Tensor entropy for uniform hypergraphs

Can Chen∗ Indika Rajapakse†

Abstract
In this paper, we develop a new notion of entropy for uniform hypergraphs, which are generalized from graphs, based on tensor theory. In particular, we employ the probability distribution of the generalized singular values, calculated from the higher-order singular value decomposition of the Laplacian tensors, to fit into the Shannon entropy formula. It is shown that this tensor entropy is an extension of von Neumann entropy for graphs. We establish results on the lower and upper bounds of the entropy and demonstrate that it is a measure of regularity, relying on the vertex degrees, path lengths, clustering coefficients and nontrivial symmetricity, for uniform hypergraphs with two simulated examples.

Keywords: uniform hypergraphs, entropy, tensor decomposition, random hypergraphs

1 Introduction

Many real world complex systems can be analyzed through a graph/network prospective. There are two classical and well-known classes of complex networks, scale-free networks and small world networks, which play a significant role in many domains such as social networks, biology, cognitive science and signal processing [1, 4, 27, 44]. The human genome is a beautiful example of complex dynamic graph. The genome-wide chromosomal conformation (Hi-C) map represents the spatial proximity of different parts of genome capturing the genome structure over time [10, 42]. When studying such dynamic graphs, one is often required to identify the pattern/couple changes including degree distribution, path lengths, clustering coefficients, etc, in the graph topology in order to capture the dynamics [25, 33, 41].

The von Neumann entropy of a graph, first introduced by Braunstein et al. [8], is a spectral measure used in structural pattern recognition. The intuition behind this measure is linking the graph Laplacian to density matrices from quantum mechanics, and measuring the complexity of the graphs in terms of the von Neumann entropy of the corresponding density matrices [32]. In addition, the measure can be viewed as the information theoretic Shannon entropy, i.e.,

\[ S = - \sum_j \eta_j \ln \eta_j, \]

where, \( \eta_j \) are the normalized eigenvalues of the Laplacian matrix of a graph such that \( \sum_j \eta_j = 1 \). Passerini and Severini observed that the von Neumann entropy of a graph tends to grow with the number of connected components, the reduction of long paths and the increase of nontrivial symmetricity, and suggested that it can be viewed as a measure of regularity [35]. They also showed that the entropy (1) is upper bounded by \( \ln (n - 1) \) where \( n \) is the number of vertices of a graph.

However, most data representation are multidimensional, and using graph models to describe them may lose information [46]. A hypergraph is a generalization of a graph in which a hyperedge can join more than two vertices [5]. Thus, hypergraphs can represent multidimensional relationship unambiguously [16]. Examples of hypergraphs include co-authorship networks, film actor/actress networks and protein-protein interaction...

∗Department of Mathematics and Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor MI (can@csmich.edu).
†Department of Computational Medicine & Bioinformatics, Medical School and Department of Mathematics, University of Michigan, Ann Arbor MI (indikar@umich.edu).
networks [34], to name a few. The authors in [40] also mention that hypergraphs require less storage space than graphs which may accelerate computation. Significantly, a hypergraph can be represented by a tensor if its hyperedges contain the same number of vertices (referred to as a uniform hypergraph). Tensors are multidimensional arrays generalized from vectors and matrices, preserving multidimensional patterns and capturing higher-order interactions and couplings within multiway data [10]. Tensor decompositions such as CANDECOMP/PARAFAC decomposition (CPD), higher-order singular value decomposition (HOSVD) and tensor train decomposition (TTD) can not only facilitate computations, but they also contain contextual interpretations regarding the target data tensors [15,31,43,45].

Hypergraph entropy has been recently explored by Hu et al. [20] and Bloch et al. [7]. In particular, the authors in [20] employ the probability distribution of the vertex degrees to fit into the Shannon entropy formula and establish its lower and upper bounds for special hypergraphs. However, such entropy contains little information regarding the underlying structural properties of hypergraphs. Based on the previous works [20,32,35], we develop a novel tensor based hypergraph entropy which is able to decipher the topological attributes of uniform hypergraphs. The key contributions of this paper are as follows:

1. We propose a new notion of entropy for uniform hypergraphs based on the HOSVD.
2. We establish results on the lower and upper bounds of the proposed entropy, and provide formula for computing the entropies for complete uniform hypergraphs.
3. We create two simulated examples to demonstrate that the proposed entropy is a measure of regularity and relies on the vertex degrees, path lengths, clustering coefficients and nontrivial symmetricity for uniform hypergraphs. A “Watts-Strogatz” model for hypergraphs is also presented.

The paper is organized into five sections. We start with the basics of tensor algebra including tensor products, tensor unfolding and higher-order singular value decomposition in section 2.1. We then propose a new form of entropy for uniform hypergraphs in section 2.2 based on the tensor algebra. We also establish results on the lower and upper bounds of the tensor entropy. In section 3, two simulated examples, including a hyperedge growth model and a “Watts-Strogatz” model for hypergraphs, are presented. Finally, we discuss some directions for future research in section 4 and conclude in section 5.

