



An axiomatization of entropy of capacities on set systems

Aoi Honda, Michel Grabisch

► To cite this version:

Aoi Honda, Michel Grabisch. An axiomatization of entropy of capacities on set systems. European Journal of Operational Research, 2008, 190 (2), pp.526-538. 10.1016/j.ejor.2007.06.033 . hal-00281598

HAL Id: hal-00281598

<https://hal.science/hal-00281598>

Submitted on 23 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An axiomatization of entropy of capacities on set systems

Aoi Honda ^{a,b} Michel Grabisch ^b

^a *Kyushu Institute of Technology*

^b *Université de Paris I — Panthéon-Sorbonne*

Abstract

We present an axiomatization of the entropy of capacities defined on set systems which are not necessarily the whole power set, but satisfy a condition of regularity. This entropy encompasses the definition of Marichal and Roubens for the entropy of capacities. Its axiomatization is in the spirit of the one of Faddeev for the classical Shannon entropy. In addition, we present also an axiomatization of the entropy for capacities proposed by Dukhovny.

Key Words : entropy, capacity, regular set system, generalized Faddeev's axiom

Corresponding author: Aoi Honda, Kyushu Institute of Technology, 680-4, Kawazu, Iizuka 820-8502, Japan, aoi@ces.kyutech.ac.jp

1 Introduction

The notion of entropy of a probability measure, as proposed by Shannon [13], is of primary importance in many application fields, e.g., information theory, operations research, classification, data mining, etc.

It was therefore natural to investigate about the entropy of capacities [1] (also called fuzzy measures [14], non-additive measures [3]), since they are a generalization of probability measures. After a first attempt by Yager [16], Marichal and Roubens [12, 11] proposed a definition having suitable properties, and which can be considered as the generalization of the Shannon entropy. Another attempt was also done by Dukhovny [4], in a different spirit. All these works considered finite universal sets, and the power set as underlying set system. In a previous paper [8], we proposed a definition of entropy able to work with capacities defined on set systems which are not necessarily the whole power set, but satisfy a regularity condition. Our definition encompasses the one of Marichal and Roubens, and also encompasses the Shannon entropy. Our entropy can be applied to capacities on lattices as well.

An important question about entropy is to find suitable axiomatizations, where axioms are enough understandable or natural. There exist many of them for the classical Shannon entropy (see, e.g., [5]). For capacities, to the best of our knowledge, there exists only the work of Kojadinovic et al. [9]. Their axiomatization is however rather complicated, due to the presence of a recursive axiom, whose meaning is hard to grasp.

Our aim in this paper is to present a new axiomatization which is more understandable in our general framework where the set system is not necessarily the whole power set, but a subset of it. Our axiomatization is in the spirit of the one proposed by Faddeev [6] for the Shannon entropy.

The paper is organized as follows. Section 2 introduces notations and necessary materials about set systems. Then Section 3 presents capacities on set systems and the definition of their entropy. The axiomatization of entropy is given in Section 4, and proofs of the main theorems are given in Section 5. Lastly, we give some final remarks on the suitability of regular set systems as a fundamental concept, and some directions for future research in Section 6.

2 Mathematical background

We begin by introducing some notations and definitions.

Throughout this paper, we consider a finite universal set $N := \{1, 2, \dots, n\}$, $n \geq 2$, and 2^N denotes the power set of N . Let us consider $\mathcal{N}$ a subcollection of 2^N which contains $\emptyset$ and N . Then we call $(N, \mathcal{N})$ (or simply $\mathcal{N}$ if no confusion occurs) a *set system*.

A set system endowed with inclusion is a particular case of a *partially ordered set* $(P, \leq)$, i.e., a set P endowed with a partial order (reflexive, an-

tisymmetric and transitive) $\leq$, as for example $(\mathcal{N}, \subseteq)$.

Let $A, B \in \mathcal{N}$. We say that A is *covered* by B , and write $A \prec B$ or $B \succ A$, if $A \subsetneq B$ and $A \subseteq C \subsetneq B$ together with $C \in \mathcal{N}$ imply $C = A$.

Definition 1 (maximal chain of set system) *Let $\mathcal{N}$ be a set system. We call $C = (c_0, c_1, \dots, c_m)$ a maximal chain of $\mathcal{N}$ if C satisfies $\emptyset = c_0 \prec c_1 \prec \dots \prec c_m = N, c_i \in \mathcal{N}, i = 0, \dots, m$.*

We denote the set of all maximal chains of $\mathcal{N}$ by $\mathcal{C}(\mathcal{N})$.

Definition 2 (totally ordered set system) *We say that $(N, \mathcal{N})$ is a totally ordered set system if for any $A, B \in \mathcal{N}$, either $A \subseteq B$ or $A \supsetneq B$.*

If $(N, \mathcal{N})$ is a totally ordered set system, then $|\mathcal{C}(\mathcal{N})| = 1$.

Definition 3 (regular set system [8]) *We say that $(N, \mathcal{N})$ is a regular set system if for any $C \in \mathcal{C}(\mathcal{N})$, the length of C is n , i.e. $|C| = n + 1$.*

Equivalently, $\mathcal{N}$ is a regular set system if and only if $|A \setminus B| = 1$ for any $A, B \in \mathcal{N}$ satisfying $A \succ B$.

3 Entropy of a capacity

Definition 4 (capacity on a set system) *Let $(N, \mathcal{N})$ be a set system. A function $v : \mathcal{N} \rightarrow [0, 1]$ is a capacity on $(N, \mathcal{N})$ if it satisfies $v(\emptyset) = 0, v(N) = 1$ and, for any $A, B \in \mathcal{N}$, $v(A) \leq v(B)$ whenever $A \subseteq B$.*

Let v be a capacity on $(N, \mathcal{N})$. For $C := (c_0, c_1, \dots, c_m) \in \mathcal{C}(\mathcal{N})$, define $p^{v,C}$ by

$$\begin{aligned} p^{v,C} &:= (p_1^{v,C}, p_2^{v,C}, \dots, p_m^{v,C}) \\ &= (v(c_1) - v(c_0), v(c_2) - v(c_1), \dots, v(c_m) - v(c_{m-1})). \end{aligned}$$

Note that $p^{v,C}$ satisfies $p_i^{v,C} \geq 0, i = 1, \dots, m$, and $\sum_{i=1}^m p_i^{v,C} = 1$.

We introduce further concepts about capacities, which will be useful for stating axioms.

Definition 5 (dual capacity) *Let v be a capacity on $(N, \mathcal{N})$. Then the dual capacity of v is defined on $\mathcal{N}^d := \{A \in 2^N \mid A^c \in \mathcal{N}\}$ by $v^d(A) := 1 - v(A^c)$ for any $A \in \mathcal{N}^d$, where $A^c := N \setminus A$.*

A Hasse diagram for $\mathcal{N}^d$ is obtained by turning upside down a Hasse diagram for $\mathcal{N}$ (cf. Fig. 1).

