Diffusion Process with Mediator on Hypergraph

Author
Yuan GAO 3035234545

Supervisor
Dr. Hubert CHAN

October 2018
Abstract

In this project, we study the spectrum of the hypergraph Laplacian and submodular transformation Laplacian. Due to the non-linearity of the hypergraph Laplacian and submodular transformation Laplacian, we resort to mathematical tools other than linear algebra to identify and characterize the spectrum of these operators. It was conjectured that the two operators possess non-trivial eigenvalues (in addition to the known second least eigenvalue)[1][2]. This project presents proofs for the conjectures over both operators.

We consider the computational problems related to the hypergraph Laplacian from two perspectives. The first is the computability of the generalized quadratic form of the hypergraph Laplacian (and hence the computability of the Rayleigh quotient of the hypergraph Laplacian), which is closely related to the concept of spectral sparsifier of hypergraphs[3]. This project will explore possible ways to improve the sparsity of the sparsifier proposed in [3]. The second is the optimization problem that naturally arises from the derivation of the spectrum of hypergraph Laplacian. The project will consider possible approximation schemes for the eigenvalue and eigenvector whose existence is proved in this project.
Acknowledgement
Contents

1 Introduction 7

2 Problem Definitions and Methodology 8
   2.1 Hypergraph Laplacian ............................................. 8
   2.2 Generalized Laplacian of Submodular Transformations ............ 9
   2.3 Least Non-trivial Eigenvalue of the Laplacian Operators .......... 9
   2.4 Sketch Proof of the Existence of Maximum Eigenvalue ........... 10
      2.4.1 Maximum Eigenvalue of Hypergraph Laplacian ............... 11
      2.4.2 Maximum Eigenvalue of Submodular Transformation Laplacian 12
      2.4.3 Comparison with the Diffusion Process Argument ............ 13
   2.5 Complexity Overhead .............................................. 14
      2.5.1 Spectral Sparsifier of Hypergraph .......................... 14
      2.5.2 Approximation of the Maximum Eigenvalue .................... 15

3 Conclusion 15

Appendices 18
List of Figures
List of Tables
Abbreviation
1 Introduction

The spectral theory studies the spectral properties of graphs and their relation to the combinatorial properties of graph. Central to this rich theory, the Cheeger’s inequality bounds the expansion or conductance of the graph by the second least eigenvalue of the Laplacian of the graph\[4\]. More specifically, for a graph $G = (V, E)$, its expansion is defined to be:

$$\phi_G := \min_{S \subseteq V} \frac{|\partial S|}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}},$$

where $\text{vol}(S)$ is the sum of the degree of vertices in $S$ and $\partial S$ is the cut of $S$. The Cheeger’s inequality\[4\] is as the following:

$$\frac{\lambda_2}{2} \leq \phi_G \leq \sqrt{2\lambda_2},$$

where $\lambda_2$ is the second least eigenvalue of the normalized Laplacian $L_G$. This inequality can be utilized in the design of various algorithms.

Recently, the Cheeger’s inequality was generalized to the case of hypergraph with notions of hypergraph expansion and hypergraph Laplacian\[1\][2]. This hypergraph Laplacian, denoted by $L_\omega$, is defined to be an operator that determines a diffusion process with mediators over hypergraphs. In \[1\][2], it is shown that this Laplacian operator $L_\omega$ possesses a second least eigenvalue $\lambda_2$ that satisfy an inequality similar to that of Cheeger’s:

$$\frac{\lambda_2}{2} \leq \phi_H \leq \sqrt{2\lambda_2},$$

where $\phi_H$ is the expansion of hypergraph. This result shows that the spectral properties of hypergraph Laplacian can indeed be linked to combinatorial properties of hypergraph. However, unlike graph Laplacians, which are symmetric linear operators, the hypergraph Laplacians are non-linear and non-smooth, which makes its spectrum hard to characterize. Hypergraph Laplacian is shown to have a trivial eigenvalue and a second least eigenvalue\[1\][2] and in this project the conjecture that the hypergraph Laplacian possesses a maximum eigenvalue is proved. Furthermore, the maximum eigenvalue of the hypergraph Laplacian is related to the approximation of the hypergraph max cut, the computation of which is NP-hard. It is thus interesting to consider the approximation of the maximum eigenvalue and its corresponding eigenvectors. One of the goals of the project would therefore be studying possible approximation algorithms of the maximum eigenvalue of the hypergraph Laplacian.

