Spanning tree modulus: deflation and hierarchical graph structure

by

Jason Clemens

B.S., Missouri Valley College, 2012

M.S., Kansas State University, 2014

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Mathematics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

The concept of discrete p-modulus provides a general framework for understanding arbitrary families of objects on a graph. The p-modulus provides a sense of "structure" of the underlying graph, with different families of objects leading to different insight into the graph's structure. This dissertation builds on this idea, with an emphasis on the family of spanning trees and the underlying graph structure that spanning tree modulus exposes.

This dissertation provides a review of the probabilistic interpretation of modulus. In the context of spanning trees, this interpretation rephrases modulus as the problem of choosing a probability mass function on the spanning trees so that two independent, identically distributed random spanning trees have expected overlap as small as possible.

A theoretical lower bound on the expected overlap is shown. Graphs that attain this lower bound are called homogeneous and have the property that there exists a probability mass function that gives every edge equal likelihood to appear in a random tree. Moreover, any nonhomogeneous graph necessarily has a homogeneous subgraph (called a homogeneous core), which is shown to split the modulus problem into two smaller subproblems through a process called deflation.

Spanning tree modulus and the process of deflation establish a type of hierarchical structure in the underlying graph that is similar to the concept of core-periphery structure found in the literature. Using this, one can see an alternative way of decomposing a graph into its hierarchical community components using homogeneous cores and a related concept: minimum feasible partitions.

This dissertation also introduces a simple greedy algorithm for computing the spanning tree modulus that utilizes any efficient algorithm for finding minimum spanning trees. A theoretical proof of the convergence rate is provided, along with computational examples.

Spanning tree modulus: deflation and a hierarchical graph structure

by

Jason Clemens

B.S., Missouri Valley College, 2012

M.S., Kansas State University, 2014

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Mathematics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Nathan Albin

Copyright

Abstract

The concept of discrete p-modulus provides a general framework for understanding arbitrary families of objects on a graph. The p-modulus provides a sense of "structure" of the underlying graph, with different families of objects leading to different insight into the graph's structure. This dissertation builds on this idea, with an emphasis on the family of spanning trees and the underlying graph structure that spanning tree modulus exposes.

This dissertation provides a review of the probabilistic interpretation of modulus. In the context of spanning trees, this interpretation rephrases modulus as the problem of choosing a probability mass function on the spanning trees so that two independent, identically distributed random spanning trees have expected overlap as small as possible.

A theoretical lower bound on the expected overlap is shown. Graphs that attain this lower bound are called homogeneous and have the property that there exists a probability mass function that gives every edge equal likelihood to appear in a random tree. Moreover, any nonhomogeneous graph necessarily has a homogeneous subgraph (called a homogeneous core), which is shown to split the modulus problem into two smaller subproblems through a process called deflation.

Spanning tree modulus and the process of deflation establish a type of hierarchical structure in the underlying graph that is similar to the concept of core-periphery structure found in the literature. Using this, one can see an alternative way of decomposing a graph into its hierarchical community components using homogeneous cores and a related concept: minimum feasible partitions.

This dissertation also introduces a simple greedy algorithm for computing the spanning tree modulus that utilizes any efficient algorithm for finding minimum spanning trees. A theoretical proof of the convergence rate is provided, along with computational examples.

Contents

List of Figures

Acknowledgments

First, I would like to express my deepest gratitude to my advisor Dr. Nathan Albin. It has been an honor to be his PhD student. I am especially thankful for the patience and commitment he made to me when I left KSU for a teaching position in Missouri. I would have not had the success that I have had with out his help and support not only with the math but also with the family and life questions that came up.

I would also like to thank Dr. Pietro Poggi-Corradini. The boundless knowledge and insight into whatever questions came up brought many successful thoughts to fruition. In addition to his knowledge, I appreciate his interest and support of my work throughout the years.

Next, I would like to thank my other committee members: Dr. Ivan Blank and Dr. Stephen Welch. They have been very supportive throughout my tenure at Kansas State. Although my studies deviated from what was initially discussed and introduce by Dr. Welch, he remained interested in the new direction my studies went, and I am very grateful. Dr. Blank always provided a much needed perspective on life and my studies and never shied away from telling me straight, which is not always comforting but always necessary.

Lastly, I wouldn't have even made it to this point without the support of my family. My parents provided me with the tools necessary to succeed, and a kick in the ass when I wasn't doing the right things. Most importantly, however, is my wife Amber. Without her undying love and support, through all the trials and tribulations that we faced, she has never wavered and never stopped believing in this dream. Finally, my three beautiful children, Matthew, Mackenzie, and Maxwell, thank you for giving your mother and I happiness, joy, and tears every day.

Dedication

To my wife

and

My children

Chapter 1

Introduction

The aim of this dissertation is to introduce key properties of spanning tree modulus on a graph and an algorithm for computing it. The approach for the algorithm is greedy, which uses the minimum spanning tree at each step. Using this approach, we show that it is possible to determine homogeneity of a graph in polynomial time. Once one is able to compute spanning tree modulus efficiently, it can be shown that this gives an approach for community detection.

1.1 Motivation

1.1.1 Why spanning trees?

Spanning trees have been used to establish results for many long-standing problems. Most of these fall into the category of minimum spanning tree (MST) problems. The first, and most notable problem is that of simply finding a minimum spanning tree of a graph.

There are several, well known, historical algorithms for finding a minimum spanning tree, three of them are Kruskal's algorithm [27], Prim's algorithm [33], and Borůvka's algorithm [31]. A thorough history of the development of these algorithms (and some others) is detailed in [21]. As with any algorithm, one of the most important properties is the running time, or how fast it can produce or approximate an answer to the problem being studied. The running time for Kruskal's algorithm and Borůvka's algorithm are $\mathcal{O}(|E|\log(|V|))$, Prim's algorithm can be as slow as $\mathcal{O}(|V|^2)$ and as fast as $\mathcal{O}(|E| + |V|\log(|V|))$ when implementing certain heap techniques, such as Fibonacci heaps [15].

There have also been some more recently developed algorithms, one by Gabow et. al., [18], that has a running time of $\mathcal{O}(|E|\log(\beta(|E|, |V|)))$ where $\beta(|E|, |V|) = \min\{i : \log^{(i)}(|E|) \leq \frac{|E|}{|V|}\}$; and another algorithm developed by Karger et.al.(see [25]) based on Borůvka's algorithm in conjunction with a random sampling technique. Using the two combined techniques, the algorithm removes edges that cannot belong to the minimum spanning tree. This algorithm has linear running time, $\mathcal{O}(|E|)$ with high probability. However, they do note that the worst case is on the order of $\min\{|V|^2, |E|\log(|V|)\}$. Also, it is worth noting that in the construction of this algorithm the restriction that the edge costs be unique was implemented.

All of these algorithms are used under the assumption that only binary comparisons between the edges can be made, known as the random-access model. If bucketing and bit manipulation of the edge weights is allowed, Fredman and Willard [16] showed that the MST problem can be solved deterministically in linear time. More recently, in 2001, Chazelle [9] showed that with the random-access model along with a soft-heap approach, with no restrictions on the edge cost values, that the MST problem could be solved in $\mathcal{O}(|E|\alpha(|E|, |V|))$ where $\alpha(|E|, |V|)$ is the functional inverse of Ackerman's function defined in [40].

One of the main results of this dissertation, is that the computational complexity of the algorithm used to compute spanning tree modulus (see section 2.8), is that of finding a minimum spanning tree on each step. Thus it relies heavily on the established algorithms for finding a minimum spanning tree.

Spanning trees have played an important role in determining and/or approximating solutions to many real world problems: from probably one of the most famous problems in the traveling-salesman problem (TSP) [22, 27, 32], to more general problems for determining the number of disjoint spanning trees of a graph [28, 30], and a hint towards the use of the MST problem to aid in cluster analysis methods [20].

1.1.2 Why modulus?

Modulus has already been shown to have a diverse set of uses, the reader is referred to [2–6]. One of the authors of these papers framed modulus with the following analogy; "modulus is like the bag for your tools, and the family of objects you wish to study are the tools in the bag, you have to decide which tool is right for the job."

Modulus, in the form of conformal modulus, originated in complex analysis, see [1, p. 81], although the more general theory of p-modulus was brought to light through the study of quasi-conformal maps [7]. In a natural way, p-modulus provides a way of quantifying the richness of a family of curves on any given domain $\Omega \in \mathbb{C}$.

Although the concept of discrete modulus is not new, the remainder of this section and the next chapter are included as a courtesy to the reader, and to make the transition more fluid from the continuous setting to the discrete setting.

