On Representations and Spectral Inequalities for Non-Uniform Hypergraphs

A Thesis

Submitted for the Degree of

Doctor of Philosophy

in the

Faculty of Engineering

by

Ashwin Guha

Computer Science and Automation

Indian Institute of Science

Bangalore – 560 012, INDIA

October 2018

Contents

Acknowledgments

I thank my research supervisor Prof. Ambedkar Dukkipati for his constant guidance and support. I am grateful to TCS Research Labs for providing financial and material support through their PhD Fellowship grant. I am especially grateful to Mr. Sachin Parkhi from TCS Labs for his assistance. My thanks also to Muni Sreenivas Pydi and Biswajit Paria for their contribution in proving Cheeger inequality for tensors.

Statement of Originality

I, **Ashwin Guha**, with SR No. **04-04-00-14-12-12-1-09150** hereby declare that the material presented in the thesis titled **'On Representations and Spectral Inequalities for Non-Uniform Hypergraphs'** represents original work carried out by me in the **Department of Computer Science and Automation** at **Indian Institute of Science** during the years **2012 – 2018**.

With my signature, I attest that:

- I have not manipulated any of the data or results.
- I have not committed any plagiarism of intellectual property. I have clearly indicated and referenced the contributions of others.
- I have explicitly acknowledged all collaborative research and discussions.
- I understand that any false claim will result in severe disciplinary action.
- I understand that the work may be screened for any form of academic misconduct.

Date: Student Signature

In my capacity as supervisor of the above-mentioned work, I certify that the above statements are true to the best of my knowledge, and I have carried out due diligence to ensure the originality of the thesis.

Advisor Name: Advisor Signature

Abstract

Spectral graph theory involves the study of the eigenvalues and eigenvectors of graph connectivity matrices such as the adjacency matrix, Laplacian or the signless Laplacian. Spectral methods have often proved to be efficient and are widely used in solving problems where the underlying objects can be represented by graphs. While graph theory has a variety of applications, it has often been observed in many real-world instances that the pair-wise relationships captured in graphs do not describe the data in its entirety. To overcome this limitation, the notion of hypergraphs was introduced. Hypergraphs are a general version of a graph, where an edge may span more than two nodes. They have been studied extensively from a combinatorial perspective. Recently there has been renewed interest in applying spectral methods to problems in hypergraphs, especially hypergraph partitioning and community detection, which have applications in machine learning. In order to employ spectral techniques, a crucial issue that needs to be addressed is the appropriate representation of the hypergraphs.

In this work, we consider various representations of hypergraphs to study their spectral properties. These representations include some matrix-based representations such as clique expansion, star expansion and simplicial complexes as well as tensor representations. We present a comparative study of the representations and study how one can extend existing results for 2-graphs to hypergraphs, in particular, to non-uniform hypergraphs. We define a Laplacian in each of the representation and study its spectrum. We also provide bounds for the largest and the second smallest eigenvalue of the Laplacians in terms of each of the representations.

One of the most important results in spectral graph theory is the Cheeger inequality, which relates the isoperimetric number of a graph and the second smallest eigenvalue of the graph Laplacian. We provide a generalized version of Cheeger inequality for non-uniform hypergraphs using the weighted clique approach as well as using a tensor approach. This is our main contribution of this work. We then compare these results with the existing Cheeger inequality for simplicial complexes. In addition, we also provide a conjecture on a generalized Mixing Lemma for simplicial complexes.

Notation

$\mathbb{Z}$	The set of integers
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
n	Number of vertices
k	Size of edge in hypergraph
$k_{\min}$	Size of smallest edge in hypergraph
$k_{\max}$	Size of largest edge in hypergraph
I_n	$n \times n$ Identity matrix
1_n	All ones vector
A	Adjacency matrix
B	Incidence matrix
D	Degree matrix
L	Laplacian matrix
$\widehat{L}$	Normalized Laplacian matrix
$\mu_1 \geq \ldots \geq \mu_n$	Eigenvalues of adjacency matrix
$\lambda_1 \leq \ldots \leq \lambda_n$	Eigenvalues of Laplacian matrix
$\mathcal{A}$	Adjacency tensor
$\mathcal{L}$	Laplacian tensor

Chapter 1

Introduction

The story of graph theory begins as a puzzle in 18-th century Konigsberg. The puzzle was to start on any of the four land masses separated by the river Pregel and walk across each of the seven bridges that spanned the grounds exactly once and return to the starting point. The leading mathematician of the day, Leonhard Euler, took it upon himself to prove the impossibility of such a tour. He further gave the necessary conditions for such a tour to be possible. Thus began the field of study which became graph theory.

Later in mid 19-th century, graph theory was studied with reference to specific applications. Spanning trees were studied by Kirchoff while describing electrical networks. Cayley enumerated all possible spanning trees while describing the structure of hydrocarbons. Subsequently in the 20-th century, graph theory developed as a separate field of study in its own right.

In a nutshell, graph theory is the study of connections. These could be any kind of connections from the physical road network to the virtual world-wide-web. Graph theory has played an important role in various disciplines, from communications to production management to software engineering to machine learning.

The traditional combinatorial graph models the connections between two entities. In many cases, such a representation proves inadequate to describe the system. For example, consider the network of researchers. One could model the problem as a 2-graph, where two authors share an edge if they have a common publication. But this network is better represented using hypergraphs, where the researchers are the vertices and each publication is an edge that spans its authors. There is a difference between knowing authors A, B and C have a paper together and knowing A, B and C have worked with each other. A hypergraph captures that difference, whereas a 2-graph does not.

Hypergraphs are a generalization of combinatorial graphs, where an edge may span two or more vertices. Formally, a hypergraph $H(V, E)$ is a collection of vertices V and edges E, where each edge $e \in E$ is a subset of V. If each edge is of the same size k, then it is called a k-uniform hypergraph or a k-graph for short. If the edges are of different sizes, then it is called a non-uniform hypergraph.

Hypergraphs have proved to be an invaluable tool in understanding large data sets. There has been a recent surge in interest to study hypergraphs. Many reasons can be attributed to this development. One is the advance in computing technology, which has allowed previously intractable problems to be solved quickly and cheaply. Second is the profusion of data in its various formats, the so-called 'big data'. Traditional graphical models no longer prove adequate to understand the complexities of large volumes of data. Thirdly, the convergence and interdependence of various scientific disciplines has enabled hypergraphs to be employed in diverse areas: from VLSI design (Karypis et al., 1999), protein interaction (Ramadan et al., 2004), to image processing (Ducournau et al., 2012), to community detection in social media (Qian et al., 2009). Hypergraphs have enormous potential to alter the landscape of data processing.

We quickly give the intuition behind the kind of problems we encounter in hypergraph theory. Consider the earlier network of researchers. Suppose we want to identify the researchers who work in the same field or in the same institution. Consider a similar problem where have a pool of cricketers and we are given no information except the playing eleven for every match. We now wish to identify the bowlers and batsmen among the players. These are examples of classification problems.

Spectral Methods

For 2-graphs, it is well-known that many problems, such as 3-colouring or maximal clique detection, are NP-Hard. One approach to tackle such problems is to employ heuristics. Another approach is to use spectral algorithms.

To each graph, we define connectivity matrices such as the adjacency matrix or the Laplacian matrix. We compute the eigenvalues and eigenvectors of these matrices, which can be used to get approximate solutions to intractable problems. For example, the Wilf Theorem and Hoffman bound provide an upper and lower bound for the chromatic number of a graph in terms of the eigenvalues of the adjacency matrix.

Spectral methods first appeared in (Hückel, 1931). Later spectral graph theory became an important branch of algebraic graph theory. Initially, the spectra of the adjacency matrices were studied. Later, the Laplacians proved to be more useful in describing the connectivity of the graphs. The Laplacians for graphs are the discrete counterparts of the Laplacians for Riemannian manifolds. Spectral graph theory has since been applied in various problems such as parallel computation (Simon, 1991), graph visualization (Biggs et al., 1993) and graph clustering (Ng et al., 2002; Von Luxburg, 2007).

Spectral clustering has been useful in machine learning. Several algorithms have been proposed based on spectral methods in the recent years (Leordeanu and Sminchisescu, 2012; Papa and Markov, 2007; Zhou et al., 2007). Typically, in these algorithms, we define an adjacency or Laplacian matrix. We then use the use the top k eigenvalues and corresponding eigenvectors to construct k clusters.

The advantage of spectral methods is that the spectrum of a matrix can be computed in polynomial time. In addition, several bounds exist on different parameters of graphs with respect to the eigenvalues, which guarantee that spectral algorithms provide a solution that is close to an optimal one. A comprehensive study of spectral graph theory can be found in (Chung and Graham, 1997; Cvetkovic et al., 2009)

Cheeger Inequality

An important result in spectral graph theory is the Cheeger inequality. It provides a guarantee on the correctness of the algorithms. The *Cheeger constant* or the *isoperimetric number* of a graph G is defined as

$$\phi(G) = \min_{0<|A|\leq\frac{|V|}{2}} \frac{|E(A, V\backslash A)|}{|A|},$$

where for every $A \subseteq V$, $E(A, V\backslash A)$ denotes the boundary edges with one vertex in A and other vertex in $V\backslash A$.

The Cheeger inequality is a two-sided bound on the isoperimetric number ϕ in terms of the second smallest eigenvalue of the Laplacian, denoted by λ. Whereas computing ϕ is NP-Hard, the eigenvalue λ can be computed easily. This demonstrates the computational advantage that is typical in spectral algorithms. The result can be stated as follows. Suppose λ denotes the second smallest eigenvalue of the Laplacian of a graph G with maximum degree $d_{\max}$ and isoperimetric number ϕ, then,

$$\frac{\lambda}{2} \leq \phi \leq \sqrt{2\lambda d_{\max}}.$$

Challenges

Before we investigate specific problems on hypergraphs, we pause to reflect on the challenges that lie ahead. The first issue is that there is no single representation for hypergraphs. As we shall see in this thesis, unlike 2-graphs for which the connectivity matrices are intuitive and straightforward, in the case of hypergraphs, there are diverse representations.

Second, many terms and parameters for 2-graphs do not have counterparts for hypergraphs. For example, consider the problem of vertex colouring. In case of graphs, one tries to colour vertices such that no edge spans vertices of same colour. In the case of hypergraphs, do we want the edges to have at least two colours or do we need vertices to have colours different from each other? Or consider the problem of edge-cuts. In graphs, there is exactly only way to break an edge, but in hypergraphs, there are many ways to break an edge. One needs to define the terms and problems clearly before studying them further.

Finally, the efficacy of spectral methods is no longer guaranteed. The key feature of spectral algorithms for graphs is that eigenvalues of a matrix can be computed easily. In the case of tensor representation of hypergraphs, computing eigenvalues of connectivity tensors is intractable. Hence spectral methods appear to confer no advantage over other approaches.

In each instance, the challenges are overcome by modifying the definition or the problem depending upon the application. For instance, in the vertex colouring problem, if the hyperedges need to have at least two colours, it is called weak vertex colouring, and if the hyperedges needs the vertices to have distinct colours, then it is called strong vertex colouring.

1.1 Contributions of this thesis

In literature, much of the work has been focussed on uniform hypergraphs. Very few papers have been devoted entirely to non-uniform hypergraphs, and typically, these papers make use of just one kind of representation. The aim of this thesis is to study different representations of non-uniform hypergraphs in terms of their spectral properties.

We analyze unweighted and undirected non-uniform hypergraphs and define Laplacians for them. We then obtain bounds on the largest and second smallest eigenvalues of these Laplacians with respect to different graph parameters. We also obtain a version of Cheeger inequality in each of these representations.

Every effort has been made to ensure this thesis is self-contained. The only requirement on the part of the reader is a working knowledge of linear algebra and some basic knowledge of combinatorial graphs. In many instances, concepts have been borrowed from diverse fields. We provide the basic terminology associated with these concepts but discuss them only to the extent to which they aid in our study of hypergraphs.

Summary of results

1. We extend the results for uniform hypergraphs to non-uniform hypergraphs under the weighted clique expansion. We give upper bounds for the largest eigenvalue and second smallest eigenvalues.

2. We obtain bounds on the analytic connectivity for non-uniform hypergraphs using tensors.

3. We obtain a version of Cheeger inequality for hypergraphs under the weighted clique expansion. We also obtain an analogous result for non-uniform hypergraphs in tensor representation.

4. We state a conjecture on Mixing Lemma for simplicial complexes.

1.2 Organization of thesis

This work is organized as follows. Chapter 2 contains basic spectral graph theory. In this chapter we state the preliminary results for 2-graphs which we extend to hypergraphs. In Chapter 3, we discuss the matrix representations of hypergraphs. We extend known bounds for 2-graphs to hypergraphs under the weighted clique expansion and provide a Cheeger inequality. It also contains a brief discussion on other kinds of matrix representations. In Chapter 4, we discuss representation of hypergraphs using simplicial complexes. Although it is essentially a matrix representation, we devote a chapter due to its significance. We discuss the recent developments regarding Cheeger inequality and Mixing Lemma for simplices. We state a conjecture on a generalized Mixing Lemma. Chapter 5 contains the results on tensor representation of hypergraphs. We define eigenvalues for tensors and extend known results for uniform hypergraphs to general hypergraphs. In particular, we obtain bounds on the analytic connectivity for non-uniform hypergraphs and Cheeger inequality for non-uniform hypergraphs. In Chapter 6, we provide some concluding remarks and point out the future directions of this area.

Chapter 2

Preliminaries and Background

In this chapter, we provide the foundation for the main results of this thesis. The purpose of this chapter is twofold. First, we provide the essential concepts and results required for this work. Second, we provide a flavour of the theorems and proof techniques for hypergraphs. The proofs discussed in this thesis often resort to approaches that are well-known for 2-graphs. The contents of this chapter thus offer much insight and intuition into hypergraphs for a beginner. A fairly knowledgeable reader may be familiar with much of the material presented here and they may skim or skip altogether the first two sections of this chapter.

Spectral graph theory is a vast field by itself and it would be impossible to give a detailed survey here. The review provided in this chapter is far from exhaustive and details only those results which pertain to this thesis. The definitions and results mentioned here may be found in any standard textbook (Godsil and Royle, 2001; Brouwer and Haemers, 2011; Chung and Graham, 1997).

This chapter contains three sections. In Section 2.1, we start with the very basic definitions of a combinatorial graph and establish the necessary notation. In Section 2.2, we recall some elementary linear algebra and list a few simple results for graphs. In Section 2.3, we discuss two key results in spectral theory, namely the Cheeger inequality and the Mixing

Lemma. The Cheeger inequality will be revisited in each of the subsequent chapters. We conclude the chapter with a short introduction to hypergraphs in Section 2.4.

2.1 Basics of Graphs

A combinatorial graph $G(V, E)$ is a collection of vertices V and edges E. An edge $e = (u, v)$ is an ordered pair of vertices in the case of directed graphs, or simply the set $\{u, v\}$ in the case of undirected graphs. Graphs without self-loops and multiple edges across same pair of vertices are called simple graphs. If uv is an edge in G, we say u is *adjacent* to v, denoted by $u \sim v$, or that u is a *neighbour* of v. The set of vertices adjacent to a vertex v is called the *neighbourhood* of v, denoted by $N(v)$. A vertex is *incident* with an edge if it is one of the two endpoints of the edge. A graph G' is a *subgraph* of G if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. It is called an *induced subgraph* if two vertices of G' are adjacent if and only if they are adjacent in G.

In the case of weighted graphs, there is a real-valued function $W : V \times V \longrightarrow \mathbb{R}$ that assigns a weight $w(u, v)$ to each edge. For most practical applications, the weight function is non-negative. One may consider an unweighted graph simply as a weighted graph with weights in $\{0, 1\}$.

$$W : V \times V \longrightarrow \{0, 1\}$$

$$W_{uv} = \begin{cases} 1 & \text{if } uv \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The *degree* of a vertex v, denoted by $d(v)$, is the number of edges containing v. For a weighted graph, the degree is the sum of weights of edges incident on the vertex, i.e. $d(v) = \sum_{uv \in E} w_{uv}$. A graph is called *regular* if all vertices have same degree. In case of directed graphs, the *outdegree* is the number of directed edges emanating from the vertex and *indegree* is the number of directed edges converging on the vertex. A *walk* is a sequence of vertices $(v_1, \ldots, v_k)$ such that v_i is adjacent to v_{i+1} for $1 \leq i < k$. A walk where all vertices are distinct is called a *path*. The *distance* between a pair of vertices u and v is the length of the shortest path connecting them. The *diameter* is the greatest distance between any pair of vertices in the graph.

A graph is said to be *connected* if there exists a path between every pair of vertices in the graph. A *connected component* is a maximal connected subgraph of G. The *vertex connectivity* of a graph is the size of the smallest set of vertices whose removal disconnects the graph. A graph is called k-connected if its vertex connectivity is greater than or equal to k. The *edge connectivity* is the size of the smallest set of edges whose removal disconnects the graphs. The vertex-connectivity of a graph is less than or equal to its edge-connectivity.

We now define the connectivity matrices of a graph. Let n be the number of vertices $|V|$ and m be the number of edges $|E|$.

Definition 1. *The* ***adjacency matrix*** *A of a graph $G(V,E)$ is an $n \times n$ matrix over $\mathbb{R}$ defined as*

$$A_{uv} = \begin{cases} 1 & \text{if } uv \in E, \\ 0 & \text{otherwise.} \end{cases}$$

For weighted graphs, A_{uv} is assigned the weight of the edge w_{uv}. This matrix is symmetric for undirected graphs. The row sum and column sum for each vertex is its degree. The adjacency matrix captures many properties of the graph. For example, for the higher powers $i \geq 2$, $(A^i)_{uv}$ is the number of paths from u to v of length i.

Another useful connectivity matrix is the incidence matrix.

Definition 2. *The* ***incidence matrix*** *B of a graph $G(V,E)$ is an $n \times m$ matrix over $\mathbb{R}$, with rows indexed by vertices and columns indexed by edges and defined as*

$$B_{ue} = \begin{cases} 1 & \text{if } u \in e, \\ 0 & \text{otherwise.} \end{cases}$$

Each row sum is the degree of the vertex. Each column sum is two. Sometimes the edges are assigned an orientation to get the *oriented incidence matrix* B', with $B'_{ue} = 1$ and $B'_{ve} = -1$ for an edge $e = uv$.

The matrix most commonly used to study a graph is the Laplacian matrix. Originally, Laplacian operators were used in differential calculus, but have since been defined in discrete settings as well. The Laplacians describe the connectivity properties of a graph better than the adjacency. This matrix is the central object of study in this thesis. In the following chapters, a Laplacian is defined for hypergraphs with respect to different representations.

Definition 3. *The* ***Laplacian matrix*** L *of a graph* $G(V, E)$ *is an* $n \times n$ *matrix over* $\mathbb{R}$*, defined as* $L = D - A$*, where* D *is the diagonal matrix containing degrees of the vertices and* A *is the adjacency matrix. Written explicitly,*

$$L_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -a_{ij} & \text{if } i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

The Laplacian is related to the oriented incidence matrix as $L = B'B'^T$. The Laplacian is symmetric, with row and column sums zero. A related but less studied matrix is the signless Laplacian Q, defined as $Q = D + A$. This matrix is also symmetric and contains non-negative entries.

Example 1. *Consider the following graph. The connectivity matrices are as follows.*

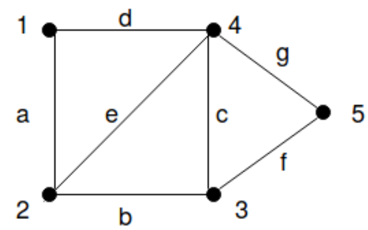

Figure 2.1: A simple graph

$A =$

	1	2	3	4	5
1	0	1	0	1	0
2	1	0	1	1	0
3	0	1	0	1	1
4	1	1	1	0	1
5	0	0	1	1	0

$B =$

	a	b	c	d	e	f	g
1	1	0	0	1	0	0	0
2	1	1	0	0	1	0	0
3	0	1	1	0	0	1	0
4	0	0	1	1	1	0	1
5	0	0	0	0	0	1	1

$L =$

	1	2	3	4	5
1	2	−1	0	−1	0
2	−1	3	−1	−1	0
3	0	−1	3	−1	−1
4	−1	−1	−1	4	−1
5	0	0	−1	−1	2

.

In many applications, normalized versions of the adjacency and Laplacian are used. The *normalized adjacency matrix* denoted by $\overline{A}$ is the row normalized matrix $D^{-1}A$. The Laplacian can be normalized in two ways. One method is to define $\overline{L} = I - D^{-1}A$. This definition is inspired by a random walk on the graph. The normalized Laplacian thus obtained is no longer symmetric. To retain the symmetry, the following definition is used.

Definition 4. *The* ***normalized Laplacian*** $\widehat{L}$ *is defined as*

$$\widehat{L} = I - D^{-1/2} A D^{-1/2}.$$

Written explicitly,

$$\widehat{L}_{ij} = \begin{cases} 1 & \text{if } i = j, \\ \frac{-1}{\sqrt{d_i d_j}} & \text{if } i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

2.2 Spectral theory for graphs

Let $M \in \mathbb{R}^{n \times n}$ be a square matrix. Suppose there exist a scalar $\lambda \in \mathbb{R}$ and a non-zero vector $x \in \mathbb{R}^n$ such that

$$Mx = \lambda x,$$

then λ is called an *eigenvalue* of M and x is called an *eigenvector* corresponding to λ. The eigenvalues are the roots of the *characteristic polynomial*, which is a polynomial in one variable of degree n defined as $\det(xI_n - M)$.