A more generalized notion of Laplacian can be defined with respect to submodular transformations\[5\]. Hypergraph Laplacians and graph Laplacians can be viewed as special cases of the generalized Laplacian by considering different submodular transformations corresponding to hypergraphs or graphs. This more abstract construction gives rise to some novel Cheeger’s inequalities under the same unifying framework. It is thus interesting to consider the spectral properties of this Laplacian. In \[5\], the generalized Laplacian is shown to possess two eigenpairs, similar to the case of hypergraph Laplacians. One possible research direction along this line would therefore be the study of other eigenpairs of the generalized Laplacians. The conjecture is that the generalized Laplacian also possesses a maximum eigenvalue.
Hence a proof for this conjecture is studied in this project.

One particular obstacle of the application of spectral theory of hypergraphs is the size of the edge set of hypergraphs. For a hypergraph with \( n \) vertices, the number of the edges of the hypergraph can be of the order \( O(2^n) \), which can cause the evaluation of the hypergraph Laplacian and the induced quadratic form intractable. In [3], a spectral sparsifier of hypergraphs is proposed, with a theoretical guarantee that the number of edges of the sparsified hypergraph is of the order \( O(n^3 \log n / \epsilon^2) \) with high probability, where \( \epsilon \) is the error of the quadratic form of the sparsified hypergraph. It might be interesting to see if it is possible to further reduce the number of the edges in the sparsified hypergraph. Application-wise speaking, as this sparsifier makes the computations involving the quadratic form of the Laplacian of dense hypergraphs feasible, experiments could be carried out to test how this sparsifier might improve the performances of algorithms that involves hypergraph Laplacian, for example semi-supervised learning on hypergraphs[6].

The remainder of this progress report is as follows. First, The report will give more explanations on the notions of hypergraph Laplacian, submodular transformation Laplacian and the related concepts. Then the methodology of manipulating these mathematical concepts will be described, with sketches of the proof strategies of several conjectures about the spectrum of these operators explained. The report will proceed to discuss the results obtained so far and identify the encountered difficulties. It will close with a conclusion, with possible directions of further work emphasized.

2 Problem Definitions and Methodology

2.1 Hypergraph Laplacian

In this project we consider edge-weighted graphs \( H = (V, E, \omega) \) where \( \omega \) is a consistent weight function. We assume that all vertex weights are positive. Given \( f \in \mathbb{R}^V \), \( f_u, u \in V \) is the coordinate corresponding to \( u \in V \). In addition, the space \( \mathbb{R}^V \) is endowed with the inner product \( \langle \cdot, \cdot \rangle_\omega \) defined as \( \langle f, g \rangle_\omega = \langle f, Wf \rangle \) where \( W \) is the diagonal matrix of the weights of the vertices. This inner product space is called the density space.

In[2], the notion of diffusion operator with mediator is defined. For each edge \( e \in E \), let \( [e] = e \cup \{0\} \), where 0 does not correspond to any vertex. A set of constants \( \beta^e \) is defined over \( [e] \) with \( \beta^e_j \) corresponding to all \( j \in [e] \). Then the following associated generalized quadratic form is considered:

\[
Q(f) := \sum_{e \in E} \omega_e \left\{ \beta^e_0 \max_{s, i \in e} (f_s - f_i)^2 + \sum_{j \in e} \beta^e_j \left[ (\max_{s \in e} f_s - f_j)^2 + (\min_{i \in e} f_i - f_j)^2 \right] \right\}.
\]

Also, for \( f \neq 0 \in \mathbb{R}^V \), \( \frac{Q(f)}{\sum_{u \in V} \omega_u f_u^2} \) is defined to be its discrepancy ratio.

