



A New Entropy for Hypergraphs

Isabelle Bloch¹ and Alain Bretto²

¹ LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France

isabelle.bloch@telecom-paristech.fr

² NormandieUnicaen, GREYC CNRS-UMR 6072, Caen, France

alain.bretto@info.unicaen.fr

Abstract. This paper introduces a new definition of entropy for hypergraphs. It takes into account the fine structure of a hypergraph by considering its partial hypergraphs, leading to an entropy vector. This allows for more precision in the description of the underlying complexity of the hypergraph. Properties of the proposed definitions are analyzed.

Keywords: Hypergraphs · Entropy ·
Algebraic and graphical structures

1 Introduction

The concept of entropy has been introduced to measure the amount of information contained or delivered by an information source. Its purpose is to model the combinatorial possibilities of the different states of a given set. Naturally, these different states are modeled by a probability distribution of their occurrences. This makes it possible to have a “relatively simple” (in terms of summarization) vision of the content of the information of the objects under study. It is therefore primarily a probabilistic theory of information. After its introduction, this notion has been developed in different parts of sciences, such as dynamic systems, computer science, stochastic theory, among others. Introducing entropy into information theory is credited to Shannon [15]. This notion has been adapted to quantum physics by von Neumann using the spectrum of the density matrix [13].

More recently this notion has been adapted to the Laplacian of a graph, thus showing the utility of this concept in graph theory and its application in particular in image analysis [7, 10–12].

Hypergraphs (generalizing the notion of graphs with higher arity of edges) are useful to model real data where relationships between different data items have to be taken into account. This includes images, where relations between pixels or regions can involve more than two elements (e.g. proximity, parallelism...). A combinatorial object such as a hypergraph can be very complicated, and the introduction of an entropy measure on such a structure is a relevant way to assess this complexity. This measure can be used in feature selection methods such as in [19] or for other tasks such as compression or similarity assessment.

Usual approaches rely on the Laplacian of hypergraphs, in a similar way as for graphs. In previous work [6], we defined similarities between hypergraphs based on mathematical morphology on hypergraphs [5] and valuations in lattices [4, 16], from which partitions and entropies can be defined, e.g. up to a morphological transformation, to introduce additional filtering and gain in robustness. In all these approaches, the entropy is defined as a single number, considering the hypergraph globally.

In this paper, we propose to define the entropy of a hypergraph as a vector, by using substructures - namely partial hypergraphs. An entropy vector associated with n random variables is a kind of generalization of Shannon's entropy, which links domains such as geometry, combinatorics and information theory [9]. In this paper we introduce a new concept of entropy called *entropy vector* associated with a hypergraph. It is based on the incidence matrix $I(H)$ associated with the hypergraph H , more precisely on $I(H)I(H)^t$, as well as all the main sub-matrices of $I(H)I(H)^t$. The defined vector conveys much more detailed information than the information measured by conventional entropy.

In Sect. 2, basic definitions on hypergraphs are recalled, and the notations used in this paper are introduced. In Sect. 3, the usual notion of entropy of a hypergraph is given, for which we show a few properties that are useful in this paper. The main contribution of this paper is introduced in Sect. 4, where a definition of entropy vector of a hypergraph is proposed, and in the next sections, where properties are analyzed. A lattice structure is introduced in Sect. 5. Since complexity can be a concern, a fast approximate calculus method is proposed in Sect. 6. Finally links with the Zeta function are suggested in Sect. 7.

2 Preliminaries

Let us first recall some useful notions on hypergraphs [3, 8]. A hypergraph H denoted by $H = (V, E = \{e_i, i = 1\dots m\})$ is defined as a pair of a finite set V (vertices) and a finite family $\{e_i, i = 1\dots m\}$ of hyperedges. In this paper we consider only simple hypergraphs (i.e. without repeated hyperedges), and E is hence a set. Hyperedges can be considered equivalently as subsets of vertices or as a relation between vertices of V . The first interpretation is considered here, and we will note $x \in e_i$ the fact that a vertex $x \in V$ belongs to the hyperedge e_i . In this paper, it is further assumed that E is non-empty and that each hyperedge is non-empty (i.e. $\forall e_i \in E, e_i \neq \emptyset$). We denote by $r(H)$ the rank of H , defined as $r(H) = \max_{e \in E} |e|$.