Let Ω be a domain in $\mathbb{C}$, and let C, D be two continua in the closure, $\bar{\Omega}$. Let $\Gamma(C, D)$ be the family of all rectifiable curves (curves with finite length) connecting C to D in Ω.

Definition 1.1.1. A *density* $\rho : \Omega \to [0, \infty)$, is a Borel measurable function, and is *admissible* for Γ, written $\rho \in \mathrm{Adm}(\Gamma)$, if

$$\int_\gamma \rho ds \geq 1 \qquad \forall \gamma \in \Gamma \tag{1.1.1}$$

This definition provides a way of assigning a cost to every curve γ in the family Γ. The modulus of Γ, more precisely the 2-modulus, is defined as follows.

$$\mathrm{Mod}_2(\Gamma) := \inf_{\rho \in \mathrm{Adm}(\Gamma)} \int_\Omega \rho^2 dA \tag{1.1.2}$$

A more general version, for p-modulus, is that for any $1 \leq p < \infty$

$$\mathrm{Mod}_p(\Gamma) := \inf_{\rho \in \mathrm{Adm}(\Gamma)} \int_\Omega \rho^p dA \tag{1.1.3}$$

The 2-modulus is illustrated through the following example.

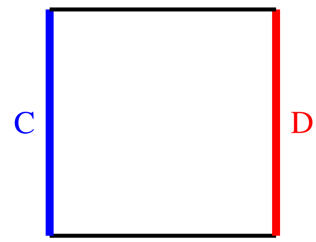


Figure 1.1: Continuous setting on a rectangle

Example 1.1.2. Consider the rectangle (see figure 1.1)

$$\Omega := \{z = x + iy \in \mathbb{C} : 0 < x < L, 0 < y < H\}$$

Let $C = \{z \in \bar{\Omega} : \text{Re}(z) = 0\}$ and $D = \{z \in \bar{\Omega} : \text{Re}(z) = L\}$, and let $\Gamma_\Omega(C, D)$ be the family of rectifiable curves in $\bar{\Omega}$ connecting C to D. To determine $\text{Mod}_2(\Gamma)$, assume $\rho \in \text{Adm}(\Gamma)$, then for all $0 < y < H$, $\gamma_y(t) := t + iy$, $t \in [0, L]$, is a curve in Γ, so from 1.1.1

$$\int_{\gamma_y} \rho ds = \int_0^L \rho(t, y) dt \geq 1.$$

Using Cauchy-Schwarz, one can obtain the following inequality

$$1 \leq \left(\int_0^L \rho(t, y) dt \right)^2 \leq L \int_0^L \rho^2(t, y) dt.$$

In other words, $\frac{1}{L} \leq \int_0^L \rho^2(t, y) dt$. Integrating over y, one obtains

$$\frac{H}{L} \leq \int_\Omega \rho^2 dA.$$

By 1.1.2 and the fact that ρ was arbitrary, $\text{Mod}_2(\Gamma) \geq \frac{H}{L}$. In the other direction, consider $\rho_0(z) = \frac{1}{L}\mathbf{1}_\Omega(z)$ and note that $\int_\Omega \rho_0^2 dA = \frac{HL}{L^2} = \frac{H}{L}$. Therefore, if $\rho_0 \in \text{Adm}(\Gamma)$, then $\text{Mod}_2(\Gamma) \leq \frac{H}{L}$. To see that ρ_0 is admissible, consider that for any $\gamma \in \Gamma$:

$$\int_0^L \frac{1}{L}|\dot{\gamma}(t)|dt \geq \frac{1}{L}\int_0^L |\text{Re}\dot{\gamma}(t)|dt \geq \frac{1}{L}(\text{Re}\gamma(1) - \text{Re}\gamma(0)) \geq 1.$$

Hence, $\text{Mod}_2(\Gamma) = \frac{H}{L}$

1.2 Outline of Dissertation

The rest of this dissertation is aimed at advancing the theory and application of p-modulus on graphs.

The first four sections of Chapter 2 are given to establish the background of the problem and basic properties developed in [2–6]. In Section 2.5, the theory is established to show that the solution to the spanning tree modulus problem on a weighted graph can be solved through the use of an unweighted multi-graph representation. Although the entire section is devoted to solving this problem, the main concept can be summarized by one of the main theorems in the section, which can be described as saying that every optimal probability mass function (pmf) on the multi-graph representation of a weighted graph $G = (V, E, \sigma)$ can be pulled back to to an optimal pmf on G. This theorem also gives a formula for determining the expected edge usages of G based on the known expected edge usages its multi-graph representation.

In Section 2.7, using the basic monotonicity property along with known solutions to the spanning tree modulus problem on homogeneous graphs, an upper and lower bound for the 2-modulus on a graph with a fixed set of vertices is given.

One of the main contributions in this dissertation is the algorithm established in Section 2.8. The algorithm is based on incrementing the usage of each edge in a minimum spanning tree. The convergence to the solution of the spanning tree modulus problem is shown to be $\mathcal{O}(\frac{1}{k})$, and is the first algorithm to have a proven rate of convergence for this problem.

Chapter 3 is devoted to showing the use of spanning tree modulus in aiding in determining the structure of a graph in terms of its community components. In Section 3.3.1 the theory of deflation developed in [3] is illustrated, showing the relationship to community detection and graph structure. The process is illustrated on Zachary's Karate Club graph [43], a common graph used to assess the accuracy of community detection algorithms.

In Section 3.3.2, a new alternative approach to graph decomposition is given that is based on minimum feasible partitions. This can be viewed as an dual approach to deflation, as it decomposes a graph in the opposite way that deflation works. Also in this section, a lower bound on the edge connectivity is established based on knowledge of the weight of the minimum feasible partition.

Chapter 2

Spanning Tree Modulus

2.1 Development of Modulus on Graphs

The use of modulus as a graph theoretic quantity was developed in [4], and further studied in [2, 3, 5, 6, 37–39]. For the readers edification, the general framework for modulus is provided. In what follows, $G = (V, E, \sigma)$ is taken to be a finite graph with vertex set V and edge set E. The graph may be directed or undirected and need not be simple. For now, assume a weighted graph with each edge assigned a corresponding weight $0 < \sigma(e) < \infty$. A graph is said to be ***unweighted***, if $\sigma(e) \equiv 1$ for every edge.

As an analogue to rectifiable curves in Example 1.1.2, modulus on a graph $G = (V, E, \sigma)$ depends on the family of objects, $\Gamma \subset 2^E$, being considered. Previous families studied include connecting families, or st-paths [4, 6], simple cycles [39], via-paths [38], and spanning trees [4, 6].

For any family Γ and object in the family, denoted by γ, one can associate a usage vector representing the amount the object uses each edge. This is made formal through the following definition.

Definition 2.1.1. Given a graph $G(V, E, \sigma)$ and a family of objects $\Gamma \subset 2^E$, then for each $\gamma \in \Gamma$ the function

$$\mathcal{N}(\gamma, \cdot) : E \to \mathbb{R}_{\geq 0}$$

measures the *usage of edge* e by γ.

Notationally, it is convenient to consider $\mathcal{N}(\gamma,\cdot)$ as a row vector $\mathcal{N}(\gamma,\cdot)\in\mathbb{R}^E_{\geq 0}$, indexed by $e\in E$. In order to avoid fringe cases, it is useful to assume that Γ is non-empty and that each $\gamma\in\Gamma$ has positive usage on at least one edge. When this is the case, we will say that Γ is *non-trivial*.

As a function of two variables, the function $\mathcal{N}$ can be thought of as a matrix in $\mathbb{R}^{\Gamma\times E}$, indexed by pairs (γ, e) with γ an object in Γ and e an edge in E. This matrix $\mathcal{N}$ is called the *usage matrix* for the family Γ.

Some examples of objects and their associated usage functions are the following.

- When $\gamma\in\Gamma$ is a walk, $\gamma = x_0\ e_1\ x_1\ \cdots\ e_n\ x_n$, one can use the traversal-counting function $\mathcal{N}(\gamma, e) =$ number times γ traverses e. In this case $\mathcal{N}(\gamma,\cdot)\in\mathbb{Z}^E_{\geq 0}$.

- For $\Gamma\subset 2^E$, for example spanning trees, one can use the characteristic function $\mathcal{N}(T, e) = \mathbb{1}_T(e) = 1$ if $e\in T$ and 0 otherwise. Here, $\mathcal{N}(\gamma,\cdot)\in\{0,1\}^E$.

In what follows, it will be useful to define the quantity:

$$\mathcal{N}_{\min} := \min_{\gamma\in\Gamma}\ \min_{e:\mathcal{N}(\gamma,e)\neq 0}\mathcal{N}(\gamma,e). \tag{2.1.1}$$

Hence, Γ non-trivial implies that $\mathcal{N}_{\min} > 0$.