As in[2], a diffusion process (with mediator) can be defined according to a set of rules. Then by considering the differentiation of the density of the measure over the vertices in the diffusion process, the Laplacian \( L_\omega \) can be defined by \( -L_\omega f = \frac{df}{dt} \), where \( t \) is the time of the diffusion process. The detailed explanation of the diffusion
process with mediator and the proof of its existence can be found in [1] and [2]. It is also shown that the Rayleigh Quotient of the hypergraph Laplacian $L_\omega$, defined as $R(f) = \frac{\langle L_\omega f, f \rangle}{\|f\|^2}$ for any $0 \neq f \in \mathbb{R}^V$, coincides with the discrepancy ratio of $f$.

\section{2.2 Generalized Laplacian of Submodular Transformations}

The notion of Laplacian can be further generalized for submodular transformations, which is the product of submodular functions over the power set of some set $V$, or equivalently $\{0,1\}^V$ [5]. In other words, a function $F : \{0,1\}^V \rightarrow \mathbb{R}^E$ is said to be a submodular transformation if each of its component function is submodular. A set function $F_e : \{0,1\}^V \rightarrow \mathbb{R}$ is a submodular function if $F_e(S) + F_e(T) \geq F_e(S \cap T) + F_e(S \cup T), \forall S,T \subset V$. For each submodular $F_e$, there exists a Lovász extension $f_e : [0,1]^V \rightarrow \mathbb{R}$ which is defined to be the convex closure of the submodular function $F_e$. More specifically, [5] considers the submodular polyhedron $P(F)$ and the base polytope $B(F)$ of $F$:

$$P(F) = \{x \in \mathbb{R}^V | \sum_{v \in S} x(v) \leq F(S), \forall S \in V\} \text{ and } B(F) = \{x \in P(F) | \sum_{v \in V} x(v) = F(V)\}.$$ 

With $P(F)$ and $B(F)$, the Lovász extension of $f_e : \mathbb{R}^V \rightarrow \mathbb{R}$ of a submodular function $F_e : \{0,1\}^V \rightarrow \mathbb{R}$ is:

$$f_e(x) = \max_{\omega \in B(F_e)} \langle \omega, x \rangle.$$ 

Consequently we define $\partial f_e = \arg \max_{\omega \in B(F_e)} \langle \omega, x \rangle$.

In [5], the notion of Laplacian operator $L_F : \mathbb{R}^V \rightarrow 2^{\mathbb{R}^V}$ is defined for the submodular transformation $F$ as follows:

$$L_F(x) := \{\sum_{e \in E} \langle \omega_e, x \rangle | \omega_e \in \partial f_e(x)\},$$

The Lovász extension $f$ of a submodular transformation is defined to be the product of Lovász extensions of each component functions of $F$. Then [5] shows that the Rayleigh quotient $R_F : \mathbb{R}^V \rightarrow \mathbb{R}$ of the generalized Laplacian $L_F$ for submodular transformation $F$ is:

$$R_F(x) := \frac{\langle x, L_F(x) \rangle}{\|x\|^2} = \frac{\|f(x)\|^2}{\|x\|^2}, \text{ for } x \neq 0,$$

where $\|\cdot\|_2$ is the 2-norm over $\mathbb{R}^V$. Graph Laplacian and hypergraph Laplacian are special cases of this generalized Laplacian. All graphs and hypergraphs can be identified with submodular transformations whose Laplacian coincide with the graph and hypergraph Laplacian respectively in a natural way.

\section{2.3 Least Non-trivial Eigenvalue of the Laplacian Operators}

It is obvious that the hypergraph Laplacian obtains a trivial eigenpair. More specifically, $1 \in \mathbb{R}^V$ is an eigenvector of the Laplacian with 0 being its corresponding eigenvalue. It is not as trivial to find other eigenpairs of the Laplacian. In [1][2], it is shown that $\lambda_2 := \min_{1 \perp f \in \mathbb{R}^V} R(f)$ is an eigenvalue with any $1 \perp f$ attaining this value being an eigenvector. The proof of $\lambda_2$ being an eigenvalue is inspired by the idea that the diffusion process over hypergraph tends to mix the measures across the
vertices and reduces the discrepancy ratio (precisely speaking, the diffusion process reduces the discrepancy ratio when the current state is not an eigenvector).