In addition, if the matrix is symmetric, i.e. $M = M^T$, as is often in the case of connectivity matrices of a graph, we can assume additional properties about the eigenvalues of M.

Theorem 2.2.1. *(Spectral Decomposition Theorem) Let $M \in \mathbb{R}^{n \times n}$ be a real and symmetric matrix. Then, there exist real numbers $\lambda_1, \ldots, \lambda_n$ and n mutually orthogonal unit vectors $u_1, \ldots, u_n$ such that u_i is an eigenvector for λ_i, for $1 \leq i \leq n$ and*

$$M = \sum_{i=1}^{n} \lambda_i u_i u_i^T = U \Lambda U^T,$$

where U is an orthonormal matrix whose columns are u_i and Λ is a diagonal matrix whose *entries are* λ_i*, for* $1 \leq i \leq n$*.*

An alternate characterization of eigenvalues is obtained through Rayleigh principle.

Definition 5. *The* ***Rayleigh quotient*** *of a non-zero vector* $x \in \mathbb{R}^n$ *with respect to a matrix* $M \in \mathbb{R}^{n\times n}$ *is defined as*

$$R_M(x) = \frac{x^T M x}{x^T x}.$$

The eigenvalues are the optimal values for the Rayleigh quotient as given by Courant-Fischer-Weyl theorem below.

Theorem 2.2.2. *Let* M *be an* $n\times n$ *real symmetric matrix with eigenvalues* $\lambda_1 \leq \ldots \leq \lambda_n$*. Then for* $1 \leq k \leq n$*,*

$$\lambda_k = \min_{U} \left\{ \max_{x\in U} \frac{x^T M x}{x^T x} \,:\, dim(U) = k \right\}$$

and

$$\lambda_k = \max_{U} \left\{ \min_{x\in U} \frac{x^T M x}{x^T x} \,:\, dim(U) = n - k + 1 \right\},$$

where U *is a subspace of* $\mathbb{R}^n$*. In particular,*

$$\lambda_1 \leq \frac{x^T M x}{x^T x} \leq \lambda_n,$$

and these bounds are attained when x *is an eigenvector.*

A real symmetric matrix M is said to be *positive definite* (or *positive semidefinite*) if the scalar $x^T M x$ is strictly positive (or non-negative) for all non-zero $x \in \mathbb{R}^n$. This condition also implies that the eigenvalues are positive (or non-negative).

We now state two important theorems from spectral theory.

Theorem 2.2.3. *(Gershgorin Theorem, (Brouwer and Haemers, 2011)) Let* A *be a complex* $n \times n$ *matrix, with entries* a_{ij}*. For* $i \in \{1, \ldots, n\}$*, let* $r_i = \sum_{j\neq i} |a_{ij}|$ *be the sum of the absolute values of the non-diagonal entries in the* i*-th row. Let* $D(a_{ii}, R_i) \subseteq \mathbb{C}$ *be a* *closed disc centered at a_{ii} with radius r_i. Then, every eigenvalue of A lies within at least one of the discs $D(a_{ii}, r_i)$, where $i \in \{1, \ldots, n\}$.*

Theorem 2.2.4. *(Perron-Frobenius Theorem, (Brouwer and Haemers, 2011)) Let A be a real $n \times n$ non-negative matrix such that for all indices i, j there is a positive integer k such that $(A^k)_{ij} > 0$. Then there is a (unique) positive real number λ_0 with the following properties:*

1. *There is a real vector x_0 with positive entries such that $Ax_0 = \lambda_0 x_0$.*
2. *λ_0 has geometric and algebraic multiplicity one.*
3. *For each eigenvalue λ of A, we have $|\lambda| \leq \lambda_0$.*
4. *Any non-negative eigenvector of A has eigenvalue λ_0.*

Let A and L be the adjacency and Laplacian matrices of a graph G as defined in the previous section. The Laplacian L is positive semidefinite, since the expression $x^T Lx$ can be rewritten as

$$x^T Lx = \sum_{uv \in E} (x_u - x_v)^2. \tag{2.1}$$

Let $\lambda_1 \leq \ldots \leq \lambda_n$ be the eigenvalues of L. From (2.1) above, we note that the smallest eigenvalue λ_1 is zero and the corresponding eigenvector is 1_n. Let $\mu_1 \geq \ldots \geq \mu_n$ be the eigenvalues of A. Since the trace of A is zero, we know that the sum of eigenvalues of A is zero. This implies that for a non-empty graph, the largest eigenvalue μ_1 is strictly positive and the smallest eigenvalue μ_n is strictly negative. In case of d-regular graphs, there is a one-to-one correspondence between λ_i and μ_i. Since $L = dI_n - A$, we have

$$\lambda_i = d - \mu_i, \text{ for all } 1 \leq i \leq n.$$

We state a few elementary results from spectral graph theory (Brouwer and Haemers, 2011). We give the spectra of a few simple graphs below. These graphs are useful in the sense that they provide an upper bound for the spectrum. In many spectral problems, these are the cases that represent the extremal cases, hence worthwhile to study. We consider graphs with n vertices. The exponent represents the multiplicity of the eigenvalue.

Proposition 2.2.5. *Let K_n denote the complete graph with all possible edges. The spectra of the adjacency and Laplacian matrices are $\{(n-1)^1, (-1)^{n-1}\}$ and $\{0^1, n^{n-1}\}$ respectively.*

Proposition 2.2.6. *Let $K_{m,n}$ denote the complete bipartite graph with partitions of sizes m and n. The spectra of the adjacency and Laplacian matrices are $\{\pm\sqrt{mn}, 0^{m+n-2}\}$ and $\{0^1, m^{n-1}, n^{m-1}, (m+n)^1\}$ respectively.*

Proposition 2.2.7. *Let S_n denote a star graph with $(n-1)$ spokes. The spectra of the adjacency and Laplacian matrices are $\{\pm\sqrt{n-1}, 0^{n-2}\}$ and $\{0^1, 1^{n-2}, n^1\}$ respectively.*

Proposition 2.2.8. *Let G be a graph with connected components $G_i, 1 \leq i \leq s$. Then, the spectra of the adjacency and Laplacian of G is the union of the spectra of corresponding matrices of G_i with the multiplicities added.*

Proposition 2.2.9. *The multiplicity of zero in Laplacian spectrum of a graph G equals the number of connected components of G.*

Proof. Suffices to show that for a connected graph, zero occurs with multiplicity exactly one. 1_n is an eigenvector corresponding to 0. $L = B'B'^T$, where B' is the oriented incidence matrix. $Lx = 0$ if and only if $B'^T x = 0$, that is, for every edge the vector x takes the same value on both endpoints. For connected graph, that means that $x = c \cdot 1_n$. □

The second smallest eigenvalue of Laplacian is called *algebraic connectivity*. It is also sometimes referred to as the spectral gap. This definition was introduced by Fiedler (1973) (also refer Lemma 3.2.9 in Chapter 3). This term measures the connectivity quantitatively as well: the larger λ_2 is, the more connected the graph. This relation is explicitly described by the Cheeger inequality in the next section.

Proposition 2.2.10. *The maximum eigenvalue λ_n is bounded by the maximum degree as follows.*

$$\lambda_n \leq 2\, d_{\max}.$$

Proof. Let x be the eigenvector of λ_n. Let x_i be the largest component. Assume $0 < x_i \leq 1$.

$$\begin{aligned}\lambda_n x_i = (Lx)_i = d_i x_i - \sum_{i \sim j} a_{ij} x_j &\leq d_i x_i + \sum_{i \sim j} |a_{ij} x_j| \\ \leq d_i x_i + \sum_{i \sim j} |a_{ij} x_i| &\leq 2 d_i x_i \leq 2\, d_{\max} x_i.\end{aligned}$$

□

Proposition 2.2.11. *The maximum eigenvalue λ_n is bounded by the sum of degrees as follows.*

$$\lambda_n \leq \max_{i \sim j} (d_i + d_j).$$

Proof. Consider $D^{-1}LD$.

$$(D^{-1}LD)_{ij} = \begin{cases} d_i & \text{if } i = j, \\ -a_{ij} \frac{d_j}{d_i} & \text{otherwise.} \end{cases}$$

Applying Gerschgorin theorem, there exists an i such that

$$|\lambda_n - d_i| \leq \sum_{i \sim j} \left| a_{ij} \frac{d_j}{d_i} \right| = \frac{\sum_j a_{ij} d_j}{\sum_j a_{ij}} \leq \max_{i \sim j} d_j.$$

Hence

$$\lambda_n \leq \max_{i \sim j} (d_i + d_j).$$

□

2.3 Key Results

One of the most important results in spectral graph theory is the Cheeger inequality, which was originally proved by Cheeger (1970) for Laplacians on Riemann manifolds. Subsequently, it was adapted for the discrete setting on graphs by Mohar (1989). The result provides a two-sided bound on the isoperimetric number or the Cheeger constant with respect to the spectral gap.

Cheeger constant is an important parameter that measures the connectivity in a graph. Graphs with high connectivity are known as expander graphs. These graphs are widely studied and have many applications in areas such as computer networks, error-correcting codes, pseudorandomness etc. Cheeger inequality is extremely useful in estimating the Cheeger constant of an expander graph. Computing Cheeger constant is NP-Hard (Mohar, 1989). However, the eigenvalues of the Laplacian can be computed efficiently. Cheeger inequality guarantees that combinatorial expansion is equivalent to spectral expansion, hence the second eigenvalue acts as a proxy for the Cheeger constant. In other words, a graph is highly connected if and only if the spectral gap is large.

Definition 6. *For a graph $G(V, E)$, the* ***Cheeger constant*** *or the* ***isoperimetric number*** *is defined as*

$$\phi(G) = \min_{0<|A|\leq\frac{|V|}{2}} \frac{|E(A, V\backslash A)|}{|A|},$$

where for every $A \subseteq V$, $E(A, V\backslash A)$ denotes the boundary edges with one vertex in A and other vertex in $V\backslash A$.

Some authors (e.g. Parzanchevski et al. (2016); Gundert and Szedlák (2014)) define it as

$$h(G) = \min_{0<|A|<|V|} \frac{|V|}{|V\backslash A|}\frac{|E(A, V\backslash A)|}{|A|}.$$

However, since the factor $\frac{|V|}{|V\backslash A|}$ lies between $1/2$ and 1, we have $\phi(G) \leq h(G) \leq 2\phi(G)$.

Sometimes the term *edge expansion* denoted by ε, is used to refer to the normalized Cheeger constant $h(G)/d$, where d is the average degree.

Theorem 2.3.1. *(Cheeger inequality) Let λ denote the second smallest eigenvalue of the Laplacian of a graph G with maximum degree $d_{\max}$ and isoperimetric number ϕ. Then,*

$$\frac{\lambda}{2} \leq \phi \leq \sqrt{2\lambda d_{\max}}.$$

In terms of h, this can be written as $\lambda \leq h \leq \sqrt{8\lambda d_{\max}}$. Equivalently,

$$\frac{\phi^2}{2d_{\max}} \leq \lambda \leq 2\phi \quad \text{or} \quad \frac{h^2}{8d_{\max}} \leq \lambda \leq h.$$

Mohar (1989) proved as slightly stronger version $\frac{\phi^2}{(2d_{max}-\lambda)} \leq \lambda$ (which holds with the exceptions of $G = K_1, K_2$ and K_3) which gives us

$$d_{max} - \sqrt{d_{max}^2 - \phi^2} \leq \lambda \leq 2\phi.$$

The isoperimetric number ϕ does not always represent the connectivity between subsets accurately. For example, consider the complete bipartite graph $K_{n,n}$. It has $\phi = n$, yet there exist sets of size $n/2$ with no edges across them. The connectivity across subsets is better described by the Mixing Lemma. It was first proposed by Alon and Chung (1988) to approximate graphs that behaved like random graphs.

Suppose $S, T \subset V(G)$ are disjoint subsets in a d-regular graph. In a random graph chosen from Erdös-Renyi model, the expected number of edges between S and T is $\frac{d}{n}|S||T|$. If the graph G is similar to a random graph, then the actual number of edges $E(S,T)$ is close to $\frac{d}{n}|S||T|$. The difference is known as *discrepancy*. Mixing Lemma bounds the discrepancy with respect to the second largest eigenvalue of the adjacency matrix.

Theorem 2.3.2. *(Mixing Lemma) For a d-regular graph $G(V,E)$ with disjoint subsets*

$S, T \subset V$ we have

$$\left| E(S,T) - \frac{d}{n}|S||T| \right| \leq \mu(G) \cdot \sqrt{|S||T|},$$

where $\mu(G) = \max\{\mu_2(A), |\mu_n(A)|\}$.

Another way to look at the result is that it shows the range of the spectrum of L. If the eigenvalues of L are concentrated, then the graph is almost random. The converse of the result has been proved by Bilu and Linial (2006). The Inverse Mixing Lemma is given below. It has since been generalized to hypergraphs by Cohen et al. (2014).

Theorem 2.3.3. *If G is an r-regular graph such that for every disjoint $S, T \subset V(G)$*

$$\left| E(S,T) - \frac{r}{n}|S||T| \right| \leq \rho|S||T|$$

for some positive real number ρ, then

$$\mu(G) = O(\rho(\log(r/\rho) + 1)).$$

2.4 Introduction to Hypergraphs

In this section, we give a brief introduction to hypergraphs. The definitions and results given here can be found in (Berge, 1984).

A hypergraph $H(V, E)$ is a collection of vertices V and edges E, where each edge $e \in E$ is a subset of V. A hypergraph is said to be *simple* if for all $e_i, e_j \in E$, $e_i \subset e_j$ implies $e_i = e_j$. A hypergraph is said to be *k-uniform* if $|e| = k$ for all $e \in E$. A vertex u is said to be *adjacent* to v if there exists an edge $e \in E$ such that $\{u, v\} \subset e$. The set of adjacent vertices of v is called *neighbourhood* of v, denoted by $N(v)$.

The *degree* of a vertex v is the number of edges containing v. A hypergraph is said to be *regular* if all vertices have the same degree.

We have the following simple relations between the size of the edge and the degrees of the vertices.

Proposition 2.4.1. *Let $H(V, E)$ be a k-uniform hypergraph. The n-tuple $d_1 \geq \ldots \geq d_n$ is the degree sequence if and only if $d_n \geq 1$ and $\sum_{i=1}^{n} d_i$ is a multiple of k.*

Proposition 2.4.2. *Given m integers $r_1, \ldots, r_m$ and n-tuple $d_1 \geq \ldots \geq d_n$, there exists a hypergraph $H(V, E)$ such that $d(v_i) = d_i$ for $i \leq n$ and $|e_j| = r_j$ for $j \leq m$ if and only if*

1. $\sum_{j=1}^{m} \min\{r_j, k\} \geq d_1 + \ldots + d_k,$ *for all* $k < n$,

2. $\sum_{j=1}^{m} r_j = d_1 + \ldots + d_n.$

Definition 7. *A hypergraph $H(V, E)$ is said to be* ***m-colourable*** *of there exists a partition $S_1, \ldots, S_m$ of V such that every edge intersects at least two classes of the partition. In other words, for all $e \in E$, $e \not\subset S_i$, for $1 \leq I \leq m$. The smallest positive integer I such that H is I-colourable is called the* ***chromatic number*** *of H, denoted by $\chi(H)$.*

Chapter 3

Matrix Representations

The simplest way to represent a hypergraph is by a matrix. The fundamental idea is to convert a hypergraph into a 2-graph, so that we may be able to apply existing results and strategies to solve hypergraph problems. The advantages of such an approach is apparent. First and foremost, it enables us to define the adjacency and Laplacian matrices for hypergraphs and directly apply spectral methods on hypergraphs with little or no modification. Secondly, matrices are compact data structures and there are many efficient algorithms to perform matrix calculations. While the focus of this thesis is on combinatorial results, one can appreciate how this type of representation may be useful in practice.

In this chapter, we introduce several matrix representations of hypergraphs. These include clique expansion, weighted clique expansion, star expansion, set-union representation, random walks and flattened tensors. Among these, we study the weighted clique expansion in depth. We provide the spectrum of complete graph, complete k-partite graph and star graph. We also provide a few bounds on the largest and second smallest eigenvalues of the Laplacian. We generalize existing connectivity bounds to non-uniform hypergraphs. Finally, we prove a Cheeger inequality for uniform and non-uniform hypergraphs under this representation.

This chapter is organized as follows. In Section 3.1, we list the various matrix representations. In Section 3.2, we present some properties of weighted clique expansion. Subsection 3.2.1 contains the preliminary results and subsection 3.2.2 contains the Cheeger inequality. We conclude the chapter with a brief discussion on the disadvantages of the weighted clique representation.

3.1 Matrix Representations

In this section we list some of the matrix-based representations. There are a few other approaches not listed here that are mostly a variation of clique or star expansion (Agarwal et al., 2006).

3.1.1 Clique Expansion

The naive and straightforward way to convert a hypergraph into a graph would be to consider a clique expansion, i.e. to consider each hyperedge as a clique in a 2-graph over the same set of vertices.

Let $H(V, E)$ be a hypergraph. We construct a 2-graph $G(V, E')$ such that for vertices $i, j \in V$, $ij \in E'(G)$ if and only if there exists $e \in E(H)$ such that $\{i, j\} \subset e$.

Example 2. *For the 3-graph with edges* $\{123, 345, 567, 287\}$*, the resulting clique-expanded 2-graph can be shown.*

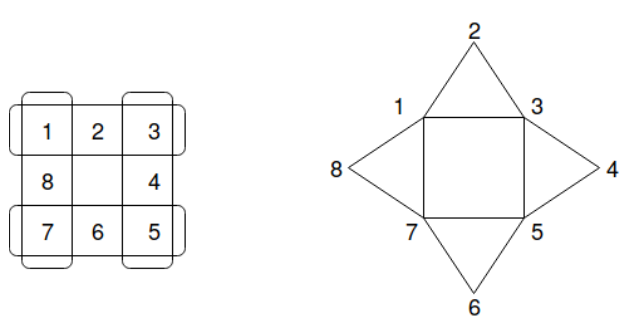

Figure 3.1: Clique expansion

The drawback of such a representation is evident. The mapping between the hypergraph and the resulting 2-graph is not injective. Many hypergraphs may give rise to the same 2-graph.

Example 3. *Consider two non-isomorphic 3-graphs* $G_1 = \{123, 124, 234\}$ *and* $G_2 = \{123, 124, 134, 234\}$. *They both give rise to the same 2-graph.*

This limitation is overcome by modifying the expansion to get a weighted graph.

3.1.2 Weighted Clique Expansion

This representation was introduced by Rodriguez (2003) and later used in (Rodriguez, 2009). Instead of an unweighted 2-graph, we construct a weighted 2-graph based on the hypergraph. For uniform hypergraphs, the 2-graph obtained is unique to the hypergraph. Most of this chapter is devoted to results pertaining to this representation.

Let $H(V, E)$ be a hypergraph. We define an $n \times n$ adjacency matrix A such that for all pairs of vertices i and j, a_{ij} is the number of edges containing i and j.

$$a_{ij} = \begin{cases} 0 & \text{if } i = j, \\ |\{e : \ i, j \in e\}| & \text{if } i \sim j, \\ 0 & \text{otherwise.} \end{cases}$$

For a 2-graph, this definition coincides with the familiar definition of adjacency matrix. We will revisit this representation in Section 3.2.

3.1.3 Clique Averaging

This description was introduced by Agarwal et al. (2005). The idea is similar to clique expansion. The aim is to convert a weighted k-graph to a weighted 2-graph. We make use of a clique incidence matrix of dimensions $\binom{n}{k} \times \binom{n}{2}$ to expand the hypergraph and map k-edges to corresponding 2-edges. This matrix is very sparse, and this property is exploited in the computation of hypergraph partitions. Typically, in clique expansion, when a k-edge is expanded as a clique of 2-edges, the weight of the k-edge is uniformly distributed across all the subedges, i.e. if e is a hyperedge and i, j are vertices, then

$$g(i,j) = \binom{n-2}{k-2}^{-1} \left(\sum_{i,j \in e} h(e) \right),$$

where $g(i, j)$ is the weight of the 2-edge (i, j) in the clique expanded graph and $h(e)$ is the weight of the k-edge e containing i and j in the original hypergraph.

In Clique Averaging, a modified function is used to distribute the weights. In addition, it has been shown that this representation differs from the usual clique expansion by multiplication of a symmetric matrix. Suppose w_H represents the weight of hyperedges and w_G represents the weight of the 2-edges in the expanded 2-graph, then

$$\Delta w_G = c_k w_H,$$

where Δ is the clique incidence matrix and c_k is some constant dependent on k. The resulting weights are related to the clique expansion weights by the following equation.

$$w_G^c = \Delta \Delta^T w_G,$$

where w_G^c is the weights obtained from clique expansion. It is claimed that this approach preserves the information otherwise lost in clique expansion.