Definition 6 (permutation of v) *Let π be a permutation on N . Then the permutation of v by π is defined on $\pi(\mathcal{N}) := \{\pi(A) \in 2^N \mid A \in \mathcal{N}\}$ by $\pi \circ v(A) := v(\pi^{-1}(A))$.*

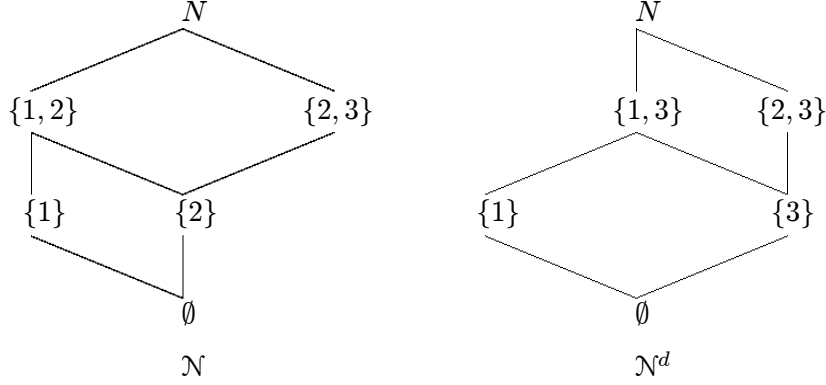


Figure 1: Dual set system

Example 7 We show an example of a permutation of a set system. Let $N := \{1, 2, 3\}$ and $\mathcal{N} := \{\emptyset, \{1\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, N\}$ and let $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Then, for instance $\pi \circ v(\{2, 3\}) := v(\pi^{-1}(\{2, 3\})) = v(\{1, 2\})$ (cf. Fig. 2). For the maximal chain $C := (\emptyset, \{2\}, \{2, 3\}, N) \in \mathcal{C}(\pi(\mathcal{N}))$, $\pi^{-1}(C) = (\pi^{-1}(\emptyset), \pi^{-1}(\{2\}), \pi^{-1}(\{2, 3\}), \pi^{-1}(N)) = (\emptyset, \{1\}, \{1, 2\}, N) \in \mathcal{C}(\mathcal{N})$.

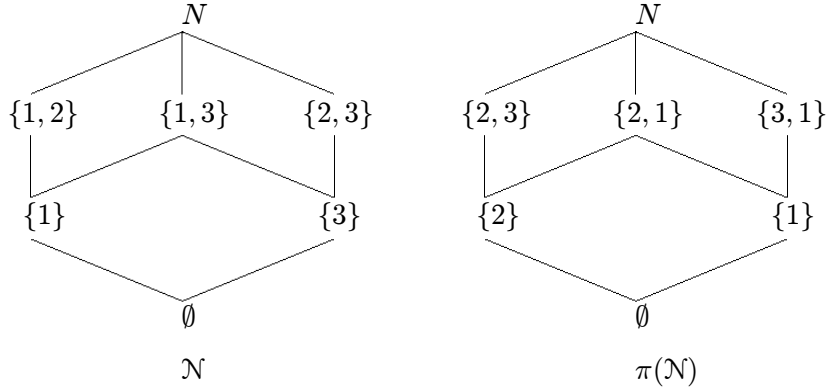


Figure 2: Permutation of set system

As it is usual for functions, we denote the *restriction* of v to some regular set system $\mathcal{N}' \subsetneq \mathcal{N}$ by $v|_{\mathcal{N}'}$, i.e., $v|_{\mathcal{N}'}(A) := v(A)$ for any $A \in \mathcal{N}'$.

Let us consider a chain of length 2 as a set system, denoted by $\mathbf{2}$ (e.g. $\{\emptyset, \{1\}, \{1, 2\}\}$), and a capacity v^2 on it. We denote by the triplet $(0, u, 1)$, $0 \leq u \leq 1$, the values of v^2 along the chain.

Definition 8 (embedding of v^2) Let v be a capacity on a totally ordered regular set system $(N, \mathcal{N})$, where $\mathcal{N} := \{c_0, \dots, c_n\}$ such that $c_{i-1} \prec c_i$, $i = 1, \dots, n$, and let $v^2 := (0, u, 1)$ be a capacity on $\mathbf{2}$. Then for $c_k \in \mathcal{N}$, v^{c_k} is called the embedding of v^2 into v at c_k , and defined on the totally ordered

regular set system $(N^{c_k}, \mathcal{N}^{c_k})$ by

$$v^{c_k}(A) := \begin{cases} v(c_{k-1}) + u \cdot (v(c_k) - v(c_{k-1})), & \text{if } A = c'_k, \\ v(c_j), & \text{if } A = c_j \text{ for } j < k \\ & \text{or } A = c''_j \text{ for } j \geq k \end{cases} \quad (3.1)$$

with $\{i_k\} := c_k \setminus c_{k-1}$, $N^{c_k} := \{1, \dots, i_{k-1}, i'_k, i''_k, i_{k+1}, \dots, n\}$ with $i'_k \neq i''_k$, $(N \setminus \{i_k\}) \cap \{i''_k, i'_k\} = \emptyset$, $c'_k := (c_k \setminus i_k) \cup i'_k$, $c''_j := (c_j \setminus i_k) \cup \{i'_k, i''_k\}$ for $j = k, \dots, n$, and $\mathcal{N}^{c_k} := \{c_0, \dots, c_{k-1}, c'_k, c''_k, c''_{k+1}, \dots, c''_n\}$.

Example 9 Let us take $N := \{1, 2, 3\}$, $\mathcal{N} := \{\emptyset, 1, 12, 123\}$ and $c_k = 12$ (“12” stands for $\{1, 2\}$, etc.). Then $i_k = 2$, which is replaced by $2'$ and $2''$. We obtain

$$\begin{aligned} N^{12} &= \{1, 2', 2'', 3\} \\ \mathcal{N}^{12} &= \{\emptyset, 1, 12', 12'2'', 12'2''3\} \\ v^{12}(1) &= v(1) \\ v^{12}(12') &= v(1) + u \cdot (v(12) - v(1)) \\ v^{12}(12'2'') &= v(12) \\ v^{12}(12'2''3) &= v(123). \end{aligned}$$

We turn now to the definition of the entropy. We first recall the classical definition of Shannon.

Definition 10 (Shannon Entropy) Let $p = (p_1, \dots, p_n)$ be a probability measure on N . Then the Shannon entropy of p is defined by

$$(S) \quad H_S(p) = H_S(p_1, \dots, p_n) := - \sum_{i=1}^n p_i \log p_i,$$

where $\log$ denoting the base 2 logarithm, and by convention $0 \log 0 := 0$.

Definition 11 (entropy of a capacity on a set system [8]) Let v be a capacity on a regular set system $(N, \mathcal{N})$. The entropy of v is defined by

$$(HG) \quad H_{HG}(v) := \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S(p^{v,C}).$$

.

We recall the definition of Dukhovny [4] for capacities defined on the power set, which we extend to our framework. We call it *minimum entropy*.

Definition 12 Let v be a capacity on a regular set system $(N, \mathcal{N})$. The minimum entropy of v is defined by

$$(D) \quad H_D(v) := \min_{C \in \mathcal{C}(\mathcal{N})} H_S(p^{v,C}).$$

As it is clear from its definition, H_D is not differentiable.

Before entering the axiomatization part, a remark on the domain of entropy functions is in order. For mathematical rigor, one should define the Shannon entropy relatively to the cardinality of the universe, hence when necessary we will use the notation H_S^n , with $|N| =: n$. More precisely, let Δ_n be the set of all probability measures on $N = \{1, \dots, n\}$ and algebra 2^N . H_S is a function defined on $\Delta := \bigcup_{n=2}^{\infty} \Delta_n$ to $[0, \infty)$, and H_S^n is its restriction to Δ_n .

Accordingly, we do the same for our definition, up to the fact that H_{HG} depends not only on N but also on the set system $\mathcal{N}$. As it will become clear hereafter, in fact only the number of maximal chains matters. But since in the axiomatization part, this cannot be known in advance, we let as superindex the set system $\mathcal{N}$, and also the cardinality n of the universe, the latter only for clarity since this can be known from the set system: $H_{HG}^{n, \mathcal{N}}$ or $H^{n, \mathcal{N}}$. We denote by $\Lambda_{n, \mathcal{N}}$ the set of all capacities defined on $(N, \mathcal{N})$, and the domain of H_{HG} is $\Lambda := \bigcup_{n=2}^{\infty} \bigcup_{\mathcal{N} \text{ on } N} \Lambda_{n, \mathcal{N}}$.