2 Method

2.1 Tensor preliminaries

We take most of the concepts and notations for tensor algebra from the comprehensive works of Kolda et al. [29,25]. A tensor is a multidimensional array. The order of a tensor is the number of its dimensions. A k-th order tensor usually is denoted by $X \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$. It is therefore reasonable to consider scalars $x \in \mathbb{R}$ as zero-order tensors, vectors $v \in \mathbb{R}^n$ as first-order tensors, and matrices $A \in \mathbb{R}^{m \times n}$ as second-order tensors. For a third-order tensor, fibers are commonly named as column ($X_{j_1j_2}$), row ($X_{j_1j_3}$) and tube ($X_{j_2j_3}$), while slices are named as horizontal ($X_{j_1:}$), lateral ($X_{j_2:}$) and frontal ($X_{j_3:}$), see Figure 1. A tensor is called cubical if every mode is the same size, i.e., $X \in \mathbb{R}^{n \times n \times \cdots \times n}$. A cubical tensor $X$ is called supersymmetric if $X_{j_1j_2\cdots j_k}$ is invariant under any permutation of the indices, and is called diagonal if $X_{j_1j_2\cdots j_k} = 0$ except $j_1 = j_2 = \cdots = j_k$.

There are several notions of tensor products. The inner product of two tensors $X, Y \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$ is defined as

$$\langle X, Y \rangle = \sum_{j_1=1}^{n_1} \cdots \sum_{j_k=1}^{n_k} X_{j_1j_2\cdots j_k} Y_{j_1j_2\cdots j_k} \tag{2}$$

leading to the tensor Frobenius norm $\|X\|^2 = \langle X, X \rangle$. The matrix tensor multiplication $X \times_p A$ along mode $p$ for a matrix $A \in \mathbb{R}^{m \times n_p}$ is defined by

$$(X \times_p A)_{j_1j_2\cdots j_{p-1}i_{p+1}\cdots j_k} = \sum_{j_p=1}^{n_p} X_{j_1j_2\cdots j_{p-1}j_p+1\cdots j_k} A_{i_{p}}, \tag{3}$$
This product can be generalized to what is known as the \textit{Tucker product}, for \( A_p \in \mathbb{R}^{m_p \times n_p} \),
\[
X \times_1 A_1 \times_2 A_2 \times_3 \cdots \times_k A_k = X \times \{ A_1, A_2, \ldots, A_k \} \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_k}.
\]

\textit{Tensor unfolding} is considered as a critical operation in tensor computations \cite{2, 11, 29, 28, 39}. The \( p \)-mode unfolding of a tensor \( X \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k} \), denoted by \( X_p \), is defined by
\[
X_{j_1 j_2 \cdots j_k} = (X_p)_{j_1 j} \quad \text{for} \quad j = 1 + \sum_{m=1}^k (j_m - 1) \prod_{i=1}^{m-1} n_i.
\]

The ranks of the \( p \)-mode unfoldings are called \textit{multilinear ranks} of \( X \), which are related to the so-called Higher-Order Singular Value Decomposition (HOSVD), a multilinear generalization of the matrix Singular Value Decomposition (SVD) \cite{6, 17}.

\textbf{Theorem 1 (HOSVD \cite{17})}. A tensor \( X \in \mathbb{C}^{n_1 \times n_2 \times \cdots \times n_k} \) can be written as
\[
X = S \times_1 U_1 \times_2 \cdots \times_k U_k,
\]
where, \( U_p \in \mathbb{C}^{n_p \times n_p} \) are unitary matrices, and \( S \in \mathbb{C}^{n_1 \times n_2 \times \cdots \times n_k} \) is a tensor of which the subtensors \( S_{j_p=\alpha} \) obtained by fixing the \( p \)-th index to \( \alpha \), have the properties of
\begin{enumerate}
  \item all-orthogonality: two subtensors \( S_{j_p=\alpha} \) and \( S_{j_p=\beta} \) are orthogonal for all possible values of \( n \), \( \alpha \) and \( \beta \) subject to \( \alpha \neq \beta \);
  \item ordering: \( \| S_{j_p=\alpha} \| \geq \cdots \geq \| S_{j_p=n_p} \| \geq 0 \) for all possible values of \( p \).
\end{enumerate}

The Frobenius norms \( \| S_{j_p=\alpha} \| \), denoted by \( \gamma_j^{(p)} \), are the \( p \)-mode singular values of \( X \).