In order to study the fine structure of an hypergraph, the following notions are important in this paper:

- A *partial hypergraph* H' of H generated by $J \subseteq \{1\dots m\}$ is a hypergraph $(V', \{e_j, j \in J\})$, where $\cup_{j \in J} e_j \subseteq V' \subseteq V$. In the sequel we will suppose that $V' = V$. This will be denoted by $H' \leq H$.
- Given a subset $V' \subseteq V$, a *subhypergraph* H' of H is the partial hypergraph $H' = (V', \{e_i, e_i \in E \mid e_i \subseteq V'\})$.

- The *induced subhypergraph* H' of H with $V' \subseteq V$ is the hypergraph defined as $H' = (V', E')$ with $E' = \{e' = \{e \cap V'\} \mid e \in E \text{ and } e \cap V' \neq \emptyset\}$. Note that if $V' = V$ and hypergraphs are considered without empty hyperedges, then $H' = H$.
- A hypergraph $H = (V, E)$ is *isomorphic* to a hypergraph $H' = (V', E')$ ($H \simeq H'$), if there exist a bijection $f : V \rightarrow V'$ and a bijection $\pi : \{1 \dots m\} \rightarrow \{1 \dots m'\}$, where $m = |E|$ and $m' = |E'|$, which induce a bijection $g : E \rightarrow E'$ (i.e. $g(e) = \{f(x) \mid x \in e\}$) such that $g(e_i) = e'_{\pi(i)}$ for all $e_i \in E$ and $e'_{\pi(i)} \in E'$. The mapping f is then called *isomorphism* of hypergraphs.

Finally, let us introduce some notations on symmetric matrices. Let $A \in \mathcal{M}_{n \times n}(\mathbb{R})$, $A = ((a_{i,j}))_{i,j \in \{1 \dots n\}}$ be a symmetric matrix on \mathbb{R} . Let $\alpha = \{\alpha_1, \dots, \alpha_t\} \subseteq \{1, \dots, n\}$. We denote by $A(\alpha; \alpha)$ the submatrix of A generated by keeping only the rows and columns of A indexed by α , i.e.

$$A(\alpha; \alpha) = ((b_{i,j}))_{i,j \in \{1 \dots t\}}, \quad b_{i,j} = a_{\alpha_i, \alpha_j}, \quad \alpha_i, \alpha_j \in \alpha.$$

Such a matrix is called *principal submatrix* of A . If $\alpha = \{1, 2, \dots, k\}$, $k \leq n$, $A(\alpha; \alpha)$ is called *leading principal submatrix* of A .

3 Entropy of a Hypergraph

Let (H, V) be a simple hypergraph with $|V| = n$, $|E| = m$. Let $I(H) = ((a_{i,j}))_{(i,j) \in \{1 \dots m\} \times \{1 \dots n\}}$ be the incidence matrix of H : $a_{i,j} = 1$, if $x_j \in e_i$ and $a_{i,j} = 0$ otherwise. We consider the matrix $m \times m$ defined as:

$$L(H) = I(H)I(H)^t = ((|e_i \cap e_j|))_{i,j \in \{1 \dots m\}}.$$

This matrix is positive semi-definite. Its eigenvalues $\lambda_i(L(H))$, $i = 1 \dots m$, are positive and can be ordered as follows:

$$0 \leq \lambda_1(L(H)) \leq \lambda_2(L(H)) \leq \dots \leq \lambda_m(L(H)).$$

Lemma 1. *The trace of the matrix $L(H)$ is the sum of the degrees $d(x)$ of the vertices x of H (i.e. the number of hyperedges that contain x):*

$$\text{Tr}(L(H)) = \sum_{x \in V} d(x).$$

Proof. $\text{Tr}(L(H)) = \sum_{i=1}^m \lambda_i = \sum_{e \in E} |e| = \sum_{x \in V} d(x)$. □

Let us define:

$$\mu_i = \frac{\lambda_i(L(H))}{\sum_{i=1}^m \lambda_i(L(H))} = \frac{\lambda_i(L(H))}{\text{Tr}(L(H))}.$$

The μ_i are the eigenvalues of the normalized matrix

$$\mathcal{L}(H) = \frac{1}{\text{Tr}(L(H))} L(H)$$

which is also positive semi-definite.