4 Axiomatization of the entropy of capacities

First, we recall Faddeev's axiomatization, which will serve as a basis for our axiomatization. In what follows, H^n is a function from Δ_n to $[0, \infty)$.

(F1) $f(x) := H^2(p, 1-p)$ is continuous on $0 \leq p \leq 1$, and there exists $p_0 \in [0, 1]$ such that $f(p_0) > 0$.

(F2) For any permutation π on N ,

$$H^n(p_{\pi(1)}, \dots, p_{\pi(n)}) = H^n(p_1, \dots, p_n).$$

(F3) If $p_n = q + r$, $q > 0$, $r > 0$, then

$$H^{n+1}(p_1, \dots, p_{n-1}, q, r) = H^n(p_1, \dots, p_n) + p_n H^2(q/p_n, r/p_n).$$

Theorem 13 (Faddeev's axiomatization of the Shannon entropy [6])
Under the condition $H^2(1/2, 1/2) = 1$, $H^n = H_S^n$ for all $n \geq 2$ holds if and only if (F1), (F2) and (F3) hold.

We introduce five axioms for the entropy of capacities. In what follows, $H^{n, \mathcal{N}}$ is a function from $\Lambda_{n, \mathcal{N}}$ to $[0, \infty)$.

(HG1) (continuity) The function $f(u) := H^{2,2}(0, u, 1)$ is continuous on $[0, 1]$, and there exists $u_0 \in [0, 1]$ such that $f(u_0) > 0$.

(HG2) (dual invariance) For any capacity $(0, u, 1)$ on $\mathbf{2}$,

$$H^{2,2}(0, u, 1) = H^{2,2}(0, 1-u, 1).$$

(HG3) (*increase by embedding*) Let v be a capacity on a totally ordered regular set system $(N, \mathcal{N})$, with $n \geq 2$. Then for any $c_k \in \mathcal{N}$ and any $v^2 := (0, u, 1)$, the entropy of v^{c_k} is

$$H^{n+1, \mathcal{N}^{c_k}}(v^{c_k}) = H^{n, \mathcal{N}}(v) + (v(c_k) - v(c_{k-1})) \cdot H^{2, 2}(0, u, 1). \quad (4.1)$$

(HG4) (*convexity*) Let $(N, \mathcal{N})$ with $n \geq 2$, $(N, \mathcal{N}_1), \dots$, and $(N, \mathcal{N}_k)$ be regular set systems on N , satisfying $\mathcal{C}(\mathcal{N}) = \mathcal{C}(\mathcal{N}_1) \cup \dots \cup \mathcal{C}(\mathcal{N}_k)$, and $\mathcal{C}(\mathcal{N}_i) \cap \mathcal{C}(\mathcal{N}_j) = \emptyset$, for all $i \neq j$. Then there exist unique $\alpha_1, \dots, \alpha_k \in]0, 1[$ satisfying $\sum_{i=1}^k \alpha_i = 1$ such that for any capacity v on $(N, \mathcal{N})$,

$$H^{n, \mathcal{N}}(v) = \alpha_1 H^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}) + \dots + \alpha_k H^{n, \mathcal{N}_k}(v|_{\mathcal{N}_k}).$$

(HG5) (*permutation invariance*) Let v be a capacity on a regular set system $\mathcal{N}$ with $n \geq 2$. Then for any permutation π on N satisfying $\pi(\mathcal{N}) = \mathcal{N}$, it holds that

$$H^{n, \mathcal{N}}(v) = H^{n, \mathcal{N}}(\pi \circ v).$$

Theorem 14 Under the condition $H^{2, 2}(0, \frac{1}{2}, 1) = 1$, $H^{n, \mathcal{N}} = H_{\text{HG}}^{n, \mathcal{N}}$ for $n \geq 2$ and any regular set system $\mathcal{N}$ on N , if and only if (HG1), (HG2), (HG3), (HG4) and (HG5) hold.

(See proof in Section 5)

For additive capacities on 2^N , H_{HG} is equal to the Shannon entropy H_S . We discuss in detail the above axioms, in the light of Faddeev's axioms.

- *continuity*

We have $f(u) = H_{\text{HG}}^{2, 2}(0, u, 1) = H_S^2(p^{(0, u, 1), C}) = H_S^2(u, 1 - u)$, where $C := (\emptyset, \{1\}, \{1, 2\})$. Therefore (HG1) corresponds to (F1).

- *dual invariance*

More generally, even where v is defined on a regular set system which is not necessarily **2**, $H_{\text{HG}}(v)$ is dual invariant.

Proposition 15 (dual invariance) For any capacity v on a regular set system, $H_{\text{HG}}(v^d) = H_{\text{HG}}(v)$ holds.

Proof Let v be a capacity on $\mathcal{N}$. For any $a \in \mathcal{N}$, $(a^c)^c = a$, hence the dual mapping is a bijection from $\mathcal{N}$ to $\mathcal{N}^d$. Then, $C := (c_0, \dots, c_n) \in \mathcal{C}(\mathcal{N})$ if and only if $C^d := (c_n^c, \dots, c_0^c) \in \mathcal{C}(\mathcal{N}^d)$, since $c_i \prec c_{i+1}$ implies $c_i^c \succ c_{i+1}^c$. Hence

$|\mathcal{C}(\mathcal{N})| = |\mathcal{C}(\mathcal{N}^d)|$. In addition, we have for any $C \in \mathcal{C}(\mathcal{N})$,

$$\begin{aligned}
H_S(p^{v,C}) &= \sum_{i=1}^n h[v(c_i) - v(c_{i-1})] \\
&= \sum_{i=1}^n h[(1 - v(c_{i-1})) - (1 - v(c_i))] \\
&= \sum_{i=1}^n h[v^d(c_{i-1}^c) - v^d(c_i^c)] \\
&= H_S(p^{v^d, C^d}),
\end{aligned}$$

where $h(x) := -x \log x$. Therefore

$$\begin{aligned}
H_{\text{HG}}(v^d) &= \frac{1}{|\mathcal{C}(\mathcal{N}^d)|} \sum_{C \in \mathcal{C}(\mathcal{N}^d)} H_S(p^{v^d, C}) \\
&= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C^d \in \mathcal{C}(\mathcal{N})} H_S(p^{v, C^d}) \\
&= H_{\text{HG}}(v). \quad \square
\end{aligned}$$

The Shannon entropy of a probability measure satisfies dual invariance, since a probability measure and its dual measure are identical.

- *increase by embedding*

Let v be a capacity on a totally ordered regular set system $\mathcal{N} = \{C := \{c_0, c_1, \dots, c_n\}\}$, where $c_{i-1} \prec c_i$, $i = 0, \dots, n$, and consider the embedding of $v^2 := (0, u, 1)$ into v at c_k . Then

$$H_{\text{HG}}(v^{c_k}) = H_S(p^{v^{c_k}, C'}),$$

where $C' := (c_0, \dots, c_{k-1}, c'_k, c''_k, c''_{k+1}, \dots, c''_n)$, with $c'_k := (c_k \setminus i_k) \cup i'_k$, $c''_j := (c_j \setminus i_k) \cup \{i'_k, i''_k\}$, $j = k, \dots, n$ (see Def. 7). And by (F3), we have

$$H_S(p^{v^{c_k}, C'}) = H_S(p^{v, C}) + (v(c_k) - v(c_{k-1})) \cdot H_S(u, 1 - u)$$

which can be rewritten as

$$H_{\text{HG}}(v^{c_k}) = H_{\text{HG}}(v) + (v(c_k) - v(c_{k-1})) \cdot H_{\text{HG}}(v^2).$$

This is exactly (HG3).