Note that in some cases, such as when Γ is the family of walks from s to t, the family may be infinite in size, meaning that there are an infinite number of rows in $\mathcal{N}$. In the case of spanning trees, Γ is finite. In general, it was shown in [6], that any family Γ with an integer-valued $\mathcal{N}$ can be replaced, without changing the modulus, by a finite subfamily. For example, if Γ is the set of all walks between two distinct vertices, the modulus can be computed by considering only simple walks, or those that do not revisit any previously visited vertex. This result implies a similar finiteness result for any family Γ whose usage matrix $\mathcal{N}$ is rational with positive entries and an upper bound on the denominators.

To draw the connection to the continuous setting discussed in section 1.1.2, consider the following analog of the density.

Definition 2.1.2. Given a graph $G(V, E, \sigma)$ and a family of objects Γ, a *density* on G is a non-negative function on the edge set: $\rho : E\to[0,\infty)$. Moreover, the total usage cost of γ is given by

$$\ell_\rho(\gamma) := \sum_{e \in E} \mathcal{N}(\gamma, e)\rho(e) = (\mathcal{N}\rho)(\gamma)$$

The value $\rho(e)$ can be thought of as the *cost of using edge* e. It is notationally useful to think of such functions as column vectors in $\mathbb{R}^E_{\geq 0}$. In linear algebra notation, $\ell_\rho(\cdot)$ is the column vector resulting from the matrix-vector product $\mathcal{N}\rho$. The analogue to the admissibility condition (1.1.1) is then defined as follows.

Definition 2.1.3. Given a graph $G(V, E, \sigma)$, a family of objects Γ, a density ρ is *admissible* for Γ if

$$\ell_\rho(\gamma) \geq 1 \qquad \forall \gamma \in \Gamma;$$

or equivalently, if

$$\ell_\rho(\Gamma) := \inf_{\gamma \in \Gamma} \ell_\rho(\gamma) \geq 1.$$

In matrix notation, ρ is admissible if

$$\mathcal{N}\rho \geq \mathbf{1},$$

where $\mathbf{1}$ is the column vector of ones and the inequality is understood to hold elementwise. The set of all admissible densities, denoted $\mathrm{Adm}(\Gamma)$, is then defined by

$$\mathrm{Adm}(\Gamma) = \left\{\rho \in \mathbb{R}^E_{\geq 0} : \mathcal{N}\rho \geq \mathbf{1}\right\} \tag{2.1.2}$$

Given any exponent $p \geq 1$, the analogue to the area integral in the continuum case, defined as the *p-energy*, is given by

$$\mathcal{E}_{p,\sigma}(\rho) := \sum_{e \in E} \sigma(e)\rho(e)^p,$$

with the weights σ playing the role of the area element dA. In the unweighted case ($\sigma \equiv 1$), we shall use the notation $\mathcal{E}_{p,1}$ for the energy. For $p = \infty$, define the unweighted and weighted

∞-*energy* respectively as

$$\mathcal{E}_{\infty,1}(\rho) := \lim_{p\to\infty} \left(\mathcal{E}_{p,\sigma}(\rho)\right)^{\frac{1}{p}} = \max_{e\in E} \rho(e)$$

and

$$\mathcal{E}_{\infty,\sigma}(\rho) := \lim_{p\to\infty} \left(\mathcal{E}_{p,\sigma^p}(\rho)\right)^{\frac{1}{p}} = \max_{e\in E} \sigma(e)\rho(e)$$

This leads to the following definition.

Definition 2.1.4. Given a graph $G = (V, E, \sigma)$, a family of objects Γ with usage matrix $\mathcal{N} \in \mathbb{R}^{\Gamma\times E}$, and an exponent $1 \le p \le \infty$, the *p-modulus* of Γ is

$$\mathrm{Mod}_{p,\sigma}(\Gamma) := \inf_{\rho\in\mathrm{Adm}(\Gamma)} \mathcal{E}_{p,\sigma}(\rho)$$

Equivalently, p-modulus corresponds to the following optimization problem

$$\begin{aligned} &\text{minimize} && \mathcal{E}_{p,\sigma}(\rho) \\ &\text{subject to} && \rho \ge 0, \quad \mathcal{N}\rho \ge \mathbf{1} \end{aligned} \tag{2.1.3}$$

where each object $\gamma \in \Gamma$ determines one inequality constraint.

For $1 < p < \infty$ a unique extremal density ρ^* always exists and satisfies $0 \le \rho^* \le \mathcal{N}_{\text{min}}^{-1}$, where $\mathcal{N}_{\text{min}}$ is defined in (2.1.1). Existence and uniqueness follows by compactness and strict convexity of $\mathcal{E}_{p,\sigma}$, see also Lemma 2.2 of [5]. The upper bound on ρ^* follows from the fact that each row of $\mathcal{N}$ contains at least one nonzero entry, which must be at least as large as $\mathcal{N}_{\text{min}}$. In the special case when $\mathcal{N}$ is integer valued, the upper bound can be taken to be 1.

The following example illustrates the analogue to Example 1.1.2 with the family of connecting walks from s to t.

Example 2.1.5 (Basic Example)**.** Let G be a graph consisting of k simple paths in parallel, see figure 2.1, each path taking r hops to connect a given vertex s to a given vertex t. Assume also that G is unweighted, that is $\sigma \equiv 1$.

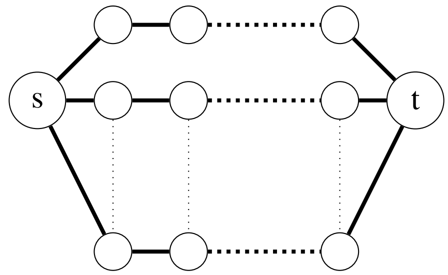


Figure 2.1: Basic example of st-paths

Let Γ be the family consisting of the k simple paths from s to t. Then $\ell(\Gamma) = r$ and the size of the minimum cut is k. A straightforward computation shows that

$$\mathrm{Mod}_p(\Gamma) = \frac{k}{r^{p-1}} \quad \text{for } 1 \le p < \infty, \qquad \mathrm{Mod}_{\infty,1}(\Gamma) = \frac{1}{r}.$$

Intuitively, when $p \approx 1$, $\mathrm{Mod}_p(\Gamma)$ is more sensitive to the number of parallel paths, while for $p \gg 1$, $\mathrm{Mod}_p(\Gamma)$ is more sensitive to the length of parallel paths.

2.2 Spanning tree modulus

For now, let $G = (V, E, \sigma)$ represent a finite, simple, weighted, connected graph with vertex set V, edge set E. In order to avoid trivial cases, it is generally assumed, unless otherwise indicated, that $|E| > 0$ (equivalently, that $|V| > 1$). When necessary to distinguish the vertices and edges of different graphs, subscript notation is employed, so that V_G refers to the vertices of G and E_G to its edges. Given a subset of edges $E' \subseteq E$, the notation $\sigma(E')$ represents the sum $\sum_{e\in E'} \sigma(e)$.

Cast in the previous chapter's notation, the family of spanning trees of G is denoted by Γ_G, or simply Γ when the underlying graph is clear from context.

The *spanning tree modulus* of G is a special case of Definition 2.1.4 with Γ the family of spanning trees and with usage matrix, $\mathcal{N}$, defined through the characteristic functions:

$$\mathcal{N}(\gamma, e) = \begin{cases} 1 & \text{if } e \in \gamma, \\ 0 & \text{if } e \notin \gamma. \end{cases}$$

Example 2.2.1. Consider the following graph, called the slashed square. The edges are labeled with their enumeration.

