Concavity, Core-concavity, Quasiconcavity: A Generalizing Framework for Entropy Measures

Arthur Américo, Pasquale Malacaria
School of Electronic Engineering and Computer Science
Queen Mary University of London
London, United Kingdom
emails: {a.passosderezende, p.malacaria}@qmul.ac.uk

2021 IEEE 34th Computer Security Foundations Symposium (CSF) | 978-1-7281-7607-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CSF51468.2021.00021

***Abstract*—We present a new generalising framework for conditional entropies, considering a limit construction over sequences of core-concave entropies, and prove that quasiconcave functions are the set of such limits. This generalising framework subsumes recently proposed frameworks for entropies in quantitative information flow, including entropies whose conditional form reflects the expected leakage and the leakage in the worst-case scenario. Thanks to the properties of the limits it is also shown that several important information theoretical properties can be proven for the generalised entropies satisfying the axioms.**

***Index Terms*—Concavity, Core-concavity, Quasiconcavity, G-concavity, Concavification, Entropy, Uncertainty, Quantitative Information Flow**

I. Introduction

In recent years there has been interest in axiomatic approaches for information leakage. in CSF 2016 Alvim et al. [1] proposed a set of axioms that characterise reasonable properties that vulnerability measures might satisfy; in particular both axioms for prior vulnerability and for posterior vulnerability are considered together with axioms establishing the relationship between prior and posterior vulnerability. Their elegant framework establishes some fundamental connection: in particular under axiom averaging for posterior vulnerability, they prove the equivalence of the following axioms: convexity, monotonicity (i.e., non-negativity of leakage), and the data-processing inequality.[1]

In the same work [1] the authors also provided an alternative axiomatization where the posterior vulnerability is not averaging but a max. In this case, under max axiom for posterior vulnerability they prove the equivalence of the following axioms: quasiconvexity, monotonicity, and the data-processing inequality.

One limitation of this seminal work is information theoretical, i.e. while it captures measures of information based on gain functions, it fails to capture most known entropies, like for example most Rényi entropies. This limitation has been addressed in the work [3] by Américo et al. which introduced an axiomatization based on core-concavity [4], [5], which is a generalization of concavity which subsumes most entropies from the literature. In [3] under a variation of the axiom averaging called η-averaging for posterior vulnerability, the following equivalence are proven: core-concavity, monotonicity, and the data-processing inequality. One interesting aspect of the axiomatization in [3] is that it goes well beyond the information leakage community and it provides a foundational work for the wider information theory community. In fact, from that axioms several well-known information theoretical properties of Shannon entropy have been generalised, for example Shannon perfect secrecy theorem and the celebrated Fano inequality.

Noteworthy is that while the averaging axioms from [1] are trivially a particular case of the axioms in [3], the max axioms case from [1] was still a distinct axiomatization.

The main contribution of this work is to close this gap, that is to present a framework generalising all axiomatizations in both [1] and [3]. This contribution is hence unifying previous axiomatic efforts, but it goes beyond this. In fact by now including quasiconcavity in the same framework as concavity and core-concavity it also shed deep insights, so far unknown, about the information theoretical properties of quasiconcavity. For example Shannon perfect secrecy theorem and the generalized Fano inequality are proven to hold for quasiconcave functions. The proofs of these generalization are based on properties of limits and the framework itself is a limit construction over sequences of core-concave entropies.

This work also connects foundation of information leakage with several works in convex analysis, optimization and microeconomics, where mathematicians have investigated when quasiconcave functions can be "concavified", i.e. when given a quasiconcave function $f(x)$, does there exist a monotonically increasing function g such that $g(f(x))$ is concave. The problem is of interest because if this is the case then one can use convex optimisation techniques to solve optimisation problems for quasiconcave functions. The question was originally asked by De Finetti [6] and a comprehensive answer has been recently provided by Connell and Rasmussen [7], [8]. We adapt several results from the concavification literature to our limit construction. The limit construction also allows for posteriors other than the ones considered in the literature so far. An example, η-geometric mean is presented in this work. Again properties of entropies with such posteriors are derived by properties of limits.

Finally another valuable contribution of this work is to provide an answer to an order theoretical problem of max based

[1] An extended journal version of this work has been later published in [2].

vulnerabilities arising in the recent work by Chatzikokolakis et al. [9].

A. Main Contributions and Outline

The following is an outline of the paper, with a summary of the main contributions.

- Section II introduces notation and some preliminaries regarding concave and Schur/quasi/core-concave functions, and Section III presents the necessary background in quantitative information flow (QIF) and on the axiomatization of conditional entropies in the literature.
- Section IV introduces *limit entropies*, a new generalising framework of entropies, and proves that the proposed conditional form is well defined. We also prove that the unconditional form of this family coincide with the set of continuous quasiconcave functions.
- In Section V we show that the family of limit entropies subsume the aforementioned generalising families from the literature, being therefore a framework that completely captures all conditional entropies used in QIF.
- Section VI proves that all limit entropies satisfy the important properties *Data Processing Inequality* (DPI) and *Conditioning Reduces Entropy* (CRE), and that symmetric limit entropies respect weaker versions of some other well-known information-theoretical properties.
- In Section VII we introduce a new conditional form based on a generalisation of the geometric mean. We prove that concave and core-concave entropies that have this conditional form are limit entropies, and therefore satisfy all the properties proven in Section VI.
- Section VIII discusses the *Blackwell-Sherman-Stein* Theorem for limit entropies, and also uses the developed framework to provide some insight into a intriguing question raised by a recent work in the area [9].
- Finally, Section IX discusses some related work in the literature, and Section X concludes and discusses future work.

II. Mathematical Preliminaries

In this paper, unless stated otherwise, we assume all functions to be continuous real valued functions over the $(n-1)$ dimensional simplex Δ_n, which is defined as

$$\Delta_n = \left\{ (p_1, \ldots, p_n) \in \mathbb{R}^n \mid \sum_i p_i = 1 \text{ and } \forall i, p_i \geq 0 \right\}.$$

Given $p \in \Delta_n$, we use p_i to denote the ith entry of p.

A. Uniform Convergence and Closure.

In this section we define convergence of real-valued functions over Δ_n, which is a fundamental concept for the results in this paper. For a complete treatment, we refer the reader to Chapter 7 of [10].

Definition 1: The sequence of functions $\{f_i\}$ is said to *converge pointwise* to f if for all $p \in \Delta_n$ and $\epsilon > 0$ there is $N \in \mathbb{N}$ such that

$$j > N \implies |f_j(p) - f(p)| < \epsilon;$$

and it is said to *converge uniformly* to f if for all $\epsilon > 0$ there is $N \in \mathbb{N}$ such that for all $p \in \Delta_n$

$$j > N \implies |f_j(p) - f(p)| < \epsilon.$$

If $\{f_i\}$ converges pointwise to f, then, no matter how large m is, there might be a p such that the distance $|f_m(p) - f(p)|$ is arbitrarily large. Uniform convergence, on the other hand, guarantees that for large m, f_m is "close" to f in the entire domain.

For our purposes, the most fundamental result about uniformly convergent sequences is that they preserve continuity. In our setting, this result can be stated as.

Theorem 1 ([10, Theorem 7.12]): If $\{f_i\}$ is a sequence of continuous functions that converges uniformly to f, then f is continuous.

To see why pointwise convergence is not a sufficiently strong assumption for Theorem 1, consider the sequence of functions over Δ_2 given by $f_i(p_1, p_2) = i^{-p_1}$. Each f_i is continuous for all i, but the sequence converges pointwise to the discontinuous function

$$f(p_1, p_2) = \begin{cases} 1, & \text{if } p_1 = 0, \\ 0, & \text{if } p_1 > 0. \end{cases}$$

Notice that the convergence is not uniform: let $\epsilon = 1/2$. For all i there is $p_1 > 0$ such that $f_i(p_1, p_2) = i^{-p_1} \geq 1/2$, and hence $|f_i(p_1, p_2) - f(p_1, p_2)| \geq 1/2$.

Given a set of functions, it is natural to consider all functions that can be obtained as limits of the functions of that set. This is what is called the *uniform closure*.

Definition 2 ([10]): Let $\mathcal{A}$ be a set of functions. The *uniform closure of* $\mathcal{A}$ is the set of all functions that are limits of uniformly convergent sequences of functions in $\mathcal{A}$.

B. Schur/quasi/core-concavity

A function $h : \Delta_n \to \mathbb{R}$ is *concave* if, $\forall p, q \in \Delta_n$ and $\lambda \in (0, 1)$,

$$h(\lambda p + (1-\lambda)q) \geq \lambda h(p) + (1-\lambda)h(q). \tag{1}$$

If the inequality above is strict for all p, q and λ, f is *strictly concave*.

A function h is *quasiconcave* if, $\forall p, q \in \Delta_n$ and $\lambda \in (0, 1)$,

$$h(\lambda p + (1-\lambda)q) \geq \min(h(p), h(q)). \tag{2}$$

And analogously, h is *strictly quasiconcave* if the inequality above is strict for all *distinct* p, q and all $\lambda \in (0, 1)$. Notice that (1) implies (2), and thus all concave functions are quasiconcave.

Contrary to concave functions, quasiconcave functions may have local maxima that are not global maxima — i.e., there might be a open neighbourhood $U \subset \Delta_n$ and a $p \in U$ such that $\forall p' \in U$, $h(p) \geq h(p')$; but $h(p) < \max_{q \in \Delta_n} h(q)$. This might be undesirable, specially for optimization problems [11].

Thus, it is sometimes interesting to consider *semistrict quasiconcavity* [12, Definition 3.11] which, under continuity assumptions, is a equivalent condition to all local optima being global optima [12, Theorem 3.37].

The function h is *semistrictly quasiconcave* if for all p, q and all $\lambda \in (0,1)$.

$$h(p) > h(q) \implies h(\lambda p + (1-\lambda)q) > h(q). \tag{3}$$

It is immediate that (1) implies (3), thus all concave functions are semistrictly quasiconcave. However, notice that (3) does not necessarily imply (2), and there are indeed *discontinuous* semistrictly quasiconcave functions that are not quasiconcave [13, Remark 2.4]. However, continuous functions are more well-behaved.[2]

Proposition 1 ([13, Theorem 2.5]): If h is semistrictly quasiconcave and continuous, then it is quasiconcave.

Concavity seems like a natural requirement for entropy measures, and in some circumstances it is equivalent to some intuitively-desirable properties [2] w.r.t. *posterior entropies*, which will be discussed in the next section. However, many widely-used entropies in the literature — such as the Rényi entropies [14] — are not concave. This led to the proposal of core-concave entropies [5], a wider class of entropies that still satisfy said intuitively-desirable properties [3].