The embedding of $(0, u, 1)$ at $c_k \in \mathcal{N}$ means that $i_k := c_k \setminus c_{k-1} \in N$ is splitted into i'_k, i''_k . In other words, i_k is the union of i'_k and i''_k . Hence (HG3) means that for capacities on totally ordered regular set systems, if elements of N are splitted, then the entropy increases according to (4.1).

- *convexity*

This axiom says that the computation of the entropy follows the structure of the set system, whatever the capacity is. If the (regular) set system can be splitted into several (regular) subsystems, provided all elements of N are present in both, and that no redundancy occurs (no common maximal chain), then the computation can also be splitted into these subsystems, each corresponding to the entropy computed on a subsystem. It is important to note that the way of splitting the computation does not depend on the capacity but only on the splitting, i.e., the α_i 's are independent of the capacity.

- *permutation invariance*

(HG5) corresponds to (F2). Note that there are set systems such that no permutation makes them invariant.

To finish this section, we consider a modification of our axiomatization so as to recover the entropy defined by Dukhovny [4]. We modify (HG4) as follows:

(HG4') *Let $(N, \mathcal{N})$, $(N, \mathcal{N}_1), \dots, (N, \mathcal{N}_k)$ be regular set systems satisfying $\mathcal{C}(\mathcal{N}) = \mathcal{C}(\mathcal{N}_1) \cup \dots \cup \mathcal{C}(\mathcal{N}_k)$. Then for any capacity v on $\mathcal{N}$,*

$$H^{n, \mathcal{N}}(v) = \min \{ H^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}), \dots, H^{n, \mathcal{N}_k}(v|_{\mathcal{N}_k}) \}.$$

Theorem 16 *Under the condition $H^{2,2}(0, \frac{1}{2}, 1) = 1$, $H^{n, \mathcal{N}} = H_D^{n, \mathcal{N}}$ for all $n \geq 2$ and all regular set systems $\mathcal{N}$ on N holds if and only if (HG1), (HG2), (HG3) and (HG4') hold.*

(see proof in Sec. 5)

5 Proof of Theorems 14 and 16

Proof of Theorem 14 (necessity) We have $H_{\text{HG}}^{2,2}(0, u, 1) = -u \log u - (1-u) \log(1-u)$, hence obviously H_{HG} satisfies (HG1).

By Proposition 15, H_{HG} satisfies (HG2).

We show that H_{HG} satisfies (HG3). Let v be a capacity on $\mathcal{N} := \{C := c_0, \dots, c_n\}$ such that $c_{i-1} \prec c_i$, $i = 1, \dots, n$. Then the capacity v^{c_k} which is embedding of $(0, u, 1)$ into v at c_k is defined on $\mathcal{N}^{c_k} = \{C' := \{c_0 \prec \dots \prec$

$c_{k-1} \prec c'_k \prec c''_k \prec \dots \prec c''_n\}\}$, with notations of Def. 7. Then we have

$$\begin{aligned}
H_{\text{HG}}^{n+1, \mathcal{N}^{c_k}}(v^{c_k}) &= H_S(p^{v^{c_k}, C'}) \\
&= \sum_{i=1}^{k-1} h[v^{c_k}(c_i) - v^{c_k}(c_{i-1})] + h[v^{c_k}(c'_k) - v^{c_k}(c_{k-1})] \\
&\quad + h[v^{c_k}(c''_k) - v^{c_k}(c'_k)] + \sum_{i=k+1}^n h[v^{c_k}(c''_i) - v^{c_k}(c''_{i-1})] \\
&= \sum_{i \neq k} h[v(c_i) - v(c_{i-1})] \\
&\quad + (v(c_k) - v(c_{k-1})) \left(h \left[\frac{v^{c_k}(c'_k) - v^{c_k}(c_{k-1})}{v(c_k) - v(c_{k-1})} \right] + h \left[\frac{v^{c_k}(c''_k) - v^{c_k}(c'_k)}{v(c_k) - v(c_{k-1})} \right] \right) \\
&\quad + h[v^{c_k}(c''_k) - v^{c_k}(c_{k-1})] \\
&= \sum_{i=1}^n h[v(c_i) - v(c_{i-1})] \\
&\quad + (v(c_k) - v(c_{k-1})) H_{\text{HG}}^{2,2} \left(0, \frac{v^{c_k}(c'_k) - v^{c_k}(c_{k-1})}{v(c_k) - v(c_{k-1})}, 1 \right) \\
&= H_{\text{HG}}^{n, \mathcal{N}}(v) + (v(c_k) - v(c_{k-1})) H_{\text{HG}}^{2,2}(0, u, 1),
\end{aligned}$$

where $u = (v^{c_k}(c'_k) - v^{c_k}(c_{k-1})) / (v(c_k) - v(c_{k-1}))$.

We show that H_{HG} satisfies (HG4). We have, for set systems $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_k$ satisfying conditions of (HG4):

$$\begin{aligned}
H_{\text{HG}}^{n, \mathcal{N}}(v) &= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S^n(p^{v, C}) \\
&= \frac{1}{|\mathcal{C}(\mathcal{N})|} \left(\sum_{C \in \mathcal{C}(\mathcal{N}_1)} H_S^n(p^{v, C}) + \dots + \sum_{C \in \mathcal{C}(\mathcal{N}_k)} H_S^n(p^{v, C}) \right) \\
&= \frac{|\mathcal{C}(\mathcal{N}_1)|}{|\mathcal{C}(\mathcal{N})|} \frac{1}{|\mathcal{C}(\mathcal{N}_1)|} \sum_{C \in \mathcal{C}(\mathcal{N}_1)} H_S^n(p^{v|_{\mathcal{N}_1}, C}) \\
&\quad + \dots + \frac{|\mathcal{C}(\mathcal{N}_k)|}{|\mathcal{C}(\mathcal{N})|} \frac{1}{|\mathcal{C}(\mathcal{N}_k)|} \sum_{C \in \mathcal{C}(\mathcal{N}_k)} H_S^n(p^{v|_{\mathcal{N}_k}, C}) \\
&= \frac{|\mathcal{C}(\mathcal{N}_1)|}{|\mathcal{C}(\mathcal{N})|} H_{\text{HG}}^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}) + \dots + \frac{|\mathcal{C}(\mathcal{N}_k)|}{|\mathcal{C}(\mathcal{N})|} H_{\text{HG}}^{n, \mathcal{N}_k}(v|_{\mathcal{N}_k})
\end{aligned}$$

and

$$\frac{|\mathcal{C}(\mathcal{N}_1)|}{|\mathcal{C}(\mathcal{N})|} + \dots + \frac{|\mathcal{C}(\mathcal{N}_k)|}{|\mathcal{C}(\mathcal{N})|} = 1.$$

We show that H_{HG} satisfies (HG5). For any $C = (c_0, c_1, \dots, c_m) \in \mathcal{C}(\pi(\mathcal{N}))$, $\pi(C) := (\pi(c_0), \pi(c_1), \dots, \pi(c_m)) \in \mathcal{C}(\mathcal{N})$, because for any $A, B \in \pi(\mathcal{N})$, $A \prec B$ implies $\pi(A) \prec \pi(B)$, and for any $C_i, C_j \in \mathcal{C}(\mathcal{N})$, $i \neq j$,

$\pi(C_i) \neq \pi(C_j)$. Consequently, we have

$$\begin{aligned}
H_{\text{HG}}^{n,\pi(\mathcal{N})}(\pi \circ v) &= \frac{1}{|\mathcal{C}(\pi(\mathcal{N}))|} \sum_{C \in \mathcal{C}(\pi(\mathcal{N}))} H_{\text{S}}^n(p^{\pi \circ v, C}) \\
&= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_{\text{S}}^n(p^{\pi \circ v, \pi(C)}) \\
&= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_{\text{S}}^n(p^{v, C}) = H_{\text{HG}}^{n,\mathcal{N}}(v).
\end{aligned}$$

(sufficiency) The proof goes as follows. In part (i), we show that (HG) holds for any totally ordered regular set system C , i.e., we show that $H^{n,C} = H_{\text{HG}}^{n,C}$ for any $n \geq 2$. In part (ii), we extend the result to any regular set system, i.e., $H^{n,\mathcal{N}} = H_{\text{HG}}^{n,\mathcal{N}}$ on Λ .