Classically, the *Shannon entropy* of the hypergraph H (see e.g. [8]) is defined as:

$$S(H) = - \sum_{i=1}^m \mu_i \log_2(\mu_i).$$

Note that other forms of entropy exist, such as the *Renyi entropy* for instance, which is defined for a hypergraph H as:

$$R_s(H) = \frac{1}{1-s} \ln\left(\sum_{i=1}^m \mu_i^s\right), \quad s \geq 0.$$

In this paper we mostly consider Shannon entropy, except in Sect. 7.

Proposition 1. *Let $H = (V, E)$ be a simple hypergraph without isolated vertex, without empty hyperedge, and with $|V| = n$ and $|E| = m$ ($m > 0$). We have the two following properties:*

- (a) $S(H) = 0$ if and only if $|E| = 1$,
- (b) $S(H) = \log_2(n) - \log_2(r(H)) = \log_2(m)$, where $r(H)$ (the rank of H) is a constant equal to $\frac{n}{m}$, if and only if H is uniform (i.e. $\forall e \in E$, $|e| = r(H)$) and the intersection of any two distinct hyperedges is empty (i.e. for all $e, e' \in E$ such that $e \neq e'$, $|e \cap e'| = 0$).

Proof. (a) Assume that $|E| = 1$. Then $L(H)$ is reduced to a scalar value, which is non-zero since $e \neq \emptyset$, and the unique normalized eigenvalue is $\mu = 1$. Hence $S(H) = -1 \log_2(1) = 0$.

Conversely, suppose that $S(H) = 0$. Since $E \neq \emptyset$ and H does not contain any empty hyperedge, then $S(H) = 0$ implies that $\forall i$, $\mu_i = 0$ or $\mu_i = 1$. Since hyperedges are not empty, $\mu_i = 0$ is not possible, and since $\sum_i \mu_i = 1$, there is a unique eigenvalue, equal to 1. Hence $|E| = 1$.

(b) Assume now that H is uniform, with $|e| = r(H)$ for each hyperedge e , and that for all $e, e' \in E$ such that $e \neq e'$, $|e \cap e'| = 0$. Note that in this case $L(H)$ is diagonal and so $\lambda_i = |e_i| = r(H)$. Since $e \cap e' = \emptyset$, $d(x) = 1$ for each vertex x and from Lemma 1, $\text{Tr}(L(H)) = \sum_{x \in V} d(x) = n$. This corresponds to a situation where vertices are uniformly distributed among the hyperedges, i.e. $r(H) = \frac{n}{m}$. Therefore we have:

$$\begin{aligned} S(H) &= - \sum_{i=1}^m \mu_i \log_2(\mu_i) \\ &= - \sum_{e \in E} \frac{|e|}{\text{Tr}(L(H))} \log_2\left(\frac{|e|}{\text{Tr}(L(H))}\right) \\ &= - \frac{mr(H)}{n} \log_2\left(\frac{r(H)}{n}\right) \\ &= - \log_2\left(\frac{r(H)}{n}\right) = \log_2(n) - \log_2(r(H)) = \log_2(m). \end{aligned}$$

Conversely suppose that $S(H) = \log_2(n) - \log_2(r(H)) = \log_2(m)$. This implies $n = mr(H)$.

Moreover from Lemma 1, from $|e| \leq r(H)$ for all e by definition of $r(H)$, and since $\sum_{x \in V} d(x) \geq n$ (no isolated vertex), we have:

$$n \leq \sum_{e \in E} |e| = \sum_{x \in V} d(x) \leq mr(H) = n.$$

It follows that $\sum_{x \in V} d(x) = n = \text{Tr}(L(H))$.

It also follows that $\sum_{e \in E} |e| = n = mr(H)$. Since $E \neq \emptyset$ and there is no empty hyperedge, we can derive that $|e| = r(H)$ for all $e \in E$. This means that H is uniform, and since $n = mr(H)$, for all $e, e' \in E$ such that $e \neq e'$, we have $|e \cap e'| = 0$. Note that in this case the matrix $L(H)$ is diagonal, and for all $i \in \{1, 2, \dots, m\}$, $\mu_i = \frac{\lambda_i}{\text{Tr}(L(H))} = \frac{r(H)}{n}$.