De Lathauwer et al. \cite{17} showed that the \( p \)-mode singular values from the HOSVD of \( X \) are the singular values of the \( p \)-mode unfoldings \( X_p \). In the following, we will use the notion of \( p \)-mode singular values as a main tool to define the tensor entropy for uniform hypergraphs.
2.2 Tensor entropy

Mathematically, a hypergraph $G = \{V, E\}$ where $V = \{1, 2, \ldots, n\}$ is the vertex set and $E = \{e_1, e_2, \ldots, e_m\}$ is the hyperedge set with $e_p \subseteq V$ for $p = 1, 2, \ldots, m$ [10, 21, 35, 22, 23, 47]. If all hyperedges contain the same number of nodes, i.e., $|e_p| = k$, $G$ is called a $k$-uniform hypergraph. Two vertices are called adjacent if they are in the same hyperedge. A hypergraph is called connected if given two vertices, there is a path connecting them through hyperedges. Given any $k$ vertices, if they are contained in one hyperedge, then $G$ is called complete. A $k$-uniform hypergraph can be represented by a $k$-th order $n$-dimensional supersymmetric adjacent tensor $A \in \mathbb{R}^{n \times n \times \cdots \times n}$ with

$$A_{j_1,j_2,\ldots,j_k} = \begin{cases} \frac{1}{(k-1)!} & \text{if } (j_1,j_2,\ldots,j_k) \in E \\ 0 & \text{otherwise} \end{cases}$$

(7)

The degree tensor $D$ of a hypergraph $G$, associated with $A$, is also a $k$-th order $n$-dimensional diagonal tensor with $D_{jj} = d$ equal to the number of hyperedges that consist of $v_j$ for $j = 1, 2, \ldots, n$. If $D_{jj} = d$ for all $j$, then $G$ is called $d$-regular. Similar to Laplacian matrices, the Laplacian tensor $L \in \mathbb{R}^{n \times n \times \cdots \times n}$ of $G$, which is supersymmetric, thus is defined as

$$L = D - A.$$  

(8)

**Definition 1.** Let $G$ be a $k$-uniform hypergraph with $n$ vertices. The tensor entropy for $G$ is defined by

$$S = -\sum_{j=1}^{n} \hat{\gamma}_j \ln \hat{\gamma}_j,$$

(9)

where, $\hat{\gamma}_j$ are the normalized $k$-mode singular values of $L$ such that $\sum_{j=1}^{n} \hat{\gamma}_j = 1$.

The convention $0 \ln 0 = 0$ is used if $\hat{\gamma}_j = 0$. The $k$-mode singular values of $L$ can be computed from the matrix SVD of the $k$-mode unfolding $L_k$, which results in a $O(n^{k+1})$ time complexity, see Algorithm 1. Since $L$ is supersymmetric, any mode unfolding of $L$ would yield the same unfolding matrix with the same singular values. Moreover, the tensor entropy can be viewed as a variation of von Neumann entropy defined for graphs, in which we regard $cL_kL_k^T$ as the density matrix for some normalization constant $c$ [9, 32, 35]. In particular, when $k = 2$, the tensor entropy is reduced to the classical von Neumann entropy for graphs. Like the eigenvalues of Laplacian matrices, the $k$-mode singular values play a significant role in identifying the structural patterns for uniform hypergraphs.

**Algorithm 1:** Computing tensor entropy from SVD

1: Given a $k$-uniform hypergraph $G$ with $n$ vertices
2: Construct the adjacent tensor $A \in \mathbb{R}^{n \times n \times \cdots \times n}$ from $G$ and compute the Laplacian tensor $L = D - A$
3: Find the $k$-mode unfolding of $L$, i.e., $L_k = \text{reshape}(L, n, n^{k-1})$
4: Compute the economy-size matrix SVD of $L_k$, i.e., $L_k = USV^T$ and let $\{\gamma_j\}_{j=1}^{n} = \text{diag}(S)$
5: Set $\hat{\gamma}_j = \frac{2}{\sum_{\ell=1}^{n} \gamma_{\ell}}$, and compute $S = \sum_{j=1}^{n} \hat{\gamma}_j \ln \hat{\gamma}_j$
6: return The tensor entropy $S$ of $G$.

**Lemma 1.** Suppose that $G$ is a $k$-uniform hypergraph with $k \geq 3$. If $L$ has a $k$-mode singular value zero, with multiplicity $p$, then $G$ contains $p$ number of non-connected vertices.

**Proof.** The result follows immediately from the definitions of Laplacian tensor and $k$-mode singular values/unfolding of $L$. 

**Proposition 1.** Suppose that $G$ is a $k$-uniform hypergraph with $k \geq 3$ and nonempty $E$. The minimum of the tensor entropy of $G$ is given by

$$S_{\text{min}} = \ln k.$$  

(10)
Proposition 2. Suppose that $G$ is a $k$-uniform hypergraph with $n$ vertices for $k \geq 3$. If $\text{mod}(n, k) = 0$, then the maximum of the tensor entropy of $G$ occurs at 1-regularity, which is given by

$$S_{\text{max}} = \ln n.$$  \hfill (11)

Proof. Since $G$ is a $k$-uniform hypergraph on $n$ vertices, the maximum multiplicity of the zero normalized $k$-mode singular value of $L$ is $n - k$ according to Lemma 1; the other normalized $k$-mode singular values of are necessarily \(\frac{1}{n-k}\). Hence, it is straightforward to show that $S_{\text{min}} = \ln k$. \hfill \Box