(i) We show by induction on n that for any v on a totally ordered regular set system (N, C) , C being a maximal chain:

$$H^{n,C}(v) = H_{\text{HG}}^{n,C} = H_{\text{S}}^n(p^{v,C}). \quad (5.1)$$

First, we show the result for $n = 2$, using (HG1), (HG2) and (HG3), in a way similar to the Faddeev's axiomatization for the Shannon entropy. Suppose that $v^{\mathbf{3}}$ is a capacity on $(N := \{1, 2, 3\}, \mathbf{3} := \{\emptyset, \{1\}, \{1, 2\}, N\})$ and $v^{\mathbf{3}}(1) := s, v^{\mathbf{3}}(\{1, 2\}) := s + t, v^{\mathbf{3}}(N) := s + t + u = 1, s, u \geq 0, t > 0$. We can regard $v^{\mathbf{3}}$ as the embedding of $v^{\mathbf{2}} = (0, \frac{t}{1-s}, 1)$ into $v = (0, s, 1)$ on $\mathcal{N} := \{\emptyset, \{1\}, \{1, 2\}\}$ at $\{1, 2\}$ (putting $i'_k = 2, i''_k = 3$ with notations of Def. 7). Then, by (HG3), we get

$$\begin{aligned}
H^{3,\mathbf{3}}(v^{\mathbf{3}}) &= H^{2,\mathcal{N}}(0, s, 1) + (1-s)H^{2,\mathbf{2}}(0, t/(1-s), 1) \\
&= H^{2,\mathcal{N}}(0, s, 1) + (1-s)H^{2,\mathbf{2}}(0, (1-s-t)/(1-s), 1) \\
&= H^{2,\mathcal{N}}(0, s, 1) + (1-s)H^{2,\mathbf{2}}(0, u/(1-s), 1), \quad (5.2)
\end{aligned}$$

where the second equality follows from (HG2) and $s + t + u = 1$. Similarly, we can also regard $v^{\mathbf{3}}$ as the embedding of $v^{\mathbf{2}} = (0, \frac{s}{s+t}, 1)$ into $v = (0, s+t, 1)$ on $\mathcal{N}' := \{\emptyset, \{1\}, \{1, 3\}\}$ at $\{1\}$ (putting $i'_k = 1, i''_k = 2$). This time, we obtain by (HG3) and $s + t + u = 1$,

$$\begin{aligned}
H^{3,\mathbf{3}}(v^{\mathbf{3}}) &= H^{2,\mathcal{N}'}(0, s+t, 1) + (s+t)H^{2,\mathbf{2}}(0, s/(s+t), 1) \\
&= H^{2,\mathcal{N}'}(0, u, 1) + (1-u)H^{2,\mathbf{2}}(0, s/(1-u), 1). \quad (5.3)
\end{aligned}$$

Hence (5.2) and (5.3) yield

$$\begin{aligned}
&H^{2,\mathcal{N}}(0, s, 1) + (1-s)H^{2,\mathbf{2}}(0, u/(1-s), 1) \\
&= H^{2,\mathcal{N}'}(0, u, 1) + (1-u)H^{2,\mathbf{2}}(0, s/(1-u), 1).
\end{aligned}$$

Since regular set systems $\mathcal{N}, \mathcal{N}'$ and $\mathbf{2}$ are isomorphic, putting $f(s) := H^{2,\mathbf{2}}(0, s, 1)$, we obtain

$$f(s) + (1-s)f(u/(1-s)) = f(u) + (1-u)f(s/(1-u)), \quad (5.4)$$

where the above expression is valid for any $s, u \in [0, 1)$ satisfying $s + u < 1$. If $s = 0, u > 0$, then

$$f(0) + f(u) = f(u) + (1 - u)f(0),$$

hence we obtain

$$f(0) = H^{2,2}(0, 0, 1) = H^{2,2}(0, 1, 1) = 0.$$

Integrating both sides of (5.4) for u from 0 to $1 - s$, where $0 \leq s < 1$, we obtain

$$\begin{aligned} (1 - s)f(s) + (1 - s)^2 \int_0^1 f(x)dx \\ = \int_0^{1-s} f(x)dx + s^2 \int_s^1 x^{-3} f(x)dx. \end{aligned} \quad (5.5)$$

By (HG1), $f(s)$ is continuous on $[0, 1]$, hence all terms except the first one of the left side of (5.5) are differentiable on $0 < s < 1$, so that $f(s)$ is also differentiable on $0 < s < 1$. Therefore differentiating (5.5) with respect to s , we have

$$\begin{aligned} (1 - s)f'(s) - f(s) - 2(1 - s) \int_0^1 f(x)dx \\ = -f(1 - s) + 2s \int_s^1 x^{-3} f(x)dx - \frac{f(s)}{s}. \end{aligned}$$

Since $f(s) = f(1 - s)$,

$$(1 - s)f'(s) = 2(1 - s) \int_0^1 f(x)dx + 2s \int_s^1 x^{-3} f(x)dx - \frac{f(s)}{s}. \quad (5.6)$$

By the same arguments as above, $f'(s)$ is differentiable on $0 < s < 1$, then differentiating (5.6) by s , we have for $0 < s < 1$,

$$f''(s) = \frac{-2}{s(1 - s)} \int_0^1 f(x)dx. \quad (5.7)$$

Integrating (5.7) twice, we obtain for $0 \leq s \leq 1$,

$$f(s) = -s \log s - (1 - s) \log(1 - s)$$

by $f(0) = 0$ and $f(1/2) = H^{2,2}(0, 1/2, 1) = 1$. Therefore for any v^2 on $\mathbf{2}$,

$$H^{2,2}(v^2) = H_S^2(p^{v^2, C}) = H_{\text{HG}}^{2,2}(v^2).$$

Suppose that (5.1) holds till n and show that it still holds for $n + 1$. Let v be a capacity on a maximal chain $C := (c_0, \dots, c_n)$. Then by (HG3), we

have, denoting by C' the chain after embedding:

$$\begin{aligned}
H^{n+1,C'}(v^{c_k}) &= H^{n,C}(v) + (v(c_k) - v(c_{k-1}))H^{2,2}(v^2) \\
&= H_S^n(p^{v,C}) + (v(c_k) - v(c_{k-1}))H^{2,2}(0, u, 1) \\
&= H_S^n(p^{v,C}) + (v(c_k) - v(c_{k-1}))H^{2,2}\left(0, \frac{v^{c_k}(c'_k) - v^{c_k}(c_{k-1})}{v(c_k) - v(c_{k-1})}, 1\right) \\
&= H_S^n(p^{v,C}) + (v(c_k) - v(c_{k-1})) \\
&\quad H_S^2\left(\frac{v^{c_k}(c'_k) - v^{c_k}(c_{k-1})}{v(c_k) - v(c_{k-1})}, \frac{v^{c_k}(c''_k) - v^{c_k}(c'_k)}{v(c_k) - v(c_{k-1})}\right) \\
&= H_S^{n+1}(p^{v^{c_k}, C'}) = H_{\text{HG}}^{n+1, C'}(v^{c_k})
\end{aligned}$$

by (3.1) and (F3), which means (5.1) holds for v^{c_k} , so that (HG) holds for any capacities on totally ordered regular set systems. From now on, the expression of the entropy on totally ordered regular set systems of size n will be denoted by $H^{n, \mathbf{n}}$.