□

This proposition shows that the entropy is closely related with the parameters of the hypergraph. Note that Case **b** implicitly assumes that $\frac{n}{m}$ is an integer. Moreover, Case **a** is a consequence of Case **b**.

A straightforward extension of this result deals with hypergraphs that may contain isolated vertices. A similar result holds by replacing n by n' , the number of non-isolated vertices (i.e. that belong to at least one hyperedge).

4 Entropy Vector Associated with a Hypergraph

Now, we built on the classical definition of hypergraph entropy to propose a new entropy, defined as a vector, based on a finer analysis of the structure of the hypergraph by considering all its partial hypergraphs.

Definition 1 (Entropy vector). Let $H = (V, E)$ be a hypergraph. For $i \leq m$ ($m = |E|$), let

$$SE_i(H) = \{S(H_i) \mid H_i = (V, E_i), H_i \leq H, |E_i| = i\}$$

be the set of entropy values of all partial hypergraphs of H whose set of hyperedges has cardinality i , arranged in increasing order.

The entropy vector of the hypergraph H is then the vector:

$$SE(H) = (SE_1(H), SE_2(H), \dots, SE_m(H))$$

with $2^m - 1$ coordinates.

Note that if $H' \leq H$ then it is easy to see that the matrix $L(H')$ is a principal submatrix of $L(H)$.

From Proposition 1, the vector $S(H)$ begins with at least m values equal to 0.

Example 1. Let us illustrate this definition on a very simple example, illustrated in Fig. 1, for a hypergraph H with three hyperedges.

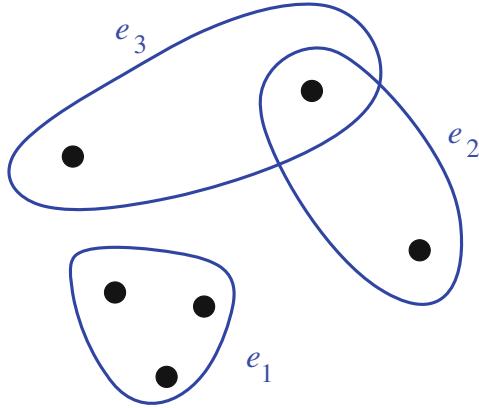


Fig. 1. A simple example of hypergraph, with three hyperedges (indicated by blue lines): e_1 contains three vertices, e_2 and e_3 both contain two vertices, with one common vertex. (Color figure online)

Let us compute the entropy vector:

- For SE_1 , there are three partial hypergraphs containing one hyperedge (e_1 , e_2 and e_3 , respectively), and $SE_1 = (0, 0, 0)$.
- For SE_2 , there are three partial hypergraphs containing two hyperedges. For (e_1, e_2) , the matrix L is equal to $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, with eigenvalues 2 and 3, and the corresponding entropy is equal to $s_1 = -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} \simeq 0.29$. The same reasoning applies for (e_1, e_3) . For (e_2, e_3) , the matrix L is equal to $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, with eigenvalues 1 and 3, and the corresponding entropy is equal to $s_2 = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} \simeq 0.24$. Then $SE_2 = (s_2, s_1, s_1) \simeq (0.24, 0.29, 0.29)$.
- For SE_3 , there is one partial hypergraph containing three hyperedges, i.e. H . The matrix L is equal to $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, with eigenvalues 1, 3 and 3, and the corresponding entropy is equal to $s_3 = -\frac{1}{7} \log_2 \frac{1}{7} - 2 \frac{3}{7} \log_2 \frac{3}{7} \simeq 0.44$. Hence $SE_3 = (s_3) \simeq (0.44)$.
- Finally, the entropy vector is

$$SE(H) = (0, 0, 0, s_2, s_1, s_1, s_3) \simeq (0, 0, 0, 0.24, 0.29, 0.29, 0.44).$$