In more details, [1][2] first showed that the diffusion process is smooth with respect to the time $t$ of the diffusion process (note that this does not imply the smoothness of the hypergraph Laplacian over the density space), which implies the existence of the right-hand derivative of the generalized quadratic form of the hypergraph. With careful computations, it is shown that

$$\frac{dR(f)}{dt} = -\frac{1}{\|f\|_2^2}(\|f\|_2^2 \|Lf\|_2^2 - \langle f, Lf \rangle) \leq 0$$

where the inequality follows from the Cauchy inequality. It follows that as the diffusion progresses for sufficiently small amount of time, the discrepancy ration will reduce if the initial state is not an eigenvector of the hypergraph Laplacian.

It should be pointed out that the submodular transformation Laplacian $L_F$ is a set-valued operator. Therefore the definition of the spectrum of $L_F$ should be stated as the following[5]:

$$(\lambda, x)$$ is an eigenpair of the $L_F$ if and only if $\lambda x \in L_F(x)$.

It is immediate that $(0, 1)$ is a trivial eigenpair of $L_F$. However, it is not entirely obvious how to generalize the concept of diffusion process over hypergraph to submodular transformations on which the proof of the existence of a non-trivial eigenpair of $L_F$ is based. Nevertheless, [5] defined a diffusion process similar to the case of hypergraph, by defining the time evolution of any vector $x \in \mathbb{R}^V$ as:

$$\frac{dx}{dt} \in -L_F(x),$$

whose well-definedness requires justifications. Essentially, [5] showed that for any initial state $x$, there exists some choices $L_t$ such that $\frac{dx}{dt} = -L_t x$ is well-defined for any time $t \geq 0$. With this diffusion process defined, by generalizing the proof that hypergraph Laplacian possesses a non-trivial eigenpair, it can be shown that $L_F$ possesses a non-trivial eigenpair as well.

It was conjectured that $\lambda^* = \max_{f \neq 0 \in \mathbb{R}^V} R(f)$ is the maximum eigenvalue of the hypergraph Laplacian, with any $x \perp 1$ attaining $\lambda^*$ being an eigenvector. A similar conjecture can be made for the submodular transformation Laplacian. However, if one tries to generalize the proof strategy based on diffusion process, s/he will have to resolve the problem of reversibility of the diffusion process. It does not seem possible to bypass this problem in the framework of diffusion process. Therefore, in the next subsection, we will study a new proof strategy for both of the Laplacian operators.

### 2.4 Sketch Proof of the Existence of Maximum Eigenvalue

I will first illustrate the main frame of the proof using the hypergraph Laplacian and then move on to the more complicated case of submodular transformation Laplacian which requires some results from linear programming on $\epsilon-$perturbation [7]. This subsection will close with a brief comparison between the proposed proof and the proof based on diffusion process.
2.4.1 Maximum Eigenvalue of Hypergraph Laplacian

For the case of hypergraph, consider any \( 0 \neq f^* \in \mathbb{R}^V \) such that \( R(f^*) = \lambda^* \) and any \( 0 \neq v \in \mathbb{R}^V \). Let \( L(R, f^*, v) \) denote the left-hand derivative of the function \( R(f^* + \epsilon v) \) with respect to \( \epsilon \) at 0 (note that the reason that left-hand derivative is considered is only for the sake of convenience), that is,

\[
L(R, f^*, v) = \lim_{\epsilon \to 0^-} \frac{R(f^* + \epsilon v) - R(f^*)}{\epsilon}.
\]

Of course, the existence of the limit requires justification. Taking for granted the existence of the limit for the moment, let’s consider the consequence of this definition. If for some \( 0 \neq v \in \mathbb{R}^V \), \( L(R, f^*, v) < 0 \), then there exists \( \epsilon < 0 \) with \( |\epsilon| \) sufficiently small such that \( R(f^* + \epsilon v) > R(f^*) \), which contradicts our definition of \( f^* \) being a maximizer of \( R \). In particular, if we can show that for \( v = -L_\omega f^*, L(R, f^*, -L_\omega f^*) < 0 \) if and only if \( L_\omega f^* \notin \text{span}(f^*) \), then we will have essentially constructed a proof by contradiction for \((\lambda^*, f^*)\) being an eigenpair.