Proposition 3. Suppose that $G$ is a complete $k$-uniform hypergraph with $n$ vertices for $k \geq 3$. The tensor entropy of $G$ is equal to

$$S_c = \frac{(1-n)\alpha}{(n-1)\alpha + \beta} \ln \frac{\alpha}{(n-1)\alpha + \beta} - \frac{\beta}{(n-1)\alpha + \beta} \ln \frac{\beta}{(n-1)\alpha + \beta} \hfill (12)$$

where,

$$\alpha = \sqrt{\frac{\Gamma(n)(\Gamma(n-k+1) + \Gamma(n))}{\Gamma(k)^2(\Gamma(n-k+1))^2} - \frac{\Gamma(n-1)}{\Gamma(k)^2\Gamma(n-k)}}, \hfill (13)$$

$$\beta = \sqrt{\frac{\Gamma(n)(\Gamma(n-k+1) + \Gamma(n))}{\Gamma(k)^2(\Gamma(n-k+1))^2} + \frac{\Gamma(n-1)\Gamma(n-1)}{\Gamma(k)^2\Gamma(n-k)}}, \hfill (14)$$

and $\Gamma(\cdot)$ is the Gamma function.

Proof. Based on the definitions of Laplacian tensor and $k$-mode singular values/unfolding, we can easily write down the matrix $L_k^T L_k \in \mathbb{R}^{n \times n}$ which is given by

$$L_k^T L_k = \begin{bmatrix} d & \rho & \rho & \ldots & \rho \\ \rho & d & \rho & \ldots & \rho \\ \rho & \rho & d & \ldots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \ldots & d \end{bmatrix},$$

where, $d = (n-1)^2 + \frac{(n-1)}{k-1} \cdot \frac{1}{(k-1)^2}$, $\rho = \frac{T_{n-k}}{(k-1)^2}$, and $T_m$ are the $k$-simplex numbers (e.g., when $k = 3$, $T_m$ are the triangular numbers). Moreover, $L_k^T L_k$ is a special matrix, in which the eigenvalues are $d - \rho$ with multiplicity $n - 1$ and $d + (n - 1)\rho$ with multiplicity 1. Hence, the result is followed immediately, and we write all expressions using the Gamma function for simplicity. \hfill \Box

For arbitrary $k$-uniform hypergraph with $n$ vertices, $S_{\text{max}}$ may not achieve the upper bound $\ln n$, and $S_c$ could be smaller than the entropies of other $d$-regular hypergraphs for the same $n$ and $k$, which is different from the graph von Neumann entropy. Nevertheless, we will show evidence that the tensor entropy is also a measure of regularity for hypergraphs. Moreover, large tensor entropy is characterized by the high uniformity of vertex degrees, short path lengths, low clustering coefficients and high level of nontrivial symmetricity, and the tensor entropy is small for hypergraphs with large cliques, long path lengths and high clustering coefficients, i.e., hypergraphs in which the vertices form a highly connected cluster, see next section.
3 Experiments

3.1 Hyperedge growth model

We consider the case where the number of vertices is fixed and new hyperedges are iteratively added to the hypergraph. Figure 2 presents the hyperedge growth evolution of a 3-uniform hypergraph with 7 vertices. Panel A and B describe the tensor entropy maximization and minimization evolutions, respectively. In addition to plotting the two entropy trajectories, we also compute some statistics of the structural properties including average shortest path length, index of dispersion of the degree distribution and average clustering coefficient of the hypergraphs during the two evolutions. If the two vertices are disconnected, we set the distance between them to be 4 for the purpose of visualization. The index of dispersion of the vertex degree distribution measures the ratio of its variance to its mean, and we define the average clustering coefficient for hypergraphs as follows:

\[
C_j = \frac{|\{e_{ijk} : v_i, v_j, v_k \in N_j, e_{ijk} \in E\}|}{\binom{|N_j|}{3}} \Rightarrow C_{avg} = \frac{1}{n} \sum_{j=1}^{n} C_j,
\]

where, \(N_j\) is the set of vertices that are immediately connected to \(v_j\). If \(|N_j| < k\), we set \(C_j = 0\).