Besides entropies, the concept of core-concave is also of interest in the study of generalised concave functions, under the name of *G-concavity* [12, Chapter 8].

Definition 3: A function $h : \Delta_n \to \mathbb{R}$ is *core-concave* if there are functions η, f such that:

- $f : \Delta_n \to \mathbb{R}$ is a continuous and concave function,
- $\eta : \mathrm{Range}(f) \to \mathbb{R}$ is a strictly increasing continuous function,[3]
- for all $p \in \Delta_n$, $h(p) = \eta(f(p))$.

As (3) holds for any concave function f, and the inequality is preserved if we apply an increasing function η to both sides, we obtain the following result.

Proposition 2 ([12, Proposition 8.1]): All core-concave functions are semistrictly quasiconcave.

As we will see on Section II-C, the converse of the above result is not true: not all semistrict quasiconcave (and hence, not all quasiconcave) functions are core-concave.

Abbreviating concavity by c.v. for and quasiconcavity by q.c.v., the discussion in this section so far may be summarized as follows, considering only continuous functions.

$$c.v \implies core\text{-}c.v. \implies semistrictly\ q.c.v. \implies q.c.v.$$

Another interesting concept related to concave functions is *Schur-concavity*, which however is of interest only for *symmetric functions* — i.e., functions for which $h(p_1, \ldots, p_n) = h\left(p_{\phi(1)}, \ldots, p_{\phi(n)}\right)$ for any permutation $\phi : \{1, \ldots, n\} \to \{1, \ldots, n\}$.

Given $p, q \in \Delta_n$, p *majorizes* q if for all $k \leq n$, $\sum_{i=1}^{k} p_{[i]} \geq \sum_{i=1}^{k} q_{[i]}$, where $(p_{[1]}, \ldots, p_{[n]})$ and $(q_{[1]}, \ldots, q_{[n]})$ are non-decreasing rearrangements of p and q. A function h is said to be *Schur-concave* iff $h(p) \leq h(q)$ whenever p majorizes q. As any symmetric quasiconcave function is Schur-concave [15, Chapter 3.C], an immediate consequence of Proposition 2 is the following:

Corollary 1: Any symmetric core-concave function is Schur-concave.

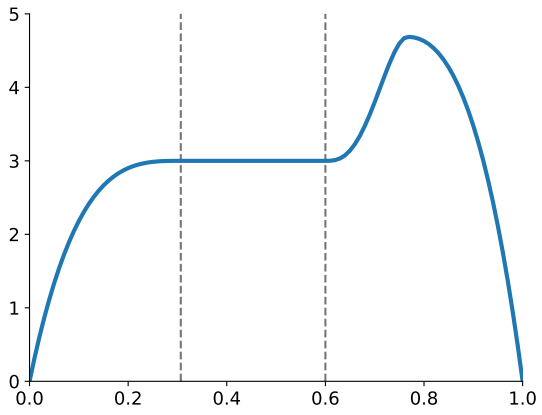

Fig. 1. A quasiconcave but not semistrictly quasiconcave function. Notice the region between the grey dotted lines, for which the function is constant.

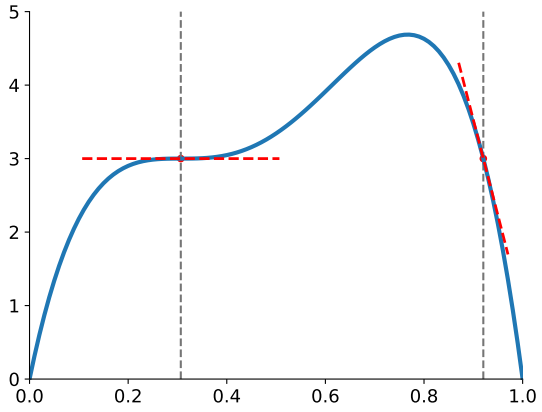

Fig. 2. Function h given in (4), a semistrictly quasiconcave but not core-concave function. The red dotted lines are the tangents of h at $p_1 = \gamma^{-1}$ and $p_1 = 3\gamma^{-1}$.

C. Are All Semistrictly Quasiconcave Functions Core-concave?

As will be discussed in Section III, quasiconcave functions and core-concave functions play a pivotal role in QIF, as unconditional forms of two different generalising families of information measures [2], [3]. As the main objective of this paper is unifying these two families, it is paramount to establish what is the relationship between core-concavity and quasiconcavity.

It is clear that there are quasiconcave functions that are not core-concave: if there is some open neighbourhood $U \subset \Delta_n$ where h is constant, but not maximal, h cannot be core-concave: in this case, for any increasing η, $\eta^{-1} \circ h$ would have a plateau at a non-maximal point, and hence $\eta^{-1} \circ h$ would not

[2]Actually, as stated in [13], *upper semicontinuity* is sufficient.

[3]Notice that, being f a continuous function over a compact, $\mathrm{Range}(f)$ is a closed interval.

be concave (See Figure 1). However, it is not obvious whether all *semistrictly concave* functions are core-concave.

This question has received particular attention in the field of microeconomics, where quasiconcavity of *utility functions* appears naturally as a necessary and sufficient requirement for rational agents that behave in some intuitively-reasonable ways [16, Lecture 4].

This question was studied by Bruno de Finetti in the paper "Sulle stratificazioni convesse" [6], in which the author identifies as one of the motivations establishing whether reasonable utility functions could always be assumed to be concave — which would imply that for all quasiconcave h there is a increasing function ζ such that $\zeta \circ h$ is concave (notice that this is equivalent to h being core-concave, by taking $\zeta = \eta^{-1}$ in Definition 3).

De Finetti proves that this is not the case, providing three counter-examples in Section 2 of [6]. To understand why this is the case, we present here an example of a semistrictly quasiconcave but not core-concave function $h : [0, 1] \to \mathbb{R}$, which is a scaled version of Example 3 in [7]. Let $\gamma = 2 + \sqrt[3]{2}$, and consider the following function, depicted in Figure 2 (notice that a function over $[0, 1]$ can be naturally interpreted as a function over Δ_2).

$$h(r) = -\gamma^4 r^4 + 6\gamma^3 r^3 - 12\gamma^2 r^2 + 10\gamma r. \tag{4}$$

Suppose there is a strictly increasing ζ such that $f = \zeta \circ h$ is concave. Now, $h(\gamma^{-1}) = h(3\gamma^{-1}) = 3$, but its derivatives yield $h'(\gamma^{-1}) = 0$ and $h'(3\gamma^{-1}) = -8\gamma$. Thus, it is not possible to have $f'(\gamma^{-1}) > 0$ unless $f'(3\gamma^{-1}) = -\infty$. As γ^{-1} is to the left of where f should attain its maximum, and $3\gamma^{-1}$ is not on the boundary of the domain, f cannot be concave.

III. Preliminaries on QIF and Information Theory

Before discussing the basic concepts of QIF, we need to introduce some notation regarding random variables. We assume all random variables to take values over *finite, nonempty* sets, and use capital letters $X, Y, \ldots$ to represent random variables taking values on the sets $\mathcal{X}, \mathcal{Y}, \ldots$. Unless stated otherwise, we index the elements of those sets by natural numbers, writing $\mathcal{X} = \{x_1, x_2, \ldots, x_{|\mathcal{X}|}\}$, $\mathcal{Y} = \{y_1, y_2, \ldots, y_{|\mathcal{Y}|}\}$, et cetera.

Given a random variable (r.v.) X with $|\mathcal{X}| = n$, we write p_X for the probability distribution $p_X(x) = \mathbf{Pr}\{X = x\}$. We also overload the notation and use p_X to denote the vector $(p_X(x_1), \ldots, p_X(x_n)) \in \Delta_n$. Given a joint random variable X, Y, we define, for each $y \in \mathcal{Y}$, $X|y$ to be a r.v. taking values in $\mathcal{X}$, with probability distribution $p_{X|y}(x) = {}^{p_{X,Y}(x,y)}/_{p_Y(y)}$. Given a function f over Δ_n, we will use $f(X)$ and $f(p_X)$ interchangeably.

A. Basic Model: Secrets, Observables and Systems

Sensitive information is modelled in QIF by a *secret*, which is represented by a r.v. X taking values on the set $\mathcal{X} = \{x_1, \ldots, x_n\}$. A *system* takes as input the value of X and, upon execution, produces an *observable behaviour* (or simply *observable*), represented by a r.v. Y. The system is modelled as a information theoretic *channel* K, which is a row stochastic matrix with rows indexed by $\mathcal{X}$ and columns indexed by $\mathcal{Y}$, $K(y|x)$ being the conditional probability of the system producing observable behaviour y when the secret value is x — i.e., $K(y|x) = p_{Y|x}(y)$. The notation $K : \mathcal{X} \to \mathcal{Y}$ indicates that $\mathcal{X}$ and $\mathcal{Y}$ are respectively the *input* and *output* sets. A channel K and a input X with probability p_X induce a joint distribution $p_{X,Y}(x, y) = p_X(x)K(y|x)$. From this joint, it is possible to compute $p_Y(y) = \sum_x p_{X,Y}(x, y)$ and $p_{X|y}(y) = \sum_x p_{X,Y}(x, y)$.

The QIF framework assumes the existence of an *adversary*, whose initial knowledge about the secret corresponds to the distribution p_X, knows the values $K(y|x)$ and is able to observe the realisation of Y. By observing an execution of the system resulting in an observable $y \in \mathcal{Y}$, the adversary is able to *update* its knowledge about the secret from the initial distribution to $p_{X|y}$. For this reason, the distribution p_X is often called the *prior*, and the distributions of the form $p_{X|y}$, the *posterior* distributions. Whenever $p_{X|y} \neq p_X$ for some $y \in \mathcal{Y}$ the system behaviour is dependent on the secret value, and there may be some leakage of information.

B. Background on Entropies and g-leakage

In order to quantify *how much* information a system leaks, we make use of *entropies*, which are real valued functions over probability distributions. Broadly speaking, an *entropy* quantifies how much uncertainty regarding the secret value a probability distribution represents, with higher values indicating that the secret is "safer" from the adversary. By comparing the uncertainty of the secret before and after the system's execution, one can quantify how much information the system leaked.