This representation can also be extended to non-uniform hypergraphs without any modification. However, it comes with a caveat that the clique averaging is not effective on all hypergraphs but only those which can be partitioned easily. In practice, this often turns out to be a reasonable assumption. Hence, this is a useful representation which provides a good approximation of a hypergraph.

3.1.4 Star Expansion

In this representation, for each hyperedge we add an extra vertex and connect it to the existing vertices that are contained in the hyperedge. This results in a bipartite 2-graph of $(|V| + |E|)$ vertices, with partitions of sizes $|V|$ and $|E|$. This representation is unique to each hypergraph, hence captures the information of the hypergraph in its entirety. Also, this definition makes no distinction between uniform and non-uniform hypergraphs. In this representation, we may potentially have to add $\binom{n}{k}$ vertices to the graph. However, in practice, the number of vertices added is much lower, typically $O(n^2)$. This approach is also used in many of the hypergraph partitioning algorithms (Zien et al., 1999). Agarwal et al. (2006), the authors prove a non-intuitive result that the eigenvectors of the normalized Laplacian for the clique expansion graph are exactly the eigenvectors of the normalized Laplacian for standard star expansion. The proof given is for k-graph, but it appears the proof can be extended to non-uniform hypergraphs as well.

Example 4. *For the 3-graph $H = \{123, 345\}$ the star expansion is given below.*

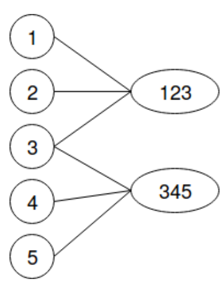

Figure 3.2: Star expansion

3.1.5 Set Union Representation

This representation was introduced by Steven Butler in his thesis (Butler, 2008). This representation has not gained much visibility, but it is an interesting approach worth noting. As with other representations, it is primarily defined for k-uniform hypergraphs. The definitions given here are extracted from Chapter 3, Section 3.5 of (Butler, 2008).

For a k-graph, we fix an integer i such that $0 < i < k$ and define an adjacency matrix

$A^{(i,k-i)}$ by indexing the rows with i-element subsets of vertices and indexing columns with $(k-i)$-element subsets of vertices. The entry a_{XY} is 1, if $X \cup Y$ forms a k-edge.

$$a_{XY} = \begin{cases} 1 & \text{if } X \cup Y \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The usual adjacency matrix for a 2-graph is simply $A^{(1,1)}$, the unique possible representation. The adjacency matrix thus obtained is no longer square, but rectangular of dimensions $\binom{n}{i}$ and $\binom{n}{k-i}$. Hence, we need to work with singular values instead of eigenvalues.

The *degree* of a *set* X is defined as $deg^{(i)}(X) = |\{Y : X \cup Y \in E\}|$. With this definition Butler proceeds to define a normalized adjacency matrix and obtains a version of Mixing Lemma for k-graphs. It is unclear how the Laplacian can be defined.

As for extending to non-uniform hypergraphs, we have two possibilities. One is to keep the definition unaltered and choose any fixed integer i, $0 < i < k_{\min}$, to obtain an adjacency matrix $A^{(i,k_{\max}-i)}$. The only difference is that X and Y need not be disjoint. As a result, a single edge may lead to more than one entry in each row.

Another possibility is to define an adjacency matrix $A^{(i,k-i)}$ for each edge size k. This maintains all the properties of k-uniform hypergraph, except we have several adjacency matrices, each corresponding to a particular edge size.

It appears that the intuition behind this representation is similar to that of coboundary operators for simplicial complexes discussed in the next chapter (See 4.1). Given the mathematical tools and considerable literature available for them, simplicial complexes are more useful compared to this representation. Even so, this matrix representation presents an interesting direction to explore in the future.

3.1.6 Random Walk Representation

For 2-graphs, the Laplacian is sometimes defined based on random walks on the graph (Chung and Graham, 1997; Lovász et al., 1993; Hoory et al., 2006). This leads to a Laplacian defined as $L = I - D^{-1}A$. A similar approach has been applied to obtain a normalized cut for hypergraphs Zhou et al. (2007).

Let B be a $|V| \times |E|$ incidence matrix. Let $d(v) = \sum_{v \in e} w(e)$, where $w(e)$ is the weight of the edge, and $\delta(e) = |e|$. Then $d(v) = \sum_{e \in E} w(e)B(v,e)$ and $\delta(e) = \sum_{v \in V} B(v,e)$.

Consider a random walk starting at vertex u. One of the hyperedges incident on u is chosen at random, with probability $w(e)/d(u)$ and one of the vertices v in the edge is chosen with probability $1/\delta(e)$. Thus, the transition probability matrix P whose entry P_{uv} denotes the walk from u to v is given by

$$P_{uv} = \sum_{e \in E} w(e) \frac{B(u,e)}{d(u)} \frac{B(v,e)}{\delta(e)}.$$

In matrix form, transition matrix can be written as

$$P = D_v^{-1} B W D_e^{-1} B^T,$$

where D_v is the diagonal matrix containing the degrees of vertices, D_e is the diagonal matrix containing the sizes of the edges and W is the diagonal matrix containing the weights of the edges. This leads to a normalized Laplacian defined as

$$\widehat{L} = I - D_v^{-1/2} B W D_e^{-1} B^T D_v^{-1/2}.$$

Spectral algorithms are applied to this normalized Laplacian to obtain clusters. A similar approach is taken in Lee et al. (2011).

3.1.7 Flattened Tensor Representation

This is not a representation of a hypergraph in itself, but an intermediate stage of computation. In many hypergraph clustering algorithms, a tensor is used to represent a hypergraph. To proceed with spectral clustering, the tensor is *flattened* or converted into a matrix by taking mode-1 product.

Definition 8. *For a tensor $\mathcal{A}$ of dimension n and order k, the* ***flattened matrix*** $\widetilde{A} \in \mathbb{R}^{n \times n^{k-1}}$ *is given by*

$$\widetilde{A}_{ij} = \mathcal{A}_{ii_2 \ldots i_m}, \ \textit{when } j = 1 + \sum_{\ell=2}^{k} (i_\ell - 1) n^{\ell - 2}.$$

Spectral algorithms such as higher-order SVD are performed on this flattened matrix to obtain the required clusters. Tensor decomposition using this idea was introduced by Govindu (2005) and has been used widely since. Tensor representation is discussed in depth in Chapter 5. Further results can be found in (Ghoshdastidar and Dukkipati, 2014; Ghoshdastidar et al., 2017; Anandkumar et al., 2016).

The following table summarizes the dimensions of the matrices in each of the representations.

Representation	Dimensions
Clique expansion	$n \times n$
Weighted clique	$n \times n$
Clique averaging	$\binom{n}{k} \times \binom{n}{2}$
Star expansion	$(n+e) \times (n+e)$
Set-union	$\binom{n}{i} \times \binom{n}{k-i}$
Random walk	$n \times n$
Flattened tensor	$n \times n^{(k-1)}$

3.2 Spectral Bounds in Weighted Clique Expansion

In this section, we study the weighted clique expansion in depth. First, we define a Laplacian matrix. For a weighted graph, the degree of a vertex is usually defined as the sum of weights of the edges incident on it. However, in this case, let the degree be defined as it is for 2-graph, which is the number of edges containing the vertex. We define $\delta_i = \sum_{j=1}^{n} a_{ij}$ as *Laplacian degree* of a vertex (Rodriguez, 2009).

Let Laplacian be defined as follows.

$$L_{ij} = \begin{cases} \delta_i & \text{if } i = j, \\ -a_{ij} & \text{otherwise.} \end{cases}$$

For a k-uniform graph, we have $\delta_i = (k-1)d_i$. For a general hypergraph, $\delta_i = \sum_{e \ni i} |e| - d_i$, which gives us the inequality, $(k_{\min} - 1)d_i \leq \delta_i \leq (k_{\max} - 1)d_i$. We may observe that the vertex with maximum degree need not be the vertex with maximum δ. For example in the graph below, $d_A > d_B$ but $\delta_B > \delta_A$.

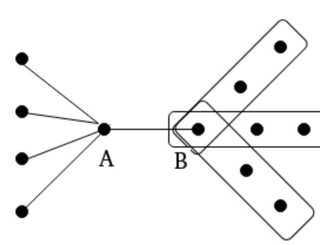

Figure 3.3: Degree and Laplacian degree

With this Laplacian, it is possible to extend several results of 2-graphs to hypergraphs.

3.2.1 Preliminary Results

We now consider the spectrum of a complete k-graph and a complete k-partite graph. The spectra for hypergraphs are similar to those of 2-graphs (Refer Chapter 2).

Proposition 3.2.1. *The Laplacian spectrum of complete k-graph with n vertices is 0 with multiplicity one and $n\binom{n-2}{k-2}$ with multiplicity $(n-1)$.*

Proof. The Laplacian matrix is as follows.

$$L_{ij} = \begin{cases} (n-1)\binom{n-2}{k-2} & \text{if } i = j, \\ -\binom{n-2}{k-2} & \text{otherwise.} \end{cases}$$

The vector 1_n^T corresponds to eigenvalue 0 and $(1,-1,0,\ldots,0)^T$ to $(1,0,\ldots,0,-1)^T$ correspond to eigenvalue $n\binom{n-2}{k-2}$. □

Example 5. *For a complete 3-graph with 5 vertices,*

$$L = \begin{pmatrix} 12 & -3 & -3 & -3 & -3 \\ -3 & 12 & -3 & -3 & -3 \\ -3 & -3 & 12 & -3 & -3 \\ -3 & -3 & -3 & 12 & -3 \\ -3 & -3 & -3 & -3 & 12 \end{pmatrix}$$

It is possible to get a similar result for a k-partite graph, i.e. a graph where the vertex set is divided into k-partitions and each edge consists of exactly one vertex from each partition.

Proposition 3.2.2. *For a complete k-partite graph with partitions of size $n_1,\ldots,n_k$, the Laplacian spectrum consists of 0 with multiplicity one, $(k-1)\left(\prod_{j=1}^{k} n_j\right)/n_i$ with multiplicity $(n_i - 1)$ for all $i = 1,\ldots,k$. The remaining $(k-1)$ eigenvalues are roots of the polynomial*

$$X^{k-1} - A_{k-2}X^{k-2} + \ldots + (-1)^{k-1}A_0,$$

where the coefficient A_i, the sum of products of roots taken $(k-1-i)$ at a time, is given *by*

$$A_i = (i+1)k^{k-2-i} \left(\prod_{j=1}^{k} n_j \right)^{k-2-i} \left(\sum_{1 \le j_1 \dots j_{i+1} \le k} n_{j_1} \dots n_{j_{i+1}} \right).$$

For example, consider a 5-partition where the sizes of the partitions are a, b, c, d, e. Then the eigenvalues are 0 with multiplicity one, $4abcd$ with multiplicity $(e-1)$, $4abce$ with multiplicity $(d-1)$, $4abde$ with multiplicity $(c-1)$, $4acde$ with multiplicity $(b-1)$ and $4bcde$ with multiplicity $(a-1)$. The remaining four eigenvalues are the roots of the polynomial

$$\begin{aligned} X^4 - 4\, S_4 X^3 + 3 \cdot 5\, S_5\, S_3 X^2 \\ - 2 \cdot 5^2\, S_5^2\, S_2 X + 1 \cdot 5^3\, S_5^3\, S_1. \end{aligned}$$

where S_i is the symmetric polynomial $S_{5,i}(a, b, c, d, e)$ for $1 \le i \le 5$.

In the case of a 3-graph, suppose the partitions are of sizes a, b, c, then the eigenvalues are $0, 2ab, 2bc, 2ac$ with multiplicities $1, (c-1), (a-1), (b-1)$ respectively. The remaining two eigenvalues are the roots of the quadratic polynomial $X^2 - BX + C$, where $B = 2(ab + bc + ac)$ and $C = 3abc(a + b + c)$. It can be verified that these roots are real and positive.

In the special case where two partitions have same size, i.e. a, a, b, the eigenvalues are 0 with multiplicity 1, $2ab$ with multiplicity $(2a-2)$, $2a^2$ with multiplicity $(b-1)$, $a(2a+b)$ and $3ab$ with multiplicities 1 each. The eigenvectors are given below.

$$\begin{aligned}
0 &: (1, \ldots, 1)^T, \\
2ab &: (1, -1, 0, \ldots, 0|0, \ldots, 0|0, \ldots, 0)^T, \ldots, (1, 0, \ldots, 0, -1|0, \ldots, 0|0, \ldots, 0)^T, \\
&: (0, \ldots, 0|1, -1, 0, \ldots, 0|0, \ldots, 0)^T, \ldots, (0, \ldots, 0|1, 0, \ldots, 0, -1|0, \ldots, 0)^T, \\
2a^2 &: (0, \ldots, 0|0, \ldots, 0|1, -1, 0, \ldots, 0)^T, \ldots, (0, \ldots, 0|0, \ldots, 0|1, 0, \ldots, 0, -1)^T, \\
a(2a+b) &: (1, \ldots, 1|1, \ldots, 1| -2a/b, \ldots, -2a/b)^T, \\
3ab &: (1, \ldots, 1| -1, \ldots, -1|0, \ldots, 0)^T.
\end{aligned}$$

The eigenvectors are similar for the general case.

Example 6. *For the complete 3-partite graph with partitions of sizes* $2, 3, 4$*, the Laplacian is given by*

$$\begin{pmatrix}
24 & 0 & -4 & -4 & -4 & -3 & -3 & -3 & -3 \\
0 & 24 & -4 & -4 & -4 & -3 & -3 & -3 & -3 \\
-4 & -4 & 16 & 0 & 0 & -2 & -2 & -2 & -2 \\
-4 & -4 & 0 & 16 & 0 & -2 & -2 & -2 & -2 \\
-4 & -4 & 0 & 0 & 16 & -2 & -2 & -2 & -2 \\
-3 & -3 & -2 & -2 & -2 & 12 & 0 & 0 & 0 \\
-3 & -3 & -2 & -2 & -2 & 0 & 12 & 0 & 0 \\
-3 & -3 & -2 & -2 & -2 & 0 & 0 & 12 & 0 \\
-3 & -3 & -2 & -2 & -2 & 0 & 0 & 0 & 12
\end{pmatrix}$$

and the eigenvalues are $0^1, 24^1, 16^2, 12^3$ *and* $B = 52$ *and* $C = 648$ *which gives the other eigenvalues* $20.7085, 31.2915$*.*

For a star k-graph with r spokes, the spectrum is obtained easily. The number of vertices in such a graph is given by $n = (k-1)r + 1$.

Proposition 3.2.3. *The Laplacian spectrum of a star* k*-graph with* r *spokes consists of*

0 and n with multiplicity one each, 1 with multiplicity $(r-1)$, and k with multiplicity $(k-2)r$.

The eigenvectors are given below.

$$\begin{aligned}
0 &: (1,\dots,1)^T,\\
n &: (n-1|-1,\dots,-1)^T,\\
1 &: (0|1,\dots,1|-1,\dots,-1|0,\dots,0)^T,\dots,(0|1,\dots,1|0,\dots,0|-1,\dots,-1)^T,\\
k &: (0|1,-1,0,\dots,0|0,\dots,0)^T,\dots,(0|1,0,\dots,0,-1|0,\dots,0)^T,\dots,\\
&: (0|0,\dots,0|1,-1,0,\dots,0)^T,\dots,(0|0,\dots,0|1,0,\dots,0,-1)^T.
\end{aligned}$$

Example 7. *For a star 4-graph with 3 spokes, the spectrum is $0^1, 1^2, 4^6, 10^1$. The Laplacian is given by*

$$\begin{pmatrix}
9 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1\\
-1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
-1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
-1 & -1 & -1 & 3 & 0 & 0 & 0 & 0 & 0 & 0\\
-1 & 0 & 0 & 0 & 3 & -1 & -1 & 0 & 0 & 0\\
-1 & 0 & 0 & 0 & -1 & 3 & -1 & 0 & 0 & 0\\
-1 & 0 & 0 & 0 & -1 & -1 & 3 & 0 & 0 & 0\\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -1 & -1\\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 3 & -1\\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 3
\end{pmatrix}$$

It is possible to get some simple bounds on the largest eigenvalue of the Laplacian. These are generalizations of the bounds for 2-graphs. The results mentioned here are valid for both uniform as well as non-uniform hypergraphs. The proofs are exactly the same as for weighted graphs.

Proposition 3.2.4. *The maximum eigenvalue λ_n is bounded by the maximum Laplacian degree as follows.*

$$\lambda_n \leq 2 \max_i \delta_i.$$

Proof. Let x be the eigenvector of λ_n. Let x_i be the largest component, i.e. $|x_i| \geq |x_j|$ for all $j = 1, \ldots, n$. Assume $0 < x_i \leq 1$.

$$\begin{aligned} \lambda x_i = (Lx)_i &= \delta_i x_i - \sum_{i \sim j} a_{ij} x_j \\ &\leq \delta_i x_i + \sum_{i \sim j} |a_{ij} x_j| \\ &\leq \delta_i x_i + \sum_{i \sim j} |a_{ij} x_i| \\ &\leq 2\delta_i x_i \leq 2 \max_i \delta_i x_i. \end{aligned}$$

□

Proposition 3.2.5. *The maximum eigenvalue λ_n is bounded by the sum of Laplacian degrees as follows.*

$$\lambda_n \leq \max_{i \sim j} (\delta_i + \delta_j).$$

Proof. Let $\Delta =$ diag$(\delta_1, \ldots, \delta_n)$. Consider $\Delta^{-1} L \Delta$.

$$(\Delta^{-1} L \Delta)_{ij} = \begin{cases} \delta_i & \text{if } i = j, \\ -a_{ij} \frac{\delta_j}{\delta_i} & \text{otherwise.} \end{cases}$$

Applying Gerschgorin theorem, there exists an i such that

$$|\lambda_n - \delta_i| \leq \sum_{i \sim j} |a_{ij} \frac{\delta_j}{\delta_i}| = \frac{\sum_j a_{ij} \delta_j}{\sum_j a_{ij}} \leq \max_{i \sim j} \delta_j.$$

Hence

$$\lambda_n \leq \max_{i\sim j}(\delta_i + \delta_j).$$

□

Let m_i denote the average degree of neighbours, i.e. $m_i = \frac{\sum_{i\sim j} d_j}{d_i}$. There are many upper bounds on the largest eigenvalue in terms of m_i. (For example, see (Aouchiche and Hansen, 2010) and the references therein).

We have the following bound for 2-graphs by Zhu (2010) that relates the largest eigenvalue with the degree and average degree of neighbours of each vertex.

Theorem 3.2.6.

$$\lambda_n \leq \max_{i\sim j}\left\{\frac{d_i(d_i+m_i)+d_j(d_j+m_j)-2\sum_{\ell\in N(i)\cap N(j)} d_\ell}{d_i+d_j}\right\},$$

where $N(i)$ is the set of vertices adjacent to i. This result holds for k-uniform hypergraphs in the same form. For non-uniform hypergraphs there is an additional factor of $\left(\frac{k_{\max}-1}{k_{\min}-1}\right)$. In order to prove the result for non-uniform case we make use of the lemma below .

Lemma 3.2.7. *(Theorem 2.3 in (Zhu, 2010)) Let $G=(V,E)$ be a simple graph. Let $f : V\times V \longrightarrow \mathbb{R}^+\cup\{0\}$ be a non-negative function that is positive on edges. Then λ_n is less than or equal to*

$$\max_{i\sim j}\left\{|N(i)\cap N(j)| + \frac{\sum_{\ell\in N(i)\setminus N(j)} f(i,\ell)+\sum_{\ell\in N(j)\setminus N(i)} f(j,\ell)}{f(i,j)}\right\}.$$

The generalization of Theorem 3.2.6 for non-uniform hypergraphs is presented below.

Theorem 3.2.8.

$$\lambda_n \leq \max_{i\sim j}\left\{\left(\frac{k_{\max}-1}{k_{\min}-1}\right)\left(\frac{d_i(d_i+m_i)+d_j(d_j+m_j)-2\sum_{\ell\in N(i)\cap N(j)} d_\ell}{d_i+d_j}\right)\right\}.$$

Proof. Substituting $f(i,j) = \delta_i + \delta_j$ in Lemma 3.2.7, we get the term

$$\begin{aligned}\sum_{\ell \in N(i) \setminus N(j)} (\delta_i + \delta_\ell) + \sum_{\ell \in N(j) \setminus N(i)} (\delta_j + \delta_\ell) \\ = \sum_{N(i)} (\delta_i + \delta_\ell) + \sum_{N(j)} (\delta_j + \delta_\ell) - \sum_{N(i) \cap N(j)} (\delta_i + \delta_j + 2\delta_\ell).\end{aligned}$$

Then

$$\begin{aligned}|N(i) \cap N(j)| + \frac{\sum_{N(i)} (\delta_i + \delta_\ell) + \sum_{N(j)} (\delta_j + \delta_\ell)}{\delta_i + \delta_j} - |N(i) \cap N(j)| \frac{(\delta_i + \delta_j)}{(\delta_i + \delta_j)} \\ - \frac{2 \sum_{N(i) \cap N(j)} \delta_\ell}{\delta_i + \delta_j} \\ = \frac{\sum_{N(i)} (\delta_i + \delta_\ell) + \sum_{N(j)} (\delta_j + \delta_\ell) - 2 \sum_{N(i) \cap N(j)} \delta_\ell}{\delta_i + \delta_j}.\end{aligned}$$

Since $\delta_i \leq (k_{\max} - 1)\, d_i$, for all $i = 1, \dots, n$, for the numerator we have,

$$\sum_{N(i)} (\delta_i + \delta_\ell) + \sum_{N(j)} (\delta_j + \delta_\ell) - 2 \sum_{N(i) \cap N(j)} \delta_\ell \tag{3.1}$$

$$\leq (k_{\max} - 1) \left(d_i(d_i + m_i) + d_j(d_j + m_j) - 2 \sum_{\ell \in N(i) \cap N(j)} d_\ell \right) \tag{3.2}$$

The denominator of the expression becomes

$$\delta_i + \delta_j \geq (k_{\min} - 1)(d_i + d_j),$$

since $\delta_i \geq (k_{\min} - 1)\, d_i$. Together with the equation (3.1) we get the required result. □

One can observe that each of the results mentioned is an improvement over the preceding bound. Not all such results, however, translate from 2-graphs to hypergraphs. For example consider the statement from Proposition 3.2.5.