Let’s denote the hypergraphs that achieve maximum (or minimum) tensor entropy at step \(j\) as \(G_{max}^{(j)}\) (or \(G_{min}^{(j)}\)) for \(j = 1, 2, \ldots, 35\). Similar to maximizing graph entropy, the vertices of \(G_{max}^{(j)}\) tend to have “almost equal” or equal degree, see Panel E. In particular, \(G_{max}^{(j)}\) usually are the \(\frac{2}{3}\)-\(j\)-regular hypergraphs for the early stages of the evolution, i.e., \(j = 7, 14\), but as the hypergraph becomes more dense and complex, it is possible that \(G_{max}^{(j)}\) will miss the regularity, i.e., \(j = 21\). Based on the evolution in Panel A, maximizing the tensor entropy will first connect all the vertices and then prefer to choose lower degree vertices with larger average geodesic distances, i.e., finding the geodesic distances between each pair in the triples and taking the means. Similarly, the average geodesic distances may lose importance if one wants to predict the next step as the hypergraph becomes more complex. Furthermore, nontrivial symmetricity seems to play a role in maximizing the tensor entropy. For example, in \(G_{max}^{(3)}\), the vertices \(\{1, 2, 6\}\) and \(\{2, 4, 6\}\) have the same average geodesic distances (both are equal to \(\frac{2}{3}\)), and the maximized tensor entropy returns the more symmetric \(G_{max}^{(4)}\). We also find that candidate hyperedges that intersect more other existing hyperedges would return higher tensor entropy, which also explains the above example. However, there is one huge disparity between the graph von Neumann entropy and the tensor entropy. The tensor entropy can temporarily decrease during the maximizing process as seen in Panel C, which never happens when maximizing the graph entropy. We observe that once the hypergraph evolution reaches some regularity or high level of nontrivial symmetricity and the next step breaks such regularity or symmetricity, the tensor entropy will decrease. In other words, for these highly regular or highly symmetric \(G_{max}^{(j)}\), the tensor entropy \(S_{max}^{(j)}\) achieves local maxima. This is also why it is hard to determine the maximum of the tensor entropy for arbitrary uniform hypergraphs. On the other hand, \(G_{min}^{(j)}\) looks very similar to the graph entropy. Minimizing the tensor entropy would result in the formation of complete sub-hypergraphs (i.e., cliques), see Panel B, E and F. In particular, we can detect “huge” jumps and drops in the next steps after completions of the sub-hypergraphs in Panel C and D, respectively. In order to make the discoveries more convincing, we repeated the same processes for hypergraphs with different number of vertices and values of \(k\), and observed similar results.

3.2 “Watts-Strogatz” model for hypergraphs

We perform an experiment on a synthetic random hypergraph \(G\) with \(n = 100\) and \(k = 4\). Similar to the Watts-Strogatz model, the initial hypergraph is 2-regular with lattice structure. Let \(q\) be the number of hyperedges added to the hypergraph in order to form cliques in every five vertices, and \(p\) be the rewiring probability of hyperedges, see Panel A in Figure 3. Then \(G^{(q)}(p)\) denotes the random hypergraph generated by the rewiring probability \(p\) for different \(q\). In particular, when \(q = 3\), the tensor entropy \(S_{max}^{(3)}(0) = 4.5527\), the average clustering coefficient \(\alpha^{(3)}(0) = 0.7571\) and the average path length \(l^{(3)}(0) = 7.0606\). The goal is
Figure 2: Panel A describes the first five stages of the unfolding entropy maximization evolution, in which \( e_1 = \{1, 2, 3\} \), \( e_2 = \{5, 6, 7\} \), \( e_3 = \{3, 4, 5\} \), \( e_4 = \{2, 4, 6\} \) and \( e_5 = \{1, 4, 7\} \). Panel B reports the first five stages of the unfolding entropy minimization process, in which \( e_1 = \{1, 2, 3\} \), \( e_2 = \{2, 3, 4\} \), \( e_3 = \{1, 2, 4\} \), \( e_4 = \{1, 3, 4\} \) and \( e_5 = \{3, 4, 5\} \). Panel C, D, E and F are the plots of the unfolding entropy, average path length, index of dispersion and average clustering coefficient with respect to the hyperedge adding steps. At the final step, the hypergraph is complete.
to explore the relations between the tensor entropy, the average clustering coefficient and path length with increasing the hyperedge rewiring probability for different \( q \). We also calculate the small world coefficient for each random hypergraph which is defined by

\[
\sigma^{(q)}(p) = \frac{\alpha^{(q)}(p) / \alpha^{(q)}_{\text{rand}}}{l^{(q)}(p) / l^{(q)}_{\text{rand}}},
\]

where, \( \alpha^{(q)} \) and \( l^{(q)} \) are the average clustering coefficient and path length of its equivalent random hypergraph, respectively [24]. For each \( G^{(q)}(p) \), we compute its tensor entropy, average clustering coefficient, average path length and small world coefficient 10 times and take the means. Since \( q = 1 \) would likely result in non-connected random hypergraphs, we do not include this case.