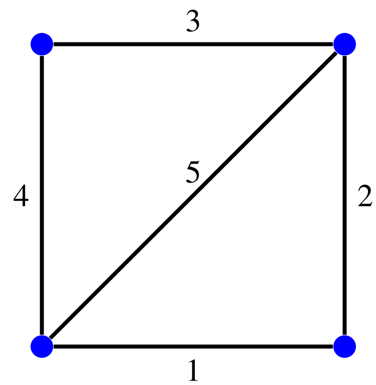


Figure 2.2: Slashed Square

The spanning trees of the slashed square are given in figure 2.3, note that the enumeration of the spanning trees does not matter. The usage matrix $\mathcal{N}$ for this enumeration of the spanning trees is then

$$\mathcal{N} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

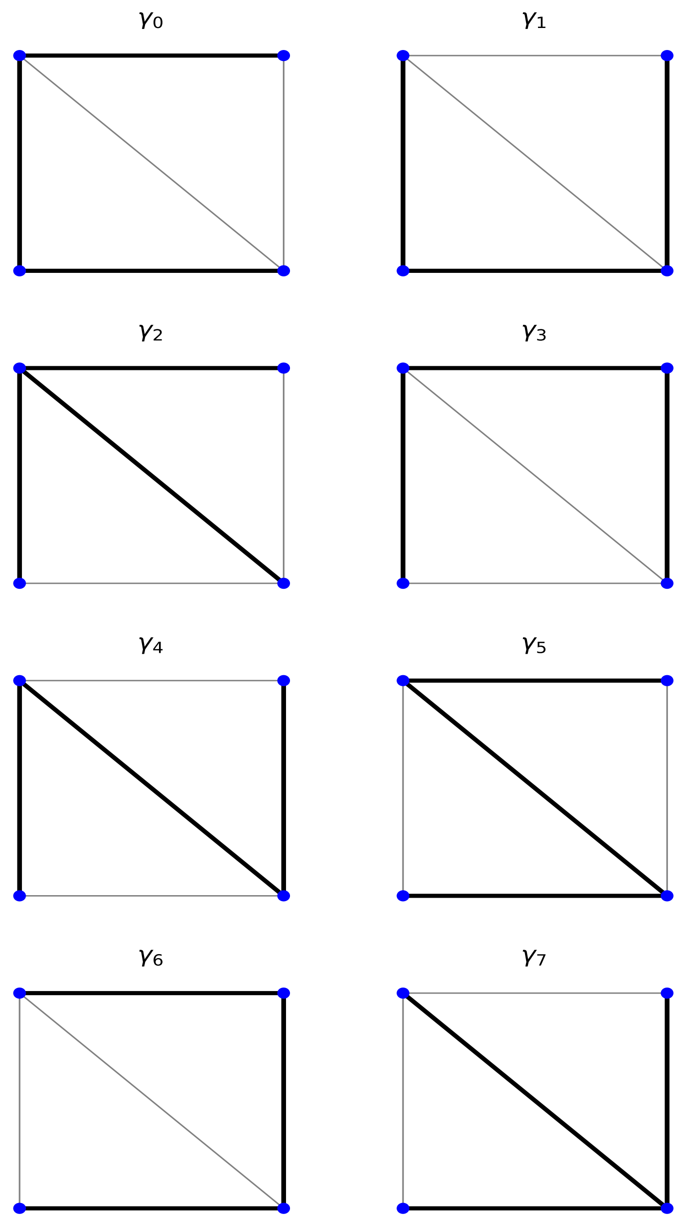


Figure 2.3: Spanning trees of the slashed square

The optimization problem (2.1.3), recast in this setting, is defined as follows.

Definition 2.2.2. Given a graph $G(V, E, \sigma)$, and the family of spanning trees, Γ, with usage matrix $\mathcal{N} \in \mathbb{R}^{\Gamma \times E}$, and an exponent $1 \leq p \leq \infty$, the p-*modulus* of Γ, $\mathrm{Mod}_{p,\sigma}(\Gamma)$, is defined as the value of following minimization problem.

$$\begin{aligned} &\text{minimize} && \mathcal{E}_{p,\sigma}(\rho) \\ &\text{subject to} && \rho \geq 0, \ \ \mathcal{N}\rho \geq 1 \end{aligned} \tag{2.2.1}$$

Written in this way, modulus can be seen to be a standard convex program, where the admissibility condition can be interpreted in the sense that every spanning tree must have total ρ cost of at least 1. Existence and uniqueness of a solution to (2.2.1) follows from the fact that if the edges are enumerated, then each density, ρ, can be thought of as vector in $\mathbb{R}^{|E|}$, where $\rho = (\rho(e_1), \rho(e_2), \ldots, \rho(e_{|E|}))$. Every tree, γ, corresponds to a row in $\mathcal{N}$, and $\ell_\rho(\gamma)$ is given by the dot product of ρ and the row of $\mathcal{N}$ in which γ corresponds to. In particular, given a spanning tree γ, the set $\{\rho \in \mathbb{R}^{|E|} : \ell_\rho(\gamma) \geq 1\}$ is a closed half-space. Hence, the admissible set, $\mathrm{Adm}(\Gamma)$, is an intersection of closed half-spaces in $\mathbb{R}^{|E|}$, therefore it is convex.

The energy function, $\mathcal{E}_{p,\sigma}(\rho)$, is a convex function of ρ when $1 < p < \infty$, hence existence and uniqueness follows from the compactness and strict convexity of the objective function and the constraint space.

When $\rho = \rho_0$ is the constant density equal to 1 on every edge, the subscript is dropped and we write $\ell(\gamma) := \ell_{\rho_0}(\gamma)$, which simply counts the number of edges in a spanning tree, hence $\ell(\gamma) = |V| - 1$, for any spanning tree $\gamma \in \Gamma$.

It is important to note here that there are $|\Gamma|$ inequality constraints, one for each spanning tree. Since $|\Gamma|$ tends to grow combinatorially in the graph size, it is generally not feasible to even enumerate all constraints for a large graph. Nonetheless, it is possible to produce a polynomial-time algorithm for this problem. This will be discussed later in Section 2.8 when discussing previous approaches to solving the spanning tree modulus problem.

2.3 Properties of p-modulus

This section will discuss several important properties of modulus that have been developed in [2, 5, 6]. However, they will be cast strictly in terms of the family of spanning trees.

Lemma 2.3.1. *Given two families of spanning trees* Γ *and* $\tilde{\Gamma}$*, such that* $\Gamma \subset \tilde{\Gamma}$*, then* $\mathrm{Adm}(\tilde{\Gamma}) \subset \mathrm{Adm}(\Gamma)$ *and hence*

$$\mathrm{Mod}_p(\Gamma) \leq \mathrm{Mod}_p(\tilde{\Gamma}) \tag{2.3.1}$$

Proof. Suppose Γ and $\tilde{\Gamma}$ are families of spanning trees such that $\Gamma \subset \tilde{\Gamma}$. Suppose $\rho \in \mathrm{Adm}(\tilde{\Gamma})$, then the ρ length, $\ell_\rho(\gamma) \geq 1$, for every tree in $\tilde{\Gamma}$ and hence for ever tree in Γ. Therefore $\rho \in \mathrm{Adm}(\Gamma)$, and since the modulus problem minimizes over the sets of admissible densities then $\mathrm{Mod}_p(\Gamma) \leq \mathrm{Mod}_p(\tilde{\Gamma})$ for all $1 \leq p \leq \infty$. □

This is known as Γ-monotonicity of modulus. In the setting of spanning trees, if edges are added to a graph, but no new nodes are added, the modulus can only increase.

Example 2.3.2. (Monotonicity) Consider Figure 2.4, all of the spanning trees of the cycle, C_4, are spanning trees of the slashed square, SS. The slashed square contains four more trees that C_4 doesn't have. The complete graph, K_4, contains all the trees of the slashed square plus an additional 8 trees. So it is clear to see that $\Gamma_{C_4} \subset \Gamma_{ss} \subset \Gamma_{K_4}$. In the figure, the edges are labeled with their corresponding edge densities, ρ^*, notice that it is the same across all three graphs.

Lemma 2.3.3. *Let* $\Gamma_1, \Gamma_2, \ldots$ *be non-empty families of spanning trees on a finite graph* G*. Then*

$$\mathrm{Mod}_p\left(\bigcup_{i=1}^{\infty} \Gamma_i\right) \leq \sum_{i=1}^{\infty} \mathrm{Mod}_p(\Gamma_i) \tag{2.3.2}$$

Proof. For every Γ_i, pick any admissible density ρ_i such that $\mathcal{E}(\rho_i) \leq \mathrm{Mod}_p(\Gamma_i) + 2^{-i}\epsilon$. Let $\rho := \left(\sum_{i=1}^{\infty} \rho_i^p\right)^{1/p}$. Then ρ is admissible for $\Gamma := \bigcup_{i=1}^{\infty} \Gamma_i$ since if $\gamma \in \Gamma$, then $\gamma \in \Gamma_i$ for some i and