Previous discussions show that to prove that \((\lambda^*, f^*)\) is an eigenpair, it is necessary to compute \( L(R, f^*, v) \) (and hence give a justification for the existence of this limit). To do this, we first consider \( L(Q, f^*, v) \), the left-hand derivative of \( Q(f^* + \epsilon v) \) with respect to \( \epsilon \) at 0, where \( Q \) is the generalized quadratic form of hypergraph. We give the following definitions and identity to aid the computation.

For any \( e \in E, \)

\[
c^l_e := \omega_e [\beta_0^s (\max_{s \in e} f^*_s - \min_{i \in e} f^*_i) + \sum_{j \in e} \beta_j^s (f^*_j - \min_{i \in e} f^*_i)],
\]

\[
c^s_e := \omega_e [\beta_0^s (\max_{s \in e} f^*_s - \min_{i \in e} f^*_i) + \sum_{j \in e} \beta_j^s (\max_{s \in e} f^*_s - f^*_j)],
\]

\[
S_e := \text{argmax}_{s \in e} f^*_s,
\]

\[
I_e := \text{argmax}_{i \in e} f^*_i
\]

and

\[
c_j := \sum_{e \in E, j \in e} \beta_j^s \omega_e (\max_{s \in e} f^*_s + \min_{i \in e} f^*_i - 2f^*_j),
\]

and write \( r := \frac{df^*}{dt} \). Certainly as one can see, these definitions are originated from the definition of the diffusion process (which will not be discussed in details here), however, they are presented here only for the sake of computation. Then I give without proof the following identity:

\[
\sum_{e \in E} c^s_e \max_{s \in S_e} r_s - \sum_{e \in E} c^l_e \min_{i \in I_e} r_i - \sum_{j \in V} c_j^r j = -\|r\|^2_\omega.
\]

With these definitions, we can show that for any \( 0 \neq v, L(Q, f^*, v) \) exists and:

\[
L(Q, f^*, v) \leq 2 \sum_{e \in E} c^s_e \max_{s \in S_e} v_s - \sum_{e \in E} c^l_e \min_{i \in I_e} v_i - \sum_{j \in V} c_j v_j.
\]

and in particular, for \( v = r = -L_\omega f^* \),

\[
L(Q, f^*, -L_\omega f^*) \leq -2\|r\|^2_\omega
\]

It implies that for any \( 0 \neq v, L(R, f^*, v) \) exists. In particular:
\[ L(R, f^*, -L_\omega f^*) \leq -\frac{1}{\| f^* \|^2} (\| f^* \|^2 \| L_\omega f^* \|^2 - \langle f^*, L_\omega f^* \rangle) \leq 0, \]

where the last inequality follows from the Cauchy inequality and the equality is satisfied if and only if \( L_\omega f^* \in \text{span}(f^*) \). As discussed above, this implies that \((\lambda^*, f^*)\) is indeed an eigenpair of \( L_\omega \).

Note that many computations are omitted in the above discussion. However, it should be pointed out that \( L(Q, f^*, v) \) is computed explicitly by the definition of left-hand side derivative, that is, an explicit formula of \( Q(f + \epsilon v) \) for \( \epsilon < 0 \) is found, which, however, is not feasible for the case of submodular transformation Laplacian. This problem is solved with the \( \epsilon \)-perturbation method in linear programming [7] (to eliminate confusion with notations, please note that in the following subsection for submodular transformation, \( \epsilon \) and \( f \) will not carry the same meaning as in the discussion for hypergraph).