The results are shown in Figure 3. In general, with increase of the rewiring probability \( p \), the tensor entropy decays for both \( q = 2 \) and 3, see Panel B. In particular, when \( p < 0.1 \), the tensor entropies for \( q = 2 \) are higher than these for \( q = 3 \). Based on the previous results, this may be because \( G^{(2)}(p) \) have lower average clustering coefficients with similar average path lengths compared to \( G^{(3)}(p) \). When \( p > 0.1 \), the tensor entropies for \( q = 3 \) is higher than these for \( q = 2 \) likely due to the fact that \( G^{(3)}(p) \) have lower average path lengths with similar average clustering coefficients compared to \( G^{(2)}(p) \). Moreover, the curves of the small world coefficient in Panel C are very similar to the one in the Watts-Strogatz model, in which it grows gradually for \( p < 0.1 \) and decreases quickly after \( p > 0.1 \). When the rewiring probability \( p \) is between 0.03 and 0.1, \( G^{(3)}(p) \) have apparent small world characteristics, e.g., \( \alpha^{(3)}(0.07) = 0.5488 \) and \( l^{(3)}(0.07) = 3.5748 \), over the 100 vertices. In addition, we find that the average clustering coefficient pattern is similar to tensor entropy. On the contrary, the average path length seems to have a different trend. It decreases faster at small \( p \) and more slowly than the average clustering coefficient at large \( p \), see Panel D. In the meantime, we also see a strictly positive correlation between the tensor entropy and the average clustering coefficient for \( q = 3 \) in Panel E.

\section{Discussion}

The \( k \)-mode singular values computed from the HOSVD of the Laplacian tensors can provide nice predictions of structural properties for uniform hypergraphs. Nevertheless, we are able to compute tensor eigenvalues directly from the Laplacian tensors, which may also contain information about the structural properties of uniform hypergraphs. The tensor eigenvalue problems of tensors were first explored by Qi [36, 37] and Lim [30] independently. The eigenvalues and eigenvectors of a \( k \)-th order supersymmetric tensor \( X \in \mathbb{R}^{n \times n \times \cdots \times n} \) are defined as follows:

\[
X \times v^{k-1} = \lambda v^{[k-1]},
\]

where, \( X \times v^{k-1} = X \times v \times v \times \cdots \times v \) and \( v^{[k-1]} \) stands for the element-wise \((k-1)\)-th power of \( v \). The eigenvalues \( \lambda \) could be complex, and if \( \lambda \) are real, we call them H-eigenvalues. There is a total of \( n(k-1)^{n-1} \) eigenpairs, and the sum of all eigenvalues are equal to \((k-1)^{n-1} \text{trace}(X)\) where \( \text{trace}(X) = \sum_{j=1}^{n} X_{jj\ldots j} \). Moreover, all the eigenvalues of \( X \) can be solved from a real characteristic polynomial defined in [36, 37]. Based on the formulation of tensor eigenvalues, we establish another new tensor entropy measure for uniform hypergraphs.

\begin{definition}
Let \( G \) be a \( k \)-uniform hypergraph with \( n \) vertices. The tensor entropy for \( G \) is defined by

\[
S = -\sum_{j=1}^{d} |\hat{\lambda}_j| \ln |\hat{\lambda}_j|,
\]

where, \( |\hat{\lambda}_j| \) are the normalized modulus of the eigenvalues of the Laplacian tensor \( L \) such that \( \sum_{j=1}^{d} |\hat{\lambda}_j| = 1 \), and \( d = n(k-1)^{n-1} \) is the total number of the eigenvalues.
\end{definition}

The convention \( 0 \ln 0 = 0 \) is used if \( |\hat{\lambda}_j| = 0 \). We repeat the first example above using the tensor eigenvalue based entropy, see Figure 4. The entropy minimization evolution trajectory is the same for the first five stages.
Figure 3: Panel A describes the cliques’ formation in the first five vertices of the hypergraph with the rewiring probability zero, in which \(e_1 = \{1, 2, 3, 4\}\), \(e_2 = \{2, 3, 4, 5\}\), \(e_3 = \{1, 2, 4, 5\}\), \(e_4 = \{1, 3, 4, 5\}\) and \(e_5 = \{1, 2, 3, 5\}\). The rest have the same patterns in every five vertices for a corresponding \(q\). Panel B plots the tensor entropy of random hypergraphs with different rewiring probabilities for different \(q\). Panel C plots the normalized small world coefficients of random hypergraphs with different rewiring probabilities for different \(q\). Panel D plots the ratio \(\alpha^{(3)}(p)/\alpha^{(3)}(0)\) and \(l^{(3)}(p)/l^{(3)}(0)\) of random hypergraphs with different rewiring probabilities. Panel E is the scatter plot between the tensor entropy and two ratios, the average clustering coefficient and the average path length for \(q = 3\).
in which cliques are formed. The maximization evolution trajectory becomes different from the fourth stage after the hypergraph is connected, in which short path lengths and high level of nontrivial symmetricity are no longer the factors that maximize the entropy. In addition, computing the eigenvalues of a tensor is an NP hard problem [19], and the spectral properties of tensors are not fully understood. The tensor eigenvalue solvers in the MATLAB Toolbox TenEig [12, 13] become computationally expensive and inaccurate when the size of Laplacian tensors are large. Hence, the tensor eigenvalue based entropy may not predict the structural properties of hypergraphs well compared to the $k$-mode singular value based entropy, but it might contain some other unknown features of hypergraphs that are required to explore in the future.