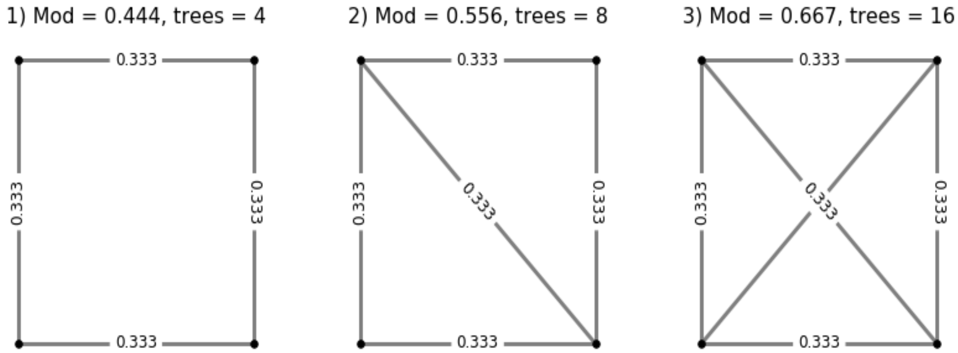


Figure 2.4: Monotonicity using 1) C_4, 2) slashed square, 3) K_4

$\rho \geq \rho_i$. Hence

$$\begin{aligned}\mathrm{Mod}_p\left(\bigcup_{i=1}^{\infty}\Gamma_i\right) &\leq \mathcal{E}(\rho) = \sum_{e\in E}\rho(e)^p = \sum_{e\in E}\sum_{i=1}^{\infty}\rho_i(e)^p = \sum_{i=1}^{\infty}\sum_{e\in E}\rho_i(e)^p \\ &= \sum_{i=1}^{\infty}\mathcal{E}(\rho_i) \leq \sum_{i=1}^{\infty}\left(\mathrm{Mod}_p(\Gamma_i) + \frac{\epsilon}{2^i}\right) = \sum_{i=1}^{\infty}\mathrm{Mod}_p(\Gamma_i) + \epsilon\end{aligned}$$

Since ϵ is arbitrary, then $\mathrm{Mod}_p(\bigcup_{i=1}^{\infty}\Gamma_i) \leq \sum_{i=1}^{\infty}\mathrm{Mod}_p(\Gamma_i)$ □

Lemma 2.3.4. *[2] Let $G = (V, E)$ be a non-trivial, connected graph and Γ the family of spanning trees on G. Let $1 < p < \infty$ and set $M := \max\{p, q\}$, where $p+q = pq$. Let ρ^* denote the extremal (optimal) density for $\mathrm{Mod}_{p,\sigma}(\Gamma)$. Then, for every $\rho \in \mathrm{Adm}(\Gamma)$:*

$$\|\rho - \rho^*\|_p^M \leq 2^{M-1}\sigma_{min}^{-M/p}\left(\mathcal{E}_{p,\sigma}(\rho)^{M/p} - \mathrm{Mod}_{p,\sigma}(\Gamma)^{M/p}\right)$$

where $\sigma_{min} = \min_{e\in E}\sigma(e)$.

Proof. Let $f(e) := \sigma(e)^{1/p}\rho(e)$ and $f^*(e) = \sigma(e)^{1/p}\rho^*(e)$. Then

$$\|f\|_p^p = \mathcal{E}_{p,\sigma}(\rho) \qquad \text{and} \qquad \|f^*\|_p^p = \mathrm{Mod}_{p,\sigma}(\Gamma).$$

Also

$$\|f - f^*\|_p^p = \sum_{e\in E} \sigma(e)|\rho(e) - \rho^*(e)|^p \geq \sigma_{\text{min}}\|\rho - \rho^*\|_p^p,$$

and, since $\mathrm{Adm}(\Gamma)$ is convex,

$$\left\|\frac{f+f^*}{2}\right\|_p^p = \sum_{e\in E}\sigma(e)\left|\frac{\rho(e)+\rho^*(e)}{2}\right|^p = \mathcal{E}_{p,\sigma}\left(\frac{\rho-\rho^*}{2}\right) \geq \mathrm{Mod}_{p,\sigma}(\Gamma).$$

Using a weak version of Clarkson's inequalities

$$\begin{aligned}\mathrm{Mod}_{p,\sigma}(\Gamma)^{M/p} + \sigma_{\text{min}}^{M/p}2^{-M}\|\rho-\rho^*\|_p^M &\leq \left\|\frac{f+f^*}{2}\right\|_p^M + \left\|\frac{f-f^*}{2}\right\|_p^M \\ &\leq \frac{\|f\|_p^M + \|f^*\|_p^M}{2} = \frac{1}{2}\left(\mathcal{E}_{p,\sigma}(\rho)^{M/p} + \mathrm{Mod}_{p.\sigma}(\Gamma)^{M/p}\right)\end{aligned}$$

□

In other words, what Lemma 2.3.4 is saying, is that if ρ is almost a minimizer for the modulus problem, then ρ must be close to ρ^*.

The next two theorems establish continuity of modulus with respect to p and σ respectfully.

Theorem 2.3.5. *[5] Let $G = (V, E, \sigma)$ be a weighted graph and Γ be the family of spanning trees on G. Then the following conditions hold:*

(1) $p \mapsto \mathcal{N}_{min}^p \mathrm{Mod}_{p,\sigma}(\Gamma)$ is decreasing

(2) $p \mapsto (\mathrm{Mod}_{p,\sigma}(\Gamma)/\sigma(E))^{\frac{1}{p}}$ is increasing, where $\sigma(E) := \sum_{e\in E}\sigma(e)$.

(3) $p \mapsto \mathrm{Mod}_{p,\sigma}(\Gamma)$ is continuous for $1 \leq p \leq \infty$

Proof. For (1), consider the fact that for any $\rho \in \mathrm{Adm}(\Gamma)$, then $0 \leq \rho \leq 1$. Letting $q > p$ then the following holds

$$\mathcal{N}_{\text{min}}^q\mathcal{E}_q(\rho) = \mathcal{N}_{\text{min}}^q\sum_{e\in E}\sigma(e)\rho(e)^q \leq \mathcal{N}_{\text{min}}^p\sum_{e\in E}\sigma(e)\rho(e)^p = \mathcal{E}_p(\rho). \tag{2.3.3}$$

To prove (2), let $\rho \in \mathrm{Adm}(\Gamma)$ and assume $1 \leq p < q < \infty$. Then using Hölder's inequality and the conjugate exponents q/p and $q/(q-p)$ then

$$\begin{aligned} \mathcal{E}_p(\rho) &= \sum_{e \in E} \sigma(e)\rho(e)^p = \sum_{e \in E} \sigma(e)^{p/q}\rho(e)^p \cdot \sigma(e)^{\frac{q-p}{q}} \\ &\leq \left(\sum_{e \in E} \sigma(e)\rho(e)^q\right)^{p/q} \left(\sum_{e \in E} \sigma(e)\right)^{\frac{q-p}{q}} = \sigma(E)^{\frac{q-p}{q}} \mathcal{E}_q(\rho)^p/q \end{aligned} \tag{2.3.4}$$

Which, after rearranging terms and taking the p^{th} root, becomes

$$\sigma(E)^{-1/p}\mathcal{E}_p(\rho)^{1/p} \leq \sigma(E)^{-1/q}\mathcal{E}_q(\rho)^{1/q}. \tag{2.3.5}$$

Picking ρ to be optimal for $\mathrm{Mod}_q(\Gamma)$ yields the desired result.

Finally to show (3), using the results of (1) if $q \downarrow p$, then

$$\limsup_{q \downarrow p} \mathrm{Mod}_{q,\sigma}(\Gamma) \leq \mathrm{Mod}_{p,\sigma}(\Gamma),$$

and from (2),

$$\liminf_{q \downarrow p} \mathrm{Mod}_{q,\sigma}(\Gamma) \geq \liminf_{q \downarrow p} \sigma(E)^{1-q/p} \mathrm{Mod}_{p,\sigma}(\Gamma)^{q/p} = \mathrm{Mod}_{p,\sigma}(\Gamma).$$

Therefore $\lim\limits_{q \downarrow p} \mathrm{Mod}_{q,\sigma}(\Gamma) = \mathrm{Mod}_{p,\sigma}(\Gamma)$.