Proposition 2. Let $H = (V, E)$ and $H' = (V', E')$ be two isomorphic hypergraphs. Then there is a permutation σ such that

$$\begin{aligned} SE(H) &= (SE_1(H), SE_2(H), SE_3(H) \dots, SE_m(H)) = \\ SE(H') &= (SE'_{\sigma(1)}(H), SE'_{\sigma(2)}(H), SE'_{\sigma(3)}(H) \dots, SE'_{\sigma(m)}(H)) \end{aligned}$$

Proof. Let $I(H)$ and $I(H')$ be the incidence matrices of H and H' , respectively. Since H and H' are isomorphic there are two permutation matrices P and Q such that

$$I(H') = PI(H)Q^t.$$

Consequently

$$I(H')^t = QI(H)^tP^t$$

and

$$I(H')I(H')^t = PI(H)Q^tQI(H)^tP^t.$$

Therefore we obtain

$$L(H') = PL(H)P^t \text{ and } \mathcal{L}(H') = P\mathcal{L}(H)P^t.$$

It is well known that if $A = PBP^t$ (with $P^{-1} = P^t$) then A and B have the same eigenvalues. The matrix P represents the isomorphism and gives rise to the permutation σ (the permutation on vertices induces a permutation on hyperedges). The isomorphism guarantees that the two hypergraphs have the same structure. Hence their sets of partial hypergraphs are in one-to-one correspondence and each partial hypergraph of H is isomorphic to a partial hypergraph of H' . The result follows.

□

5 Partial Ordering and Lattice Structure

In this section we further analyze the properties of partial ordering on hypergraphs and on vector entropy, which result in lattice structures.

We first recall results from [5].

Definition 2. Let \mathcal{H} be the set of isomorphism classes of hypergraphs. A partial order \leq_f on \mathcal{H} is defined as: $\forall H', H \in \mathcal{H}$, $H' \leq_f H$ if $\exists V' \subseteq V$ such that H' is isomorphic (by f) to the subhypergraph of H induced by V' .

It is clear that \leq_f is a partial order relation. We denote by $=_f$ the corresponding equality, and by $<_f$ the corresponding strict ordering.

Hereafter, we will denote a class by a representative hypergraph H in this class.

Proposition 3 ([5]). The structure (\mathcal{H}, \leq_f) is a complete lattice. The supremum of any $H_1 = (V_1, E_1), H_2 = (V_2, E_2)$ is $\sup\{H_1, H_2\} = H_1 \vee H_2 = (V_1 \cup V_2, E_1 \cup E_2)$, and the infimum $\inf\{H_1, H_2\}$ is the maximum common induced subhypergraph (and their extensions to any family).

Let us now move to the vector entropy. Let \mathcal{V} be a set of vectors. A partial order is defined on \mathcal{V} as follows. For $x = (x_1, x_2, \dots, x_k) \in \mathcal{V}$ and $y = (y_1, y_2, \dots, y_t) \in \mathcal{V}$, $x \leq y$ if $\forall i \leq \min(k, t)$, $x_i \leq y_i$. Note that this relation is equivalent to the usual Pareto ordering, if the shortest length vector (say x , i.e. $k \leq t$) is completed by $t - k$ components equal to 0. The set \mathcal{V} endowed with this partial ordering is called an *ordered vector set*.

Proposition 4. *The set $SE_{\mathcal{H}} = \{SE(H) \mid H \in \mathcal{H}\}$ is an ordered vector set.*

Note that while each $SE_i(H)$ has its values increasingly ordered, this is not the case for $SE(H)$. Proposition 4 means that a partial ordering can be defined on $SE_{\mathcal{H}}$, as defined above.

Proposition 5. *The set $SE_{\mathcal{H}}$ endowed with the partial ordering \leq is a lattice.*

Proof. It is clear that the partial ordering \leq (Pareto-like ordering) gives rise to a supremum (least upper-bound) and an infimum (greatest lower bound) for any finite family of $SE(H)$. The supremum is computed as the component-wise maximum and the infimum as the component-wise minimum. It is bounded by 0, obtained for $H = (V, E = \{e\})$ from Proposition 1, so any infinite family has also an infimum.