In reality, hypergraphs like co-authorship networks and protein-protein interaction networks exist in a very large scale, which would result in large Laplacian tensors, and computing matrix SVD to the corresponding unfolding matrices will definitely fail. For large graphs, one may use quadratic approximation to estimate von Neumann entropy [14, 32]. Hence, it is necessary to develop approximation forms for the tensor entropy with aid of tensor decompositions like CANDECOMP/PARAFAC decomposition and tensor train decomposition. These decomposition methods can help reveal hidden patterns and redundancies which reduces storage effort and enables efficient computations. For example, matrix SVD can be efficiently done in the tensor train format for large and sparse dataset [10]. It is therefore worthwhile to explore the roles of tensor decompositions in computing or approximating the tensor entropy for large uniform hypergraphs.

Furthermore, the effective resistance for hypergraphs is a very important topic for future work. In [18], the definition of the effective graph resistance is the sum of the effective resistances over all pairs of vertices, and the authors in [26] also prove that the effective graph resistance can be written in terms of the graph Laplacian eigenvalues, i.e.,

$$R_G = n \sum_{j=2}^{n} \frac{1}{\eta_j},$$

(19)

where, $n$ the total number of vertices and $\eta_j$ are the eigenvalues of the Laplacian matrix corresponding to a connected graph $G$. It will be interesting to establish similar relationships using $k$-mode singular values or eigenvalues of the Laplacian tensors to describe the robustness of uniform hypergraphs.
5 Conclusion

In this paper, we proposed a new notion of entropy for uniform hypergraphs based on the tensor higher-order singular value decomposition. The $k$-mode singular values of Laplacian tensors provide nice interpretations regarding the structural properties of hypergraphs. We found that the tensor entropy depends on the vertex degrees, path lengths, clustering coefficients and nontrivial symmetricity. Additionally, we investigated the lower and upper bounds of the entropy and provided the entropy formula for complete uniform hypergraphs. As mentioned in section 4, some approximation forms of the tensor entropy and the idea of effective resistance are required for large scale uniform hypergraphs and hypergraph robustness. Furthermore, controllability and influenceability of hypergraphs are also important for future research.

References

[1] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. Proceedings of the National Academy of Sciences, 97(21):11149–11152, 2000.

[2] B. W. Bader and T. G. Kolda. Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. ACM Transactions on Mathematical Software, 32(4):635–653, December 2006.

[3] A. Banerjee, A. Char, and B. Mondal. Spectra of general hypergraphs. Linear Algebra and its Applications, 518:14–30, 2017.

[4] A. Barabasi and E. Bonabeau. Scale-free networks. Scientific American, 288, 2003.

[5] C. Berge. Hypergraphs, Combinatorics of Finite Sets, Third Edition. North-Holland, Amsterdam, 1989.

[6] G. Bergqvist and E. G. Larsson. The higher-order singular value decomposition: Theory and an application [lecture notes]. IEEE Signal Processing Magazine, 27(3):151–154, May 2010.

[7] I. Bloch and A. Bretto. A new entropy for hypergraphs. In M. Couprie, J. Cousty, Y. Kenmochi, and N. Mustafa, editors, Discrete Geometry for Computer Imagery, pages 143–154, Cham, 2019. Springer International Publishing.

[8] S. Braunstein, S. Ghosh, T. Mansour, S. Severini, and R. C. Wilson. Some families of density matrices for which separability is easily tested. Physical Review A, 73:012320, 2006.

[9] S. Braunstein, S. Ghosh, and S. Severini. The laplacian of a graph as a density matrix: A basic combinatorial approach to separability of mixed states. Annals of Combinatorics, 10, 07 2004.

[10] C. Chen, A. Surana, A. Bloch, and I. Rajapakse. Data-guided model reduction for multilinear time invariant systems. submitted.

[11] C. Chen, A. Surana, A. Bloch, and I. Rajapakse. Multilinear time invariant system theory. In 2019 Proceedings of the Conference on Control and its Applications, pages 118–125.

[12] L. Chen, L. Han, T.-Y. Li, and L. Zhou. Teneig, 2015. version 2.0.

[13] L. Chen, L. Han, and L. Zhou. Computing tensor eigenvalues via homotopy methods. SIAM Journal on Matrix Analysis and Applications, 37(1):290–319, 2016.

[14] P.-Y. Chen, L. Wu, S. Liu, and I. Rajapakse. Fast incremental von neumann graph entropy computation: Theory, algorithm, and applications. In K. Chaudhuri and R. Salakhutdinov, editors, Proceedings of the 36th International Conference on Machine Learning, volume 97 of Proceedings of Machine Learning Research, pages 1091–1101, Long Beach, California, USA, 09-15 Jun 2019. PMLR.

[15] A. Cichocki, N. Lee, I. Oseledets, A. Phan, Q. Zhao, and D. P. Mandic. Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 2 Applications and Future Perspectives. now, 2017.
[16] J. Cooper and A. Dutle. Spectra of uniform hypergraphs. *Linear Algebra and its Applications*, 436(9):3268–3292, 2012.