(ii) We show that for a regular set system which is not totally ordered, (HG) holds. 2^N has $n!$ maximal chains $C_1, \dots, C_{n!}$. By (HG4), there exist unique $\alpha_{C_1}, \dots, \alpha_{C_{n!}} > 0$, $\alpha_{C_1} + \dots + \alpha_{C_{n!}} = 1$, such that for any capacity v on 2^N it holds

$$H^{n, 2^N}(v) = \alpha_{C_1} H^{n, \mathbf{n}}(v|_{C_1}) + \dots + \alpha_{C_{n!}} H^{n, \mathbf{n}}(v|_{C_{n!}}). \quad (5.8)$$

Let $u_1, \dots, u_n > 0$, and put $H_S(u) := u \log u + (1 - u) \log(1 - u)$. First we define a capacity ν on 2^N by

$$\nu(A) := \begin{cases} 0, & |A| < 1, \\ u_i, & A = \{i\}, i = 1, \dots, n, \\ 1, & |A| > 1. \end{cases}$$

Then

$$H^{n, 2^N}(\nu) = \left(\sum_{C \ni \{1\}} \alpha_C \right) H_S^n(u_1) + \dots + \left(\sum_{C \ni \{n\}} \alpha_C \right) H_S^n(u_n).$$

For any permutation π on N and any chain C_i , we have $(\pi \circ v)|_{C_i} = v|_{\pi^{-1}(C_i)}$, hence

$$\begin{aligned}
H^{n, 2^N}(\pi \circ \nu) &= \alpha_{C_1} H^{n, \mathbf{n}}((\pi \circ \nu)|_{C_1}) + \dots + \alpha_{C_{n!}} H^{n, \mathbf{n}}((\pi \circ \nu)|_{C_{n!}}) \\
&= \alpha_{C_1} H^{n, \mathbf{n}}(\nu|_{\pi^{-1}(C_1)}) + \dots + \alpha_{C_{n!}} H^{n, \mathbf{n}}(\nu|_{\pi^{-1}(C_{n!})}) \\
&= \left(\sum_{C \ni \{1\}} \alpha_C \right) H_S^n(u_{\pi^{-1}(1)}) + \dots + \left(\sum_{C \ni \{n\}} \alpha_C \right) H_S^n(u_{\pi^{-1}(n)}).
\end{aligned}$$

By (HG5), $H^{n, 2^N}(\nu) = H^{n, 2^N}(\pi \circ \nu)$ holds for any permutation π on N , and any $u_1, \dots, u_n > 0$, so that we obtain

$$\left(\sum_{C \ni \{1\}} \alpha_C \right) = \dots = \left(\sum_{C \ni \{n\}} \alpha_C \right) = \frac{1}{n}.$$

Next we consider the set system $\mathcal{N}_1 := \{A \in 2^N \mid A \subsetneq \{1\}, A \supsetneq \{1\} \text{ or } A = \{1\}\}$, which is a regular set system and has $(n-1)!$ maximal chains. Without loss of generality, we assume $\mathcal{C}(\mathcal{N}_1) = \{C_1, \dots, C_{(n-1)!}\} \subsetneq \mathcal{C}(2^N)$. Then by (HG4), there exist unique $\beta_{C_1}, \dots, \beta_{C_{(n-1)!}} > 0$, $\beta_{C_1} + \dots + \beta_{C_{(n-1)!}} = 1$, such that for any capacity v on 2^N

$$H^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}) = \beta_{C_1} H^{n, \mathbf{n}}(v|_{C_1}) + \dots + \beta_{C_{(n-1)!}} H^{n, \mathbf{n}}(v|_{C_{(n-1)!}}). \quad (5.9)$$

Since $\mathcal{C}(2^N) = \mathcal{C}(\mathcal{N}_1) \cup \{C_{(n-1)!+1}\} \cup \dots \cup \{C_n\}$ and $\mathcal{C}(\mathcal{N}_1) \cap \{C_{(n-1)!+1}\} \cap \dots \cap \{C_n\} = \emptyset$, by applying (HG4) again, there exist $\beta, \beta_{C_{(n-1)!+1}}, \dots, \beta_{C_n!}$, $\beta + \beta_{C_{(n-1)!+1}} + \dots + \beta_{C_n!}$ such that for any capacity v on 2^N

$$H^{n, 2^N}(v) = \beta H^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}) + \beta_{C_{(n-1)!+1}} H^{n, \mathbf{n}}(v|_{C_{(n-1)!+1}}) + \dots + \beta_{C_n!} H^{n, \mathbf{n}}(v|_{C_n!}).$$

Substituting (5.9) in the above and comparing with (5.8), we get in particular $\alpha_{C_i} = \beta \beta_{C_i}$ for $i = 1, \dots, (n-1)!$. Since $C_i \ni \{1\}, i = 1, \dots, (n-1)!$,

$$\sum_{i=1}^{(n-1)!} \beta \beta_{C_i} = \sum_{i=1}^{(n-1)!} \alpha_{C_i} = \sum_{C \ni \{1\}} \alpha_C = \frac{1}{n},$$

which implies $\beta = \frac{1}{n}$. Let us take ν on 2^N such that

$$\nu|_{\mathcal{N}_1}(A) := \begin{cases} 0, & |A| < 2, \\ u_i, & A = \{1, i\}, i = 2, \dots, n, \\ 1, & |A| > 2. \end{cases}$$

Then

$$H^{n, \mathcal{N}_1}(\nu|_{\mathcal{N}_1}) = \left(\sum_{C \ni \{1\}, \{1, 2\}} \beta_C \right) H_S^n(u_2) + \dots + \left(\sum_{C \ni \{1\}, \{1, n\}} \beta_C \right) H_S^n(u_n),$$

and by (HG5), $H^{n, \mathcal{N}_1}(\nu|_{\mathcal{N}_1}) = H^{n, \pi(\mathcal{N}_1)}(\pi \circ (\nu|_{\mathcal{N}_1}))$ holds for any permutation π on N satisfying $\pi(\mathcal{N}_1) = \mathcal{N}_1$ (i.e., permutations on $N \setminus \{1\}$), so that we obtain

$$\sum_{C \ni \{1\}, \{1, 2\}} \beta_C = \dots = \sum_{C \ni \{1\}, \{1, n\}} \beta_C = \frac{1}{n-1}, \quad (5.10)$$

which implies for $i = 2, \dots, n$,

$$\sum_{C \ni \{1\}, \{1, i\}} \alpha_C = \sum_{C \ni \{1\}, \{1, i\}} \beta \beta_C = \frac{1}{n(n-1)}.$$

Applying the same method for $\mathcal{N}_i := \{A \in 2^X \mid A \subsetneq \{i\}, A \supsetneq \{i\} \text{ or } A = \{i\}\}$, $i = 2, \dots, n$, and

$$\nu|_{\mathcal{N}_i}(A) := \begin{cases} 0, & |A| < 2, \\ u_j, & A = \{i, j\}, i \neq j, \\ 1, & |A| > 2, \end{cases}$$

we obtain for every i and j , $i \neq j$,

$$\sum_{C \ni \{i\}, \{i,j\}} \alpha_C = \sum_{C \ni \{i\}, \{i,j\}} \beta \beta_C = \frac{1}{n(n-1)}.$$