Similarly, if $q \uparrow p$, then from (1)

$$\liminf_{q \uparrow p} \mathrm{Mod}_{q,\sigma}(\Gamma) \geq \mathrm{Mod}_{p,\sigma}(\Gamma),$$

and from (2),

$$\limsup_{q \uparrow p} \mathrm{Mod}_{q,\sigma}(\Gamma) \leq \limsup_{q \uparrow p} \sigma(E)^{1-q/p} \mathrm{Mod}_{p,\sigma}(\Gamma)^{q/p} = \mathrm{Mod}_{p,\sigma}(\Gamma)$$

Hence, $\lim_{q \to p} \mathrm{Mod}_{q,\sigma}(\Gamma) = \mathrm{Mod}_{p,\sigma}(\Gamma)$. $\square$

Theorem 2.3.6. *[2] Let $G = (V, E, \sigma)$ be a finite, weighted graph and Γ be the nonempty, finite, family of spanning trees. Fix $1 < p < \infty$ and let ρ^*_σ be the extremal density for $\operatorname{Mod}_{p,\sigma}(\Gamma)$. Then*

(1) the map $\phi : \mathbb{R}^E_{>0} \to \mathbb{R}$ given by $\phi(\sigma) := \operatorname{Mod}_{p,\sigma}(\Gamma)$ is Lipschitz continuous;

(2) the extremal density ρ^*_σ is also continuous in σ;

(3) the map ϕ is concave;

(4) the map ϕ is differentiable, and the partial derivatives satisfy

$$\frac{\partial \phi}{\partial \sigma(e)} = \rho^*_\sigma(e)^p \qquad \forall e \in E$$

Proof. To show (1), fix $\sigma_1, \sigma_2 \in \mathbb{R}^E_{>0}$ and assume that $\operatorname{Mod}_{p,\sigma_2}(\Gamma) \leq \operatorname{Mod}_{p,\sigma_1}(\Gamma)$. Recall that for any optimal density $\rho^* \leq \mathcal{N}^{-1}_{\text{min}}$, therefore

$$\begin{aligned}
|\operatorname{Mod}_{p,\sigma_1}(\Gamma) - \operatorname{Mod}_{p,\sigma_2}(\Gamma)| &\leq \mathcal{E}_{p,\sigma_1}(\rho^*_{\sigma_2}) - \operatorname{Mod}_{p,\sigma_2}(\Gamma) \\
&= \mathcal{E}_{p,\sigma_1}(\rho^*_{\sigma_2} - \mathcal{E}_{p,\sigma_2}(\rho^*_{\sigma_2}) \\
&= \sum_{e\in E}(\sigma_1(e) - \sigma_2(e))\rho^*_{\sigma_2}(e)^p \\
&\leq \mathcal{N}^{-p}_{\text{min}} \|\sigma_1 - \sigma_2\|_1
\end{aligned}$$

A similar argument holds when $\operatorname{Mod}_{p,\sigma_1}(\Gamma) \leq \operatorname{Mod}_{p,\sigma_2}(\Gamma)$.

To show (2), by Lemma 2.3.4 and $M \geq 2$,

$$\|\rho^*_{\sigma_2} - \rho^*_{\sigma_1}\|^M_p \leq 2^{M-1}\sigma^{-M/p}_{1,\text{min}} \left(\mathcal{E}_{p,\sigma_1}(\rho^*_{\sigma_2})^{M/p} - \operatorname{Mod}_{p,\sigma_1}(\Gamma)^{M/p}\right).$$

So it is enough to show that

$$\mathcal{E}_{p,\sigma_1}(\rho^*_{\sigma_2}) \longrightarrow \operatorname{Mod}_{p,\sigma_1}(\Gamma) \qquad \text{as } \sigma_2 \to \sigma_1$$

Since

$$\begin{aligned}
\operatorname{Mod}_{p,\sigma_1}(\Gamma) &\le \mathcal{E}_{p,\sigma_1}(\rho^*_{\sigma_2}) = \sum_{e\in E} \sigma_1(e)\rho^*_{\sigma_2}(e)^p \\
&= \sum_{e\in E} \sigma_2(e)\rho^*_{\sigma_2}(e)^p + \sum_{e\in E}(\sigma_1(e)-\sigma_2(e))\rho^*_{\sigma_2}(e)^p \\
&\le \operatorname{Mod}_{p,\sigma_2}(\Gamma) + \mathcal{N}_{\text{min}}^{-p}\|\sigma_1-\sigma_2\|_1
\end{aligned}$$

and the last line converges to $\operatorname{Mod}_{p,\sigma_1}(\Gamma)$ by continuity of the map in part (1), the desired result is reached.

To show (3), fix $\sigma_0, \sigma_1 \in \mathbb{R}^E_{>0}$ and $t \in [0,1]$. Let ρ^*_t be extremal for $\sigma_t := t\sigma_1 + (1-t)\sigma_0$. Then since $\rho^*_t \in \operatorname{Adm}(\Gamma)$,

$$\begin{aligned}
t\operatorname{Mod}_{p,\sigma_1}(\Gamma) + (1-t)\operatorname{Mod}_{p,\sigma_0}(\Gamma) &\le t\mathcal{E}_{p,\sigma_1}(\rho^*_t) + (1-t)\mathcal{E}_{p,\sigma_0}(\rho^*_t) \\
&= \sum_{e\in E}\sigma_t(e)\rho^*_t(e)^p = \operatorname{Mod}_{p,\sigma_t}(\Gamma).
\end{aligned}$$

This gives concavity.

Finally, to show (4), fix $\tau \in \mathbb{R}^E$ and let $\epsilon > 0$. Set $\sigma_\epsilon := \sigma + \epsilon\tau$. Then notice that for small enough ϵ, $\sigma_\epsilon \in \mathbb{R}^E_{>0}$. Let ρ^*_ϵ be the extremal density corresponding to σ_ϵ. Then for any $\rho \in \mathbb{R}^E_{\ge 0}$:

$$\mathcal{E}_{p,\sigma_\epsilon}(\rho) = \sum_{e\in E}(\sigma(e)+\epsilon\tau(e))\rho(e)^p = \mathcal{E}_{p,\sigma}(\rho) + \epsilon\mathcal{E}_{p,\tau}(\rho).$$

Therefore,

$$\operatorname{Mod}_{p,\sigma_\epsilon}(\Gamma) = \mathcal{E}_{p,\sigma_\epsilon}(\rho^*_\epsilon) = \mathcal{E}_{p,\sigma}(\rho^*_\epsilon) + \epsilon\mathcal{E}_{p,\tau}(\rho^*_\epsilon) \ge \operatorname{Mod}_{p,\sigma}(\Gamma) + \epsilon\mathcal{E}_{p,\tau}(\rho^*_\epsilon),$$

and

$$\operatorname{Mod}_{p,\sigma_\epsilon}(\Gamma) \le \mathcal{E}_{p,\sigma_\epsilon}(\rho^*_0) = \operatorname{Mod}_{p,\sigma}(\Gamma) + \epsilon\mathcal{E}_{p,\tau}(\rho^*_0).$$

Hence

$$\mathcal{E}_{p,\tau}(\rho^*_\epsilon) \le \frac{\phi(\sigma+\epsilon\tau)-\phi(\sigma)}{\epsilon} \le \mathcal{E}_{p,\tau}(\rho^*_0).$$

By part (2), $\rho_\epsilon^* \to \rho_0^*$ as $\epsilon \to 0$. So the directional derivative of ϕ in the direction of τ is:

$$D_\tau(\phi) = \sum_{e\in E} \tau(e)\rho_\sigma^*(e)^p.$$

This implies that all the directional derivatives are continuous, and thus it follows that ϕ is differentiable. $\square$

2.4 Probabilistic Interpretation

Letting $p = 2$ and $\sigma \equiv 1$, then (2.2.1) becomes the following quadratic program:

$$\begin{aligned} &\text{minimize} \qquad \rho^T\rho \\ &\text{subject to} \qquad \mathcal{N}\rho \geq 1 \end{aligned} \tag{2.4.1}$$

with corresponding dual problem

$$\begin{aligned} &\text{maximize} \qquad \lambda^T\mathbf{1} - \frac{1}{4}\lambda^T\mathcal{N}\mathcal{N}^T\lambda \\ &\text{subject to} \qquad \lambda \geq 1 \end{aligned} \tag{2.4.2}$$

According to [4, Theorem 3.5], if Γ is a minimal subfamily of spanning trees, then $\operatorname{rank}(\mathcal{N}) = |\Gamma|$ and the optimizer, λ^*, to (2.4.2) is unique.

Alternatively, by letting $\mathcal{P}(\Gamma)$ represent the set of probability mass functions (pmfs) on the set Γ, the set of vectors $\mu \in \mathbb{R}^\Gamma_{\geq 0}$ with the property that $\mu^T\mathbf{1} = 1$. The optimization problem can be restated as follows.

Theorem 2.4.1. *[4] Let Γ be the family of spanning trees on a finite graph G. Then the 2-modulus problem can be written*

$$\operatorname{Mod}_2(\Gamma)^{-1} = \min_{\mu\in\mathcal{P}(\Gamma)} \mu^T\mathcal{N}\mathcal{N}^T\mu \tag{2.4.3}$$

Moreover, any optimal pmf μ^* is related to the optimal density ρ^* for $\mathrm{Mod}_2(\Gamma)$ as follows:

$$\frac{\rho^*(e)}{\mathrm{Mod}_2(\Gamma)} = \left(\mathcal{N}^T\mu^*\right)(e) \qquad \forall e \in E. \tag{2.4.4}$$

In this form, the 2-modulus problem can be interpreted as minimum overlap problem. Consider the following definition.