For 2-graphs we have,

$$\lambda_n \leq \max_{i \sim j}(d_i + d_j).$$

However, this statement does not extend to hypergraphs in the form

$$\lambda_n \leq \max_{i_1 \ldots i_r \in E}(d_{i_1} + \ldots + d_{i_r}).$$

The counterexample is as follows.

Example 8. *Consider the graph* $H = \{123, 124, 235, 345\}$. $\lambda_n = 8.23 > 8 = \{d_2 + d_3 + d_5\}$.

A generalized result may exist in a different representation. For example, a similar upper bound for second smallest eigenvalue can be extended using a tensor representation. (See Theorem 5.3.1 in Chapter 5)

We now present some connectivity results for general hypergraphs. Let ∂S denote the edge boundary of S, that is, the set of edges with at least one vertex within S and one vertex outside of S, which is defined as $\partial S = E(S, V \backslash S)$ [1]. Various connectivity parameters can be defined in terms of ∂S. Some results have been provided for uniform hypergraphs by Rodriguez (2009). We provide the same for non-uniform hypergraphs.

First, we prove the following lemma.

Lemma 3.2.9. *For any* $S \subset V$*,*

$$\frac{4\lambda_2 |S|(n - |S|)}{n k_{\max}^2} \leq |\partial S| \leq \frac{\lambda_n |S|(n - |S|)}{n(k_{\min} - 1)}.$$

Proof. The bound holds for extreme cases $S = \emptyset$ and $S = V$. Let S be a proper subset of

[1] A modified definition is given in Chapter 4. Conceptually they are equivalent.

V. Let W_S be the indicator vector of S,

$$W_S = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\sum_{i \in V} \sum_{j \in V} (w_i - w_j)^2 = 2|S|(n - |S|)$. We have the following results by Fiedler (1973).

$$\lambda_2 = 2n \min \left\{ \frac{\sum_{i \sim j} a_{ij}(w_i - w_j)^2}{\sum_{i \in V} \sum_{j \in V} (w_i - w_j)^2} : w \neq c \cdot 1_n \text{ for } c \in \mathbb{R} \right\}.$$

$$\lambda_n = 2n \max \left\{ \frac{\sum_{i \sim j} a_{ij}(w_i - w_j)^2}{\sum_{i \in V} \sum_{j \in V} (w_i - w_j)^2} : w \neq c \cdot 1_n \text{ for } c \in \mathbb{R} \right\}.$$

Thus we have,

$$\lambda_2 \leq \frac{n \sum_{i \sim j} a_{ij}(w_i - w_j)^2}{|S|(n - |S|)} \leq \lambda_n.$$

In the sum $(w_i - w_j)^2$, only the boundary edges contribute to the sum. Let $e \in \partial S$ such that $|e \cap S| = k$. The edge e contributes $k(|e| - k)$ to the sum. The maximum value of this function over all the edges in E is $k_{\max}^2/4$, and the minimum value is $(k_{\min} - 1)$. Substituting we get the required inequality. □

The following results directly follow from Lemma 3.2.9 above. We prove a bound on the edge-density of a hypergraph. We obtain a lower bound in terms of λ_2 and an upper bound in terms of λ_n.

Theorem 3.2.10. *Let the edge-density of a set $S \subset V$ be defined as $\rho(S) = \frac{|\partial S|}{|S|(n-|S|)}$. Then,*

$$\frac{4\lambda_2}{n k_{\max}^2} \leq \rho(S) \leq \frac{\lambda_n}{n(k_{\min} - 1)}.$$

We also have an upper bound for the max-cut of a hypergraph.

Theorem 3.2.11. *Let max-cut of H be defined as* $\max\{|\partial S| : S \subset V\}$. *Then,*

$$\textit{max-cut}(H) \leq \frac{n\lambda_n}{4(k_{\min}-1)}.$$

Proof. From Lemma 3.2.9, we have

$$|\partial S| \leq \frac{\lambda_n |S|(n-|S|)}{n(k_{\min}-1)}. \tag{3.3}$$

The expression $|S|(n-|S|)$ attains maximum value at $|S| = n/2$ when n is even and $|S| = (n-1)/2$ when n is odd. The maximum value is $\frac{n^2}{4}$ when n is even and $\frac{n^2-1}{4}$ when n is odd. Hence $\frac{n^2}{4}$ is an upper bound for the expression $|S|(n-|S|)$. Substituting in (3.3) for the set S which gives the max-cut, we get

$$\begin{aligned}\text{max-cut}(H) &\leq \frac{n^2}{4}\frac{\lambda_n}{n(k_{\min}-1)} \\ &\leq \frac{n\lambda_n}{4(k_{\min}-1)}.\end{aligned}$$

□

3.2.2 Cheeger Inequality

We now present the Cheeger inequality in this representation. For 2-graphs we have from Chapter 2,

$$\frac{\phi^2}{2d_{\max}} \leq \lambda_2 \leq 2\phi.$$

Let the *isoperimetric number* be defined as $\phi(H) = \min_{S\subset V}\left\{\frac{|\partial S|}{|S|} : |S| \leq n/2\right\}$. The upper bound for λ_2 in terms of $\phi(H)$ can be obtained easily as follows.

Theorem 3.2.12.

$$\phi(H) \geq \frac{2\lambda_2}{k_{\max}^2}.$$

This once again follows from Lemma 3.2.9.

We now proceed to derive an equivalent lower bound for λ_2 using existing results for weighted graphs. First, let us consider the case for k-graphs.

Let H be a k-graph and let G_w be the weighted 2-graph we obtain upon the expansion. Suppose $d_{\max}(H)$ is the largest degree in H. Then $\delta_{\max}(G_w) = (k-1)d_{\max}(H)$. For weighted 2-graphs, the boundary of a subset $S \subset V$ is defined as the sum of weights of the boundary edges. Considering $S \subset V(H)$ as a subset of $V(G_w)$, we have $|\partial S|(G_w) = \sum_{\substack{i\in S\\ j\notin S}} a_{ij} \geq (k-1)|(\partial S)(H)|$.

$$\frac{|\partial S|(G_w)}{|S|} \geq (k-1)\frac{|\partial S|(H)}{|S|}.$$

In particular, this is true for the subset that attains the isoperimetric number.

$$\phi(G_w) \geq (k-1)\phi(H). \tag{3.4}$$

For weighted 2-graphs, we have the result in (Friedland and Nabben, 2002; Berman and Zhang, 2000),

$$\lambda_2(M^{-1}L) \geq \frac{1}{\max_i m_i}\left(\frac{\phi^2}{2\delta_{\max}}\right)$$

where M is a diagonal matrix $\mathrm{diag}(m_1, \ldots, m_n)$. When $M = I_n$, we have from (Mohar, 1997)

$$\lambda_2(L) \geq \frac{\phi^2}{2\delta_{\max}}. \tag{3.5}$$

Substituting G_w in (3.5) we get

$$\begin{aligned}\lambda_2(G_w) &\geq \frac{(\phi(G_w))^2}{2\delta_{\max}(G_w)}\\ &= \frac{(k-1)^2(\phi(H))^2}{2(k-1)d_{\max}} \quad \text{from(3.4)}.\end{aligned}$$

Thus for the hypergraph H, we have a lower bound second smallest eigenvalue .

$$\lambda_2(H) \geq \frac{(k-1)(\phi(H))^2}{2d_{\max}}. \tag{3.6}$$

For non-uniform hypergraphs,

$$\delta_{\max}(G_w) \leq (k_{\max}-1)d_{\max}(H)$$
$$\phi(G_w) \geq (k_{\min}-1)\phi(H).$$

We then have

$$\lambda_2 \geq \frac{(k_{\min}-1)^2\phi^2}{2(k_{\max}-1)d_{\max}}. \tag{3.7}$$

We can now give a Cheeger inequality for weighted clique expansion. From Theorem 3.2.12 and (3.7) we have the following theorem.

Theorem 3.2.13.

$$\frac{(k_{\min}-1)^2\phi^2}{2(k_{\max}-1)d_{\max}} \leq \lambda_2 \leq \frac{k_{\max}^2\phi}{2}. \tag{3.8}$$

Example 9. *Consider the 4-graph* $\{1234, 1345\}$*.* $k = 4$ *and* $d_{\max} = 2$*. The Laplacian* L *is given by*

$$\begin{pmatrix} 6 & -1 & -2 & -2 & -1 \\ -1 & 3 & -1 & -1 & 0 \\ -2 & -1 & 6 & -2 & -1 \\ -2 & -1 & -2 & 6 & -1 \\ -1 & 0 & -1 & -1 & 3 \end{pmatrix}$$

with spectrum $0, 3, 5, 8^2$*.* $\lambda_2 = 3$*. The minimum* $\phi(H) = 1$*, attained by the singleton* $\{2\}$*.*

$$\frac{3\cdot 1}{2\cdot 2} < 3 < \frac{4^2\cdot 1}{2}.$$

Note that $\{2, 5\}$ also gives $\phi = 1$, which may lead us to infer incorrectly that vertices 2 and 5 are tightly connected, while in fact there is no edge spanning these vertices. This occurs because $\{1, 3, 4\}$ are tightly connected. This anomaly is due to the small size of vertex set and is unlikely to occur in real-life situations.

3.2.3 Discussion

It appears that the introduction of terms $k_{\min}$ and $k_{\max}$ necessarily slackens the bounds. The equality is attained only in k-uniform hypergraphs.

Example 10. *Consider the following 3-graph* $G = \{123, 234, 456, 156\}$. *The spectrum of*

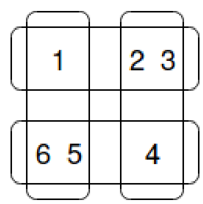

Figure 3.4: A 3-uniform graph

the Laplacian is $0^1, 2^1, 4^1, 6^3$. *The set* $A = \{1, 4\}$ *attains the upper bound of Lemma 3.2.9.*

$$|\partial A| = 4 = \left\lfloor \frac{6 \cdot 2 \cdot 4}{2 \cdot 6} \right\rfloor.$$

The set $B = \{1, 2, 3\}$ *attains the lower bound.*

$$|\partial B| = 2 = \left\lceil \frac{4 \cdot 2 \cdot 3 \cdot 3}{8 \cdot 6} \right\rceil.$$

Now consider the following non-uniform graph $G' = \{123, 456, 34, 1256\}$. *The spectrum*

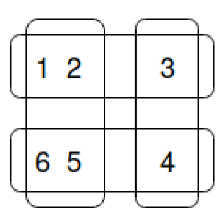

Figure 3.5: A non-uniform graph

of the Laplacian is $0^1, 3^2, 6^1, 7^2$. *The set* $A = \{1, 2, 4\}$ *attains the maximum.*

$$|\partial A| = 4 < \left\lfloor \frac{7 \cdot 3 \cdot 3}{1 \cdot 6} \right\rfloor .$$

Similarly, the singleton set $B = \{3\}$ *attains the minimum.*

$$|\partial B| = 2 > \left\lceil \frac{4 \cdot 3 \cdot 1 \cdot 5}{4 \cdot 4 \cdot 6} \right\rceil .$$

Note that if we had used $k_{\min}^2$ *in the denominator instead of* $k_{\max}^2$*, the lemma would no longer be true.*

Another crucial property in this representation that does not translate from uniform to non-uniform hypergraphs is the uniqueness. Consider the two different hypergraphs below.

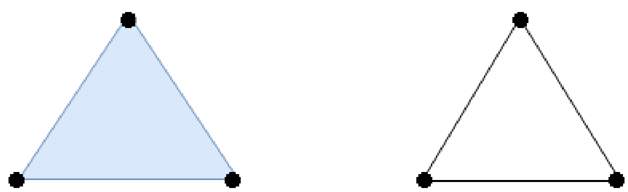

Figure 3.6: A single 3-edge and three 2-edges

They both give rise to the same matrix.

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Hence it is not possible to distinguish between the two hypergraphs using the adjacency matrix from this representation. This runs counter to our fundamental principle that the 3-edge is different from a triangle made of 2-edges. The idea of weighted clique expansion was introduced to circumvent this problem in the first place.

Also, this raises an apparent contradiction in our results proved above, as to how the results could be valid if the underlying matrices do not provide a true description. Since the adjacency matrices are same, the Laplacian and its eigenvalues will be identical to both the graphs, but the connectivity properties differ for both graphs. This can be explained when we observe that the quantity $|\partial S|$ differs in each case. For example, consider the singleton set $\{1\}$ in each of the two graphs. In the 3-graph, $|\partial S|$ consists of a single 3-edge, whereas in the 2-graph $|\partial S|$ is two (2-edges). Hence, the results are valid, though their applications may differ depending on the kind of hypergraph in question.

Chapter 4

Simplicial Complex

Simplicial complexes are higher dimensional solids and they are useful in visualizing hypergraphs. These solids are endowed with a certain algebraic structure that can be exploited to understand the combinatorial properties of hypergraphs. Simplicial complexes lie at the intersection of various disciplines of mathematics. One might study simplicial complexes as objects in geometry, or topology, or algebra. Our approach is through combinatorics.

A simplicial complex consists of smaller building blocks known as simplices. One may picture a simplex as a higher dimensional triangle or tetrahedron, with each vertex connected to every other vertex. In a hypergraph, viewing each edge as a simplex, the simplicial complex thus obtained uniquely represents the hypergraph. A 2-edge would correspond to a line segment, a 3-edge to a triangle and so on. This way a simplicial complex captures the the structure of the hypergraph more faithfully compared to other matrix representations. For example, a 3-edge as a triangle offers more insight than three pairs of 2-edges obtained in clique expansion.

The faces across different dimensions are connected using boundary maps and their dual known as coboundary maps. These provide the complex with an algebraic structure or homology. The Laplacians of graphs can be generalized to hypergraphs by defining them using boundary maps. Using the spectral properties of these Laplacians, we hope to infer the combinatorial properties of the hypergraph. As in other representations, the results of graphs do not always translate smoothly in the hypergraph setting. In that case, we modify the definitions to obtain similar meaningful results.

In this chapter, we provide some spectral results for hypergraphs based on a representation using simplicial complexes. Although simplicial complexes have been well-studied, their application to spectral hypergraph theory appears to be fairly recent (Horak and Jost, 2013). Many significant results have been established in this area (Steenbergen et al., 2014). In this thesis, we shall restrict our attention to two major results, namely the Cheeger inequality and Mixing Lemma. We provide an alternative proof to obtain the spectrum of a complete simplex from first principles. We state a conjecture on a generalized Mixing Lemma.

This chapter is organized as follows. In Section 4.1, we provide the basic definitions and terminologies. These are standard definitions that can be found in any textbook on algebraic topology (Munkres, 1984; Hatcher, 2002). The additional notations given in this chapter are borrowed from (Horak and Jost, 2013; Gundert and Szedlák, 2014; Parzanchevski, 2017). In Section 4.2, we provide the spectrum of a complete simplex. Section 4.3 contains the recent developments in literature concerning Cheeger inequality for simplicial complexes. Section 4.4 contains a discussion on Mixing Lemma for hypergraphs. We conclude the chapter by providing a conjecture on Mixing Lemma for simplicial complexes.

4.1 Preliminaries

Let V be a finite set. An *abstract simplicial complex* X on a finite set V is a collection of subsets of V, which is closed under inclusion, *i.e.* $\sigma \subset \tau \in X$ implies $\sigma \in X$. An element $\sigma \in X$ is called a *simplex* or *face* of X. The *dimension* of a simplex σ denoted by $\dim \sigma$ is $|\sigma| - 1$. 0-faces correspond to the elements of V and are usually called vertices. The dimension of the complex X is the maximal dimension of its faces, *i.e.* $\dim X = \max_{\sigma \in X} \dim \sigma$.

The collection of all i-faces of simplicial complex X is denoted by X_i. The faces that are maximal under inclusion are called *facets*. A complex X is called *pure* if all facets have the same dimension. This corresponds to a uniform hypergraph where all edges are of same size.

Two i-faces are said to be *lower adjacent* if they both share a common $(i-1)$-face and are said to be *upper adjacent* if they both are part of a common $(i+1)$-face. Note that upper adjacency implies lower adjacency. The *degree* of an i-face is the number of $(i+1)$-faces containing it, *i.e.* $\deg \sigma = |\{\tau \supset \sigma : \dim \tau = \dim \sigma + 1\}|$.

The collection of all simplices of dimension at most i is called the *i-skeleton* of X denoted by $X^{(i)}$. We define $X_{-1} = \{\emptyset\}$. The *complete k-simplex* denoted by K_n^k contains all possible i-faces for all $i \leq k$. The *link* of a face σ in X is defined as $\mathrm{lk}(\sigma, X) = \{\tau \in X : \sigma \cup \tau = X, \sigma \cap \tau = \emptyset\}$.

By assuming a linear order on the set of vertices, it is possible to give an *orientation* to each face of X. An oriented face $\sigma = \{v_0, \ldots, v_i\}$ is denoted by $[v_0, \ldots, v_i]$. Two orderings of the vertices are equivalent if they differ by an even permutation, and opposite otherwise. For an i-face $\tau = [v_0, \ldots, v_i]$ and $(i-1)$-face σ, the *oriented incidence number* is defined as $[\tau : \sigma] = (-1)^j$, where $\sigma = \tau \setminus v_j$ for some $j = 0, \ldots, i$ if $\sigma \subset \tau$ and zero otherwise, *i.e.* if $\sigma \not\subseteq \tau$.

Given a k-complex X, for all $i = 0, \ldots, k$, let $C_i(X, \mathbb{F})$ denote the set of linear combinations of i-faces over a field $\mathbb{F}$. The vector space C_i is known as the *i-th chain group* of X over $\mathbb{F}$. Usually, the chain group is defined over $\mathbb{R}$ or $\mathbb{Z}_2$. The set of i-faces X_i forms a basis for C_i, and the dimension of C_i is $|X_i|$.

The dual space of C_i which is the set of linear functions from C_i to $\mathbb{F}$, is called the *i-th co-chain group*, denoted by $C^i(X, \mathbb{F})$, for all $i = 0, \ldots, k$. It is possible to map each element

$\tau \in X_i$ to a unique function $e_\tau \in C^i$ defined as follows:

$$e_\tau(\sigma) = \begin{cases} 1 & \text{if } \tau = \sigma, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, there is a one-to-one correspondence between i-faces and functionals acting on them. Functions of the form e_τ are known as *elementary cochains*. Clearly, $\dim C^i = \dim C_i$. We define the one-dimensional vector space $C^{-1}(X, \mathbb{F})$ to be generated by the identity function on the empty simplex. Hence, $C^{-1}(X, \mathbb{F}) \cong \mathbb{F}$.

Definition 9. *The coboundary operator* $\delta_i : C^i(X, \mathbb{F}) \longrightarrow C^{i+1}(X, \mathbb{F})$ *is the linear function given by*

$$(\delta_i f)(\tau) = \sum_{\sigma \in X_i} [\tau : \sigma] f(\sigma),$$

for all $\tau \in X_{i+1}, f \in C^i$ *and* $-1 \leq i \leq k$.

Explicitly, this means

$$(\delta_i f)([v_0, \dots, v_i]) = \sum_{j=0}^{i} (-1)^j [v_0, \dots, \hat{v}_j, \dots, v_i],$$

where $\hat{v}_j$ denotes that the vertex v_j has been omitted.

Each coboundary operator connects the i-faces to $(i+1)$-faces forming a sequence which is called the *augmented cochain complex*. We define $\delta_i = 0$, for all other values of i to bookend the sequence.

$$0 \longrightarrow C^{-1} \dots \longrightarrow C^{i-1} \xrightarrow{\delta_{i-1}} C^i \xrightarrow{\delta_i} C^{i+1} \longrightarrow \dots C^k \longrightarrow 0.$$

We can define an inner product on the vector spaces C^i as follows.

$$\langle f, g \rangle = \sum_{\tau \in X_i} f(\tau) g(\tau),$$

for all $0 \leq i \leq k$.