The choice of which entropy to use it is a nontrivial problem and it is usually context-sensitive, depending on the interests and capabilities of the adversary, and on properties of the system being modelled. The most famous of such quantity, specially in the field of information theory, is *Shannon Entropy* [17], which measures the expected amount of questions of the type "is the secret a member of $\mathcal{X}' \subset \mathcal{X}$?" an optimal adversary would need to identify the secret value. It is defined as $H_1(X) = -\sum p(x_i) \log(p(x_i))$. [4]

A very common scenario in security settings is that of a brute-force attack, i.e. when the adversary can check a big number of secrets in a sequential manner. *Guessing entropy*, introduced by Massey [18], is very useful for measuring uncertainty in such scenarios. It is defined as $H_{guess}(X) = \sum i \cdot p(x_{[i]})$, in which $\mathcal{X} = \{x_{[1]}, \ldots, x_{[n]}\}$ is an enumeration of $\mathcal{X}$ such that $i < j \Rightarrow p(x_{[i]}) \geq p(x_{[j]})$.

Min-entropy, defined as $H_\infty(X) = -\log \max_i p(x_i)$, had its usage in QIF first proposed by Smith [19]. By reflecting the value of maximum probability, min-entropy is a useful entropy for scenarios in which the adversary has only one opportunity to guess the value of the secret correctly.

Introduced by Alvim et al. [20], the g-leakage framework is a generalization of entropies, predicated in the use of gain

[4] All logarithms are in base 2.

functions $g : \mathcal{W} \times \mathcal{X} \to \mathbb{R}$, in which $\mathcal{W}$ is a set of *actions* the adversary can take. The value $g(w, x)$ reflects the gain the adversary obtains by choosing action w when the secret value is x. The g-vulnerability is defined as

$$V_g(X) = \max_{w \in \mathcal{W}} \sum_i g(w, x_i) p(x_i).$$

It reflects the expected gain of an adversary that chooses the optimal action. Notice that this quantity is measuring the opposite of entropies: it is greater the less uncertain the secret is. This is inconsequential, as one can easily define a entropy by simply defining, for example, a *g-entropy* by taking $H_g = -\log V_g$, or a *g-uncertainty* by taking $U_g = -V_g$.

The g-leakage framework is a powerful generalising tool: it can be used to recover min-, Shannon and guessing entropy [21]; and it has been recently shown by Alvim et al. [2] that g-vulnerabilities are capable of expressing any nonegative continuous and convex function over Δ_n.

The *Rényi entropie*s, proposed by Alfred Rényi in [14], are a family of entropies generalises both Shannon and min-entropy as limit cases. Given $\alpha \geq 0$ and $\alpha \neq 1$, the Rényi entropy of order α is defined as

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \|p_X\|_\alpha,$$

where $\|p\|_\alpha = \left(\sum_i p_i^\alpha\right)^{1/\alpha}$ is the α norm.[5] Shannon and min-entropy can be recovered from the above definition by taking $\alpha \to 1$ and $\alpha \to \infty$, respectively.

C. Conditional Forms and Information Leakage

Given an entropy H, such as the ones introduced in Section III-B, it is immediate to compute the uncertainty $H(X)$ associated with the prior distribution. To obtain how much information the system leaks, we must compare the value $H(X)$ with the adversary's uncertainty after the realisation of the system's observable behaviour Y. This is usually achieved by considering a *conditional* (also called *posterior*) form of entropy, which are usually obtained from the distribution p_Y over the outputs and the conditional distribution $p_{X|y}$. The conditional form of H is denoted by $H(X|Y)$. The information leaked by the system can then be computed by the *H-mutual information*

$$I_H(X;Y) = H(X) - H(X|Y). \tag{5}$$

If H is Shannon entropy, the leakage of the system is simply the well-known mutual information $I_{H_1}(X;Y) = H_1(X) - H_1(X|Y)$ [22]. The conditional forms of the first four entropies in Section III-B are defined as

$$\begin{aligned}
H_1(X|Y) &= \sum_{y \in \mathcal{Y}^+} p(y) H_1(X|y), \\
H_{guess}(X|Y) &= \sum_{y \in \mathcal{Y}^+} p(y) H_{guess}(X|y), \\
H_\infty(X|Y) &= -\log \sum_{y \in \mathcal{Y}^+} p(y) \max_i p_{X|y}(x_i), \text{ and} \\
H_g(X|Y) &= -\log \sum_{y \in \mathcal{Y}^+} p(y) V_g(X|y);
\end{aligned}$$

where $\mathcal{Y}^+ = \{y \in \mathcal{Y} \mid p(y) > 0\}$.

The conditional forms of Shannon and guessing entropies are easily justified, as they are simply the expected value of the uncertainty of the adversary after the execution of the system. The same intuitive approach, however, cannot be applied to min-entropy, g-entropy or to (most of) the Rényi entropy family, as the resulting conditional form would not, in general, respect *Conditioning Reduces Entropy* (`CRE`). This means that the adversary could lose information by observing the result of the system, which disqualify those as reasonable conditional forms. A more thorough discussion of `CRE` is postponed to Section III-D.

There is currently no agreed conditional form for the Rényi entropy family [23], [24]. In this work, we consider two versions. The first, known as the Arimoto-Rényi conditional entropy [25], is given by

$$H_\alpha(X|Y) = \frac{\alpha}{1-\alpha} \log \sum_{y \in \mathcal{Y}^+} p(y) \left\|p_{X|y}\right\|_\alpha, \tag{6}$$

while the second, the Hayashi-Rényi conditional form introduced in [26], is given by

$$H'_\alpha(X|Y) = \frac{1}{1-\alpha} \log \sum_{y \in \mathcal{Y}^+} p(y) \left\|p_{X|y}\right\|_\alpha^\alpha. \tag{7}$$

These conditional forms behave in very different ways. The form (6) is specially interesting as it has the desirable property that $\lim_{\alpha \to 1} H_\alpha(X|Y) = H_1(X|Y)$ [25] and $\lim_{\alpha \to \infty} H_\infty(X|Y)$ [24]. While the form (7) coincides with Shannon conditional entropy when α tends to 1, we have [24]

$$\lim_{\alpha \to \infty} H'_\alpha(X|Y) = -\log \left(\max_{y \in \mathcal{Y}^+,\, x \in \mathcal{X}} p_{\mathcal{X}|y}(x) \right) \neq H_\infty(X|Y).$$

D. Axioms of Entropy Measures

As we have seen, different settings call for different entropies. This is specially true for QIF, as each entropy measure models particularities of the system and the interests of an adversary. Given the myriad of possible entropy definitions, a natural question is to establish which are the entropies that "make sense". To this end, Alvim et al. [2] used an axiomatic approach to determine exactly those entropies that satisfy some intuitively desirable properties.

Before we discuss their results, we need to make the definition of "entropy", which has been somewhat obscure until here, a bit more precise.

[5] We use this definition even when $\|\cdot\|_\alpha$ is not a norm, i.e., when $0 < \alpha < 1$.

Definition 4: Let $n > 0$. An *entropy* (over Δ_n) is a quantity H over r.v.s with two associated forms.

- A *unconditional form* $H(X)$, defined for all r.v.s X for which $|\mathcal{X}| = n$. The unconditional form can be seen as the function $p_X \mapsto H(X)$ over Δ_n.
- A *conditional form* $H(X|Y)$, defined for all pairs of r.v.s (X, Y) for which $|\mathcal{X}| = n$.

Moreover we assume the unconditional form to be continuous. We usually omit the dimension of the simplex, as it is usually irrelevant - all our results hold for any choice of $n > 0$.

Definition 4 is quite ample, and many entropies satisfying this definition are undesirable, not representing a intuitive notion of uncertainty about X, or, for the conditional form, of uncertainty about X when the value of Y is known.

In order to characterize exactly which entropies are reasonable (in their words, "bring order to the zoo"), Alvim et al. took an axiomatic approach. Their axioms concern properties of the conditional and unconditional forms, and also relationships between them.[6]

In the following paragraphs, we give a detailed presentation of Alvim et al.'s [2] proposed axioms and results.

1) Axioms Regarding the Structure of Conditional Forms: It is reasonable to expect that the conditional form of an entropy should be expressible in terms of its unconditional form. As discussed in Section III-C, it is usually assumed that the conditional form $H(X|Y)$ will be related to the quantities $\{H(X|y)\}_{y\in\mathcal{Y}}$, and also to the distribution p_Y. The most natural way this can be done is simply by taking the expected value of $\{H(X|y)\}_{y\in\mathcal{Y}}$, according to p_Y.

*Definition 5 (*AVG *axiom):* H satisfies *Averaging (*AVG*)* if its conditional form is defined as:

$$H(X|Y) = \sum_{y\in\mathcal{Y}^+} p(y)H(X|y).$$

In security scenarios, the AVG axiom might be too forgiving, and in some contexts one may prefer to evaluate the *maximum* amount of information that the system may leak. Some well known notions of privacy, like differential privacy, use this maximum leakage approach. In such scenarios, a more fitting conditional form would be one that evaluates the uncertainty in a "worst-case" scenario.

*Definition 6 (*MIN *axiom):* H satisfies *Minimum (*MIN*)* if its conditional form is defined as:

$$H(X|Y) = \min_{y\in\mathcal{Y}^+} H(X|y).$$

2) Axioms Reflecting Reasonable Information-theoretical Properties: The next two axioms are foundational information-theoretic properties (see, e.g., Chapter 2 of [22]). The first captures the idea that "information cannot hurt". That is, whatever the uncertainty about X is, it cannot increase by learning a side information Y.

[6]As they studied measures of *certainty*, Alvim et al.'s original axioms are dual versions of the ones presented here: they use *convexity*, *quasiconvexity* and *maximum* instead of *concavity*, *quasiconcavity* and *minimum*.

*Definition 7 (*CRE *axiom):* H satisfies *Conditioning reduces entropy (*CRE*)*[7] if, for all r.v. X,Y,

$$H(X|Y) \leq H(X),$$

with equality holding if X and Y are independent.

The second axiom relates to the idea that one cannot decrease uncertainty by cleverly manipulating data. For example, digitally processing a footage of a person will not create more information about the subject than what was already available in the original footage (to the dismay of many TV criminalists).[8]

*Definition 8 (*DPI *axiom):* H satisfies *Data-processing inequality (*DPI*)* if, for all r.v. X,Y,Z such that $X \to Y \to Z$ (i.e., if X and Z are conditionally independent given Y),

$$H(X|Y) \leq H(X|Z).$$

3) Axioms Regarding Properties of Unconditional Forms: Finally, the last two axioms relate properties of the unconditional forms as functions over Δ_n.