□

Proposition 6. *Let $H = (V, E)$ and $H' = (V', E')$. If $H' \leq_f H$, then $SE(H') \leq SE(H)$.*

Proof. Let $H = (V, E)$ and $H' = (V', E')$ such that $H' \leq_f H$, with $|E| = m$ and $|E'| = m'$. Two cases arise:

1. If $H' =_f H$ then, by Proposition 2 there is a permutation σ such that $SE_{\sigma}(H) = SE(H)$.
2. If $H' <_f H$ then by definition there is an induced subhypergraph $H'' = (V'', E'')$ of H which is isomorphic to H' . From Proposition 2, $\mathcal{L}(H')$ and $\mathcal{L}(H'')$ have the same eigenvalues, and moreover, since f is an isomorphism between H' and H'' , there is a permutation σ (which comes from f) such that

$$\forall i, 1 \leq i \leq m', SE_{\sigma(i)}(H') \leq SE_i(H'').$$

Since $H' <_f H$ then $m' < m$, hence $2^{m'} - 1 < 2^m - 1$. Hence, by adding $2^m - 1 - (2^{m'} - 1)$ components equal to 0 in $(SE_{\sigma(1)}(H), SE_{\sigma(2)}(H), \dots, SE_{\sigma(m')}(H))$ at a good place we obtain a vector with $2^m - 1$ components such that

$$SE(H') \leq SE(H).$$

□

Let $H = (V, E)$ and $H' = (V', E')$ be two hypergraphs such that $V = \{x_1, x_2, \dots, x_n\}$, $V' = \{x'_1, x'_2, \dots, x'_{n'}\}$, $E = \{e_1, e_2, \dots, e_m\}$, $E' = \{e'_1, e'_2, \dots, e'_{m'}\}$. Let $I(H) = ((a_{i,j}))_{(i,j) \in \{1\dots m\} \times \{1\dots n\}}$ and $I(H') = ((a'_{i,j}))_{(i,j) \in \{1\dots m'\} \times \{1\dots n'\}}$. Let $H'' = H \cup H' = (V \cup V', E \cup E')$. Its incidence matrix is built as follows: add rows and columns with 0's in $I(H)$ (respectively $I(H')$) for vertices and hyperedges present in H' and not in H (respectively present in H and not in H'), so as to get two matrices of size $m'' \times n''$ whose generic terms are still denoted $a_{i,j}$ and $a'_{i,j}$. Then the matrix $I(H'')$ has m'' rows corresponding to $E \cup E'$,

n'' columns corresponding to $V \cup V'$, and coefficients $a''_{i,j} = \max(a_{i,j}, a'_{i,j})$, for $1 \leq i \leq m'', 1 \leq j \leq n''$.

We define $H \cap H' = (V \cap V', E \cap E')$ in the same way by suppressing rows and columns in $I(H)$ and $I(H')$ corresponding to non-common vertices and edges, and by replacing max by min to define the coefficients of the resulting incidence matrix.

Clearly both $L(H \cup H')$ and $L(H \cap H')$ are well defined.

6 An Algorithm for Approximate Calculus of the Entropy Vector

When the size of the hypergraph increases, the entropy vector may become very costly. Suggestions to reduce the complexity would be to disregard partial hypergraphs with very few hyperedges (bringing reduced information), or in contrast with near to m hyperedges (and hence less relevant). Another way would be to order the hyperedges in order of increasing cardinality and to take the leading principal matrices.

In this section, we propose approximations that alleviate two main drawbacks to calculate the entropy vectors.

The first drawback is the calculation of the eigenvalues. Since entropy is a mean value of the information, we can take an approximate value. Recall that $\log(x) = \log(1 + (x - 1)) \simeq x - 1$ for $-1 \leq x - 1 \leq 1$. Since $\mu_i \in [0, 1]$, and hence $\mu_i - 1 \in [-1, 0]$, $S(H) = -\sum_{i=1}^m \mu_i \log_2(\mu_i)$ can be approximated by:

$$\begin{aligned} -\sum_{i=1}^m \mu_i(1 - \mu_i) &= \text{Tr}(\mathcal{L}(H)) - \text{Tr}(\mathcal{L}(H)^2) \\ &= 1 - \frac{1}{(\sum_{e \in E} |e|)^2} \left(\sum_{e \in E} |e|^2 + \sum_{e \in E} \sum_{e' \in A_e} |e \cap e'|^2 \right) \quad (1) \end{aligned}$$

where $A_e = \{e' \in E \mid e' \neq e \text{ and } e \cap e' \neq \emptyset\}$. This expression is easy to compute.