Definition 10. *The boundary operator* $\partial_i : C^i(X, \mathbb{F}) \longrightarrow C^{i-1}(X, \mathbb{F})$ *is the dual operator of the coboundary operator* δ_{i-1} *such that*

$$\langle \partial_i f, g \rangle_{C^{i-1}} = \langle f, \delta_{i-1} g \rangle_{C^i},$$

for $f \in C^i$ *and* $g \in C^{i-1}$.

$$0 \longleftarrow C^{-1} \ldots \longleftarrow C^{i-1} \underset{\partial^i}{\longleftarrow} C^i \underset{\partial^{i+1}}{\longleftarrow} C^{i+1} \longleftarrow \ldots C^k \longleftarrow 0.$$

The boundary operator can also be viewed to act on the i-faces as a function that determines which $(i-1)$-faces form the boundary of an i-face. This is similar to the incidence matrix of a 2-graph.

$$\begin{aligned} \partial_i : C_i &\longrightarrow C_{i-1} \\ [v_0, \ldots, v_i] &\longmapsto \sum_{j=0}^{i} (-1)^j [v_0, \ldots, \hat{v}_j, \ldots, v_i], \end{aligned}$$

where $\hat{v}_j$ denotes that the vertex v_j has been omitted.

Example 11. *The boundary matrix* $\partial_2 = \delta_1^*$ *of the complete simplex containing all 2-edges and 3-edges over* $\{1, 2, 3, 4\}$ *is given below.*

$$\partial_2 = \left(\begin{array}{c|cccc} & 123 & 124 & 134 & 234 \\ \hline 12 & 1 & 1 & 0 & 0 \\ 13 & -1 & 0 & 1 & 0 \\ 23 & 1 & 0 & 0 & 1 \\ 14 & 0 & -1 & -1 & 0 \\ 24 & 0 & 1 & 0 & -1 \\ 34 & 0 & 0 & 1 & 1 \end{array} \right)$$

We define $Z^i = \ker\, \delta_i$ and $B^i = \mathrm{im}\ \delta_{i-1}$. Similarly, for the boundary operator we define $Z_i = \ker\, \partial_i$ and $B_i = \mathrm{im}\ \partial_{i+1}$. It can be seen that $\delta_i \circ \delta_{i-1} = 0 = \partial_{i+1} \circ \partial_i$, for all $0 \leq i \leq k-1$. For example, in the simplex in Example 11 we can see that

$$\partial_1\partial_2([123]) = \partial_1([23] - [13] + [12]) = [2] - [3] - ([1] - [3]) + [1] - [2] = 0.$$

Hence $B^i \subseteq Z^i$ and $B_i \subseteq Z_i$.

Definition 11. *The i-th homology group and i-th cohomology group of X over $\mathbb{F}$ are defined as*

$$H^i(X; \mathbb{F}) = Z^i(X; \mathbb{F})/B^i(X; \mathbb{F}),$$

$$H_i(X; \mathbb{F}) = Z_i(X; \mathbb{F})/B_i(X; \mathbb{F}),$$

for $0 \leq i \leq k$.

This definition is slightly modified to accommodate the case $i = -1$ to obtain a *reduced cohomology*. This is simply a convention from algebraic topology to tie up the loose end, and since the cohomology and reduced cohomology groups are identical for all positive i, in this thesis we shall only refer to the cohomology group.

We have the following *Hodge decomposition* theorem by Eckmann (1944).

Theorem 4.1.1. *For the terms defined as above $H_i(X; \mathbb{R}) \cong H^i(X; \mathbb{R})$ and the vector space $C^i(X; \mathbb{R})$ can be decomposed as*

$$C^i(X; \mathbb{R}) = H_i(X; \mathbb{R}) \oplus B^i(X; \mathbb{R}) \oplus B_i(X; \mathbb{R}).$$

Thus we have the following relations between the subspaces.

$$Z^i = H^i \oplus B^i = (B_i)^{\perp}.$$

$$Z_i = H_i \oplus B_i = (B^i)^{\perp}.$$

Alternatively, one may also define $H_i = Z^i \cap Z_i$, as done by Gundert and Szedlák (2014), since these definitions are equivalent. The orthogonality between each pair of the subspaces in the decomposition may be checked as below. For $x \in H^i, y \in B^i, z \in B_i$, we have

$$\begin{aligned}
\langle x, y\rangle &= \langle x, \delta_{i-1} v\rangle = \langle \partial_i x, v\rangle_{C^{i-1}} = 0, \text{ since } x \in \ \ker \partial_i. \\
\langle x, z\rangle &= \langle x, \partial_{i+1} w\rangle = \langle \delta_i x, w\rangle_{C^{i+1}} = 0, \text{ since } x \in \ \ker \delta_i. \\
\langle y, z\rangle &= \langle \delta_{i-1} v, \partial_{i+1} w\rangle = \langle \delta_i \delta_{i-1} v, w\rangle_{C^{i+1}} = 0, \text{ since } \delta_i \delta_{i-1} = 0.
\end{aligned}$$

For the remainder of this section, we assume the underlying field is the set of real numbers $\mathbb{R}$ unless mentioned otherwise. Consider the vector spaces

$$C^{i-1} \underset{\partial^i}{\overset{\delta_{i-1}}{\rightleftarrows}} C^i \underset{\partial^i}{\overset{\delta_i}{\rightleftarrows}} C^{i+1}.$$

Definition 12. *For a k-complex X, the upper, lower and full Laplacian operators on C^i, denoted by L_i^{up}, L_i^{dn}, L_i, are defined as*

$$\begin{aligned}
L_i^{up} &= \partial_{i+1} \delta_i = \delta_i^* \delta_i, \\
L_i^{dn} &= \delta_{i-1} \partial_i = \delta_{i-1} \delta_{i-1}^*, \\
L_i &= L_i^{up} + L_i^{dn} = \delta_i^* \delta_i + \delta_{i-1} \delta_{i-1}^*,
\end{aligned}$$

for $0 \leq i \leq k-1$.

We can define the degree, upper and lower adjacency operators denoted by D_i, $A_i^{\sim}$ and

$A_i^{\pitchfork}$ on C^i as

$$(D_i f)(\sigma) = \deg\,(\sigma) f(\sigma),$$
$$(A_i^{\sim} f)(\sigma) = \sum_{\sigma' \sim \sigma} f(\sigma'),$$
$$(A_i^{\pitchfork} f)(\sigma) = \sum_{\sigma' \pitchfork \sigma} f(\sigma'),$$

where $\sigma \pitchfork \sigma'$ denotes that i-faces σ and σ' are lower adjacent, *i.e.* share a common $(i-1)$-face and induce the same orientation on it, and $\sigma \sim \sigma'$ denotes that i-faces σ and σ' are upper adjacent *i.e.* $\sigma \pitchfork \sigma'$ and $\sigma \cup \sigma' \in X_{i+1}$. Then the Laplacian can be written explicitly in terms of the adjacency operators as

$$L_i^{up} = D_i - A_i^{\sim}$$
$$f(\sigma) \longmapsto \deg\,(\sigma) f(\sigma) - \sum_{\sigma' \sim \sigma} f(\sigma').$$
$$L_i^{dn} = (i+1)I + A_i^{\pitchfork}$$
$$f(\sigma) \longmapsto (i+1) f(\sigma) + \sum_{\sigma' \pitchfork \sigma} f(\sigma').$$
$$L_i = L_i^{up} + L_i^{dn}$$
$$f(\sigma) \longmapsto (\deg\,\sigma + i + 1) f(\sigma) + \sum_{\substack{\sigma' \pitchfork \sigma \\ \sigma' \not\sim \sigma}} f(\sigma').$$

This definition is consistent with that of Laplacians of graphs, which is L_0^{up}.

From the definition of L_i^{up}, L_i^{dn}, L_i, we can see that the matrices are self-adjoint and positive semi-definite. This implies all eigenvalues of these matrices are non-negative. In addition, since for any linear operator T the spectra of TT^* and T^*T differ only in the multiplicity of zero, L_i^{up} and L_{i+1}^{dn} have the same non-zero eigenvalues. Hence, it is sufficient to study L_i^{up}. Also, it must be noted that the spectrum of the Laplacians do not depend on the orientation of the simplices.

Example 12. *The upper Laplacian matrix $L_1^{up} = \delta_1^* \delta_1$ of the complete simplex in Example*

11 is given below.

$$L_1^{up} = \left(\begin{array}{c|cccccc} & 12 & 13 & 23 & 14 & 24 & 34 \\ \hline 12 & 2 & -1 & 1 & -1 & 1 & 0 \\ 13 & -1 & 2 & -1 & -1 & 0 & 1 \\ 23 & 1 & -1 & 2 & 0 & -1 & 1 \\ 14 & -1 & -1 & 0 & 2 & -1 & -1 \\ 24 & 1 & 0 & -1 & -1 & 2 & -1 \\ 34 & 0 & 1 & 1 & -1 & -1 & 2 \end{array}\right)$$

4.2 Spectrum of complete simplex

In this section we provide the spectrum of the complete simplex K_n^k, which is the simplex on n vertices containing all possible k-faces.

Before we give the spectrum of a complete simplex, we make a small observation on the size of the edge cover for a complete simplex. Suppose we have a complete simplicial complex and we wish to count the minimum number of $(k+1)$-edges (or k-faces) needed so that every k-edge (or $(k-1)$-face) is contained in at least one of the $(k+1)$-edges. This is a reformulation of the well-known edge cover problem for 2-graphs.

For a fixed k, let $C(n,k)$ denote the size of such a cover for a complete simplicial complex with n vertices. The number of edges to be covered is $\binom{n}{k}$, which is of complexity $\Theta(n^k)$.[1]

We obtain a recurrence relation for $C(n,k)$ by reasoning as follows. $C(n+1,k)$ denotes the number of $(k+1)$ edges needed to include all k-edges over $(n+1)$ vertices. Let P be a fixed vertex. $C(n,k)$ covers all k-edges not containing P. For the k-edges containing P, we need at most $C(n,k-1)$ k-edges, where one of the vertices is P.

$$C(n+1,k) \leq C(n,k) + C(n,k-1).$$

[1]There exist constants a and b such that $an^k \leq \binom{n}{k} \leq bn^k$

This recurrence relation has the same asymptotic growth rate as $\binom{n+1}{k}$. We also have

$$\binom{n}{k}/(k+1) \leq C(n,k) \leq \binom{n}{k}.$$

Thus $C(n,k)$ is of complexity $\Theta(n^k)$ for a fixed k.

The complete simplex K_n^k corresponds to a complete $(k+1)$-uniform hypergraph on n vertices. Recall that for a complete 2-graph K_n, the spectrum of the Laplacian consists of 0 and n with multiplicities 1 and $(n-1)$ respectively. We generalize the result to higher dimensions and show that the Laplacian of a complete simplex exhibits a similar property. A similar result is given in (Gundert and Wagner, 2012) with a more elegant proof. Also refer (Brouwer and Haemers, 2011)(Section 3. 12).

Proposition 4.2.1. *For all* $0 \leq i \leq k$, *Rank(*$L_i^{up}(K_n^k)$*)=* $\binom{n-1}{i+1}$.

Proof. Since $L_i^{up} = \partial_{i+1}\partial_{i+1}^*$, Rank($L_i^{up}$)= Rank($\partial_{i+1}$). Let $\partial^{(n)}$ denote the $(i+1)$-boundary matrix for n vertices. We prove that rank of $\partial^{(n)}$ is $\binom{n-1}{i+1}$ using induction on number of vertices. For base case $n = i+2$, the number of $(i+1)$-faces is one. The rank is $\binom{i+2-1}{i+1}$ which is one, hence the statement holds.

Assume the statement is true for n vertices. Fix a vertex O in the simplex. Adding a new vertex P adds $\binom{n}{i}$ new i-faces and $\binom{n}{i+1}$ new $(i+1)$-faces, for all $1 \leq i < k$, of which $\binom{n-1}{i}$ i-faces contain O and P, and $\binom{n-1}{i+1}$ $(i+1)$-faces contain P but not O.

Let $\{\sigma_1, \ldots, \sigma_{\binom{n-1}{i}}\}$ denote the columns of $\partial^{(n+1)}$ corresponding to the $(i+1)$-faces containing both O and P. Each such face contains a unique i-face containing neither O nor P. Hence they are linearly independent. The matrix corresponding to $\partial^{(n+1)}$ is of the form:

$$\partial^{(n+1)} = \left(\begin{array}{c|ccc|c} \text{w/o } P & \sigma_1 & \cdots & \sigma_{\binom{n-1}{i}} & \text{with } P \text{ not } O \\ \hline \partial^{(n)} & & & & \\ \hline 0 & -1 & -1 & -1 & \\ 0 & 1 & 0 & 0 & \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & 1 & \end{array}\right)$$

Let T be the basis for columns of $\partial^{(n)}$. By induction hypothesis, $|T| = \binom{n-1}{i+1}$. By abuse of notation we will allow them to represent columns of $\partial^{(n+1)}$ as well. Consider column v of $\partial^{(n+1)}$. If the $(i+1)$-face of v does not contain P, then $v \in \text{span}(T)$. If v contains P and O, then $v \in \{\sigma_1, \ldots, \sigma_{\binom{n-1}{i}}\}$. If v contains P but not O, then it consists of $(i+1)$ i-faces of which i contain P and one does not. But every such i-face is part of σ_j for some $j = 1, \ldots, \binom{n-1}{i}$. Hence, v can be written as a linear combination of those columns.

$$v = \sum_{v \cap \sigma_j \neq \emptyset} \sigma_j + \tau,$$

where τ is a i-face that does not contain P. However, $\tau \in$ span(T), since the simplex contains every $(i+1)$-face, for $i < k$. Therefore, $T \cup \{\sigma_1, \ldots, \sigma_{\binom{n-1}{i}}\}$ forms a basis for column space of $\partial^{(n+1)}$. This implies for all $0 \leq i < k$

$$\begin{aligned} \text{Rank}(\partial^{(n+1)}) &= \text{Rank}(\partial^{(n)}) + \binom{n-1}{i} \\ &= \binom{n-1}{i+1} + \binom{n-1}{i} \\ &= \binom{n}{i+1}. \end{aligned}$$

□

Proposition 4.2.2. *For all $0 \leq i \leq k$, n is an eigenvalue of L_i^{up} with multiplicity $\binom{n-1}{i+1}$.*

Proof. It suffices to show that $L_i^{up} v = nv$, for all column vectors of ∂_{i+1}. Consider the matrix L_i^{up}.

$$(L^{up})_{\sigma,\tau} = \begin{cases} n-k-1 & \text{if } \sigma = \tau, \\ 1 & \text{if } \sigma \sim \tau \text{ and they have same orientation,} \\ -1 & \text{if } \sigma \sim \tau \text{ and they have opposite orientation,} \\ 0 & \text{otherwise.} \end{cases}$$

Let v be a column of ∂_{i+1} corresponding to $(i+1)$-face formed by $\sigma_1, \ldots, \sigma_i$.

$$v_\tau = \begin{cases} (-1)^{t+1} & \text{if } \tau = \sigma_t, t = 1, \ldots, i, \\ 0 & \text{otherwise,} \end{cases}$$

where σ, τ are i-faces.

$$(L^{up}v)_\tau = L^{up}_{\tau\sigma_1} v_{\sigma_1} + \ldots + L^{up}_{\tau\sigma_i} v_{\sigma_i}. \tag{4.1}$$

For the i-face σ_1 we have

$$(L^{up}v)_{\sigma_1} = L^{up}_{\sigma_1\sigma_1} v_{\sigma_1} + \ldots + L^{up}_{\sigma_1\sigma_k} v_{\sigma_k} = (n-k-1) + \underbrace{(-1)(-1) + \ldots + (1)(1)}_{(k-1)\text{ terms}} = n.$$

For every other i-face τ, the sum in the expression (4.1) is zero. Hence,

$$(L^{up}v)_\tau = \begin{cases} n(-1)^{t+1} & \text{if } \tau = \sigma_t, t = 1, \ldots, i, \\ 0 & \text{otherwise,} \end{cases}$$
$$= nv_\tau.$$

□

Propositions 4.2.1 and 4.2.2 directly give us the following result.

Theorem 4.2.3. *For all* $0 \le i \le k$ *spec*$(L_i^{up}(K_n^k))$ *consists of* 0 *and* n *of multiplicities*

$\binom{n-1}{i}$ *and* $\binom{n-1}{i+1}$ *respectively.*

Example 13. *For the complete simplex* K_4^3 *in Example 11, the spectrum is* $\{0^3, 4^3\}$.

4.3 Cheeger Inequality

While simplicial complexes have been studied extensively in topology, there has also been work that extend several existing results for graphs to hypergraphs. Horak and Jost (2013) provide a general definition for the Laplacian using weighted inner products on C^i. The definition for Laplacian in Definition 12 and normalized Laplacian arise out of specific weight functions, namely the constant function 1 and deg σ respectively. They also prove a bound on the largest eigenvalue of the Laplacian (Theorem 3.2 in (Horak and Jost, 2013)).

$$\lambda_{\max}(L_i^{up}) \leq (i+2) \max_{\sigma \in X_i} \deg \sigma,$$
$$\lambda_{\max}(\widehat{L}_i^{up}) \leq (i+2),$$

where $\widehat{L}$ denotes the normalized Laplacian. This is the extension of the result $\lambda_{\max}(G) \leq 2d_{\max}$ for graphs.

Various attempts have been made to extend Cheeger inequality for simplicial complexes. The Laplacian L_{k-1} contains many trivial zeros. Hence the spectral gap λ is defined in terms of the eigenvalues that occur in the subspace Z_{k-1}, as defined in Hodge decomposition theorem. It has been shown that the inequality does not hold for simplicial complexes in general.

Gundert and Wagner (2012) provide a counterexample on $\mathbb{Z}_2$-cohomology using a random generating model, which has a non-trivial spectral gap but Cheeger constant zero, i.e. $\lambda > h = 0$. Steenbergen et al. (2014) proved that there does not exist a Cheeger inequality for the cochain complex. They prove a stronger assertion that there are no constants c, m_1, m_2 such that $\lambda \leq c\phi^{m_1}$ or $\frac{\phi^{m_2}}{d_{\max}} \leq \lambda$. However, they were able to provide a Cheeger inequality on the chain complex for some special cases. Under certain assumptions on the orientation of a k-complex and the degree of the $(k-1)$-skeleton, they were able to show

$$\frac{h^2}{2k+1} \leq \lambda \leq h.$$

Parzanchevski et al. (2016) modify the definition of $h(X)$ as

$$h(X) = \min_{\substack{V=\uplus A_i \\ A_i \neq \emptyset}} \frac{|V||F(A_0,\dots,A_k)|}{|A_0|\dots|A_k|},$$

where $A_0,\dots,A_k$ form a partition on V and $F(A_0,\dots,A_k)$ represents the set of k-faces with exactly one vertex in each A_i, for $i = 0,\dots,k$. They proved the lower bound of $h(X)$ for complexes with full $(k-1)$-skeleton, i.e. $h \leq \lambda$. This result has been proved for k-complexes without the assumption of complete skeleton by Gundert and Szedlák (2014). It has also been shown by Parzanchevski et al. (2016) that the upper bound of h in terms of λ is not possible.

4.4 Mixing Lemma

In this section we provide a brief description of the existing results on Mixing Lemma. We also state a conjecture for a generalized Mixing Lemma.

Recall from Chapter 2, that the Mixing Lemma for 2-graphs provides a measure for the randomness of the graph. We use the eigenvalue of the adjacency rather than that of the Laplacian. Stated formally, the result is as follows.

Theorem 4.4.1. *For a d-regular graph $G(V,E)$ with disjoint subsets $S,T \subset V$ we have*

$$\left| E(S,T) - \frac{d|S||T|}{n} \right| \leq \mu(G) \cdot \sqrt{|S||T|},$$

where $\mu(G) = \max\{\mu_2(A), |\mu_n(A)|\}$.

The condition for the sets to be disjoint is not necessary. If S and T are not disjoint, the edges within the intersection are counted twice.

In a sense, this counting problem is the dual of the colourability problem. In vertex colourability, we partition the vertices such that no edge lies wholly within any partition. In this case, we are given a partition and we wish to count the edges across the partition. The Mixing Lemma is a relaxed version of this problem, where we do not require a partition but merely two subsets.

In the case of k-graphs, the equivalent counting problem would be to estimate the number of k-edges that span a k-partition $S_1, \ldots, S_k \subset V$.

The result has been extended to hypergraphs in several ways. One version using tensors was given by Friedman and Wigderson (1995). Another version using a different representation was given by Butler (2008) (See Section 3.1 in Chapter 3). A result similar to Mixing Lemma was proved for 3-graphs using the notion of quasi-randomness by Gowers (2006). This has been improved for simplicial complexes of two dimensions by Gundert (2013).

Parzanchevski et al. (2016) provide a Mixing Lemma for a complex with complete skeleton. For a complex of dimension k and average degree d, we have

$$\left| F(A_0, \ldots, A_k) - \frac{d}{n} |A_0| \ldots |A_k| \right| \leq \rho \cdot (|A_0| \ldots |A_k|)^{\frac{k}{k+1}},$$

where $A_0, \ldots, A_k$ are disjoint subsets of V and $F(A_0, \ldots, A_k)$ denotes the number of edges with one vertex in each A_i, for $i = 0, \ldots, k$, and ρ is the maximum absolute value of the non-trivial eigenvalues of $dI - L^{up}$. The result has been extended by Parzanchevski (2017) by removing the assumption of complete skeleton. Essentially, the result states that assuming the complex is a (d_j, ϵ_j)-expander in lower dimensions, then it is an expander in dimension k as well.