Definition 2.4.2. Let the *overlap matrix* for Γ be the matrix $C := \mathcal{N}\mathcal{N}^T$. The entries of C correspond to pairs of spanning trees (γ_i, γ_j) and

$$C(\gamma_i, \gamma_j) = \sum_{e\in E} \mathcal{N}(\gamma_i, e)\mathcal{N}(\gamma_j, e)$$

Going back to Example 2.2.1, the overlap matrix in this example is given by

$$\mathcal{N}\mathcal{N}^T = C = \begin{bmatrix} 3 & 2 & 2 & 2 & 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 2 & 2 & 1 & 2 & 2 \\ 2 & 1 & 3 & 2 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 3 & 2 & 1 & 2 & 1 \\ 1 & 2 & 2 & 2 & 3 & 1 & 1 & 2 \\ 2 & 1 & 2 & 1 & 1 & 3 & 2 & 2 \\ 2 & 2 & 1 & 2 & 1 & 2 & 3 & 2 \\ 1 & 2 & 1 & 1 & 2 & 2 & 2 & 3 \end{bmatrix}$$

The probabilistic interpretation of the 2-modulus problem follows from the fact that given any $\mu \in \mathcal{P}(\Gamma)$, μ defines a random variable $\underline{\gamma}$ such that $\mu(\gamma) = \mathbb{P}_\mu(\underline{\gamma} = \gamma)$, which defines a probability that $\underline{\gamma}$ takes the value $\gamma \in \Gamma$. This gives rise to a very important quantity in moving forward.

Definition 2.4.3. Let $\mu \in \mathcal{P}(\Gamma)$ with Γ the family of spanning trees on G, the *edge usage probability*, η is given by

$$\eta(e) := \left(\mathcal{N}^T\mu\right)(e) = \sum_{\gamma\in\Gamma} \mathcal{N}(\gamma, e)\mu(\gamma) = \mathbb{E}_\mu\left[\mathcal{N}(\underline{\gamma}, e)\right]. \tag{2.4.5}$$

An alternative view point is that η determines the expected edge usage of an edge e being in a random spanning tree γ chosen with probability μ. The uniform pmf, denoted μ_0, is defined as

$$\mu_0(\gamma) = \frac{1}{|\Gamma|} \qquad \forall \gamma \in \Gamma. \tag{2.4.6}$$

Again returning to Example 2.2.1, consider the pmf μ_0 that assigns the uniform probability to all spanning trees. Notice that the four outside edges show up in exactly 5 of the 8 spanning trees, while the diagonal edges shows up in 4.

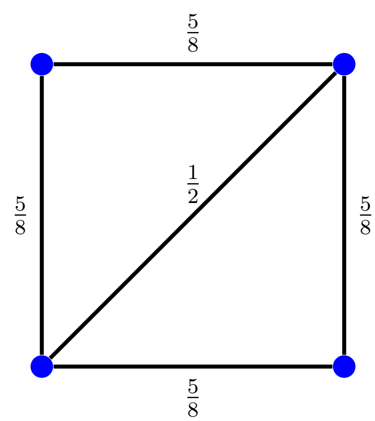


Figure 2.5: Slashed Square with edge probabilities base on μ_0

In other words, η measures how many times an edge shows up in a random spanning tree chosen with probability μ. Moreover, this means that $\rho^*(e)$ is proportional to the expected edge usage of e with respect to an optimal pmf μ^* which minimizes the energy $\eta^T\eta = \mu^T\mathcal{N}\mathcal{N}^T\mu$. More than this, in light of (2.4.4), for $\mathrm{Mod}_2(\Gamma)$ this means that ρ^* is parallel to η^*, where η^* is the optimal solution to (2.4.2).

Combining this with definition 2.4.2, this can be interpreted as an expectation.

$$\begin{aligned} \mu^T\mathcal{N}\mathcal{N}^T\mu &= \sum_{\gamma,\gamma'\in\Gamma} C(\gamma,\gamma')\mu(\gamma)\mu(\gamma') \\ &= \sum_{\gamma,\gamma'\in\Gamma} C(\gamma,\gamma')\mathbb{P}_\mu(\underline{\gamma}=\gamma,\underline{\gamma'}=\gamma') = \mathbb{E}_\mu\left[C(\underline{\gamma},\underline{\gamma'})\right] \end{aligned} \tag{2.4.7}$$

where C is the matrix defined in Definition 2.4.2. Equation (2.4.7) gives rise to the *minimum* *expected overlap* (MEO) problem introduced in [3] as yet another interpretation of the 2-modulus problem.

$$\begin{aligned} &\text{minimize} \quad \mathbb{E}_\mu|\underline{\gamma}\cap\underline{\gamma}'| \\ &\text{subject to} \quad \mu \in \mathcal{P}(\Gamma_G). \end{aligned} \tag{2.4.8}$$

This gives the motivation for a lot of the work that has followed regarding spanning tree modulus. The fact that this means that spanning tree modulus is in effect equivalent to finding the minimum overlap of spanning trees, provides important insight into the construction of the algorithm presented in section 2.8.2.

2.5 Weighted Graphs

Theorem 2.3.6 provides the intuition to consider unweighted multi-graphs for the remainder of the theory, however in practice, this may not be ideal, as the number of edges could be increased extremely high, and when the running time of the algorithms (see Section 2.8) depends on the number of edges, this can pose efficiency issues.

However, the theory in this section provides a connection between the weighted and unweighted 2-modulus problems, using the continuity established by Theorem 2.3.6.

Suppose $G = (V, E, \sigma)$ is a weighted graph with weights $\sigma \in \mathbb{R}^E_{>0}$, and assume there is a method of approximating $\mathrm{Mod}_{p,\sigma}(\Gamma)$ within some preset tolerance ϵ_{tol} (see Section 2.8). By Theorem 2.3.6 (1), it is safe to approximate by $\sigma \in \mathbb{Q}^E_{>0}$; by Theorem 2.3.6 (2) the optimal ρ is continuous in σ.

Lemma 2.5.1. *Given a weighted graph $G = (V, E, \sigma)$, with $\sigma \in \mathbb{Q}^E_{>0}$, and the family of spanning trees Γ. Then for any $s > 0$, modulus satisfies*

$$\mathrm{Mod}_{p,s\sigma}(\Gamma) = s\,\mathrm{Mod}_{p,\sigma}(\Gamma),$$

with exactly the same optimal density ρ.

Proof. The proof follows from definition 2.2.1, and noting that

$$\begin{aligned}\mathcal{E}_{p,s\sigma}(\rho) &= \sum_{e\in E} s\sigma(e)\rho(e)^p \\ &= s\sum_{e\in E}\sigma(e)\rho(e)^p \\ &= s\mathcal{E}_{p,\sigma}(\rho)\end{aligned}$$

Taking the infimum in $\rho \in \mathrm{Adm}(\Gamma)$ yields the desired result. □

In Lemma 2.5.1, consider s to be taken as the greatest common denominator among all $\sigma(e)$, then by approximation, one may replace any weighted graph with a graph with positive integer weights.

Definition 2.5.2. Let $G = (V, E, \sigma)$ be a weighted graph with $\sigma \in \mathbb{Z}^E_{>0}$, let $G' = (V, E')$ be the *unweighted representation* of G, defined as follows: For every edge $e \in E$ connecting two nodes x and y in G, there corresponds $\sigma(e)$ edges in E' connecting x and y in G'.

In other words, G' is an unweighted multi-graph, where the edge weights $\sigma(e)$ dictate the multiplicity of edges in G' that correspond to the same two nodes. This is illustrated in the following example.

Example 2.5.3. Consider figures 2.6 and 2.7, notice the edges with weight 2 are doubled, and the edge with weight 3 is tripled.

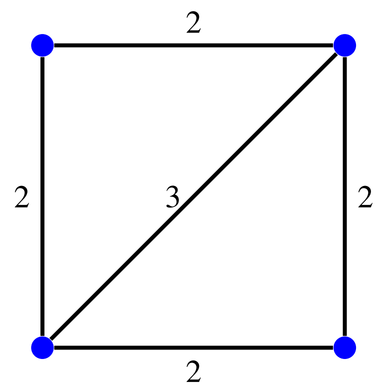


Figure 2.6: Slashed Square, G, with integer weights

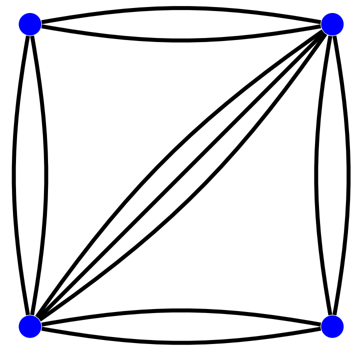


Figure 2.7: Unweighted representation G'

Definition 2.5.4. Let $G = (V, E, \sigma)$ and $G' = (V, E')$ be as in definition 2.5.2 and let $\varphi : E' \to E$ be the map that takes each of the $\sigma(e)$ edges of E' induced by $e \in E$ back onto e. In other words, $E = \varphi(E')$ and

$$|\varphi^{-1}(e)| = \sigma(e) \tag{2.5.1}$$

for every edge $e \in E$. This map is called the *weighting map* of G.