The second drawback is that we have $2^m - 1$ principal matrices. We can take only the leading principal matrices, hence, we have to manage $m - 1$ matrices and we get an entropy vector with $m - 1$ components. In this case the order of the hyperedges in the matrix $\mathcal{L}(H)$ is important: if we permute two hyperedges in $\mathcal{L}(H)$, for instance the first one with the last one, we may change $SE(H)$. This comes from Eq. 1. Consequently we have to find an appropriate order on hyperedges. We sort hyperedges when building $\mathcal{L}(H)$ so that e is before e' (denoted by $e \preceq e'$) if:

- $|e| < |e'|$, or
- $|e| = |e'|$ and $\sum_{a \in A_e} |e \cap a| \leq \sum_{a' \in A_{e'}} |e' \cap a'|$.

The case where $|e| = |e'|$ and $\sum_{a \in A_e} |e \cap a| = \sum_{a' \in A_{e'}} |e' \cap a'|$ is not important because it does not change Eq. 1, hence we can choose to put either e before e' or e' before e . Relation \preceq is a pre-order (both reflexive and symmetric relation). After pre-ordering $\mathcal{L}(H)$, the vector $SE(H)$ is canonically associated to this matrix.

7 Zeta Function and Entropy

In this section, we suggest some links between the Zeta function and the notion of hypergraph entropy.

The Zeta spectral function plays an important role in areas where linear operators are often present [17], that is in most fields of physics. It has also been introduced on graphs and in image analysis [11, 18]. Define now the *spectral Zeta function* of a hypergraph $H = (V, E)$, with $|E| = m$, as:

$$\zeta_H(s) = \text{Tr}(\mathcal{L}(H)^{-s}) = \sum_{i=1, \mu_i \neq 0}^m \mu_i^{-s}$$

In a similar way as for entropy, we can define the *spectral Zeta vector function*.

It is easy to show that the derivative of ζ with respect to s can be expressed as:

$$\zeta'_H(s) = \frac{\zeta_H(s)}{ds} = - \sum_{i=1}^m \mu_i^{-s} \ln(\mu_i).$$

Consequently we have:

$$\begin{aligned} \zeta'_H(-1) &= - \sum_{i=1}^m \mu_i \ln(\mu_i) = \ln(2)S(H), \\ \zeta'_H(0) &= - \sum_{i=1}^m \ln(\mu_i) = -\ln(\prod_{i=1}^m \mu_i) = -\ln(\det(\mathcal{L}(H))). \end{aligned}$$

The following result also holds.

Proposition 7. *Let $H = (V, E)$ be a simple hypergraph, with $|E| = m$, and strictly positive eigenvalues of $\mathcal{L}(H)$. The Zeta function is related to the Renyi entropy by the following equation:*

$$\forall s \geq 0, \quad \zeta_H(-s) = e^{(1-s)R_s(H)}$$

Proof. We have, for $s \geq 0$:

$$(1-s)R_s(H) = \ln\left(\sum_{i=1}^m \mu_i^s\right)$$

and

$$e^{(1-s)R_s(H)} = \sum_{i=1}^m \mu_i^s = \zeta_H(-s)$$

□

These relations show how to relate spectro-analytic theory to information theory. This is of great importance, especially in dynamic networks seen as a series of hypergraphs. Indeed the notion of dynamical Zeta functions are an important tool to analyze chaotic dynamical systems [2, 14] and should be effective in quantifying and tracking the evolution of entropy vectors, by applying the results in this section to the entropy values of all partial hypergraphs.

8 Conclusion

In this paper, a new measure of entropy for hypergraphs was proposed, defined as a vector. It generalizes Shannon entropy, representing a richer information on the complexity of a hypergraph, taking into account all its partial hypergraphs, hence its sub-structures. Algebraic properties were proved, as well as some links with the Zeta function.