Theorem 4.4.2. *If a k-dimensional complex X is a (d_j, ϵ_j)-expander for all $j = 0, \ldots, k-$*

1, *and* $A_0, \dots, A_k$ *are disjoint subsets of* V, *then*

$$\left| F(A_0, \dots, A_k) - \frac{d_0 \dots d_{k-1}}{n^k} |A_0| \dots |A_k| \right| \leq c_k d_0 \dots d_{k-1} (\epsilon_0 + \dots + \epsilon_{k-1}) \max |A_i|,$$

where c_k *depends only on* k.

The existing results count the number of k-edges spanning k-disjoint subsets. We believe the question makes sense even if we considered fewer than k sets. Suppose we were to consider a k-edge spanning m subsets, with $m < k$. Then, some of the subsets will contain more than one vertex of the edge. In that case, we simply count those subsets repeatedly. In other words, we feel that Theorem 4.4.2 holds even when the subsets A_i are not distinct.

For example, consider a complex of dimension 2. Let S and T be disjoint subsets of V. Let c be the cell density of the complex. Then we expect the following statement to hold.

$$\left| F(S,T) - c\binom{S}{2}\binom{T}{1} \right| \leq \rho \cdot S^{4/3} T^{2/3}$$
$$\left| F(S,T) - c\binom{S}{1}\binom{T}{2} \right| \leq \rho \cdot S^{2/3} T^{4/3}$$
$$|F(S,T) - E(S,T)| \leq \rho \cdot \max\{S, T\}$$

where $F(S,T)$ is the number of 3-edges across S and T, and $E(S,T)$ is the expected number of 3-edges, which is $c\left(\binom{S}{2}\binom{T}{1} + \binom{S}{1}\binom{T}{2}\right)$.

For a general k-graph we expect the following statement to be true.

Conjecture 4.4.1. *Let* $A_1, \dots, A_m$ *be disjoint subsets of* V, *with* $m < k$. *Let* $F(A_1, \dots, A_m)$ *denote the number of* k*-edges* e *such that* $e \cap A_i \neq \emptyset$ *and* $e \subset \bigcup_i A_i$, *for* $i = 1, \dots, m$. *Let* c *be the cell density of the complex,* d *the average degree and* ρ *denote the maximum* *non-trivial eigenvalue of* $dI - L^{up}$*. Then*

$$\left| F(A_1, \ldots, A_m) - c \sum_{\substack{s_i > 0 \\ \sum s_i = k}} \binom{A_1}{s_1} \ldots \binom{A_m}{s_m} \right| \leq \rho \max |A_i|.$$

Chapter 5

Tensor Representations

Tensors are an intuitive way to represent uniform hypergraphs. Put simply, tensors are multi-dimensional arrays that generalize matrices. Tensors have been well-studied and are widely used in physics and engineering. In computer science, tensors are ideal for representing data with several features. For example, in image processing, every image can be represented by a tensor, where each pixel has several attributes such as co-ordinates, RGB values and so on.

Mathematically, a tensor is a multilinear function between vector spaces. Under a suitable basis, the function can be denoted by a hypermatrix which is a multi-dimensional array. Often, the hypermatrix is identified with the tensor and the terms are used interchangeably.

A k-uniform hypergraph can be represented using an order k tensor. Unlike the weighted clique representation, this describes the hypergraph completely without any loss of information. Just as in graphs, it is possible to infer many properties of the hypergraph from the tensor. Several combinatorial results exist using tensors. In this chapter we study the spectral properties of a hypergraph using a tensor representation.

However, first, one needs to define an eigenvalue for a tensor. Lim (2005) and Qi (2005) independently introduced the concept of eigenvalue for a tensor and proceeded to develop a spectral theory for tensors. Some properties of spectra of matrices no longer hold. For example, there may be more than n eigenvalues for a tensor of dimension n. Nonetheless, it is possible to extend many of the results for tensors.

While tensors may seem a natural approach to understand uniform hypergraphs, one of its disadvantages is the large size of the array that makes it cumbersome to store and process. In addition, computing the eigenvalues of a tensor is NP-Hard. Tensor computation is a vast area in its own right. Some results on tensor decomposition may be found in (Kolda and Bader, 2009).

In this chapter, we provide some bounds on analytic connectivity α of a hypergraph, which is the analogue of the second smallest eigenvalue of the Laplacian matrix. We obtain a bound with respect to the degrees of vertices. We also prove a lower bound with respect to the diameter of a hypergraph. Finally, we prove Cheeger inequality to non-uniform hypergraphs, which is the most important result of this chapter.

This chapter is organized as follows. Section 5.1 contains the basic definitions we require. In Section 5.2 we survey the existing literature and provide some necessary preliminary lemmas. The original results are given in Section 5.3. We prove two bounds on the second eigenvalue of the Laplacian tensor and prove a version of Cheeger inequality for non-uniform hypergraphs. We conclude the chapter with a brief discussion in Section 5.4

5.1 Preliminaries

In this section, we introduce some of the basic terminology for tensors and state the existing results for hypergraphs. The following definitions can be found in (Qi, 2005; Cooper and Dutle, 2012). Some of the exposition given here can be found in (Ghoshdastidar, 2016).

Definition 13. *A **tensor** $\mathcal{T}$ of **order** (p, q) is a multilinear map of the form*

$$\mathcal{T} : \underbrace{V^* \times \ldots \times V^*}_{p \text{ times}} \times \underbrace{V \times \ldots \times V}_{q \text{ times}} \longrightarrow \mathbb{F}$$

where V is a vector space over field $\mathbb{F}$ and V^* is the dual space of V.

If the dimension of the vector space V is n, then under a fixed basis, the tensor can be represented as a collection of n^k elements over $\mathbb{F}$, where $k = p + q$. Such a collection arranged as a multidimensional array of the form $n \times \ldots \times n$ (k times) is a *hypermatrix*. In this thesis the underlying field will be $\mathbb{R}$ or $\mathbb{C}$. A hypermatrix is said to be *symmetric* if entries corresponding to a fixed set of indices is same, i.e. if $t_{i_1 i_2 \ldots i_k} = t_{i_{\sigma(1)} i_{\sigma(2)} \ldots i_{\sigma(k)}}$ for all $\sigma \in \mathfrak{S}_k$, where $\mathfrak{S}_k$ is the set of permutations on a sets of size k. For $k = 2$, this agrees with the usual definition of a symmetric matrix. (Some authors, e.g. Qi (2005); Kofidis and Regalia (2002), refer to this condition as supersymmetry.)

The *trace* of a tensor denotes the sum of the diagonal entries, i.e. $\text{Trace}(\mathcal{T}) = \sum_{i=1}^{n} t_{i \ldots i}$.

Definition 14. *Let $\mathcal{T}$ be a hypermatrix over $\mathbb{R}$ of order k and dimension n and $U \in \mathbb{R}^{r \times n}$. The **mode-**$m$ product of $\mathcal{T}$ and U is a tensor of order k denoted by $\mathcal{T} \times_m U$, whose size is n along all directions except the m-th one, for which the dimension is r. The entries of $\mathcal{T} \times_m U$ are given by*

$$(\mathcal{T} \times_m U)_{i_1 \ldots i_{m-1} j i_{m+1} \ldots i_k} = \sum_{i_m = 1}^{n} t_{i_1 \ldots i_{m-1} i_m i_{m+1} \ldots i_k} u_{j i_m},$$

for $i_1, \ldots, i_{m-1}, i_{m+1}, \ldots, i_k \in \{1, \ldots n\}$ and $j \in \{1, \ldots, r\}$.

Suppose we took the mode-m product of the tensor with a row vector along each direction, the result would be a scalar. Viewing the tensor $\mathcal{T}$ as a k-linear functional, for any k vectors $u_1, \ldots, u_k \in \mathbb{R}^n$, we have

$$(u_1, \ldots, u_k) \longmapsto \mathcal{T} \times_1 u_1^T \times_2 \ldots \times_k u_k^T.$$

Then, one could consider the following optimization problem.

$$\underset{u:\|u\|_p=1}{\text{maximize}} \quad \mathcal{T} \times_1 u^T \times_2 \ldots \times_k u^T. \tag{5.1}$$

This formulation is similar to the optimization problem for a matrix $A \in \mathbb{R}^{n\times n}$, and a vector $u \in \mathbb{R}^n$.

$$\underset{u:\|u\|_2=1}{\text{maximize}} \quad u^T A u.$$

The solution to this problem gives the largest eigenvalue of the matrix A. Additional constraints on the argument vector u gives the Rayleigh principle that leads to the other eigen pairs of the matrix. Similarly, the maximum of (5.1) would correspond to eigenvalues and the arguments would correspond to the eigenvectors of the tensor $\mathcal{T}$. This motivates the following definition.

For a hypermatrix $\mathcal{T}$ over $\mathbb{R}$ of order k and dimension n, and a vector $u \in \mathbb{R}^n$, the product $\mathcal{T}u^{k-1}$ denotes a vector in $\mathbb{R}^n$, whose j-th component is given as

$$(\mathcal{T}u^{k-1})_j = \sum_{i_2,i_3\ldots i_k=1}^{n} t_{ji_2,i_3\ldots i_k} u_{i_2} \ldots u_{i_k}.$$

Definition 15. *Let $\mathcal{T}$ be a hypermatrix over $\mathbb{R}$ of order k and dimension n. We say $\lambda \in \mathbb{C}$ is an* ***eigenvalue*** *and $x \in \mathbb{C}^n$ is an* ***eigenvector*** *of $\mathcal{T}$ if for all $j \in \{1, \ldots, k\}$,*

$$(\mathcal{T}x^{k-1})_j = \sum_{i_2,i_3\ldots i_k=1}^{n} t_{ji_2,i_3\ldots i_k} x_{i_2} \ldots x_{i_k} = \lambda x_j^{k-1},$$

where x_j^{k-1} represents the j-th component raised to exponent $(k-1)$. In addition, if $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$, then they are called ***H-eigenvalue*** *and* ***H-eigenvector*** *respectively.*

Definition 16. *Let $\mathcal{T}$ be a hypermatrix over $\mathbb{R}$ of order k and dimension n, we say $\lambda \in \mathbb{C}$*

is an ***E-eigenvalue*** *and* $x \in \mathbb{C}^n$ *is an* ***E-eigenvector*** *of* $\mathcal{T}$ *if they satisfy the following*

$$\mathcal{T}x^{k-1} = \lambda x \qquad and$$
$$\sum_{i=1}^{n} x_i^2 = 1.$$

In addition, if $\lambda \in \mathbb{R}$ *and* $x \in \mathbb{R}^n$*, then they are called* ***Z-eigenvalue*** *and* ***Z-eigenvector*** *respectively.*

It can be seen that substituting $k = 2$ yields the familiar definition of eigenvalue in both cases. Note that in (5.1), taking ℓ_k-norm we get the definition for H-eigenvalues and taking ℓ_2-norm we get the definition for Z-eigenvalue. While multiplying an H-eigenvector by a constant preserves its property, it is not so in the case of Z-eigenvalue. Also, other concepts such as determinant, characteristic polynomial have their counterparts for tensors as well. For instance, the idea of hyperdeterminant for tensors appears to have been proposed by Cayley (1845) in mid-19th century. Since then it has been redefined by Gelfand et al. (2008). Some of the properties of the eigenvalues are given below from (Qi, 2005).

Theorem 5.1.1. *For a hypermatrix* $\mathcal{A}$ *of order* m *and dimension* n *over* $\mathbb{C}$*, the following properties hold.*

1. *A number* $\lambda \in \mathbb{C}$ *is an eigenvalue of* $\mathcal{A}$ *if and only if it is a root of the characteristic polynomial.*

2. *The number of eigenvalues of* $\mathcal{A}$ *is* $n(m-1)^{n-1}$ *. Their product is equal to* $det(\mathcal{A})$.

3. *The sum of all the eigenvalues of* $\mathcal{A}$ *is* $(m-1)^{n-1} \cdot tr(\mathcal{A})$.

4. *If* A *is diagonal, then* A *has* n *H-eigenvalues, which are its diagonal elements with multiplicity* $(m-1)^{n-1}$ *and with corresponding unit vectors as their H-eigenvectors.*

In the recent years, many papers have been published on spectra of tensors. One useful result is the generalization of Perron-Frobenius theorem (Chang et al., 2008; Yang and Yang, 2010).

These results hold for all tensors. In this thesis we are interested in applying tensors within the context of hypergraphs. Just as in 2-graphs, the tensors we use will have many special properties, such as symmetry, distribution of eigenvalues and so on. We begin by introducing the adjacency and Laplacian tensors which are defined in a natural way.

Definition 17. *For a k-hypergraph $H(V,E)$ with n vertices, the (normalized)* ***adjacency hypermatrix*** *$\mathcal{A}(H)$ of order k and dimension n is defined as*

$$a_{i_1 i_2 \dots i_k} = \frac{1}{(k-1)!}\begin{cases} 1 & \text{if } \{i_1 i_2 \dots i_k\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The normalization factor of $1/(k-1)!$ is added to ensure that each edge contributes a sum of one in each slice of the hypermatrix. This helps us generalize many of the properties of spectra of graphs to hypergraphs in a natural way. The *degree* of a vertex is defined as the number of edges that contain it, which is identical to the definition in a 2-graph. The spectrum of the adjacency matrix provides many insights into the hypergraph and has been studied in (Friedman and Wigderson, 1995; Cooper and Dutle, 2012). In this work, we study the spectrum of the Laplacian hypermatrix, which is the more useful.

Definition 18. *For a k-hypergraph $H(V,E)$ with n vertices, the* ***Laplacian hypermatrix****, $\mathcal{L}(H)$ is defined as $\mathcal{L}(H) = \mathcal{D}(H) - \mathcal{A}(H)$, where $\mathcal{D}(H)$ is the degree hypermatrix with diagonal entries $d_{i\dots i} = deg\ (v_i)$ and non-diagonal entries as zero.*

This definition has been extended to non-uniform hypergraphs by Banerjee et al. (2017) as follows.

Definition 19. *Let $H(V,E)$ be a non-uniform hypergraph with size of the largest edge denoted by $k_{\max}$. Then the* ***adjacency hypermatrix*** *$\mathcal{A}$ of H is a hypermatrix of order $k_{\max}$ and dimension n defined such that for each edge $\{v_{p_1}, \dots, v_{p_k}\} \in E$ of cardinality $k \le k_{\max}$*

$$a_{i_1 \dots i_{k_{\max}}} = \frac{k}{\Omega},$$

where

$$\Omega = \sum_{\substack{m_j \geq 1 \\ \sum m_j = k}} \frac{k_{\max}!}{m_1! \dots m_k!},$$

and $i_1, \dots, i_{k_{\max}}$ are chosen in all possible ways from $\{p_1, \dots, p_k\}$ with each element chosen at least once. All other entries of the hypermatrix are zero.

The Laplacian of a non-uniform hypergraph can be defined as usual as $\mathcal{L}(H) = \mathcal{D}(H) - \mathcal{A}(H)$. Once Laplacian has been defined, it is possible to defined a normalized Laplacian. As in the case of 2-graphs, there are two different ways to induce normalization (Refer Chapter 2). We use the definition that gives a symmetric hypermatrix (See Definition 3.12 in (Banerjee et al., 2017)).

Definition 20. *Let $H(V, E)$ be a non-uniform hypergraph with size of the largest edge denoted by $k_{\max}$. Then the **normalized Laplacian hypermatrix** $\widehat{\mathcal{L}}$ of H is a hypermatrix of order $k_{\max}$ and dimension n defined such that for each edge $\{v_{p_1}, \dots, v_{p_k}\} \in E$ of cardinality $k \leq k_{\max}$*

$$l_{i_1 \dots i_{k_{\max}}} = -\frac{k}{\Omega} \prod_{j=1}^{k_{\max}} \frac{1}{\sqrt[k_{\max}]{d(v_{i_j})}},$$

where

$$\Omega = \sum_{\substack{m_j \geq 1 \\ \sum m_j = k}} \frac{k_{\max}!}{m_1! \dots m_k!},$$

and $i_1, \dots, i_{k_{\max}}$ are chosen in all possible ways from $\{p_1, \dots, p_k\}$ with each element chosen at least once. The diagonal entries are 1 and all other entries of the hypermatrix are zero.

This definition provides with us many results for the spectra of the normalized Laplacian analogous to graphs. For ease of readability, let $m = k_{\max}$ denote the size of the largest edge.

Let $e = \{v_{p_1}, \dots, v_{p_k}\} \in E(H)$ be an edge. By $x_m(e)$ we denote the expression

$$x_m(e) = \sum x_{i_1} \dots x_{i_m},$$

where $i_1, \ldots, i_m$ are chosen in all possible ways from $\{p_1, \ldots, p_k\}$ with each element chosen at least once. Let $\mathcal{L}(e)x^m$ denote a homogeneous polynomial of degree m in k variables such that

$$\mathcal{L}(e)x^m = \sum_{j=1}^{k} x_{i_j}^m - \frac{k}{\Omega} x_m(e).$$

$\mathcal{L}x^m = \sum_{e \in E} \mathcal{L}(e)x^m$ denotes the summation over all edges.

5.2 Previous results

In this section we review the existing results for tensors. One of the important results in spectral theory is the Perron-Frobenius Theorem. It is applicable to the adjacency matrix of 2-graphs. A similar result has been extended for tensors relatively recently (Friedland et al., 2013).

As mentioned in Theorem 5.1.1, the adjacency and Laplacian tensors may have several eigenvalues. The eigenvalues of interest are those that reflect the inherent properties of the hypergraphs. These are the non-negative H-eigenvectors which are referred as H^+-eigenvalues. The properties of H^+-eigenvalues are studied in (Qi, 2014). It can be shown that for $k \geq 3$, the Laplacian has at least $(n+1)$ H^+-eigenvalues. The behaviour of these eigenvalues is similar to those of 2-graphs. The bounds of largest eigenvalues of Laplacian hold for k-graphs. We list some of the results for H^+-eigenvalues (Qi, 2014).

Theorem 5.2.1. *(Qi, 2014) For a k-graph with $k \geq 3$, the following results hold for the Laplacian $\mathcal{L}$.*

1. *For all H-eigenvalues λ, $0 \leq \lambda \leq 2d_{\max}$.*

2. *For $j = 1, \ldots, n$, d_j is a H^+-eigenvalue of $\mathcal{L}$ with H-eigenvector $\mathbf{e}_j$.*

3. The largest Laplacian H^+-eigenvalue is $d_{\max}$. We have

$$d_{\max} = \max\left\{ \mathcal{L}x^k \ : x \in \mathbb{R}^n_+, \sum_{i=1}^{n} x_i^k = 1 \right\}.$$

Recall that for a graph, the second eigenvalue of the Laplacian λ_2 is given by the Rayleigh quotient as

$$\lambda_2(G) = \min_{x \perp 1} \frac{x^T L x}{x^T x}.$$

The parameter that best represents the second smallest eigenvalue of a 2-graph is that of analytic connectivity. It is inspired by algebraic connectivity of 2-graphs, which was introduced by Fiedler (1975) as the second eigenvalue of the Laplacian. (Also refer Lemma 3.2.9 of Chapter 3).

Lemma 5.2.2. *(Fiedler, 1975) For a graph G with n vertices,*

$$\lambda_2(G) = 2n \min\left\{ \frac{\sum_{(v_i,v_j)\in E} (x_i - x_j)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} (x_i - x_j)^2} : x \neq c.\mathbf{1}_n, \ \textit{for all } c \in \mathbb{R} \right\}.$$

Qi (2014) introduced the concept of analytic connectivity for a uniform hypergraph, usually denoted by $\alpha(H)$.

Definition 21. *The* ***analytic connectivity*** *of a k-uniform hypergraph H is defined as*

$$\alpha(H) = \min_{j=1,\ldots,n} \min\left\{ \mathcal{L}x^k \ : x \in \mathbb{R}^n_+, \sum_{i=1}^{n} x_i^k = 1, x_j = 0 \right\}.$$

As in 2-graphs, $\alpha(H)$ captures the connectivity of hypergraphs.

Theorem 5.2.3. *(Qi, 2014) A k-graph H is connected if and only if the analytic connectivity $\alpha(H) > 0$.*

In addition, we also have the following results for $\alpha(H)$.

Theorem 5.2.4. *(Qi, 2014) For a k-graph H, we have* $0 \leq \alpha(H) \leq d_{\min}$.

Lemma 5.2.5. *For a complete k-graph on n vertices,* $\alpha(K_n^k) = \binom{n-2}{k-2}$.