*Definition 9 (*CV *axiom):* H satisfies *Concavity (*CV*)* if its unconditional form is concave over Δ_n.

*Definition 10 (*QCV *axiom):* H satisfies *Quasiconcavity (*QCV*)* if its unconditional form is quasiconcave over Δ_n.

These axioms were justified for their own sake by Alvim et al. in [2], but this is predicated on the conditional form axioms AVG and MIN. However, the CV and QCV axioms are important because they provide a complete, and straightforward characterizations for the entropies in relation to the other axioms, as will be seen in the next section.

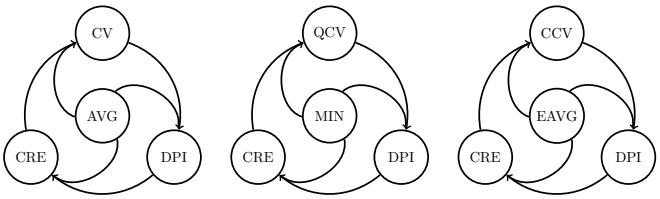

Fig. 3. Implication graphs of axioms schemes in the literature. The AVG set of axioms (left) and the MIN set of axioms (middle) were proposed by [2]. The EAVG set of axioms [3] (right) generalises the AVG one, when one takes η to be the identity function.

4) Relationships Between the Axioms: The results obtained by Alvim et al. relating the above axioms may be expressed as follows:

Theorem 2 ([2, Propositions 14, 15, 16 and 18]): If H satisfies AVG, then CRE, DPI and CV are equivalent.

Theorem 3 ([2, Propositions 15, 20, 22 and 23]): If H satisfies MIN, then CRE, DPI and QCV are equivalent.

In words, any entropy whose conditional form is obtained according to AVG respects the important properties CRE and DPI if, and only if, its unconditional form is concave; and

[7]Note that this property is named *monotonicity* in [2].

[8]However, postprocessing may make the information easier to be understood by humans. For example, noise reduction algorithms may make the words in an old recording of a speech more easily discernible.

a similar result holds for MIN and QCV. The results are summarised in Figure 3.

As mentioned in Section III-B, the set of (unconditional) g-vulnerabilities was proven, in the same work, to coincide with that of convex functions over Δ_n. Therefore, one important result relating to QIF is that any reasonable entropy measure that satisfies AVG can be represented by (a dual of a) g-vulnerability.

5) An Extension of Alvim et al.'s Axioms: One problem with the above of axioms is that they seem to leave out many entropies that are widely used in the literature. For example, as seen Section III-C, min-entropy does not fall in either AVG nor MIN, being therefore not contemplated in their results. A similar fate befalls several other entropies, such as Rényi entropies of different conditional forms, Tsallis [27] and Sharma-Mittal entropies [28] (see [3], [5] for a more throughout discussion).

This problem was recently addressed [3] by relaxing the AVG axiom to a more generalised averaging axiom. This encompasses the entropies mentioned above, and generalises the AVG part of Alvim et al.'s results. Those results are based on *core-concave entropies*, first introduced in [5].

Definition 11: Let $F : \Delta_n \to \mathbb{R}$ and $\eta : \mathrm{Range}(F) \to \mathbb{R}$ be continuous functions, and η be strictly increasing. We say that H is *given by* the pair (η, F), and write $H = (\eta, F)$, if the unconditional form of H is $H(X) = \eta(F(X))$.

Notice that this definition does not a priori exclude any continuous unconditional form, as it is possible to reconstruct any arbitrary H by taking $F(X) = H(X)$ and η to be the identity function.

Definition 12: An entropy $H = (\eta, F)$ is *core-concave* if F is concave

Notice that the unconditional form of a core-concave entropy is a core-concave function, as per Definition 3.

The new axioms introduced in [3] are the following

Definition 13: Let $H = (\eta, F)$. H is said to satisfy

CCV **(Core-concavity)**: if $H = (\eta, F)$ is core-concave (as in Definition 12).

EAVG **(η-averaging)**: if, given r.v. X,Y, the conditional form is defined as:

$$H(X|Y) = \eta\left(\sum_{y \in \mathcal{Y}^+} p(y) F(X|y)\right).$$

A similar result to those obtained by Alvim et al. was proved for this new set of axioms, also depicted in Figure 3.

Theorem 4 ([3]): If $H = (\eta, F)$ satisfies EAVG, then CRE, DPI and CCV are equivalent.

Notice that this result recovers Theorem 2, when one takes η to be the identity function.

The generalising family given by CCV and EAVG solve the aforementioned problem, capturing all conditional forms described in Section III-C. Some choices of (η, F) that recover those conditional forms are summarised on Table I

TABLE I
SOME CHOICES OF (η, F) RECOVERING THE CONDITIONAL FORMS IN SECTION III-C

Entropy	$\eta(r)$	$F(p)$	
H_1	r	$H_1(p)$	
H_{guess}	r	$H_{guess}(p)$	
H_∞	$-\log -r$	$-\max_i p_i$	
H_g	$-\log -r$	$-V_g(p)$	
H_α	$\frac{\alpha}{1-\alpha}\log(r)$	$\|p\|_\alpha$	(if $0 < \alpha < 1$)
	$\frac{\alpha}{1-\alpha}\log(-r)$	$-\|p\|_\alpha$	(if $\alpha > 1$)
H'_α	$\frac{1}{1-\alpha}\log(r)$	$\|p\|_\alpha^\alpha$	(if $0 < \alpha < 1$)
	$\frac{1}{1-\alpha}\log(-r)$	$-\|p\|_\alpha^\alpha$	(if $\alpha > 1$)

IV. A UNIFYING FRAMEWORK FOR ENTROPIES

As discussed in Section III-D, the recent literature regarding the axiomatic treatment of information measures w.r.t. posterior entropy suggests three classes of entropies, according to the form of the posterior:

- entropies which satisfy CV and AVG,
- entropies which satisfy CCV and EAVG,
- entropies which satisfy QCV and MIN.

In this section, we will denote the sets of these entropies by $\mathcal{C}_{\texttt{AVG}}$, $\mathcal{H}_{\texttt{EAVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$, respectively. Each element of both $\mathcal{C}_{\texttt{AVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$ is uniquely identified by the function that gives its unconditional form, whereas an element of $\mathcal{H}_{\texttt{EAVG}}$ can be uniquely identified by a pair (η, F). Notice that $\mathcal{C}_{\texttt{AVG}} \subsetneq \mathcal{H}_{\texttt{EAVG}}$, as any given entropy in $\mathcal{C}_{\texttt{AVG}}$ is equivalent to a pair $(\mathrm{id}, F) \in \mathcal{H}_{\texttt{EAVG}}$ where F is a concave function and $\mathrm{id}(r)$=r is the identity function.

The objective of this section is to define a new set of entropies, $\mathcal{Q}$, such that $\mathcal{H}_{\texttt{EAVG}} \subset \mathcal{Q}$ and $\mathcal{Q}_{\texttt{MIN}} \subset \mathcal{Q}$. In other words, we will present a new definition that subsumes all aforementioned entropies. After that, we will prove that the entropies in $\mathcal{Q}$ satisfy some fundamental information-theoretical properties, which are thus also proofs for $\mathcal{H}_{\texttt{EAVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$.

A. Entropies as Limits of Sequences

The set $\mathcal{Q}$ will be defined as limits of sequences of core-concave entropies that satisfy EAVG.

Definition 14: Let $\{H^i = (\eta_i, F_i)\}_i$ be a sequence in $\mathcal{H}_{\texttt{EAVG}}$, such that $\eta_i \circ F_i$ converges uniformly. We define the limit of $\{H^i\}$ to be the entropy H defined as

- $H(X) = \lim_{i\to\infty} \eta_i(F_i(p_X))$,
- $H(X|Y) = \limsup_{i\to\infty} H^i(X|Y)$.

We denote by $\mathcal{Q}$ the set of all entropies which are the limit of a sequence of entropies in $\mathcal{H}_{\texttt{EAVG}}$ whose unconditional form is uniformly convergent. We refer to the entropies in $\mathcal{Q}$ by the name of *limit entropies*.

Henceforth, we assume all sequences of entropies to be of elements of $\mathcal{H}_{\texttt{EAVG}}$ and that their unconditional form is uniformly convergent. One might wonder whether the conditional form above is not well defined, as the limit superior ($\limsup$) of a sequence might be ∞ or $-\infty$. The next result guarantees that this is not the case.

Proposition 3: Let $\{H^i = (\eta_i, F_i)\}$ be a sequence in $\mathcal{H}_{\texttt{EAVG}}$ and suppose $\{\eta_i \circ F_i\}$ converges uniformly. Then, the limit H of $\{H^i\}$ satisfies $|H(X|Y)| < \infty$ for all X, Y.

Proof: Because $\eta_i \circ F_i$ is continuous and converges uniformly on a compact set, there is $\gamma > 0$, $N \in \mathbb{N}$ such that

$$i > N \implies \quad \forall p \in \Delta_n, \quad |\eta_i(F_i(p))| < \gamma.$$

As H^i satisfies EAVG and η_i is strictly increasing for all i, we have for all X, Y,

$$\inf_{p\in\Delta_n} \eta_i(F_i(p)) \leq H^i(X|Y) \leq \sup_{p\in\Delta_n} \eta_i(F_i(p)).$$

Thus, for $i > N$,

$$|H^i(X|Y)| \leq \max\left(-\inf_{p\in\Delta_n} \eta_i(F_i(p)), \sup_{p\in\Delta_n} \eta_i(F_i(p))\right) \leq \gamma.$$

■

The use of limit superior instead of regular limit on Definition 14 is necessary as the latter may not exist, even though $\eta_i \circ F_i$ converges uniformly. Take, for example, the sequence $\{H^i = (\eta_i, F_i)\}$ given by,

$$\eta_i(x) = \begin{cases} \frac{i}{1-i}\log(-x) & \text{if } i \text{ is odd,} \\ \frac{1}{1-i}\log(-x) & \text{if } i \text{ is even.} \end{cases}$$

$$F_i(x) = \begin{cases} -\|p\|_i & \text{if } i \text{ is odd,} \\ -\|p\|_i^i & \text{if } i \text{ is even.} \end{cases}$$

In both cases, $\eta_i \circ F_i$ converges uniformly to the unconditional form of min entropy. However, their conditional form under EAVG is alternately the Arimoto-Rényi and Hayashi-Rényi conditional forms. Thus, as discussed in Section III-C, the odd subsequence converges to conditional min-entropy, while the even subsequence converges to $-\log\max_{x,y} p_{X|y}(x)$.