### 2.4.2 Maximum Eigenvalue of Submodular Transformation Laplacian

Recall the definition of \( f \) as the Lovász extension of a submodular transformation \( F \) and \( L_F \) as the Laplacian of the submodular transformation \( F \). Also, recall that for any \( x \in \mathbb{R}^V \), \( R_F(x) = \| f(x) \|_2^2 \). We note briefly in the preceding subsection that the main difficulty for the discussion of submodular transformation is that \( \| f(x) \|_2^2 \) is difficult to handle. Precisely speaking, it is not very clear how to compute the following expression explicitly for \( \delta \in \mathbb{R} \), with \( |\delta| \) sufficiently small and any \( v \in \mathbb{R}^V \):

\[ \| f(x + \delta v) \|_2^2 - \| f(x) \|_2^2 = \sum_{e \in E} f_e(x + \delta v)^2 - f_e(x)^2. \]

This subsection will therefore focus on \( f_e(x + \delta v)^2 - f_e(x)^2 \) for some \( e \in E \). A natural guess is that there exists some operators \( L_x \) such that for any \( \delta \in \mathbb{R} \), with \( |\delta| \) sufficiently small and any \( v \in \mathbb{R}^V \), \( f_e(x)^2 = \langle x, L_x x \rangle \) and \( f_e(x + \delta v)^2 = \langle x + \delta v, L_x (x + \delta v) \rangle \). To this end, I rewrite the definition of \( f_e(x) \) as linear program such that \( f_e(x) \) equals to:

\[
\begin{align*}
\text{Maximize} & \quad \omega^T x \\
\text{s.t.} & \quad 1_s^T \omega \leq F_e(S), \forall S \subset V \\
& \quad 1_V^T \omega = F_e(V)
\end{align*}
\]

Suppose we order all elements except for \( \emptyset \) and \( V \) in the power set of \( V \) and mark them as \( S_1, S_2, \ldots, S_{2^{|V| - 2}} \), then equivalently, \( f_e(x) \) can be defined by the dual problem of the above linear program as \( P_e(x) \):

\[
\begin{align*}
\text{Minimize} & \quad F_e^T y \\
\text{s.t.} & \quad Ay = x \\
& \quad y \geq 0
\end{align*}
\]

where \( A = \begin{pmatrix} 1_s_1 & \cdots & 1_s_{2^{|V| - 2}} & 1_v & -1_v \end{pmatrix} \in \mathbb{Z}^{|V| \times 2^{|V|}} \) is of full row rank and \( F_e = \begin{pmatrix} F_e(S_1) & \cdots & F_e(S_{2^{|V| - 2}}) & F_e(V) & -F_e(V) \end{pmatrix}^T \in \mathbb{R}^{2^{|V|}} \). Then naturally, we have transformed the discussion on \( f_e(x) \) and \( f_e(x + \delta v) \) into the discussion on the
Suppose that $B$ is an optimal basis of $(P_e(x))$, then the goal is to justify that $B$ is also an optimal basis of $(P_e(x + \delta v))$ for $|\delta|$ sufficiently small. At first sight, this is rather trivial due to the principle of sensitivity analysis. However, one will run into problem because of the possible degeneracy of $(P_e(x))$ and will notice that $B$ might not be optimal for $(P_e(x + \delta v))$, no matter how small $|\delta|$ is. Then the main difficulty here is to avoid degeneracy in the discussion.

In [7], an $\epsilon-$perturbation method is analyzed. It stated that there exists $\epsilon \in \mathbb{R}^V$ such that the linear programs $(P_e(x, \epsilon))$ and $(P_e(x + \delta v, \epsilon))$ are equivalent to $(P_e(x))$ and $(P_e(x + \delta v))$ respectively, in terms of the feasibility and boundedness, and any optimal bases of $(P_e(x, \epsilon))$ and $(P_e(x + \delta v, \epsilon))$ are optimal bases in $(P_e(x))$ and $(P_e(x + \delta v))$ respectively, and most importantly, $(P_e(x, \epsilon))$ and $(P_e(x + \delta v, \epsilon))$ are non-degenerate, all under the condition that $P_e(x)$ and $(P_e(x + \delta v))$ are all integer-valued. The definitions of the linear program $(P_e(x, \epsilon))$ can be viewed in [7] and will not be stated here. Note that the linear program $(P_e(x))$ are real-valued, hence the results of [7] are not immediately applicable in our discussion. However, we can modify the arguments in [7] slightly so that they work for the real-valued linear programs as well. The details will be omitted here.