Next, we consider the set system $\mathcal{N}_{1,2} := \{A \in 2^X \mid A \subsetneq \{1\}, A \supsetneq \{1,2\} \text{ or } A = \{1\}, \{1,2\}\}$, and we assume w.l.o.g. that $\mathcal{C}(\mathcal{N}_{1,2}) := \{C_1, \dots, C_{(n-2)!}\}$. Applying (HG4) twice, there exist unique positive $\gamma_{C_1}, \dots, \gamma_{C_{(n-2)!}}, \gamma, \gamma_{C_{(n-2)!+1}}, \dots, \gamma_{C_{(n-1)!}}, \gamma_{C_1} + \dots + \gamma_{C_{(n-2)!}} = 1$, and $\gamma + \gamma_{C_{(n-2)!+1}} + \dots + \gamma_{C_{(n-1)!}} = 1$, such that for any capacity v on 2^N it holds

$$\begin{aligned} H^{n, \mathcal{N}_{1,2}}(v|_{\mathcal{N}_{1,2}}) &= \gamma_{C_1} H^{n, \mathbf{n}}(v|_{C_1}) + \dots + \gamma_{C_{(n-2)!}} H^{n, \mathbf{n}}(v|_{C_{(n-2)!}}) \\ H^{n, \mathcal{N}_1}(v|_{\mathcal{N}_1}) &= \gamma H^{n, \mathcal{N}_{1,2}}(v|_{\mathcal{N}_{1,2}}) + \gamma_{C_{(n-2)!+1}} H^{n, \mathbf{n}}(v|_{C_{(n-2)!+1}}) + \dots \\ &\quad + \gamma_{C_{(n-1)!}} H^{n, \mathbf{n}}(v|_{C_{(n-1)!}}). \end{aligned}$$

Substituting the first equation into the second and comparing with (5.9), we get in particular $\beta_{C_i} = \gamma \gamma_{C_i}$, hence $\alpha_{C_i} = \beta \gamma \gamma_{C_i}$, for $i = 1, \dots, (n-2)!$. Taking a capacity ν on 2^N such that

$$\nu|_{\mathcal{N}_{1,2}}(A) := \begin{cases} 0, & |A| < 3, \\ u_i, & A = \{1, 2, i\}, i = 3, \dots, n, \\ 1, & |A| > 3, \end{cases}$$

we have

$$H^{n, \mathcal{N}_{1,2}}(\nu|_{\mathcal{N}_{1,2}}) = \left(\sum_{\substack{C \ni \{1\}, \{1,2\}, \\ \{1,2,3\}}} \gamma_C \right) H_S^n(u_3) + \dots + \left(\sum_{\substack{C \ni \{1\}, \{1,2\}, \\ \{1,2,n\}}} \gamma_C \right) H_S^n(u_n).$$

By (HG5), using π on N satisfying $\pi(\mathcal{N}_{1,2}) = \mathcal{N}_{1,2}$ (i.e., the permutations on $N \setminus \{1, 2\}$), we obtain

$$\sum_{\substack{C \ni \{1\}, \{1,2\}, \\ \{1,2,3\}}} \gamma_C = \dots = \sum_{\substack{C \ni \{1\}, \{1,2\}, \\ \{1,2,n\}}} \gamma_C = \frac{1}{n-2},$$

which implies for $i = 3, \dots, n$, and using (5.10), we have

$$\sum_{\substack{C \ni \{1\}, \{1,2\}, \\ \{1,2,i\}}} \alpha_C = \frac{1}{n(n-1)(n-2)}.$$

Applying the same method for $\mathcal{N}_{i,j} := \{A \in 2^N \mid A \subsetneq \{i\}, A \supsetneq \{i, j\} \text{ or } A = \{i\}, \{i, j\}\}, i \neq j$, we obtain for every i, j, k satisfying $i \neq j, i \neq k, j \neq k$,

$$\sum_{\substack{C \ni \{i\}, \{i,j\} \\ \{i,j,k\}}} \alpha_C = \sum_{\substack{C \ni \{i\}, \{i,j\} \\ \{i,j,k\}}} \alpha \beta \gamma_C = \frac{1}{n(n-1)(n-2)}.$$

Repeating a similar process up to $\mathcal{N}_{i_1, \dots, i_{n-1}} := \{A \in 2^N \mid A \subsetneq \{i_1\}, A \supsetneq \{i_1, \dots, i_{n-1}\} \text{ or } A = \{i_1\}, \{i_1, i_2\}, \dots, \{i_1, \dots, i_{n-1}\}\}$, we finally obtain $\alpha_{C_i} = \frac{1}{n!}$, $i = 1, \dots, n!$, that is,

$$H^{n, 2^N}(v) = \frac{1}{n!} \sum_{C \in \mathcal{C}(2^N)} H^{n, \mathbf{n}}(v|_C).$$

Consider a regular set system $\mathcal{N} \subsetneq 2^N$. We can write 2^N as the union of all its $n!$ maximal chains:

$$2^N = C_1 \cup C_2 \cup \dots \cup C_{n!}.$$

For commodity, let us call $C_1, \dots, C_k$ the chains which are not in $\mathcal{N}$, hence

$$\mathcal{C}(2^N) = \{C_1\} \cup \dots \cup \{C_k\} \cup \mathcal{C}(\mathcal{N}).$$

and

$$\{C_1\} \cap \dots \cap \{C_k\} \cap \mathcal{C}(\mathcal{N}) = \emptyset.$$

By (HG4), we can find unique positive $\alpha_1, \dots, \alpha_k, \alpha_{k+1}$ with $\sum_{i=1}^{k+1} \alpha_i = 1$, such that

$$H^{n, 2^N}(v) = \alpha_1 H^{n, 1}(v|_{C_1}) + \dots + \alpha_k H^{n, 1}(v|_{C_k}) + \alpha_{k+1} H^{n, \mathcal{N}}(v|_{\mathcal{N}}).$$

Similarly, it exists unique positive $\beta_{k+1}, \dots, \beta_{n!}$ with $\sum_{i=k+1}^{n!} \beta_i = 1$, such that

$$H^{n, \mathcal{N}}(v|_{\mathcal{N}}) = \beta_{k+1} H^{n, 1}(v|_{C_{k+1}}) + \dots + \beta_{n!} H^{n, 1}(v|_{C_{n!}}).$$

Substituting in the above leads to

$$\begin{aligned} H^{n, 2^N}(v) &= \alpha'_1 H^{n, 1}(v|_{C_1}) + \dots + \alpha'_k H^{n, 1}(v|_{C_k}) \\ &\quad + \alpha'_{k+1} H^{n, 1}(v|_{C_{k+1}}) + \dots + \alpha'_{n!} H^{n, 1}(v|_{C_{n!}}), \end{aligned}$$

with $\alpha'_i = \alpha_i$ for $i = 1, \dots, k$, and $\alpha'_i = \alpha_{k+1} \beta_i$ for $i = k+1, \dots, n!$. We have proved that $\alpha'_1 = \alpha'_2 = \dots = \alpha'_{n!} = \frac{1}{n!}$, for $i = 1, \dots, n!$. It follows that $\frac{k}{n!} + \alpha_{k+1} = 1$, hence

$$\alpha_{k+1} = \frac{n! - k}{n!} = \frac{|\mathcal{C}(\mathcal{N})|}{n!}$$

by definition of $C_{k+1}, \dots, C_{n!}$. Now, $\alpha_{k+1} \beta_i = \frac{1}{n!}$ for $i = k+1, \dots, n!$, hence $\beta_i = \frac{1}{|\mathcal{C}(\mathcal{N})|}$, for $i = k+1, \dots, n!$. Finally we get

$$\begin{aligned} H^{n, \mathcal{N}}(v|_{\mathcal{N}}) &= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{i=k+1}^{n!} H^{n, 1}(v|_{C_i}) \\ &= \frac{1}{|\mathcal{C}(\mathcal{N})|} \sum_{C \in \mathcal{C}(\mathcal{N})} H_S(p^{v, C}) \end{aligned}$$

using (5.1). □

Proof of Theorem 16 (necessity) When v is a capacity on a totally ordered regular set system, $H_D(v) = H_{HG}(v)$ holds. Hence (D) implies (HG1), (HG2) and (HG3) by Theorem 14.