Of significance is the Cheeger inequality for uniform hypergraphs, where the isoperimetric number of the hypergraph is bounded by $\alpha(H)$. As in other representations, these concepts were defined for uniform hypergraphs, but can be applied to non-uniform hypergraphs without any modification. The definition for $\alpha(H)$ also extends to a general hypergraphs in a natural way (Banerjee et al., 2017). We use the definition to extend the results for uniform hypergraphs by Li et al. (2017) to non-uniform hypergraphs. It must be noted that unlike the case of matrices of 2-graphs, calculating $\alpha(H)$ is non-trivial. However, it is possible to obtain reliable estimates efficiently (Cui et al., 2016).

Some of the results for uniform hypergraphs have been extended to non-uniform hypergraphs.

Theorem 5.2.6. *(Banerjee et al., 2017) Let* μ *be an H-eigenvalue of* $\mathcal{A}$ *. Then* $|\mu| \leq d_{\max}$.

Proof. Let $\mathcal{A}$ be of order m and dimension n. Let μ be a H-eigenvalue with eigenvector $x = (x_1, x_2, \ldots, x_n)$. Let $x_p = \max\{|x_1|, \ldots, |x_n|\}$. Without loss of any generality we can assume that $x_p = 1$.

$$\begin{aligned}
|\mu| = |\mu x_p^{m-1}| &= \left| \sum_{i_2 \ldots i_m = 1}^{n} a_{pi_2 \ldots i_m} x_{i_2} \ldots x_{i_m} \right| \\
&\leq \sum_{i_2 \ldots i_m = 1}^{n} |a_{pi_2 \ldots i_m}||x_p|^{m-1} = d(v_p) \\
&\leq d_{\max}.
\end{aligned}$$

□

In particular, for a d-regular graph this implies the following property.

Theorem 5.2.7. *(Banerjee et al., 2017) Let* $H = (V, E)$ *be a d-regular hypergraph with* *n vertices. Then, $\mathcal{A}$ has an H-eigenvalue d.*

Proof. For all $i = 1, \ldots, n$, we have $(\mathcal{A} \cdot \mathbf{1})_i = \sum_{i_2 \ldots i_m = 1}^{n} a_{ii_2 \ldots i_m} = d$. □

As in 2-graphs and uniform graphs, the eigenvalues of Laplacian are bounded by twice the maximum degree and the analytic connectivity reflects the connectivity.

Theorem 5.2.8. *(Banerjee et al., 2017) Let $\mathcal{L}$ be the Laplacian hypermatrix of a general hypergraph H. Then, $0 \leq \lambda \leq 2d_{\max}$, where λ is an H-eigenvalue of $\mathcal{L}$.*

Theorem 5.2.9. *(Banerjee et al., 2017) A general hypergraph H is connected if and only if $\alpha(H) > 0$.*

The results mentioned here hold for the combinatorial Laplacian. The normalized Laplacian has also been studied in literature. The spectrum of the normalized Laplacian lies in the interval $[0, 2]$. Since the objective of the normalized Laplacian is to mitigate the effects of vertices with extremely large or small degrees, the spectral bounds are independent of the degrees in the hypergraph. This property is useful in many applications such as partitioning. We state one result on the spectrum of normalized Laplacian below.

Theorem 5.2.10. *(Banerjee et al., 2017) Let H be a general hypergraph with edge of largest size $k_{\max}$. Let $\widehat{\mathcal{L}}$ be the normalized Laplacian hypermatrix of H. Suppose $m(\lambda)$ represents the algebraic multiplicity of $\lambda \in spec(\widehat{\mathcal{L}})$. Then, $\sum_{\lambda} m(\lambda)\lambda = n(k_{\max} - 1)^{n-1}$.*

We now state the results for uniform hypergraphs which we wish to extend to general hypergraphs.

Theorem 5.2.11. *(Li et al., 2017) Let H be k-graph with more than one edge. Then*

$$\alpha(H) \leq \min \left\{ \frac{d(v_{i_1}) + d(v_{i_2}) + \ldots + d(v_{i_k}) - k}{k} : \{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\} \in E \right\}.$$

We have the following result that gives a lower bound on the diameter of a graph in terms of the second eigenvalue of Laplacian. We rewrite it to give a lower bound on the second eigenvalue instead.

Theorem 5.2.12. *(Mohar, 1991) For a 2-graph G with n vertices,*

$$\lambda_2(G) \geq \frac{4}{n \cdot diam(G)}.$$

Proof. For each pair (u, v) choose the shortest path P_{uv}. Let x be the eigenvector of λ_2. We can show from Lemma 5.2.2 that

$$2n \sum_{uv \in E} (x_u - x_v)^2 = \lambda_2 \sum_{u \in V} \sum_{v \in V} (x_u - x_v)^2,$$

since $x \perp 1_n$ implies $\sum_{v \in V} x_v = 0$ and $\sum_{uv \in E} (x_u - x_v)^2 = \lambda_2 \sum_{v \in V} x_v^2$.

$$\begin{aligned}(x_u - x_v)^2 &= [(x_u - x_{v_1}) + (x_{v_1} - x_{v_2}) + \dots (x_{v_{k-1}} - x_v)]^2 \\ &\leq \ \mathrm{dist}(u, v) \sum_{e \in P_{uv}} \delta^2(e),\end{aligned}$$

where $\delta^2(e)$ denotes $(x_a - x_b)^2$ for edge $e = (a, b)$. Define indicator function χ_{uv} as

$$\chi_{uv} = \begin{cases} 1 & \text{if } e \in P_{uv}, \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned}\sum_{u \in V} \sum_{v \in V} (x_u - x_v)^2 &\leq \sum_{u \in V} \sum_{v \in V} \ \mathrm{dist}(u, v) \sum_{e \in E} \delta^2(e) \chi_{uv}(e) \\ &= \sum_{e \in E} \delta^2(e) \sum_{u \in V} \sum_{v \in V} \ \mathrm{dist}(u, v) \chi_{uv}(e).\end{aligned}$$

We have

$$\sum_{u \in V} \sum_{v \in V} \chi_{uv}(e) \leq \frac{2n^2}{4} = \frac{n^2}{2},$$

and since $\mathrm{dist}(u, v) \leq \mathrm{diam}(G)$, this implies,

$$2n \sum_{uv \in E} (x_u - x_v)^2 = 2n \sum_{e \in E} \delta^2(e) \leq \lambda_2 \ \mathrm{diam}(G) \cdot \frac{n^2}{2} \sum_{e \in E} \delta^2(e),$$

which gives the result

$$\lambda_2 \geq \frac{4}{n \cdot \ \mathrm{diam}(G)}.$$

□

The above result has been extended to uniform hypergraphs by Li et al. (2017).

Theorem 5.2.13. *(Li et al., 2017) Let H be a k-graph. Then*

$$\alpha(H) \geq \frac{4}{n^2(k-1) diam(H)}.$$

The definition of isoperimteric number ϕ is identical to the definition in 2-graphs.

$$\phi(H) = \min \left\{ \frac{E(S, V \backslash S)}{|S|} : 0 < |S| \leq \frac{|V|}{2} \right\}.$$

Cheeger inequality for uniform hypergraphs is given as follows (Li et al., 2017).

Theorem 5.2.14. *(Li et al., 2017) For a k-graph H with $k \geq 3$,*

$$(d_{\max} - \sqrt{d_{\max}^2 - \phi^2}) \leq \alpha(H) \leq \frac{k}{2}\phi.$$

5.3 Results for general hypergraphs

In this section we prove the corresponding results for non-uniform hypergraphs. The inequalities are similar to that of uniform hypergraphs, except for an additional factor of $\left(\frac{k_{\min}}{k_{\max}}\right)$.

First, we prove an upper bound of the analytic connectivity that is analogous to Theorem 5.2.11. The statement remains unchanged, except the size of the edge is no longer constant.

Theorem 5.3.1. *Let H be a non-uniform hypergraph with more than one edge. Then*

$$\alpha(H) \leq \min\left\{\frac{d(v_{i_1}) + d(v_{i_2}) + \ldots + d(v_{i_k}) - k}{k} : \{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\} \in E\right\}.$$

Proof. Let us denote the size of the largest edge $k_{\max}$ by m for ease of notation. Let $e_0 = \{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\} \in E$. Define a vector $x \in \mathbb{R}^n_+$ such that

$$x_i = \begin{cases} k^{-m} & \text{if } v_i \in e_0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\sum_{i=1}^{m} x_i^m = k(\frac{1}{k^{1/m}})^m = 1$. $\mathcal{L}(e_0)x^m = \sum_{i=1}^{m} x_i^m - \frac{k}{\Omega} x_m(e_0) = 0$.

$$\begin{aligned} \alpha(H) &\leq \mathcal{L}x^m \\ &= \sum_{e \in E} \mathcal{L}(e)x^m \\ &= \sum_{e \in E \setminus \{e_0\}} \mathcal{L}(e)x^m + \mathcal{L}(e_0)x^m \\ &= (d(v_{i_1}) - 1)(1/k) + (d(v_{i_2}) - 1)(1/k) + \ldots + (d(v_{i_k}) - 1)(1/k) \\ &= \frac{d(v_{i_1}) + d(v_{i_2}) + \ldots + d(v_{i_k}) - k}{k}. \end{aligned}$$

□

We state a lemma involving the arithmetic and geometric mean of a sequence, which is required to prove Theorems 5.3.3 and 5.3.4 .

Lemma 5.3.2. *(Li et al., 2017) Let $a = (a_1, \ldots, a_n) \in \mathbb{R}^n_+$. Let $A = (a_1 + \ldots + a_n)/n$*

and $G = (a_1 \dots a_n)^{1/n}$. *Then*

$$A - G \geq \frac{1}{n(n-1)} \sum_{1 \leq i < j \leq n} (\sqrt{a_i} - \sqrt{a_j})^2. \tag{5.2}$$

$$A - G \geq \frac{1}{n} \sum_{j=1}^{\lfloor n/2 \rfloor} (\sqrt{b_j} - \sqrt{b_{n+1-j}})^2. \tag{5.3}$$

where $b_j = a_{\sigma(j)}$, *for* $j = 1, \dots, n$ *and* σ *is a permutation of the set* $\{1, \dots, n\}$.

We now give a lower bound on analytic connectivity in terms of the diameter of the hypergraph. This generalizes the results of Theorems 5.2.12 and 5.2.13. The statement replaces the size of edge k with the size of largest edge $k_{\max}$and has an additional factor of $\frac{k_{\min}}{k_{\max}}$.

Theorem 5.3.3. *Let* H *be a general hypergraph. Then*

$$\alpha(H) \geq \frac{4k_{\min}}{n^2 k_{\max}(k_{\max}-1) diam(H)}.$$

Proof. Let us denote the size of the largest edge $k_{\max}$ by m for ease of notation. Let $x = (x_1, x_2, \dots, x_n)$ be the vector achieving $\alpha(H)$. Assume $x_n = 0$. Define a 2-graph H^* with vertex set $V(H)$ and $u \sim v$ in H^* if and only if $\{u, v\} \subset e \in E$. In other words H^* is the clique expansion of H. We know that $\text{diam}(H) = \text{diam}(H^*)$. For edge $e = \{v_{i_1}, \dots, v_{i_k}\} \in E$, consider ℓ_1 copies of x_1^m, ℓ_2 copies of x_2^m and ℓ_k copies of x_k^m, where $\ell_j \geq 1$ and $\sum \ell_j = m$. From Lemma 5.3.2 we have,

$$\begin{aligned}
&\frac{\ell_1 x_1^m + \ell_2 x_2^m + \dots \ell_k x_k^m}{m} - x_1^{\ell_1} x_2^{\ell_2} \dots x_k^{\ell_k} \\
&\geq \frac{1}{m(m-1)} [(x_1^{m/2} - x_2^{m/2})^2 + (x_1^{m/2} - x_3^{m/2})^2 + \dots + (x_{k-1}^{m/2} - x_k^{m/2})^2] \\
&= \frac{1}{m(m-1)} \sum_{1 \leq i < j \leq k} (x_i^{m/2} - x_j^{m/2})^2.
\end{aligned}$$

Summing over different values of $\ell_1, \ell_2, \ldots, \ell_k$ we get

$$
\begin{aligned}
\frac{\Omega}{k}\left(\sum_{i=1}^{k} x_i^m\right) - x_m(e) &\geq \frac{\Omega}{m(m-1)} \sum_{1\leq i<j\leq k} (x_i^{m/2} - x_j^{m/2})^2 \\
\frac{1}{k}\left(\sum_{i=1}^{k} x_i^m - \frac{k}{\Omega} x_m(e)\right) &\geq \frac{1}{m(m-1)} \sum_{1\leq i<j\leq k} (x_i^{m/2} - x_j^{m/2})^2 \\
\mathcal{L}(e)x^m &\geq \frac{k}{m}\frac{1}{(m-1)} \sum_{1\leq i<j\leq k} (x_i^{m/2} - x_j^{m/2})^2 \\
&\geq \frac{k_{\min}}{m}\frac{1}{(m-1)} \sum_{1\leq i<j\leq k} (x_i^{m/2} - x_j^{m/2})^2.
\end{aligned}
$$

Thus for every edge $e \in E(H)$, the expression $\mathcal{L}(e)x^m$ is lower bounded. For ease of notation, let $y = x^{(m/2)}$. It is possible to write the expression in terms of the edges $e \in E(H*)$.

$$
\begin{aligned}
\alpha &= \sum_{e\in E(H)} \mathcal{L}(e)x^m \\
&\geq \frac{k_{\min}}{m}\frac{1}{(m-1)} \sum_{(v_iv_j)\in E(H^*)} (x_i^{m/2} - x_j^{m/2})^2 \\
&= \frac{k_{\min}}{m}\frac{1}{(m-1)} \sum_{(v_iv_j)\in E(H^*)} (y_i - y_j)^2 \\
&= \frac{k_{\min}}{m}\frac{1}{(m-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} (y_i - y_j)^2 \frac{\sum_{(v_iv_j)\in E(H^*)} (y_i - y_j)^2}{\sum_{i=1}^{n}\sum_{j=1}^{n} (y_i - y_j)^2} \\
&\geq \frac{\lambda_2(H^*)}{2n(m-1)}\frac{k_{\min}}{m} \sum_{i=1}^{n}\sum_{j=1}^{n} (y_i - y_j)^2 \qquad \text{from Lemma 5.2.2.}
\end{aligned}
$$

The expression $\sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2$ can be bounded by a constant as follows.

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}(y_i - y_j)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} y_i^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} y_j^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \\
&= 2n\left(\sum_{i=1}^{n} y_i^2\right) - 2\left(\sum_{i=1}^{n-1} y_i\right)^2 \qquad \text{(since } y_n = 0\text{)} \\
&\geq 2n - 2(n-1)\left(\sum_{i=1}^{n} y_i^2\right) \qquad \text{(from Cauchy-Schwarz)} \\
&= 2.
\end{aligned}
$$

We have

$$\alpha \geq \frac{\lambda_2(H^*)}{2n(m-1)} \frac{k_{\min}}{m} \cdot 2.$$

From Theorem 5.2.12, $\lambda_2(H^*) \geq \frac{4}{\operatorname{diam}(H^*) \cdot n}$. Therefore,

$$\alpha \geq \left(\frac{k_{\min}}{k_{\max}}\right) \frac{4}{n^2 (k_{\max} - 1) \cdot \operatorname{diam}(H^*)}.$$

□

We prove the Cheeger inequality for general hypergraphs in this representation.

Theorem 5.3.4. *For a non-uniform hypergraph H,*

$$\left(\frac{k_{\min}}{k_{\max}}\right)\left(d_{\max} - \sqrt{d_{\max}^2 - \phi^2}\right) \leq \alpha(H) \leq \frac{k_{\max}}{2}\phi.$$

Proof. Let m the size of the largest edge. First let us show the upper bound on α.

Suppose $S \subset V$ gives the isoperimetric number ϕ. Let $y = (y_1, \ldots, y_n) \in \mathbb{R}_+^n$ be the vector defined as follows.

$$y_i = \begin{cases} \frac{1}{|S|^{1/m}} & \text{if } v_i \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Let $t_S(e) = |\{v : v \in e \cap S\}|$ be the number of vertices of e in S and $t(S) = \frac{\sum_{e\in\partial S} t_S(e)}{|\partial S|}$.

$$t(S) + t(\overline{S}) = \frac{\sum_{e\in\partial S}(t_S(e) + t_{\overline{S}}(e))}{|\partial S|} \leq m. \tag{5.4}$$

Summing over all edges in the hypergraph,

$$\alpha \leq \mathcal{L}y^m = \sum_{e\in S} \mathcal{L}(e)y^m + \sum_{e\in\overline{S}} \mathcal{L}(e)y^m + \sum_{e\in\partial S} \mathcal{L}(e)y^m. \tag{5.5}$$

If $e = \{v_1, \ldots, v_k\} \subset \overline{S}$, $\mathcal{L}(e)y^m = \sum y_i^m - \frac{k}{\Omega} y_m(e) = 0$. If $e = \{v_1, \ldots, v_k\} \subset S$,

$$\begin{aligned}\mathcal{L}(e)y^m &= \sum_{i=1}^{m} y_i^m - \frac{k}{\Omega} y_m(e) \\ &= \frac{k}{|S|} - \frac{k}{\Omega} \sum \frac{1}{|S|} \cdot 1 \\ &= 0.\end{aligned}$$

Therefore only edges in the boundary contribute to the sum of (5.5).

$$\begin{aligned}\alpha &\leq \sum_{e\in\partial S} \sum_{v_i \in e\cap S} y_i^m \\ &= \sum_{e\in\partial S} \frac{t_S(e)}{|S|} = \frac{1}{|S|} t(S)|\partial S| \\ &= t(S) \cdot \phi.\end{aligned}$$

Similarly we can get $\alpha \leq t(\overline{S}) \cdot \phi$. Adding them we get $2\alpha \leq (t(S) + t(\overline{S}))\phi$. Combining with (5.4), we get $\alpha \leq (k_{\max}\phi)/2$.

To prove the lower bound, suppose $x = (x_1, \ldots, x_n)$ achieves α. For each edge $e = \{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\}$ assume $x_{i_1} \leq x_{i_2} \leq \ldots \leq x_{i_k}$ by rearranging the vertices. We define a 2-graph $\widetilde{H}$ whose vertex set is same as H and edges are such that

$$E(\widetilde{H}) = \cup_{e\in E(H)} \{v_{i_j} v_{i_{k+1-j}} : j = 1, \ldots, \lfloor k/2 \rfloor\}.$$

Then

$$\alpha = \sum_{e=\{v_{i_1},\ldots,v_{i_k}\}} \left(\sum_{j=1}^{k} x_{i_j}^m - \frac{k}{\Omega} \sum_{\substack{\ell_i \geq 1 \\ \sum \ell_i = m}} x_{i_1}^{\ell_1} \ldots x_{i_k}^{\ell_k} \right). \tag{5.6}$$

Consider ℓ_1 copies of x_1^m, ℓ_2 copies of x_2^m and ℓ_k copies of x_k^m, where $\ell_j \geq 1$ and $\sum \ell_j = m$. Applying Lemma 5.3.2 on A.M.-G.M. permutations we get

$$\frac{\ell_1 x_1^m + \ell_2 x_2^m + \ldots \ell_k x_k^m}{m} - x_1^{\ell_1} x_2^{\ell_2} \ldots x_k^{\ell_k} \geq \frac{1}{m} \sum_{i=1}^{\lfloor m/2 \rfloor} (\sqrt{b_i} - \sqrt{b_{m+1-i}})^2 \tag{5.7}$$

where $b_1, \ldots, b_m$ is any permutation of the variables x_i. In particular consider the assignment $b_1 = x_{i_1}^m, b_2 = x_{i_2}^m, \ldots, b_{\lfloor k/2 \rfloor} = x_{i_{k/2}}^m, b_{m+1-\lfloor k/2 \rfloor} = x_{i_{k+1-\lfloor k/2 \rfloor}}^m, \ldots, b_m = x_{i_k}^m$, and $b_{\lfloor k/2 \rfloor + 1}, \ldots, b_{m - \lfloor k/2 \rfloor}$ are assigned to any of the remaining $(m-k)$ variables. Then

$$\begin{aligned}
\sum_{i=1}^{\lfloor m/2 \rfloor} (\sqrt{b_i} - \sqrt{b_{m+1-i}})^2 &\geq \sum_{i=1}^{\lfloor k/2 \rfloor} (\sqrt{b_i} - \sqrt{b_{m+1-i}})^2 \quad (\text{ since } k \leq m) \\
&= \sum_{j=1}^{\lfloor k/2 \rfloor} (\sqrt{x_{i_j}^m} - \sqrt{x_{i_{k+1-j}}^m})^2
\end{aligned}$$

Using the above inequality with (5.7), and summing over all possible values of $l_1, \ldots, l_k$ as in proof of Theorem 5.3.3 we get

$$\begin{aligned}
\alpha &\geq \sum_{e=\{v_{i_1},\ldots,v_{i_k}\} \in E(H)} \left(\frac{k}{m} \sum_{j=1}^{\lfloor k/2 \rfloor} (\sqrt{x_{i_j}^m} - \sqrt{x_{i_{k+1-j}}^m})^2 \right) \\
&\geq \frac{k_{\min}}{k_{\max}} \sum_{e=\{v_{i_1},\ldots,v_{i_k}\} \in E(H)} \left(\sum_{j=1}^{\lfloor k/2 \rfloor} (\sqrt{x_{i_j}^m} - \sqrt{x_{i_{k+1-j}}^m})^2 \right) \\
&= \frac{k_{\min}}{k_{\max}} \sum_{\{v_i,v_j\} \in E(\widetilde{H})} (\sqrt{x_i^m} - \sqrt{x_j^m})^2.
\end{aligned}$$