The consequence is that there is indeed an optimal basis $B$ of $(P_e(x))$ such that $B$ is also an optimal basis of $(P_e(x + \delta v))$. Hence for $\delta$ with $|\delta|$ sufficiently small and any $x, v \in \mathbb{R}^V$, there exists an $\omega_{e,x} \in \mathbb{R}^V$ such that $f_e(x) = \langle \omega_{e,x}, x \rangle$ and $f_e(x + \delta v) = \langle \omega_{e,x}, x + \delta v \rangle$. With this result, the computation of $f_e(x + \delta v)^2 - f_e(x)^2$ is much clearer than before. The remaining arguments for the proof that $L_F$ possesses another eigenpair is simple and similar to that of the case of hypergraph.

### 2.4.3 Comparison with the Diffusion Process Argument

In this subsection, I will focus on the case of hypergraph Laplacian and adopt the notations for it. As mentioned in the above subsections, the main difficulty of applying the diffusion process argument to the maximum eigenvalues of the hypergraph Laplacian is the reversibility of the diffusion process. Intuitively speaking, if $0 \neq f \in \mathbb{R}^V$ is not an eigenvector of the hypergraph Laplacian, then reversing the diffusion process will give a state that has larger discrepancy ratio (value of the Rayleigh quotient). Such argument could lead to a proof that $(\lambda^*, f^*)$ is an eigenpair of hypergraph Laplacian, if the diffusion process is indeed (locally) reversible. However, it is possible that $f^*$ can only be an initial state of the diffusion process, which disputes the reversibility of the diffusion.

In comparison, the argument presented in this project bypasses the diffusion process and study the hypergraph Laplacian solely as an operator over $\mathbb{R}^V$. However, one might notice that the study of $L(R, f^*, -L_\omega f^*)$ is essentially "simulating" the reverse of diffusion process. Such an argument gives more freedom in the manipulation of $L$ than the diffusion process argument since we are not confined to the time evolution of the diffusion process, instead we can now look at the "evolution" in any direction.
2.5 Complexity Overhead

It has already been established that the Laplacians do possess a maximum eigenvalue. However, to really make these results useful, it is necessary to deal with the computational issue arises around the spectrum of the Laplacians. In this section, two major computational issues about the spectral properties of the Laplacians will be discussed. The first is on the computation of the generalized quadratic form, which is center to all the computational problems on the spectral properties. The second is about the approximation of the maximum eigenvalue of the the Laplacian. For simplicity, this section will only consider hypergraph Laplacians.

2.5.1 Spectral Sparsifier of Hypergraph

The size of the edge set of hypergraphs can be of order \(2^n\). Therefore, in the worst-case analysis, the evaluation of the hypergraph Laplacian and Rayleigh quotient (or the generalized quadratic form) is computationally intractable. It is therefore important to consider the approximation of the generalized quadratic form of a hypergraph with the generalized quadratic form of another hypergraph with a polynomial-sized edge set. Formally, a subgraph \(H'\) of the hypergraph \(H\) is an \(\epsilon\)-spectral sparsifier of \(H\) if:

\[
(1 - \epsilon)Q_H(f) \leq Q_{H'}(f) \leq (1 + \epsilon)Q_H(f)
\]

for every \(f \in \mathbb{R}^V\). To make the sparsifier useful, it is important that the size of the edge set of \(H'\) is small. A randomized algorithm is proposed in [3] that can generate a sparsifier with \(O(n^3 \log n / \epsilon^2)\) edges with high probability for any hypergraph, which makes the computation of various problems that involves the quadratic form of hypergraphs feasible. This algorithm is a simple (the analysis of the algorithm is certainly not simple) sampling algorithm that samples each edge of \(H\) with a probability. This sparsifier is of order \(O(n^3 \log n / \epsilon^2)\) [3]. However, a hypergraph with edges that contains no more than 3 vertices has no more than \(O(n^3)\) edges, which means that the proposed sparsifier will not sparsify this hypergraph at all. Hence it would be interesting to consider sparsifiers that might give better theoretical guarantees. An obvious possible direction for improvement is to design better sampling probabilities for each edge. However, smaller sampling probability might cause worse approximation ratio of the generalized quadratic form. For now, it is not very clear how we might tackle this issue.