We show that (D) satisfies (HG4'). We have, with set systems $\mathcal{N}_1, \dots, \mathcal{N}_k$ satisfying conditions of (HG4'),

$$\begin{aligned} H_D^{n,\mathcal{N}}(v) &= \min_{C \in \mathcal{C}(\mathcal{N})} H_S^n(p^{v,C}) \\ &= \min \left\{ \min_{C \in \mathcal{C}(\mathcal{N}_1)} H_S^n(p^{v,C}), \dots, \min_{C \in \mathcal{C}(\mathcal{N}_k)} H_S^n(p^{v,C}) \right\} \\ &= \min \left\{ H_D^{n,\mathcal{N}_1}(v_1), \dots, H_D^{n,\mathcal{N}_k}(v_k) \right\} \end{aligned}$$

with usual notations.

(sufficiency) When v is a capacity on a totally ordered regular set system, $H_D(v) = H_{HG}(v)$ holds. Hence (HG1), (HG2) and (HG3) imply (D) by the same way of the proof of Theorem 14.

Next, let v be a capacity on $\mathcal{N}$ which is regular and not totally ordered, and write $\mathcal{N}$ as the union of all its maximal chains: $\mathcal{N} = C_1 \cup \dots \cup C_k$. Then by application of (HG4') we have

$$\begin{aligned} H^{n,\mathcal{N}}(v) &= \min \{ H^{n,C_1}(v|_{C_1}), \dots, H^{n,C_k}(v|_{C_k}) \} \\ &= \min \{ H_S^n(p^{v|_{C_1}, C_1}), \dots, H_S^n(p^{v|_{C_k}, C_k}) \} \\ &= \min_{C \in \mathcal{C}(\mathcal{N})} H_S^n(p^{v,C}), \\ &= H_D^{n,\mathcal{N}}(v). \end{aligned}$$

□

6 Applications and concluding remarks

We have proposed an axiomatization of the entropy of capacities defined on regular set systems, in a spirit very close to the axiomatization of Faddeev for the Shannon entropy. The key axiom which is additional is (HG4), binding the computation of entropy to the structure of the set system.

Essentially, two comments are of interest here. The first one concerns the nature of the 5 axioms we propose. Axioms (HG1), (HG3) and (HG5) clearly correspond to the three axioms of Faddeev for the Shannon entropy. Axiom (HG2) is necessary in our framework, but is implicit for the Shannon entropy. Put differently, (HG2) too is characteristic of Shannon entropy. Hence, only axiom (HG4) is specific to our construction; it is able to deal with the structure of the set system, provided it is regular. It is responsible for the general form of the entropy, that is, a weighted average of classical Shannon entropies along maximal chains. This suggests the following: as other entropies have been proposed for probabilities, together with axiomatizations, it would be possible to carry the present framework for axiomatizing these other entropies, just replacing axioms (HG1), (HG2), (HG3) and (HG5) by axioms which are specific to these entropies.

Another comment concerns the usefulness of regular set systems. Would it be possible to consider yet more general structures, i.e., set systems which are not regular? Formally, Definition 11 can be applied to any set system, since maximal chains are always defined, and thus probability distributions $p^{v,C}$ too. A fundamental problem is however that these chains need not to be of the same length, whose direct consequence is that the probability distribution would be undefined for some elements of N . For example, take $\mathcal{N}' := \{\emptyset, \{1\}, \{1, 2, 3\}\}$, which is totally ordered but not regular, and let v be a capacity defined on $\mathcal{N}'$. Computing the entropy of v by Definition 11 leads to:

$$H(v) = -v(\{1\}) \log v(\{1\}) - (v(\{1, 2, 3\}) - v(\{1\})) \log(v(\{1, 2, 3\}) - v(\{1\})).$$

Clearly, the quantities $v(\{1\})$, $v(\{1, 2, 3\}) - v(\{1\})$ cannot be considered as the values of a probability distribution on N . Another troublesome fact is the following. We know that classically the entropy takes its maximal value when v is an additive uniform capacity, that is, which satisfies $v(A) = |A|/n$. But for v on $\mathcal{N}'$, $H(v)$ takes the maximal value when $v(\{1\}) = 1/2$ which is not the additive uniform capacity.

It seems then, as far as maximal chains are considered as fundamental ingredients, that regular set systems are the most general structure we can consider. It is important to note that regular set systems encompass many well-known ordered structures. This is studied in details in [10], and it turns out that distributive lattices, set lattices, convex geometries and anti-matroids are strictly included into regular set systems. The following simple example shows that a regular set system is not necessarily a lattice: take $N = \{1, 2, 3, 4\}$ and $\mathcal{N} := \{\emptyset, \{1\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 4\}, \{3, 4\}, \{1, 2, 3\}, \{1, 3, 4\}, N\}$. Then, $\{1\}$ and $\{3\}$ have no supremum. Now, the case of lattices is obviously a particular case of interest. We refer the reader to [8] where we have examined the application of our definition of entropy to the case where capacities are defined on lattices (not necessarily set lattices).

7 Acknowledgment

We address all our deepest thanks to the anonymous reviewers for a careful reading of the manuscript and the detection of several flaws in the previous version of the axiomatization.

References

- [1] Choquet G. Theory of capacities. Annales de l'Institut Fourier 1953; 5; 131-295.
- [2] Davey B.A, Priestley H.A. Introduction to lattices and order: Cambridge University Press; 1990.

- [3] Denneberg D. Non-Additive Measure and Integral: Kluwer Academic; 1994.
- [4] Dukhovny A. General entropy of general measures. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems; 2002; 10; 213-225.
- [5] Ebanks B, Sahoo P, Sander W. Characterization of information measures: World Scientific; 1998.
- [6] Faddeev D.K. The notion of entropy of finite probabilistic schemes (Russian). Uspekhi Mat. Nauk 1956; 11; 15-19.
- [7] Grabisch M. An axiomatization of the Shapley value and interaction index for games on lattices. SCIS-ISIS 2004, 2nd Int. Conf. on Soft Computing and Intelligent Systems and 5th Int. Symp. on Advanced Intelligent Systems, Yokohama, Japan, September 2004.
- [8] Honda A, Grabisch M. Entropy of capacities on lattices. Information Sciences 2006; 176; 3472-3489.
- [9] Kojadinovic I, Marichal J.-L, Roubens M. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. Information Sciences 2005; 172; 131-153.
- [10] Lange F., Grabisch M. Values on regular games under Kirchhoff's laws. Submitted to *SIAM J. of Discrete Mathematics*.
- [11] Marichal J.-L. Entropy of discrete Choquet capacities. Eur. J. of Oper. Res. 2002; 137; 612-624.
- [12] Marichal J.-L, Roubens M. Entropy of discrete fuzzy measure. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 2000; 8; 625-640.
- [13] Shannon C.E. A mathematical theory of communication. Bell System Tech. Journ. 1948; 27; 374-423, 623-656.
- [14] Sugeno M. Theory of fuzzy integrals and its applications. Doctoral Thesis; Tokyo Institute of Technology; 1974.
- [15] Xie L, Grabisch M. The Core of Capacities on Lattices. Proc. of the Int. Fuzzy Systems Association World Congress (IFSA); Beijing; China, July 2005; 461-466.
- [16] Yager R.R. On the entropy of fuzzy measures. IEEE Transaction on Fuzzy Systems, 2000; 8; 453-461.