1.2(b) can be treated as a simplistic form of a chain rule for min-entropy. However, its significant weakness is that the inequality does not depend on the random properties of Y but its actual size λ. We illustrate this by another simple example.

Example 1.3 *(Two blocks almost half entropy).* Let X, Y be two random variables distributed over $\{0,1\}^n$ with joint distribution of min-entropy $\mathbf{H}_\infty((X,Y)) = n$. Then, Lemma 1.2(b) gives us a trivial estimate $\widetilde{\mathbf{H}}_\infty(X|Y) \ge \mathbf{H}_\infty((X,Y)) - |Y| = 0$ regardless of the distribution of Y.

Nevertheless, if we condition (X_1, X_2) given in Example 1.1 with a random variable Z defined by the formula

$$Z = i \iff X_i = e,$$

(e is defined inside Example 1.1) then the variable X_{3-i} has conditional min-entropy $\mathbf{H}_\infty(X_{3-i}|Z=i) = n$. Therefore, there exists a certain additional "knowledge" Z which allows us to extract almost the whole min-entropy from the pair (X_1, X_2). Namely, the event

$$\mathbf{H}_\infty(X_1|Z=z) + \mathbf{H}_\infty(X_2|Z=z) \ge \mathbf{H}_\infty(X_1, X_2) - 1$$

holds with probability 1 over the random choice of $z \leftarrow Z$. This suggests that the way to obtain the full version of a chain rule is the additional conditioning. This is exactly what we do in this paper: for a pair of random variables and $\varepsilon > 0$ we exhibit an random variable Z_ε (depending solely on Y) such that

$$\Pr_{y \leftarrow Y}\left(\mathbf{H}_\infty(X|Y=y) + \mathbf{H}_\infty(Y|Z_\varepsilon) \le (1-\varepsilon)\cdot\mathbf{H}_\infty((X,Y))\right)$$

is negligible.

Our main technical results are Lemma 3.1 (bivariate case) and Lemma 3.5 (general case). Similar approach was used in certain different applications and is classically called *spoiling knowledge*.

1.2. Previous research and the statement of the main result

Previous research concerning chain rule is not restricted to the paper [1] mentioned above. For instance, the authors of [2] and [3] prove that random (sufficiently large) subtuple of some set of variables with high min-entropy must preserve some significant amount of this entropy. Our result can be viewed as a generalization of this fact. More precisely, from our reasoning we get that some *specific* (not random) subtuple preserves some significant portion of min-entropy (see Corollary 3.5). Moreover in [3] the authors try to deal with the problem of chain rule for min-entropy but need to make big effort to get some complicated workaround since they do not have any quasi chain rule for min-entropy in hand. They show a simplified version of their result and give a short brief proof based on chain rule for classical Shannon entropy. (Different examples of various "workarounds" may be also found in [4–6]). Another important previous result is the following lemma [7, Lemma A.1], which we state in the form from [8].

Lemma 1.4 *(Lemma 4.2 (Min-Entropy-Splitting Lemma) in [8]). Let $\epsilon \ge 0$ and let X_0, X_1 be random variables (over possibly different alphabets) with $\mathbf{H}^\epsilon_\infty(X_0X_1) \ge \alpha$. Then, there exists a binary random variable C over $\{0,1\}$ such that $\mathbf{H}^\epsilon_\infty(X_{1-C}C) \ge \alpha/2$.*

This is a very interesting result that shows that it is possible to extract partial min-entropy from a pair of variables. However, the authors justify high min-entropy of just *one* variable from the pair. In our main result (Lemma 3.1) we get significantly more, dealing with *both* variables at once. More precisely, we prove the following