B. Unconditional Forms of Limit Entropies

As discussed in Section III-D, the set of functions that are the unconditional forms of some entropy in $\mathcal{H}_{\texttt{EAVG}}$ coincide with the set of core-concave functions [3], and similarly for $\mathcal{Q}_{\texttt{MIN}}$ and the set of quasiconcave functions [2]. From Definition 14, however, it is not clear what set of functions coincide with the unconditional forms of the entropies in $\mathcal{Q}$. The objective of this section is characterising this set.

As one can take constant sequences in Definition 4, this family definitely contains all core-concave functions. But what else may be obtained by uniformly convergence sequences of core-concave functions?

It turns out that this family is exactly the set of continuous quasiconcave functions. This is a direct consequence of a result obtained in a recent work by Connell and Rasmussen [7]. We quote their result below, changing only the definitions to match those of our work.

Proposition 4 ([7, Corollary 3]): Suppose h is any quasiconcave function (possibly discontinuous). There exists a sequence of continuous strictly quasiconcave and core-concave functions h^i which converge to h as $i \to \infty$ pointwise almost everywhere, and uniformly on compact sets if h^i is continuous.

We are particularly interested in the last part of the statement, as we assume all entropies to have a continuous unconditional form over Δ_n, which is a compact set. This leads us to an interesting result.

Theorem 5: The set of continuous quasiconcave functions over Δ_n is the uniform closure (as in Definition 2) of the set of continuous core-concave functions over Δ_n.

Proof: That the set of continuous quasiconcave functions is contained in the uniform closure of the set of core-concave functions is directly implied by Proposition 4. Conversely, suppose $\{h^i\}$ is a uniformly convergent sequence of core-concave functions. As any core-concave is quasiconcave, for any $p_1, p_2 \in \Delta_n$, and $\lambda \in [0, 1]$ we have

$$h^i(\lambda p_1 + (1-\lambda)p_2) \geq \min(h^i(p_1), h^i(p_2)).$$

Thus, by taking $i \to \infty$ on the inequality above, we see that the limit of the sequence $\{h^i\}$ is quasiconcave. ■

An immediate consequence of this Theorem is the following Corollary.

Corollary 2: All entropies in $\mathcal{Q}$ have a continuous quasiconcave unconditional form. Moreover, for each continuous quasiconcave function h, there is some entropy $H \in \mathcal{Q}$ such that $H(X) = h(p_X)$.

V. Existing Entropies as Subsets of $\mathcal{Q}$

The next step is to prove that the generalised families of entropies introduced in Section III-D are subsumed by the family of limit entropies — that is, that $\mathcal{H}_{\texttt{EAVG}} \subset \mathcal{Q}$ and $\mathcal{Q}_{\texttt{MIN}} \subset \mathcal{Q}$.

Proving that $\mathcal{H}_{\texttt{EAVG}} \subset \mathcal{Q}$ is straightforward.

Proposition 5: $\mathcal{H}_{\texttt{EAVG}} \subset \mathcal{Q}$

Proof: Let $H = (\eta, F) \in \mathcal{H}_{\texttt{EAVG}}$ and define a constant sequence in $\mathcal{H}_{\texttt{EAVG}}$ by taking $(\eta_i, F_i) = (\eta, F)$ for all $i \in \mathbb{N}$. Then, H is the limit of (η_i, F_i), and thus $H \in \mathcal{Q}$. ■

Proving that $\mathcal{Q}_{\texttt{MIN}} \subset \mathcal{Q}$ is not as straightforward. From Corollary 2, we know that the unconditional forms of the entropies in $\mathcal{Q}$ actually contemplate all continuous quasiconcave functions. However, we still must show that, for any quasiconcave unconditional form, we can obtain a conditional entropy that satisfies MIN. We start by introducing the following result, which is a stronger version of Theorem 4 in [3], and which follows from the proof therein.

Theorem 6 ([3, Theorem 4]): Let $H_M = (\eta, F) \in \mathcal{H}_{\texttt{EAVG}}$, and let $\sigma = \sup_{p\in\Delta_n} F(p)$. Let $\{(\eta_i, F_i)\}_{i\in\mathbb{N}}$ be the sequence given by

$$\eta_i(x) = \eta\left(-(-x)^{1/i} + \sigma\right) \quad \text{and} \quad F_i(p) = -\left(-F(p) + \sigma\right)^i.$$

Then, $(\eta_i, F_i) \in \mathcal{H}_{\texttt{EAVG}}$ for all $i \in \mathbb{N}$. Moreover, for all X, Y,

$$\min_y H_M(X|y) = \lim_{i\to\infty} \eta_i\left(\sum_{y\in\mathcal{Y}^+} p(y)F_i(X|y)\right).$$

Given this result, we are able to obtain a sequence of entropies whose limit satisfy MIN, for any quasiconcave unconditional form.

Theorem 7: $\mathcal{Q}_{\texttt{MIN}} \subset \mathcal{Q}$

Proof: Let $H_M \in \mathcal{Q}_{\texttt{MIN}}$. It suffices to prove that there is a sequence $\{(\eta_i, F_i)\}_{i\in\mathbb{N}}$ in $\mathcal{H}_{\texttt{EAVG}}$ such that: 1) $\eta_i \circ F_i(p)$ converges uniformly to $H_M(p)$; and, 2) for all r.v.s X, Y,

$$H_M(X|Y) = \min_y H_M(X|y) = \lim_{i\to\infty} \eta_i\left(\sum_{y\in\mathcal{Y}^+} p(y)F_i(X|y)\right). \tag{8}$$

By Corollary 2, there is a sequence of entropies $H_j = (\eta_j, F_j)_{j\in\mathbb{N}}$ in $\mathcal{H}_{\texttt{EAVG}}$ which converges uniformly to the unconditional form of H_M.

Let $i \in \mathbb{N}$ and X, Y be r.v.s. Because H_j converges uniformly to H_M and min is a continuous function, $\min_y H_j(X|y)$ also converges to $\min_y H_M(X|y)$ as $j \to \infty$. Therefore, there is $N_i \in \mathbb{N}$ such that $j > N_i$ implies

$$\left|\min_y H_j(X|y) - \min_y H_M(X|y)\right| < \frac{1}{2i}.$$

For each j, let $\sigma_j = \sup_{p\in\Delta_n} F(p)$ and define a sequence $\{\eta_{j,k}, F_{j,k}\}_k$ by

$$\eta_{j,k}(x) = \eta_j\left(-(-x)^{1/k} + \sigma_j\right) \quad \text{and}$$
$$F_{j,k}(p) = -\left(-F_j(p) + \sigma_j\right)^k.$$

Then, Theorem 6 implies that there is $N_{i,j} \in \mathbb{N}$ such that, for all $k > N_{i,j}$,

$$\left|\eta_{j,k}\left(\sum_y p(y)F_{j,k}(X|y)\right) - \min_y H_j(X|y)\right| < \frac{1}{2i}.$$

Now, for each $i \in \mathbb{N}$, choose $j_i > N_i$ and $k_i > N_{i,j_i}$, and define $\eta'_i = \eta_{j_i,k_i}$, $F'_i = F_{j_i,k_i}$. Let $\epsilon > 0$. Then, for all $i > 1/\epsilon$,

$$\left|\eta'_i\left(\sum_y p(y)F'_i(X|y)\right) - \min_y H_M(X|y)\right| < \frac{1}{2i} + \frac{1}{2i} < \epsilon.$$

Thus, $\{\eta'_i, F'_i\}_i$ satisfies (8). Moreover, notice that $\eta'_i(F'_i(p)) = \eta_{j_i,k_i}(F_{j_i,k_i}(p)) = \eta_{j_i}(F_{j_i}(p))$, and thus the sequence of functions $\{\eta'_i \circ F'_i\}_i$ is a subsequence of $\{\eta_j \circ F_j\}_j$, and hence it converges uniformly to the unconditional form of H_M. ∎

VI. Information Theoretical Properties of Limit Entropies

In this section, we establish some fundamental information-theoretical properties of the entropies in $\mathcal{Q}$ — and hence, obtaining the same properties for $\mathcal{H}_{\texttt{EAVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$ as a by-product.

The properties discussed in this section have been proved separately for $\mathcal{H}_{\texttt{EAVG}}$ in [3], and could be adapted for $\mathcal{Q}_{\texttt{MIN}}$. Hence, the novelty here is to provide unified proofs that serves both cases. Moreover, in case in the future one finds other interesting entropy families subsumed in $\mathcal{Q}$, the proofs here will guarantee that they satisfy these properties as well.

Because limit superior preserves inequalities [10, Theorem 3.19], the proofs of these results is a quite straightforward modification of those in [3], with Definition 14.

A. DPI and CRE

Proposition 6: All entropies in $\mathcal{Q}$ satisfy CRE and DPI.

Proof: Let $H \in \mathcal{Q}$ be the limit of a sequence $\{(\eta_i, F_i)\}_i$ in $\mathcal{H}_{\texttt{EAVG}}$.

(CRE) From Theorem 4, each (η_i, F_i) satisfies CRE; that is, for all i and all r.v.s X, Y

$$\eta_i\left(\sum_y p(y)F_i(p_{X|y})\right) \le \eta_i(F_i(p_X))$$

By taking the limit superior in both sides, we arrive at $H(X|Y) \le H(X)$.

(DPI) Again from Theorem 4, each (η_i, F_i) satisfies DPI; that is, for all i and all r.v.s X, Y, Z such that $X \to Y \to Z$, we have

$$\eta_i\left(\sum_y p(y)F_i(p_{X|y})\right) \le \eta_i\left(\sum_y p(z)F_i(p_{X|z})\right).$$

Again, by taking the limit superior in both sides, we arrive at $H(X|Y) \le H(X|Z)$. ∎

B. Properties Requiring Symmetry and Expansibility

The results of this section require the entropies to be both *symmetric* and *expansible*.

An entropy $H \in \mathcal{Q}$ is called *symmetric* if, for all $(p_1, \ldots, p_n) \in \Delta_n$ and all bijections $\phi : \{1, \ldots, n\} \to \{1, \ldots, n\}$, we have $H(p_1, \ldots, p_n) = H(p_{\phi(1)}, \ldots, p_{\phi(n)})$.

As any symmetric quasiconcave function is Schur-concave [15, Chapter 3.C], Corollary 2 implies that any symmetric entropy in $\mathcal{Q}$ has a Schur-concave unconditional form.