[17] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000.

[18] W. Ellens, F. Spieksma, P. V. Mieghem, A. Jamakovic, and R. Kooij. Effective graph resistance. *Linear Algebra and its Applications*, 435(10):2491–2506, 2011. Special Issue in Honor of Dragos Cvetkovic.

[19] C. J. Hillar and L.-H. Lim. Most tensor problems are np-hard. *J. ACM*, 60(6):45:1–45:39, Nov. 2013.

[20] D. Hu, X. L. Li, X. G. Liu, and S. G. Zhang. Extremality of graph entropy based on degrees of uniform hypergraphs with few edges. *Acta Mathematica Sinica, English Series*, 35(7):1238–1250, Jul 2019.

[21] S. Hu and L. Qi. The eigenvectors associated with the zero eigenvalues of the laplacian and signless laplacian tensors of a uniform hypergraph. *Discrete Applied Mathematics*, 169:140–151, 2014.

[22] S. Hu and L. Qi. The laplacian of a uniform hypergraph. *Journal of Combinatorial Optimization*, 29(2):331–366, Feb 2015.

[23] S. Hu, L. Qi, and J. Xie. The largest laplacian and signless laplacian h-eigenvalues of a uniform hypergraph. *Linear Algebra and its Applications*, 469:1–27, 2015.

[24] M. D. Humphries and K. Gurney. Network ‘small-world-ness’: A quantitative method for determining canonical network equivalence. *PLOS ONE*, 3(4):1–10, 04 2008.

[25] M. Karsai, N. Perra, and A. Vespignani. Time varying networks and the weakness of strong ties. *Sci Rep*, 4, 2015.

[26] D. Klein and M. Randic. Resistance distance. *Journal of Mathematical Chemistry*, 12:81–95, 1993.

[27] J. Kleinberg. Navigation in a small world. *Nature*, 406, 2000.

[28] T. G. Kolda. Multilinear operators for higher-order decompositions. 2006.

[29] T. G. Kolda and B. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[30] L.-H. Lim. Singular values and eigenvalues of tensors: A variational approach. volume 2005, pages 129–132, 01 2006.

[31] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos. Mpca: Multilinear principal component analysis of tensor objects. *IEEE Transactions on Neural Networks*, 19(1):18–39, Jan 2008.

[32] G. Minello, L. Rossi, and A. Torsello. On the von Neumann entropy of graphs. *Journal of Complex Networks*, 7(4):491–514, 11 2018.

[33] N. D. Monnig and F. G. Meyer. The resistance perturbation distance: A metric for the analysis of dynamic networks. *Discrete Applied Mathematics*, 236:347–386, 2018.

[34] M. Newman. *Networks: An Introduction*. OUP Oxford, 2010.

[35] F. Passerini and S. Severini. The von neumann entropy of networks. 2008.

[36] L. Qi. Eigenvalues of a real supersymmetric tensor. *Journal of Symbolic Computation*, 40(6):1302–1324, 2005.

[37] L. Qi. Eigenvalues and invariants of tensors. *Journal of Mathematical Analysis and Applications*, 325(2):1363–1377, 2007.

[38] L. Qi. $H^+_h$-eigenvalues of laplacian and signless laplacian tensors. *Communications in Mathematical Sciences*, 12, 03 2013.

[39] S. Ragnarsson and C. Van Loan. Block tensor unfoldings. *SIAM J. Matrix Analysis Applications*, 33:149–169, 2012.
[40] I. Rajapakse and M. Groudine. On emerging nuclear order. *The Journal of Cell Biology*, 192(5):711–721, 2011.

[41] S. Ranshous, S. Shen, D. Koutra, S. Harenberg, C. Faloutsos, and N. F. Samatova. Anomaly detection in dynamic networks: A survey. *WIREs Comput. Stat.*, 7(3):223–247, May 2015.

[42] T. Ried and I. Rajapakse. The 4d nucleome. *Methods*, 123:1–2, 2017. The 4D Nucleome.

[43] W. Wang, V. Aggarwal, and S. Aeron. Principal component analysis with tensor train subspace. *Pattern Recognition Letters*, 122:86–91, 2019.

[44] D. Watts and S. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393, 1998.

[45] A. H. Williams, T. H. Kim, F. Wang, S. Vyas, S. I. Ryu, K. V. Shenoy, M. Schnitzer, T. G. Kolda, and S. Ganguli. Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron*, 98(6):1099–1115.e8, 2018.

[46] M. M. Wolf, A. M. Klinvex, and D. M. Dunlavy. Advantages to modeling relational data using hypergraphs versus graphs. In *2016 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–7, Sep. 2016.

[47] J. Xie and A. Chang. On the z-eigenvalues of the adjacency tensors for uniform hypergraphs. *Linear Algebra and its Applications*, 439(8):2195–2204, 2013.