Proceeding as in proof of k-graphs in (Li et al., 2017), let

$$M = \sum_{\{v_i,v_j\} \in E'} (y_i - y_j)^2, \text{ where } E' = E(\widetilde{H})$$

and $y_i = \sqrt{x_i^m}$. From Cauchy-Schwarz inequality we have

$$M \geq \frac{(\sum_{\{v_i,v_j\}\in E'} |y_i^2 - y_j^2|)^2}{\sum_{\{v_i,v_j\}\in E'} (y_i + y_j)^2}. \tag{5.8}$$

Let $w_0(=0) < w_1 < \ldots < w_h$ be the distinct values of y_i, for $i = 1, \ldots, n$. For $j = 0, \ldots, h$, let $V_j = \{v_i \in V : y_i \geq w_j\}$. The summation can be split across the distinct values in the vector.

$$\begin{aligned}
\sum_{\{v_i,v_j\}\in E'} |y_i^2 - y_j^2| &= \sum_{r=1}^{h} \sum_{\substack{\{v_i,v_j\}\in E' \\ v_i \in V_r \\ v_j \notin V_r}} (y_i^2 - y_j^2) = \sum_{r=1}^{h} \sum_{\substack{\{v_i,v_j\}\in E' \\ y_i = w_r \\ y_j = w_s \\ s<r}} (w_r^2 - w_s^2) \\
&= \sum_{r=1}^{h} \sum_{\substack{\{v_i,v_j\}\in E' \\ y_i = w_r \\ y_j = w_s \\ s<r}} (w_r^2 - w_{r-1}^2) + (w_{r-1}^2 - w_{r-2}^2) + \ldots + (w_{s+1}^2 - w_s^2) \\
&= \sum_{r=1}^{h} \sum_{\substack{\{v_i,v_j\}\in E' \\ v_i \in V_r}} \sum_{v_j \notin V_r} (w_r^2 - w_{r-1}^2)
\end{aligned}$$

For each edge $e \in \partial(V_j)$, let $\delta_j(e) = \min\{|V_j \cap e|, |\overline{V_j} \cap e|\}$. Let $\delta(V_j) = \min\{\delta_j(e) : e \in \partial(V_j)\}$ and $\delta(H) = \min_{j=0,\ldots,h} \delta(V_j)$. Then, the expression $\sum_{\{v_i,v_j\}\in E'} |y_i^2 - y_j^2|$ can be bounded as follows.

$$\begin{aligned}
\sum_{\{v_i,v_j\}\in E'} |y_i^2 - y_j^2| &\geq \sum_{r=1}^{h} \delta(V_r)|\partial V_r|(w_r^2 - w_{r-1}^2) \\
&\geq \sum_{r=1}^{h} \delta(H)\phi(H)|V_r|(w_r^2 - w_{r-1}^2) \\
&= \delta(H)\phi(H)(|V_h|(w_h^2 - w_{h-1}^2) + \ldots + |V_1|(w_1^2 - w_0^2)) \\
&= \delta(H)\phi(H)((|V_h| - |V_{h-1}|)w_h^2 + \ldots + (|V_1| - |V_2|)w_1^2) \\
&= \delta(H)\phi(H)\sum_{i=1}^{n} y_i^2
\end{aligned}$$

The denominator of (5.8) can be bounded in terms of the maximum degree.

$$\begin{aligned}
\sum_{\{v_i,v_j\}\in E'} (y_i + y_j)^2 &= 2\sum_{v_i,v_j\in E'} (y_i^2 + y_j^2) - \sum_{v_i,v_j\in E'} (y_i - y_j)^2 \\
&\leq 2\sum_{i=1}^{n} d(v_i)y_i^2 - \sum_{v_i,v_j\in E'} (y_i - y_j)^2 \\
&\leq 2d_{\max}(\widetilde{H})\sum_{i=1}^{n} y_i^2 - \sum_{v_i,v_j\in E'} (y_i - y_j)^2 \\
&= 2d_{\max}(\widetilde{H}) - M \\
&\leq 2d_{\max}(H) - M.
\end{aligned}$$

$$M \geq \frac{\delta(H)^2\phi^2}{2d_{\max} - M} \geq \frac{\phi^2}{2d_{\max} - M},$$

solving which we get $M \geq d_{\max} - \sqrt{d_{\max}^2 - \phi^2}$. Substituting in (5.8) we get

$$\alpha \geq \left(\frac{k_{\min}}{k_{\max}}\right)(d_{\max} - \sqrt{d_{\max}^2 - \phi^2}).$$

□

We can relax the lower bound to obtain a neater expression below that is often used in practice.

$$\left(\frac{k_{\min}}{k_{\max}}\right)\frac{\phi^2}{2d_{\max}} \leq \alpha \leq \frac{k_{\max}}{2}\phi.$$

5.4 Discussion

We conclude this chapter with a brief discussion on the results. Tensor representation for hypergraphs has been well-studied compared to others. This is partly due to the natural way in which tensors lend themselves for application in hypergraphs. Hence many results have been proved, at least with respect to uniform hypergraphs. The emphasis of this thesis has been in generalizing the results to non-uniform hypergraphs.

The biggest challenge one faces while studying general hypergraphs compared to uniform hypergraphs is the density of the connectivity tensors. For most practical applications, the tensors tend to be sparse. However, in the case of general hypergraphs, a single edge adds several entries to the tensors.

Another challenge is the computational overhead. The spectrum of the tensor cannot be computed in polynomial time. Performing these computations on tensors might be infeasible, particularly in cases where the number of vertices and the order of the tensors is large. Fortunately, tensors are ideal candidates for parallel processing. Recent advances in parallel computation have made tensors a suitable representation for hypergraphs.

Chapter 6

Conclusions

We conclude this thesis with some final remarks on the topic. In this chapter, we review the results discussed in this thesis in Section 6.1. In Section 6.2, we explore some of the future directions of research that emerge from this thesis.

6.1 Review of results

In this thesis, we have studied different representations for non-uniform hypergraphs. Broadly, these representations can be classified as matrix-based and tensor representations. Within matrix representations, we have observed several different approaches. This thesis has focussed on the weighted clique expansion and the representation using simplicial complexes. We have been able to state and prove many familiar results from 2-graphs to non-uniform hypergraphs.

In Chapter 3, we extended the weighted clique expansion to general hypergraphs and described the spectrum of a complete graph, a complete k-partite graph and a star graph. We obtained upper bounds on the largest eigenvalue of the Laplacian with respect to degrees and average degree of neighbours of a hypergraph. We obtained bounds on connectivity parameters such as edge-density and max-cut in terms of second smallest eigenvalue and the largest eigenvalue. We also proved a Cheeger inequality for uniform and non-uniform hypergraphs.

In Chapter 4, we studied simplicial complexes. We gave an alternate proof to obtain the spectrum of a complete simplex. We studied the existing results on Cheeger inequality in this representation. Also, we provided a conjecture on a generalized Mixing Lemma for simplicial complexes.

In Chapter 5, we discussed the tensor representation of hypergraphs. We studied the existing definition for eigenvalues. We obtained results for analytic connectivity of general hypergraphs. We proved an upper bound with respect to degrees and a lower bound in terms of the diameter of a hypergraph. Finally, we proved Cheeger inequality for non-uniform hypergraphs.

Recall that the Cheeger constant for graphs and hypergraphs is defined as $\phi = \min \frac{|E(S,V\setminus S)|}{|S|}$ and the modified Cheeger constant $h(X) = \min \frac{|V||F(A_0,\ldots,A_k)|}{|A_0|\ldots|A_k|}$. The spectral gap is denoted by λ, which is the second smallest eigenvalue of the Laplacian. The analytic connectivity for tensors α is defined as

$$\alpha(H) = \min_{j=1,\ldots,n} \min \left\{ \mathcal{L}x^k \ : x \in \mathbb{R}^n_+, \sum_{i=1}^{n} x_i^k = 1, x_j = 0 \right\}.$$

The result across different representations are tabulated below.

Representation	Cheeger Inequality
2-graph	$\frac{\phi^2}{2d_{\max}} \leq \lambda \leq 2\phi$
Weighted Clique	$\frac{(k_{\min}-1)^2\phi^2}{2(k_{\max}-1)d_{\max}} \leq \lambda \leq \frac{k_{\max}^2\phi}{2}$
Simplicial	$\frac{h^2}{2k+1} \leq \lambda \leq h$
Tensor	$\left(\frac{k_{\min}}{k_{\max}}\right) \frac{\phi^2}{2d_{\max}} \leq \alpha \leq \frac{k_{\max}}{2}\phi$

6.2 Future Directions

Improving the bounds

The most striking feature of the bounds is the slackness. As mentioned earlier, this slackness cannot be eliminated entirely. A factor of $(k_{\min}/k_{\max})$ is present in most bounds and removing it necessarily collapses the general hypergraph to a uniform hypergraph. However, the bounds may be improved by replacing with similar expressions. For example, perhaps we may be able to substitute $(k_{avg}/k_{\max})$ instead of $(k_{\min}/k_{\max})$. The results given here provide firm theoretical guarantees. However, the bounds may be improved further.

Proving the conjecture

An open problem contained in this thesis is the conjecture on Mixing Lemma for simplicial complexes (in Section 4.4). The reason this generalization may be an improvement is that the existing result counts the number of edges satisfying a specific condition, namely that the edge spans k sets. First, this does not directly translate for a general hypergraph. Second, the number of edges between *any* subsets might be a better measure indicative of randomness in a hypergraph. We believe the statement can be proved in the future.

Hypergraph partitioning

The key motivation for studying spectral bounds for hypergraphs is to develop clustering algorithms. Several spectral algorithms exist for hypergraph partitioning. Many algorithms currently used require some assumptions. For example, the number of clusters needs to be known a priori. It would be interesting to see how the results given here will help in clustering general hypergraphs. Also, an important question which has been ignored in this thesis is determining the appropriate representation for each application. We have avoided any discussion on the computational aspects. These questions must be explored further.

Weighted and directed hypergraphs

This work has examined unweighted and undirected hypergraphs. A natural direction to proceed is to extend these results to weighted and directed hypergraphs. Note that we already have the notion of orientation in simplicial complexes. But the feature of direction of a hyperedge appears to be different altogether. Some work has already been done for directed hypergraphs. It would be interesting to see how Laplacians behave in this case.

Real-world examples

In this thesis we have tried to cover the broadest variety of hypergraphs without any assumption on the number of vertices or the size of the edges. We have implicitly assumed that n, $k_{\min}$ and $k_{\max}$ are independent of each other. However, in most practical applications, the largest and the smallest edge may not differ by much. It is quite possible that the size of the edge is dependent on the number of vertices. The application of the results in such contexts can be investigated further.

We have studied hypergraphs at their most fundamental level in this thesis. Many of the results given here follow naturally from intuition, but have been rigorously proved nonetheless. We hope this work, the results herein as well as the new research problems that arise, prove to be a useful addition to the literature on hypergraphs.

Bibliography

Agarwal, S., K. Branson, and S. Belongie (2006). Higher order learning with graphs. In *Proceedings of the 23rd international conference on Machine learning*, pp. 17–24. ACM.

Agarwal, S., J. Lim, L. Zelnik-Manor, P. Perona, D. Kriegman, and S. Belongie (2005). Beyond pairwise clustering. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, Volume 2, pp. 838–845. IEEE.

Alon, N. and F. R. Chung (1988). Explicit construction of linear sized tolerant networks. *Discrete Mathematics 72*(1-3), 15–19.

Anandkumar, A., P. Jain, Y. Shi, and U. N. Niranjan (2016). Tensor vs. matrix methods: Robust tensor decomposition under block sparse perturbations. In *Artificial Intelligence and Statistics*, pp. 268–276.

Aouchiche, M. and P. Hansen (2010). A survey of automated conjectures in spectral graph theory. *Linear Algebra and its Applications 432*(9), 2293–2322.

Banerjee, A., A. Char, and B. Mondal (2017). Spectra of general hypergraphs. *Linear Algebra and its Applications 518*, 14–30.

Berge, C. (1984). *Hypergraphs: combinatorics of finite sets*, Volume 45. Elsevier.

Berman, A. and X.-D. Zhang (2000). Lower bounds for the eigenvalues of Laplacian matrices. *Linear Algebra and its Applications 316*(1-3), 13–20.

Biggs, N., N. L. Biggs, and E. N. Biggs (1993). *Algebraic graph theory*, Volume 67. Cambridge university press.

Bilu, Y. and N. Linial (2006). Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica 26*(5), 495–519.

Brouwer, A. E. and W. H. Haemers (2011). *Spectra of graphs.* Springer Science & Business Media.

Butler, S. K. (2008). *Eigenvalues and structures of graphs.* Ph. D. thesis, UC San Diego.

Cayley, A. (1845). *On the theory of linear transformations.* E. Johnson.

Chang, K.-C., K. Pearson, and T. Zhang (2008). Perron-Frobenius theorem for nonnegative tensors. *Communications in Mathematical Sciences 6*(2), 507–520.

Cheeger, J. (1970). A lower bound for the smallest eigenvalue of the Laplacian. *Problems in analysis*, 195–199.

Chung, F. R. and F. C. Graham (1997). *Spectral graph theory.* Number 92. American Mathematical Soc.

Cohen, E., D. Mubayi, P. Ralli, and P. Tetali (2014). Inverse expander mixing for hypergraphs. *arXiv preprint arXiv:1407.2285*.

Cooper, J. and A. Dutle (2012). Spectra of uniform hypergraphs. *Linear Algebra and its Applications 436*(9), 3268–3292.

Cui, C., Z. Luo, L. Qi, and H. Yan (2016). Computing the analytic connectivity of a uniform hypergraph. *arXiv preprint arXiv:1611.01372*.

Cvetkovic, D., S. Simic, and P. Rowlinson (2009). *An introduction to the theory of graph spectra.* Cambridge University Press.

Ducournau, A., A. Bretto, S. Rital, and B. Laget (2012). A reductive approach to hypergraph clustering: An application to image segmentation. *Pattern Recognition 45*(7), 2788–2803.

Eckmann, B. (1944). Harmonische funktionen und randwertaufgaben in einem komplex. *Commentarii Mathematici Helvetici 17*(1), 240–255.

Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak mathematical journal 23*(2), 298–305.

Fiedler, M. (1975). A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal 25*(4), 619–633.

Friedland, S., S. Gaubert, and L. Han (2013). Perron–Frobenius theorem for nonnegative multilinear forms and extensions. *Linear Algebra and its Applications 438*(2), 738–749.

Friedland, S. and R. Nabben (2002). On Cheeger-type inequalities for weighted graphs. *Journal of Graph Theory 41*(1), 1–17.

Friedman, J. and A. Wigderson (1995). On the second eigenvalue of hypergraphs. *Combinatorica 15*(1), 43–65.

Gelfand, I. M., M. Kapranov, and A. Zelevinsky (2008). *Discriminants, resultants, and multidimensional determinants.* Springer Science & Business Media.

Ghoshdastidar, D. (2016). *Consistency of Spectral Algorithms for Hypergraphs under Planted Partition Model.* Ph. D. thesis, Indian Institute of Science.

Ghoshdastidar, D. and A. Dukkipati (2014). Consistency of spectral partitioning of uniform hypergraphs under planted partition model. In *Advances in Neural Information Processing Systems*, pp. 397–405.

Ghoshdastidar, D., A. Dukkipati, et al. (2017). Consistency of spectral hypergraph partitioning under planted partition model. *The Annals of Statistics 45*(1), 289–315.

Godsil, C. and G. Royle (2001). *Algebraic graph theory.* Springer.

Govindu, V. M. (2005). A tensor decomposition for geometric grouping and segmentation. pp. 1150–1157. IEEE.

Gowers, W. T. (2006). Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combinatorics, Probability and Computing 15*(1-2), 143–184.

Gundert, A. (2013). *On expansion and spectral properties of simplicial complexes.* Ph. D. thesis, ETH Zurich.

Gundert, A. and M. Szedlák (2014). Higher dimensional Cheeger inequalities. In *Proceedings of the thirtieth annual symposium on Computational geometry*, pp. 181. ACM.

Gundert, A. and U. Wagner (2012). On Laplacians of random complexes. In *Proceedings of the twenty-eighth annual symposium on Computational geometry*, pp. 151–160. ACM.

Hatcher, A. (2002). *Algebraic topology.* Cambridge University Press.

Hoory, S., N. Linial, and A. Wigderson (2006). Expander graphs and their applications. *Bulletin of the American Mathematical Society 43*(4), 439–561.

Horak, D. and J. Jost (2013). Spectra of combinatorial Laplace operators on simplicial complexes. *Advances in Mathematics 244*, 303–336.

Hückel, E. (1931). Quantentheoretische beiträge zum benzolproblem. *Zeitschrift für Physik 70*(3-4), 204–286.

Karypis, G., R. Aggarwal, V. Kumar, and S. Shekhar (1999). Multilevel hypergraph partitioning: applications in vlsi domain. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems 7*(1), 69–79.

Kofidis, E. and P. A. Regalia (2002). On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM Journal on Matrix Analysis and Applications 23*(3), 863–884.

Kolda, T. G. and B. W. Bader (2009). Tensor decompositions and applications. *SIAM review 51*(3), 455–500.

Lee, J., M. Cho, and K. M. Lee (2011). Hyper-graph matching via reweighted random walks.

Leordeanu, M. and C. Sminchisescu (2012). Efficient hypergraph clustering. In *Artificial Intelligence and Statistics*, pp. 676–684.

Li, W., J. Cooper, and A. Chang (2017). Analytic connectivity of k-uniform hypergraphs. *Linear and Multilinear Algebra 65*(6), 1247–1259.

Lim, L.-H. (2005). Singular values and eigenvalues of tensors: a variational approach. In *Computational Advances in Multi-Sensor Adaptive Processing, 2005 1st IEEE International Workshop on*, pp. 129–132. IEEE.

Lovász, L. et al. (1993). Random walks on graphs: A survey. *Combinatorics, Paul erdos is eighty 2*(1), 1–46.

Mohar, B. (1989). Isoperimetric numbers of graphs. *Journal of Combinatorial Theory, Series B 47*(3), 274–291.

Mohar, B. (1991). Eigenvalues, diameter, and mean distance in graphs. *Graphs and combinatorics 7*(1), 53–64.

Mohar, B. (1997). Some applications of Laplace eigenvalues of graphs. In *Graph symmetry*, pp. 225–275. Springer.

Munkres, J. R. (1984). *Elements of algebraic topology*, Volume 2. Addison-Wesley Menlo Park.

Ng, A. Y., M. I. Jordan, and Y. Weiss (2002). On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pp. 849–856.

Papa, D. A. and I. L. Markov (2007). Hypergraph partitioning and clustering.

Parzanchevski, O. (2017). Mixing in high-dimensional expanders. *Combinatorics, Probability and Computing 26*(5), 746–761.

Parzanchevski, O., R. Rosenthal, and R. J. Tessler (2016). Isoperimetric inequalities in simplicial complexes. *Combinatorica 36*(2), 195–227.

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

Qi, L. (2014). H+-eigenvalues of Laplacian and signless Laplacian tensors. *Communications in Mathematical Sciences 12*(6), 1045–1064.

Qian, R., K. Zhang, and G. Zhao (2009). Hypergraph partitioning for community discovery in complex network. In *2009 International Conference on Web Information Systems and Mining*, pp. 64–68. IEEE.

Ramadan, E., A. Tarafdar, and A. Pothen (2004). A hypergraph model for the yeast protein complex network. In *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International*, pp. 189. IEEE.

Rodriguez, J. (2003). On the Laplacian spectrum and walk-regular hypergraphs. *Linear and Multilinear Algebra 51*(3), 285–297.

Rodriguez, J. (2009). Laplacian eigenvalues and partition problems in hypergraphs. *Applied Mathematics Letters 22*(6), 916–921.

Simon, H. D. (1991). Partitioning of unstructured problems for parallel processing. *Computing systems in engineering 2*(2), 135–148.

Steenbergen, J., C. Klivans, and S. Mukherjee (2014). A Cheeger-type inequality on simplicial complexes. *Advances in Applied Mathematics 56*, 56–77.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing 17*(4), 395–416.

Yang, Y. and Q. Yang (2010). Further results for Perron–Frobenius theorem for nonnegative tensors. *SIAM Journal on Matrix Analysis and Applications 31*(5), 2517–2530.

Zhou, D., J. Huang, and B. Schölkopf (2007). Learning with hypergraphs: Clustering, classification, and embedding. In *Advances in neural information processing systems*, pp. 1601–1608.

Zhu, D. (2010). On upper bounds for Laplacian graph eigenvalues. *Linear Algebra and its Applications 432*(11), 2764–2772.

Zien, J. Y., M. D. Schlag, and P. K. Chan (1999). Multilevel spectral hypergraph partitioning with arbitrary vertex sizes. *IEEE Transactions on computer-aided design of integrated circuits and systems 18*(9), 1389–1399.