Future work aims at exploring further the proposed notion, including additional properties and applications, in particular for image description. For instance, as suggested in [6], a hypergraph could be built from an image considering pixels or regions resulting from an over-segmentation as vertices, and hyperedges could be defined from relations (e.g. neighborhood, grey-levels or colors, spatial relations between regions...). Looking at the sub-structures of the hypergraph would then provide a description of the image complexity which would be finer and more precise than a global description.

For concrete applications to be tractable, the computation cost may be an issue for even moderately large hypergraphs, since the size of the entropy vector grows exponentially with the number of hyperedges. We suggested a few ways to address this issue in Sect. 6. Other could be developed, for instance by choosing randomly partial hypergraphs, or by looking at specific patterns in the hypergraph.

Another extension of hypergraph entropy to sequences of entropy values was proposed in [1]. The approach we proposed here is different since all the partial hypergraphs are considered, instead of centroid expansion subgraphs as in this earlier work. In future work it would be also interesting to compare both approaches based on their respective properties, as well as on concrete examples.

Another direction of research consists in extending the proposed notion to other forms of entropy (such as Renyi entropy), to weighted hyperedges (for instance using the distance between vertices), to weighted terms in the matrix $L(H)$ (e.g. using a distance between two hyperedges e_i and e_j to weight $|e_i \cap e_j|$), or more generally to attributed hypergraphs.

References

1. Bai, L., Escolano, F., Hancock, E.R.: Depth-based hypergraph complexity traces from directed line graphs. *Pattern Recogn.* **54**, 229–240 (2016)
2. Baladi, V.: Dynamical zeta functions. In: Branner, B., Hjorth, P. (eds.) *Real and Complex Dynamical Systems*. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol. 464, pp. 1–26. Springer, Dordrecht (1995). https://doi.org/10.1007/978-94-015-8439-5_1
3. Berge, C.: *Hypergraphs*. Elsevier Science Publisher, The Netherlands (1989)
4. Birkho, G.: *Lattice Theory*, vol. 25, 3rd edn. American Mathematical Society, Providence (1979)
5. Bloch, I., Bretto, A.: Mathematical morphology on hypergraphs, application to similarity and positive kernel. *Comput. Vis. Image Underst.* **117**(4), 342–354 (2013)

6. Bloch, I., Bretto, A., Leborgne, A.: Robust similarity between hypergraphs based on valuations and mathematical morphology operators. *Discrete Appl. Math.* **183**, 2–19 (2015)
7. Brauunstein, S.L., Ghosh, S., Severini, S.: The Laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states. *Ann. Comb.* **10**(3), 291–317 (2006)
8. Bretto, A.: Hypergraph Theory: An Introduction. Mathematical Engineering. Springer, New York (2013). <https://doi.org/10.1007/978-3-319-00080-0>
9. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, Hoboken (2012)
10. Dehmer, M., Mowshowitz, A.: A history of graph entropy measures. *Inf. Sci.* **181**(1), 57–78 (2011)
11. Han, L., Escolano, F., Hancock, E.R., Wilson, R.C.: Graph characterizations from von Neumann entropy. *Pattern Recogn. Lett.* **33**(15), 1958–1967 (2012)
12. Lin, H., Zhou, B.: On the von Neumann entropy of a graph. *Discrete Appl. Math.* **247**, 448–455 (2018)
13. von Neumann, J.: Mathematische Grundlagen der Quantenmechanik, vol. 38. Springer, New york (1955)
14. Shanker, O.: Graph zeta function and dimension of complex network. *Mod. Phys. Lett. B* **21**(11), 639–644 (2007)
15. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423 (1948)
16. Simovici, D.: Betweenness, metrics and entropies in lattices. In: 38th IEEE International Symposium on Multiple Valued Logic, ISMVL 2008, pp. 26–31 (2008)
17. Voros, A.: Spectral zeta functions. *Adv. Stud. Pure Math.* **21**, 327–358 (1992)
18. Xiao, B., Hancock, E.R., Wilson, R.C.: Graph characteristics from the heat kernel trace. *Pattern Recogn.* **42**(11), 2589–2606 (2009)
19. Zhang, Z., Hancock, E.: Hypergraph based information-theoretic feature selection. *Pattern Recogn. Lett.* **33**, 1991–1999 (2012)