The next proposition is quite natural, but important for our next results.

Proposition 7: If $H \in \mathcal{Q}$ is symmetric, it is the limit of a sequence $\{(\eta_i, F_i)\}_i$, where each F_i is symmetric.

Proof: Let $H \in \mathcal{Q}$ be symmetric, and $\{(\eta_i, F_i)\}_i$ be a sequence whose limit is H. Fix $i \in \mathbb{N}$ and let Φ be the set of all permutations over $\{1, \ldots, n\}$. For each permutation $\phi \in \Phi$ define $F_i^\phi(p_1, \ldots, p_n) = F_i\left(p_{\phi(1)}, \ldots, p_{\phi(n)}\right)$. Finally, let

$$F_i^*(p) = \frac{1}{n!}\sum_{\phi\in\Phi} F_i^\phi(p).$$

Being a finite sum of concave functions, F_i^* is itself concave. Moreover by construction it is symmetric. We claim that $\{(\eta_i, F_i^*)\}$ has H as limit. In fact, for each $p \in \Delta_n$

$$\min_{\phi\in\Phi} \eta_i \left(F_i^\phi(p)\right) \leq \eta_i \left(F_i^*(p)\right) \leq \max_{\phi\in\Phi} \eta_i \left(F_i^\phi(p)\right).$$

As H being symmetric implies that $\lim \eta_i \left(F_i^\phi(p)\right) = H(p)$ for all ϕ, we have $\lim_{i\to\infty} \eta_i(F_i^*(p)) = H(p)$. Similarly, for each X, Y,

$$\min_{\phi\in\Phi} \eta_i \left(\sum_y p(y) F_i^\phi(X|y)\right) \leq \eta_i \left(\sum_y p(y) F_i^*(Y|y)\right) \text{ and}$$
$$\eta_i \left(\sum_y p(y) F_i^*(Y|y)\right) \leq \max_{\phi\in\Phi} \eta_i \left(\sum_y p(y) F_i^\phi(X|y)\right).$$

And thus the conditional form is also preserved. ∎

The other condition is expansibility. In information theory, this usually means that one can "append" a distribution with outcomes with probability equal to 0, without changing the value of the entropy. Although the idea in this section in similar — to provide results relevant to expansible entropies — defining them on our framework is unnecessarily complicated.

Instead, when comparing entropies of distributions of different sizes, we will henceforth use the following convention: we "append" the shortest distribution with the necessary number of 0s and assume all quantities are over the same simplex. For example, if $|\mathcal{X}| = 2$ and $|\mathcal{Y}| = 3$; the inequality $H(p(x_1), p(x_2)) \leq H(q(y_1), q(y_2), q(y_3))$ actually means $H(p(x_1), p(x_2), 0) \leq H(q(y_1), q(y_2), q(y_3))$ and both sides of inequality are defined over Δ_3.

Notice that the results using this convention correctly contemplate expansible entropies — for example, all results below hold for Rényi entropies with the regular definition.

The next properties have been proved to hold for expansible, symmetric entropies in $\mathcal{H}_{\texttt{EAVG}}$ in Section V of [3]. We use those results and Proposition 7 to extend it to all entropies in $\mathcal{Q}$. These properties have an important and natural meaning in the context of information theory and information security. The following is an informal overview:

1) AIE (additional information increases entropy): this says that the entropy of a joint random variable (X, Y) is higher than the entropy of any of its constituents X, Y.
2) weak subadditivity: subadditivity expresses that the entropy of a joint random variable (X, Y) is upper bounded by the sum of the entropies of its constituents, i.e. $H(X, Y) \leq H(X) + H(Y)$. Subadditivity doesn't holds for general entropies like Rényi entropies but we show a weaker version holding for all entropies in $\mathcal{Q}$.
3) perfect secrecy: we show that Shannon's celebrated result about encryption holds for all entropies in $\mathcal{Q}$.
4) generalized Fano inequality: Fano inequality is the basic result connecting Shannon entropy with Bayes error. We show a generalized Fano inequality which is valid for all entropies in $\mathcal{Q}$.

1) AIE (additional information increases entropy):

Proposition 8: Let $H \in \mathcal{Q}$ be symmetric and expansible. Then:

1) $\forall X, Y \ H(X) \leq H(X, Y)$,
2) $\forall X, Y, Z \ H(X|Z) \leq H(X, Y|Z)$.

Proof: Those inequalities have been proven for symmetric and expansible $\mathcal{H}_{\texttt{EAVG}}$ [3, Proposition 5]. Let $H \in \mathcal{Q}$. Invoking Proposition 7, there is a sequence of symmetric $\{H^i = (\eta_i, F_i)\}_i$ whose limit is H. By taking the limit as $i \to$ for $H^i(X)$, $H^i(X, Y)$ in inequality 1), and the limit superior for $H^i(X|Z)$, $H^i(X, Y|Z)$ in inequality 2), we obtain the result for $H \in \mathcal{Q}$. ∎

2) Subadditivity: The subadditivity property $H(X, Y) \leq H(X) + H(Y)$ does not hold for all entropies in $\mathcal{Q}$. As a counterexample consider the following joint distribution over (X, Y), with $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2\}$

$$p_{(X,Y)}(x_1, y_1) = 0, \qquad p_{(X,Y)}(x_1, y_2) = 1/4,$$
$$p_{(X,Y)}(x_2, y_1) = 1/4, \qquad p_{(X,Y)}(x_2, y_2) = 1/2.$$

And one can check that for any Rényi entropy H_α with $\alpha > 1.61$, we have $H_\alpha(X, Y) > H_\alpha(X) + H_\alpha(Y)$.

Proposition 9: Let $H \in \mathcal{Q}$ be symmetric. Then $H(X, Y) \leq H(\tilde{p})$ where $\tilde{p}$ is the $|\mathcal{X}||\mathcal{Y}|$-sized distribution given by $\tilde{p}(x, y) = p_X(x)/|\mathcal{Y}|$, for all $x \in \mathcal{X}, y \in \mathcal{Y}$.

Proof: Again, this result has been proven for all entropies in $\mathcal{H}_{\texttt{EAVG}}$ [3, Proposition 6]. By Proposition 7, there is a sequence $\{H^i = (\eta_i, F_i)\}_i$ whose limit is H, and the result is obtained by taking the limit as i goes to infinity on the inequality $H^i(X, Y) \leq H^i(\tilde{p})$. ∎

3) Perfect Secrecy: Shannon's perfect secrecy theorem was generalized to symmetric entropies in $\mathcal{H}_{\texttt{EAVG}}$ in [3, Proposition 7], by extending an argument from [24].

Suppose $\mathcal{M}$ is a set of plaintext messages, $\mathcal{C}$ one of ciphertext and $\mathcal{K}$ of keys, and associate with them the r.v.s M, C and K, respectively. A symmetric encryption scheme is a pair of functions (e, d) such that the encryptor $e : \mathcal{M} \times \mathcal{K} \to \mathcal{C}$ associates one ciphertext for each choice of plaintext and key, and the decryptor $d : \mathcal{C} \times \mathcal{K} \to \mathcal{M}$ reverses the process. The pair (e, d) is said to satisfy *perfect secrecy* if C and M are independent r.v., and it satisfies *perfect correctness* if it does not make any decryption errors, that is

$$p(m, k|c) = \begin{cases} 0 & \text{if } d(c, k) \neq m, \\ p(k|c) & \text{if } d(c, k) = m. \end{cases}$$

Proposition 10: Let $H \in \mathcal{Q}$ be symmetric. If the scheme (e, d) satisfies perfect secrecy and correctness, then the following holds: $H(M) \leq H(K)$.

Proof: Invoking Proposition 7, there is a sequence of symmetric $\{H^i = (\eta_i, F_i)\}_i$ whose limit is H. The inequality $H(M) \leq H(K)$. has been proven for all symmetric entropies in $\mathcal{H}_{\texttt{EAVG}}$ [3, Proposition 7]. Therefore, $H^i(M) \leq H^i(K)$, and the result follows by taking the limit $i \to \infty$ in both sides of the inequality. ∎

4) Bounds in terms of probability of error: One of the most celebrated inequalities in information theory is the Fano's inequality, which bounds the conditional entropy $H(X|Y)$ in terms of the probability of error — that is, the probability of an optimal adversary guessing the value of X wrongly given the knowledge of Y.

Fano's inequality does not hold in general for entropies other than Shannon entropy. However, it is possible to obtain the following generalisation.

Let X be a r.v.. The probability of error e_X is, of course, $e_X = 1 - \max_{x\in\mathcal{X}} p_X(x)$. Given a joint r.v. (X,Y), the expected probability of error $\hat{e}$ is $\hat{e} = \sum_y p(y)e_{X|y}$ where $\hat{e}_{X|y} = 1 - \max_{x\in\mathcal{X}} p_{X|y}(x)$.

Proposition 11: (Fano's generalization) Let $H \in \mathcal{Q}$ be symmetric.

1) $H(X) \leq H\left(1-e_X, \frac{e_X}{n-1}, \ldots, \frac{e_X}{n-1}\right)$.
2) $H(X|Y) \leq H\left(1-\hat{e}, \frac{\hat{e}}{n-1}, \ldots, \frac{\hat{e}}{n-1}\right)$.

Proof: Those inequalities have been proven for $\mathcal{H}_{\texttt{EAVG}}$ [3, Proposition 8]. Invoking Proposition 7, there is a sequence of symmetric $\{H^i = (\eta_i, F_i)\}_i$ whose limit is H. By substituting H for H^i and taking the limit as $i \to$ in both sides in inequality 1), and the limit superior in inequality 2), we obtain the result for all symmetric $H \in \mathcal{Q}$. ∎

Proposition 11 is a generalisation of Fano's inequality:

$$H_1(X|Y) \leq H_1(\hat{e}, 1-\hat{e}) + \hat{e}\log(n-1),$$

as it reduces exactly to the above expression when Shannon entropy (H_1) is chosen. In fact, we have

$$\begin{aligned} H_1(X|Y) \leq & H_1(1-\hat{e}, \frac{\hat{e}}{n-1}, \ldots, \frac{\hat{e}}{n-1}) \\ = & (1-\hat{e})\log(\frac{1}{1-\hat{e}}) \\ & + (n-1)\frac{\hat{e}}{n-1}\log(\frac{n-1}{\hat{e}}) \\ = & H_1(\hat{e}, 1-\hat{e}) + \hat{e}\log(n-1). \end{aligned}$$

VII. Other Conditional Forms Derivable from $\mathcal{Q}$

One way of understanding the relationship between the subfamilies $\mathcal{C}_{\texttt{AVG}}$, $\mathcal{H}_{\texttt{EAVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$ to the wider $\mathcal{Q}$ is as a restriction imposed on the conditional form: $\mathcal{C}_{\texttt{AVG}}$ is exactly the subset of entropies of $\mathcal{Q}$ that respect AVG, and the same is true for $\mathcal{H}_{\texttt{EAVG}}$ and $\mathcal{Q}_{\texttt{MIN}}$; and EAVG and MIN, respectively.