Another possible direction of research might be to consider the cut sparsifier. Formally, a subgraph \(H''\) of the hypergraph \(H\) is an \(\epsilon\)-cut sparsifier of \(H\) if:

\[
(1 - \epsilon)Q_{H''}(f) \leq Q_H(f) \leq (1 + \epsilon)Q_{H''}(f)
\]

for every \(f \in \mathbb{R}^V\) such that \(f = 1_S\) for some \(S \subseteq V\). That is, for cut-sparsifiers, we only consider those \(f\)'s that are indicator vectors for the subsets of the \(V\). \(H''\) is called the cut-sparsifier because for any \(S \subseteq V, Q_H(1_S)\) is equal to the cut size of \(S\) in \(H\). Clearly, an \(\epsilon\)-sparsifier is an \(\epsilon\)-cut sparsifier, while the converse is not true. Therefore, it might be easier to design cut sparsifiers of small edge set than spectral sparsifier.

Due to the fact that many applications of the hypergraph spectral theory requires
computations of the generalized Laplacian, the spectral sparsifier is an extremely useful tool. One particular example of the possible application of the sparsifier might be on the semi-supervised learning on hypergraphs[6], which in principle can be accelerated by spectral sparsifiers[3]. In [6], a learning task on (directed) hypergraph is considered. Labels $\hat{x}^*(u)$ are given for some vertices $u \in L \subset V$ and the task is to predict the labels of vertices in $V \setminus L$. This problem is then transformed into an optimization problem with the generalized quadratic form as the objective. A subgradient method is proposed to solve this optimization problem[6]. Since clearly this optimization problem involves the generalized quadratic form, applying the spectral sparsifier of hypergraph in this case should be able to speed up the computations significantly, especially when the hypergraph model is dense[3]. Then an experimental work of this project will be on finding dense hypergraph models in the scenario of semi-supervised learning and experiment the spectral sparsifier with it.

2.5.2 Approximation of the Maximum Eigenvalue

The eigenpair $(\lambda^*, f^*)$ is closely related to the max-cut of hypergraphs, the computation of which is NP-hard. Hence knowing the maximum eigenvalue and its corresponding eigenvector might be useful for the approximation of the max cut, which requires an approximation algorithm for the maximum of the Rayleigh quotient of hypergraphs. Therefore, one of the objectives of this project will be on the approximation of this eigenpair.

In [1], a semi-definite programming (SDP) relaxation and a rounding algorithm is proposed for the approximation of the procedural minimizers of the discrepancy ratio of hypergraphs. The first two eigenpairs of the hypergraph Laplacian are in fact the first two procedural minimizers of the discrepancy ratio. One possible idea for the approximation of the maximum eigenvalue of the hypergraph Laplacian might be formulating a similar SDP relaxation and rounding algorithm. However, it is not known if a similar SDP relaxation is possible since the two approximation problems are subtly different. This project will investigate into this idea and perhaps explore various techniques on solving SDP.

3 Conclusion

This project studies both the theoretical and computational aspect of the spectral theory of hypergraphs and submodular transformations. The study of these theories can connect to various areas of interests, including approximation schemes and optimization theory. Besides the theoretical flavor in the project, it is anticipated that the project’s output might benefit our understandings of practical problems, including semi-supervised learning and semi-definite programming.

This report described sketch proofs for the conjectures on the spectrum of the Laplacians defined in the previous section. The proof touches upon various interesting mathematical theories, including some theoretical results on linear programming. We also considered some computational issues on the spectral theories, namely the spectral sparsifier of hypergraph and the approximation of the maximum eigenvalue of the hypergraph Laplacian. We identified that it is difficult to reduce the com-
plexity overhead induced by these computational problems and the study of them might require some major improvements on the mathematical tools used.

As is clear from the previous section, there are a plenty of room for future works. The approximation of the maximum eigenvalue is still an open problem. An SDP relaxation of this problem would induce a max/max optimization situation, which is rather tricky. Future work might look into possible ways that could bypass this issue. Improving the theoretical guarantees of the spectral sparsifier is also a major direction of future work. The central problem is the design of better sampling probability. This is a challenging question due to the requirement of maintaining a good approximation ratio of the generalized quadratic form.
References


Appendices

Appendix