A natural question is whether there are any other interesting conditional forms that could be considered by our framework. In this section we answer this question positively, by showing a subfamily of limit entropies whose conditional form generalises the *geometric mean*. As a result of our generalising approach, this new family immediately inherits all properties proven in Section VI.

A. Entropies Satisfying Geometric and η-geometric Mean

Definition 15: Given an entropy H as in Definition 4, we say that it satisfies *geometric mean* (GM) if

$$H(X|Y) = \prod_y \left(H(X|y)\right)^{p(y)}.$$

And, given an entropy $H = (\eta, F)$, we say that it satisfies *η-geometric mean* (EGM) if

$$H(X|Y) = \eta\left(\prod_y \left(F(X|y)\right)^{p(y)}\right).$$

Notice that EGM generalises GM by taking $\eta = \mathrm{id}$ and $H(X) = F(X)$.

We have the following result.

Theorem 8: Let $H = (\eta, F)$ satisfy CCV and EGM and suppose F is nonegative. Then, $H \in \mathcal{Q}$.

Proof: To prove the result, we need to provide a sequence $\{(\eta_i, F_i)\}_i$ of entropies in $\mathcal{H}_{\texttt{EAVG}}$ such that 1) $\{\eta_i \circ F_i\}$ converges uniformly to the unconditional form of H and 2) for all X, Y, we have

$$\begin{aligned} H(X|Y) &= \eta\left(\prod_y \left(F(X|y)\right)^{p(y)}\right) \\ &= \lim_{i\to\infty} \eta_i\left(\sum_{y\in\mathcal{Y}^+} p(y)F_i(X|y)\right). \end{aligned} \tag{9}$$

For each i, let $F_i = F^{1/i}$. Then, for all $p, q \in \Delta_n$ and $\lambda \in [0,1]$, we have

$$\begin{aligned} F_i(\lambda p + (1-\lambda)q) &= \left(F(\lambda p + (1-\lambda)q)\right)^{1/i} \\ &\geq \left(\lambda F(p) + (1-\lambda)F(q)\right)^{1/i} \\ &\geq \left(\lambda F(p)\right)^{1/i} + (1-\lambda)\left(F(q)\right)^{1/i} \\ &= \lambda F_i(p) + (1-\lambda)F_i(q), \end{aligned}$$

where the first inequality follows from concavity of F and $r \mapsto r^{1/i}$ being increasing when r is nonegative, and the second from concavity of $r \mapsto r^{1/i}$ when r is nonegative. Thus F_i is concave.

Define η_i by taking $\eta_i(r) = \eta((r)^i)$. As $r \mapsto r^i$ is increasing for $r \geq 0$, η_i is the composition of two increasing functions, and therefore increasing. Now, $\eta_i(F_i(p)) = \eta(F(p))$, thus $\{\eta_i \circ F_i\}$ trivially converges uniformly to H.

Next, we prove the convergence for the conditional form. First, we introduce the following result from [29, Section 2.3]: let $\lambda_1, \ldots, \lambda_n$ and $p_1, \ldots, p_n$ be nonegative real numbers, and suppose $\sum_i p_i = 1$. Then

$$\lim_{\alpha\to 0}\left(\sum_i p_i\lambda_i^\alpha\right)^{1/\alpha} = \prod_i \lambda_i^{p_i}. \tag{10}$$

Using the result above, we have:

$$\begin{aligned}
&\lim_{i\to\infty} \eta_i \left(\sum_y p(y) F_i(X|y) \right) \\
&= \lim_{i\to\infty} \eta \left(\left(\sum_y p(y) \left(F(X|y)^{1/i} \right) \right)^i \right) \\
&= \eta \left(\lim_{i\to\infty} \left(\sum_y p(y) \left(F(X|y)^{1/i} \right) \right)^i \right) \\
&= \eta \left(\prod_y F(X|y)^{p(y)} \right) = H(X|Y),
\end{aligned}$$

where the third line follows from continuity of η, and the fourth from (10). ∎

Theorem 8 allows us to define two interesting entropy families.

Definition 16: We define $\mathcal{H}^+_{\texttt{EGM}}$ to be the set of all entropies $H = (\eta, F)$ that satisfy EGM, CCV, and for which F is nonegative. Analogously, we define $\mathcal{C}^+_{\texttt{GM}}$ to be the set of all entropies satisfying GM, CV, and that have a nonegative unconditional form.

As a corollary of Theorem 8, it is immediate that $\mathcal{H}^+_{\texttt{EGM}} \subset \mathcal{Q}$ and, by taking $\eta = \mathrm{id}$, we also obtain $\mathcal{C}^+_{\texttt{GM}} \subset \mathcal{Q}$. Therefore, both families satisfy DPI, CRE, and all the properties in Section VI.

As a final remark, we note that the nonegativity requirement over $\mathcal{H}^+_{\texttt{EGM}}$ does not limit the unconditional form of these entropies.

Proposition 12: Given any core-concave function h over Δ_n, there is a entropy in $\mathcal{H}_{\texttt{EGM}}$ whose unconditional form coincides with h.

Proof: Let $h = (\eta, f)$ be a core-concave function, and let $\sigma = \inf_{p\in\Delta_n} f(p)$. Define a entropy $H' = (\eta', F')$ by taking $\eta'(r) = \eta(r + \sigma)$, $F'(p) = f(p) - \sigma$ and $H'(X|Y)$ satisfying EGM. Then, $H' \in \mathcal{H}^+_{\texttt{EGM}}$ and the unconditional form of H' coincides with h. ∎

VIII. Blackwell-Sherman-Stein Theorem for Limit Entropies and Implications to Preorders

Given a pair of channels $K_1 : \mathcal{X} \to \mathcal{Y}_1$, $K_2 : \mathcal{X} \to \mathcal{Y}_2$, one problem of interest in QIF is whether K_2 is robustly *better* than K_1 — that is, whether for all entropies and all distributions on the input, K_1 leaks at least as much information as K_2. Given a family of entropies $\mathcal{A}$, we write $K_1 \geq_{\mathcal{A}} K_2$ whenever the above statement is true for this family. More specifically, $K_1 \geq_{\mathcal{A}} K_2$ means that $H(X|Y_1) \leq H(X|Y_2)$ for all $H \in \mathcal{A}$ and all p_X. This relation defines a *preorder* over channels that share a same input.

An important result both in QIF and in the field of *Statistical Decision Making* is the *Blackwell-Sherman-Stein* (BSS) Theorem [30]–[32]. It connects the preorders defined above to the notion of *degradedness* [33], also known in the QIF community as *refinement* [34]. We say that a channel $K_2 : \mathcal{X} \to \mathcal{Y}_2$ is *degraded from* a channel $K_1 : \mathcal{X} \to \mathcal{Y}_2$ — and write $K_1 \geq_{\mathrm{d}} K_2$ — if there is a channel $W : \mathcal{Y}_1 \to \mathcal{Y}_2$ such that $K_2(y_2|x) = \sum_{y_1} K_1(y_1|x) W(y_2|y_1)$. In words, a channel K_2 is degraded from K_1 when it is the result of *postprocessing* the output of K_1 by feeding it to another channel W.

The BSS Theorem was independently proven in QIF within the g-leakage framework, taking the following form.

Theorem 9 ([34, Theorems 8,9]): Let $\mathcal{H}_G$ be the set of all g-entropies with $\mathrm{Range}(g) \subset [0, 1]$. Then $K_1 \geq_{\mathrm{d}} K_2 \iff K_1 \geq_{\mathcal{H}_G} K_2$.

This result can immediately be generalised for the wider set $\mathcal{Q}$.

Theorem 10 (BSS for $\mathcal{Q}$): $K_1 \geq_{\mathrm{d}} K_2 \iff K_1 \geq_{\mathcal{Q}} K_2$.

Proof: Sufficiency follows from Proposition 6, and necessity follows from Theorem 9, since for all g, $H_g \in \mathcal{Q}$. ∎

As an interesting consequence, for any subset of $\mathcal{Q}$ the degradedness relation is some kind of "minimal order" that implies less information leakage. In particular, we have the following corollary.

Corollary 3: $K_1 \geq_{\mathrm{d}} K_2 \implies K_1 \geq_{\mathcal{Q}_{\texttt{MIN}}} K_2$.

Incidentally, this problem has been recently studied by Chatzikokolakis et al. in [9]. Their Theorem 3 states not only the same result as in Corollary 3, but also proves that the converse is not true.

A. An Insight Into a Recently Observed Property

An intriguing property of the set $\mathcal{Q}_{\texttt{MIN}}$ arose in a proof of a recent paper by Chatzikokolakis et al. [9]. Therein, in the proof of Theorem 2, it is seen that whenever $K_1 \not\geq_{\mathcal{Q}_{\texttt{MIN}}} K_2$ it is always possible to find a concave function F such that

$$\min_{y_1\in\mathcal{Y}_1} F(X|y_1) > \min_{y_2\in\mathcal{Y}_2} F(X|y_2). \tag{11}$$

This is perhaps a unexpected phenomenon: even though the order is predicated on the set of all quasiconcave functions, there is always a *concave* witness whenever the order does not hold. Why should it be the case that whenever channels are ordered with regards to all concave entropies in $\mathcal{Q}_{\texttt{MIN}}$, they are also ordered with regards to the whole set?

Theorem 7 provides us a simple answer to this question. If $K_1 \not\geq_{\mathcal{Q}_{\texttt{MIN}}} K_2$, there is $H \in \mathcal{Q}_{\texttt{MIN}}$ such that

$$\min_{y_1\in\mathcal{Y}_1} H(X|y_1) > \min_{y_2\in\mathcal{Y}_2} H(X|y_2).$$

As $\mathcal{Q}_{\texttt{MIN}} \subset \mathcal{Q}$, H is the limit of a sequence of entropies in $\mathcal{H}_{\texttt{EAVG}}$. Thus, the inequality above guarantees that there is $H' = (\eta', F') \in \mathcal{H}_{\texttt{EAVG}}$ for which

$$\begin{aligned}
&\min_{y_1\in\mathcal{Y}_1} H'(X|y_1) > \min_{y_2\in\mathcal{Y}_2} H'(X|y_2) \\
\iff &\min_{y_1\in\mathcal{Y}_1} \eta'(F'(X|y_1)) > \min_{y_2\in\mathcal{Y}_2} \eta'(F'(X|y_2)) \\
\iff &\min_{y_1\in\mathcal{Y}_1} F'(X|y_1) > \min_{y_2\in\mathcal{Y}_2} F'(X|y_2).
\end{aligned}$$

And thus F' is a concave function satisfying (11).

IX. Related Work

Generalising frameworks have a long history in the information theory literature. Already in 1961, Alfred Rényi proposed the Rényi entropy family [14], which generalises Shannon entropy [17] by relaxing one axiom that characterises it. The axiomatic approach has been particularly fruitful in the field of information theory, and these efforts are neatly summarised in a survey paper by Csiszár [35].

The field of quantitative information flow has been very active in the last couple of decades, and the intricacies of the field implied that many different entropy measures were considered, often in the same work [19], [36], [37]. This prompted the use of generalising frameworks [5], [20], [21], which allows one to reason about security of systems in a more robust sense [34], [38], and design systems that are optimal for a wide range of entropies [5], [39], [40].

As conditional entropies play a fundamental role in QIF, an axiomatic approach to these frameworks was considered by Alvim et al. [2], and later expanded to account for some entropies not used in QIF, but commonly used in the information theory literature [3]. This work builds upon this recent effort by the community, providing a generalising framework for conditional entropies that completely generalises those in [2] and [3], while still being meaningful, in the sense that all entropies in the generalising family respect some desirable properties. Incidentally, this framework also overlap with recent work regarding preorder over channels [9], as discussed in Section VIII.

Finally, the study of core-concave functions — or, more specifically, of which quasiconcave functions can be *concavifiable* — has a vast literature motivated by the field of Microeconomics, wherein quasiconcave functions appear naturally as reasonable tools when studying consumer preference relations [16, Lecture 4]. In 1949, Bruno de Finetti [6] proved that not all quasiconcave functions could be *concavified* — that is, transformed into a concave function by a increasing function. Since then, there has been considerable effort into studying the concavification of preference relations (and hence, of quasiconcave functions) [7], [41]–[44]. For a succinct presentation on the topic, we refer the reader to [12, Chapter 8].

X. Conclusions and Future Work

Several previous axiomatic approaches to information leakage have been here unified in a general framework, that of limit entropies. Limit entropies are based on a limit construction of core-concave entropies satisfying EAVG, and are shown to coincide with quasiconcave functions on the unconditional form, and to subsume different conditional forms in the literature.

Because of the limit construction several information theoretical and information leakage properties can be proven for quasiconcave entropies, perfect secrecy and a generalization of Fano inequality. This guarantees that any other entropies subsumed in this framework, such as the new proposed families of entropies in Section VII with a η-geometric mean conditional form, will satisfy not only the important DPI and CRE properties but also all the information theoretical properties in Section VI.

The limit construction here introduced, based on recent results — yet a long tradition of research — in microeconomics and convex analysis, opens the door for further investigations about possible connections between microeconomics theory, information leakage and information theory.

References

[1] M. S. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith, "Axioms for information leakage," in *Proc. of CSF*, 2016, pp. 77–92.

[2] ——, "An axiomatization of information flow measures," *Theoretical Computer Science*, vol. 777, pp. 32 – 54, 2019, in memory of Maurice Nivat, a founding father of Theoretical Computer Science - Part I.

[3] A. Américo, M. Khouzani, and P. Malacaria, "Conditional entropy and data processing: an axiomatic approach based on core-concavity," *IEEE Transactions on Information Theory*, 2020.

[4] M. Khouzani and P. Malacaria, "Relative perfect secrecy: Universally optimal strategies and channel design," in *IEEE 29th Computer Security Foundations Symposium, CSF 2016, Lisbon, Portugal, June 27 - July 1, 2016*, 2016, pp. 61–76.

[5] ——, "Generalised Entropies and Metric-Invariant Optimal Countermeasures for Information Leakage under Symmetric Constraints," *IEEE Transactions on Information Theory*, 2018.

[6] B. De Finetti, "Sulle stratificazioni convesse," *Annali di Matematica Pura ed Applicata*, vol. 30, no. 1, pp. 173–183, 1949.

[7] C. Connell and E. Rasmusen, "Concavifying the quasiconcave," August 2012. [Online]. Available: https://ssrn.com/abstract=1907180

[8] ——, "Concavifying the quasiconcave," *Journal of Convex Analysis*, vol. 24, no. 4, pp. 1239–1262, 2017.

[9] K. Chatzikokolakis, N. Fernandes, and C. Palamidessi, "Comparing systems: Max-case refinement orders and application to differential privacy," in *2019 IEEE 32nd Computer Security Foundations Symposium (CSF)*, 2019, pp. 442–457.

[10] W. Rudin, *Principles of Mathematical Analysis*. McGraw-hill New York, 1964.

[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[12] M. Avriel, W. E. Diewert, S. Schaible, and I. Zang, *Generalized Concavity*. Society for Industrial and Applied Mathematics, 2010.

[13] S. Karamardian, "Strictly quasi-convex (concave) functions and duality in mathematical programming," *Journal of Mathematical Analysis and Applications*, vol. 20, no. 2, pp. 344 – 358, 1967.

[14] A. Rényi, "On Measures of Entropy and Information," in *Proc. 4th Berkeley Symposium on Mathematics, Statistics, and Probability*, 1961, pp. 547–561.

[15] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of majorization and its applications*. Mathematics In Science And Engineering, Academic Press, 1979, vol. 143.

[16] A. Rubinstein, *Lecture Notes in Microeconomic Theory: The Economic Agent*, 2nd ed. Princeton University Press, 2012.

[17] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 625–56, 1948.

[18] J. L. Massey, "Guessing and entropy," in *Proc. IEEE Int. Symposium on Information Theory (ISIT)*, June 1994, p. 204.

[19] G. Smith, "On the foundations of quantitative information flow," in *Proc. 12th Int. Conf. Foundations of Software Science and Computational Structures (FOSSACS)*, ser. LNCS, vol. 5504. Springer, 2009, pp. 288–302.

[20] M. S. Alvim, K. Chatzikokolakis, C. Palamidessi, and G. Smith, "Measuring information leakage using generalized gain functions," in *Proc. IEEE 25th Computer Security Foundations Symposium (CSF)*, 2012, pp. 265–279.

[21] M. S. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith, "Additive and multiplicative notions of leakage, and their capacities," in *Proc.IEEE 27th Computer Security Foundations Symposium (CSF)*. IEEE, 2014, pp. 308–322.

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. J. Wiley & Sons, Inc., 2006.

[23] A. Teixeira, A. Matos, and L. Antunes, "Conditional rényi entropies," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4273–4277, July 2012.

[24] M. Iwamoto and J. Shikata, "Information theoretic security for encryption based on conditional rényi entropies," in *Information Theoretic Security*, C. Padró, Ed. Cham: Springer International Publishing, 2014, pp. 103–121.

[25] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," *Topics in information theory*, 1977.

[26] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3989–4001, 2011.

[27] C. Tsallis, "Possible generalization of boltzmann-gibbs statistics," *Journal of statistical physics*, vol. 52, no. 1-2, pp. 479–487, 1988.

[28] B. Sharma and D. Mittal, "New non-additive measures of entropy for discrete probability distributions," *Journal of Mathematical Science (Soc. Math. Sci., Calcutta, India*, vol. 10, pp. 28–40, 1975.

[29] G. Hardy, J. Littlewood, K. M. R. Collection, G. Pólya, D. Littlewood, and G. Pólya, *Inequalities*, ser. Cambridge Mathematical Library. Cambridge University Press, 1952.

[30] D. Blackwell, "The comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman, Ed. Berkeley: Univ. of California Press, 1951, pp. 93–102.

[31] S. Sherman, "On a theorem of hardy, littlewood, polya, and blackwell," *Proceedings of the National Academy of Sciences*, vol. 37, no. 12, pp. 826–831, 1951.

[32] D. Blackwell, "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, vol. 24, no. 2, pp. 265–272, 1953.

[33] T. M. Cover, "Broadcast channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 2–14, 1972.

[34] A. McIver, C. Morgan, G. Smith, B. Espinoza, and L. Meinicke, "Abstract channels and their robust information-leakage ordering," in *Proc. 3rd Int. Conf. Principles of Security and Trust (POST)*, ser. LNCS, vol. 8414. Springer, 2014, pp. 83–102.

[35] I. Csiszár, "Axiomatic characterizations of information measures," *Entropy*, vol. 10, no. 3, pp. 261–273, 2008.

[36] P. Malacaria, "Algebraic foundations for quantitative information flow," *Mathematical Structures in Computer Science*, vol. 25, no. 2, p. 404–428, 2015.

[37] B. Köpf and D. Basin, "An information-theoretic model for adaptive side-channel attacks," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, ser. CCS '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 286–296.

[38] D. M. Smith and G. Smith, "Tight bounds on information leakage from repeated independent runs," in *Proc. IEEE 30th Computer Security Foundations Symposium*, Aug 2017, pp. 318–327.

[39] M. Khouzani and P. Malacaria, "Leakage-Minimal Design: Universality, Limitations, and Applications," in *Proc.IEEE 30th Computer Security Foundations Symposium (CSF)*. IEEE, 2017, pp. 305–317.

[40] A. Américo, M. Khouzani, and P. Malacaria, "Deterministic Channel Design for Minimum Leakage," in *Proc. IEEE 32nd Computer Security Foundations Symposium (CSF)*, 2019, pp. 428–441.

[41] Y. Kannai, "Concavifiability and constructions of concave utility functions," *Journal of Mathematical Economics*, vol. 4, no. 1, pp. 1 – 56, 1977.

[42] M. K. Richter and K.-C. Wong, "Concave utility on finite sets," *Journal of Economic Theory*, vol. 115, no. 2, pp. 341 – 357, 2004.

[43] Y. Kannai, "When is individual demand concavifiable?" *Journal of Mathematical Economics*, vol. 40, no. 1, pp. 59 – 69, 2004, aggregation, Equilibrium and Observability in honor of Werner Hildenbrand.

[44] ——, "Remarks concerning concave utility functions on finite sets," *Economic Theory*, vol. 26, no. 2, pp. 333–344, 2005.