With the unweighted representation G' of G and the corresponding weighting map φ in place, one can replace the minimum expected overlap (MEO) problem 2.4.8 on G by equivalent problems on G'. The key is in interpreting σ as edge multiplicity and modifying all definitions accordingly.

For example, when an object $\gamma \in \Gamma$ uses an edge $e \in E$ with $\sigma(e) > 1$, on can interpret this as if the object has a choice to use one of $\sigma(e)$ identical copies of the edge. Since the goal is to spread out the edge usage as much as possible, the natural thing to do is give each possible choice equal probability. If two random objects in G share an edge e, then there is a $\sigma(e)^{-1}$ probability that both will pick the same copy of the associated multi-edge. Using this intuition, we get the following weighted MEO problem, denoted $\text{MEO}_{\sigma^{-1}}$

$$\begin{aligned} \text{minimize} \quad & \mathbb{E}_\mu(C(\underline{\gamma} \cap \underline{\gamma}')) \\ \text{subject to} \quad & \mu \in \mathcal{P}(\Gamma). \end{aligned} \tag{2.5.2}$$

Where the weighted overlap matrix $C(\gamma, \gamma')$ is given by

$$C(\gamma, \gamma') := \sum_{e \in E} \sigma(e)^{-1} \mathcal{N}(\gamma, e) \mathcal{N}(\gamma', e)$$

Another example of how to use this interpretation to modify the definitions is to consider the uniform pmf μ_0, this should be replaced by the σ-weighted pmf μ_σ that selects a spanning tree $\gamma \in \Gamma$ with probability

$$\mu_\sigma(\gamma) = \frac{\prod\limits_{e \in \gamma} \sigma(e)}{\sum\limits_{\gamma' \in \Gamma} \prod\limits_{e \in \gamma'} \sigma(e)}.$$

In the following, assume that $G = (V, E, \sigma)$ is an integer weighted multigraph, $G' = (V, E')$ is its unweighted representation, ϕ is the corresponding weighting map, and Γ is the family of spanning trees on G.

Definition 2.5.5. Let $\gamma \in \Gamma$. Every right inverse, τ, of φ defines a *lift*, $\gamma' = \tau(\gamma)$. γ' is an object on G' with usage matrix defined as

$$\mathcal{N}(\gamma', e') := \mathcal{N}(\gamma, \varphi(e')) \mathbb{1}_{e' = \tau(e)} \tag{2.5.3}$$

In other words, τ transfers the usage on each edge $e \in E$ to exactly one $e' \in \varphi^{-1}(e)$. When γ can be identified with a subset of E, as is the case for spanning trees, $\gamma' = \tau(\gamma) \subset E'$. For more general families, $\tau(\gamma)$ should be interpreted more symbolically: each $\gamma \in \Gamma$ and right inverse τ gives rise to a $\gamma' \in \Gamma'$, where Γ and Γ' should be interpreted simply as an indicator set, coupled with their corresponding usage matrices $\mathcal{N}$.

For an object $\gamma \in \Gamma$, the set of all lifts of γ is denoted

$$\varphi^\sharp \gamma := \{\tau(\gamma) : \tau \text{ is a right inverse of } \varphi\}. \tag{2.5.4}$$

with usage matrix defined as in (2.5.3)

Definition 2.5.6. Given a weighted graph $G = (V, E, \sigma)$ and its unweighted representation G', the *lifted family* Γ', is defined as

$$\Gamma' = \varphi^\sharp \Gamma = \bigcup_{\gamma \in \Gamma} \varphi^\sharp \gamma \tag{2.5.5}$$

with usage matrix defined as in (2.5.3).

For general families, it is unclear as to whether $\gamma' \in \Gamma'$ will be of the same family of objects as $\gamma \in \Gamma$. However, for spanning trees this is not the case, as is made clear through the following theorem.

Theorem 2.5.7. *If $\Gamma = \Gamma_G$ is the family of spanning trees, then $\Gamma' = \Gamma'_G$.*

Proof. Since the edge usages in this case are $0/1$-valued, it is convenient to identify each object with the corresponding edge set. Let $\gamma' \in \Gamma'$. Then there exists $\gamma \in \Gamma_G$ such that $\gamma' \in \varphi^\sharp\gamma$. By the definition of the lift, $|\gamma'| = |\gamma| = |V| - 1$. Moreover, γ' inherits connectedness from γ and spans V because γ does. Thus, γ' is a spanning tree.

On the other hand, if $\gamma' \in \Gamma_{G'}$, then by a similar argument $\gamma = \varphi(\gamma') \in \Gamma_G$. Moreover, it is possible to construct a right inverse τ such that $\gamma' = \tau(\gamma)$; we simply need to ensure that each $e \in \gamma$ pulls back to the correct $e' \in \gamma'$. ($\tau(e)$ can be chosen arbitrarily for any $e \in E \setminus \gamma$.)

□

Lemma 2.5.8. *Let $f : E' \to \mathbb{R}$ be an arbitrary function, let $\gamma \in \Gamma$, let τ be a right inverse of φ and let $\gamma' = \tau(\gamma) \in \varphi\sharp\gamma$. Then, for any $e \in E$,*

$$\sum_{e' \in \varphi^{-1}(e)} \mathcal{N}(\gamma', e') f(e') = \mathcal{N}(\gamma, e) f(\tau(e)).$$

Proof. This follows using the fact that τ is a right inverse of φ and (2.5.3). □

Theorem 2.5.9. *Let $1 < p < \infty$. The modulus problem for Γ and $\Gamma' = \varphi^\sharp\Gamma$ are related as follows*

$$\operatorname{Mod}_{p,\sigma}(\Gamma) = \operatorname{Mod}_p(\Gamma'). \tag{2.5.6}$$

Moreover, if ρ is optimal for the former and ρ' for the latter, then

$$\rho'(e') = \rho(\varphi(e')). \tag{2.5.7}$$

Proof. First, suppose $\rho' \in \operatorname{Adm}(\Gamma)$ is optimal. Define $\rho \in \mathbb{R}^E$ as

$$\rho(e) = \min_{e' \in \varphi^{-1}(e)} \rho'(e')$$

and select a right inverse $\tau : E \to E'$ with the property that

$$\rho'(\tau(e)) = \rho(e) \qquad \forall e \in E.$$

That is, for each $e \in E$, choose an edge $e' \in \varphi^{-1}(e)$ where ρ' is smallest.

Let $\gamma \in \Gamma$ and let $\gamma' = \tau(\gamma)$. Then, by Lemma 2.5.8,

$$\begin{aligned}\ell_\rho(\gamma) &= \sum_{e \in E} \mathcal{N}(\gamma, e)\rho(e) = \sum_{e \in E} \mathcal{N}(\gamma, e)\rho'(\tau(e)) = \sum_{e \in E} \sum_{e' \in \varphi^{-1}(e)} \mathcal{N}(\gamma', e')\rho'(e) \\ &= \sum_{e' \in E'} \mathcal{N}(\gamma', e')\rho'(e') = \ell_{\rho'}(\gamma') \geq 1.\end{aligned}$$

Thus, $\rho \in \operatorname{Adm}(\Gamma)$. Moreover, by (2.5.1),

$$\begin{aligned}\operatorname{Mod}_{p,\sigma}(\Gamma) \leq \mathcal{E}_{p,\sigma}(\rho) &= \sum_{e \in E} \sigma(e)\rho(e)^p = \sum_{e \in E} |\varphi_{-1}(e)|\rho(e)^p = \sum_{e \in E} \sum_{e' \in \varphi^{-1}(e)} \rho(e)^p \\ &\leq \sum_{e \in E} \sum_{e' \in \varphi^{-1}(e)} \rho'(e')^p = \sum_{e' \in E'} \rho'(e')^p = \operatorname{Mod}_p(\Gamma').\end{aligned}$$

To establish the opposite inequality, let $\rho \in \operatorname{Adm}(\Gamma)$ be optimal for $\operatorname{Mod}_{p,\sigma}(\Gamma)$ and define $\rho'(e') = \rho(\varphi(e'))$. Let $\gamma' \in \Gamma'$. By definition, $\gamma' = \tau(\gamma)$ for some right inverse τ of φ. So, using Lemma 2.5.8, one can see that

$$\ell_{\rho'}(\gamma') = \sum_{e' \in E'} \mathcal{N}(\gamma', e')\rho'(e') = \sum_{e \in E} \sum_{e' \in \varphi^{-1}(e)} \mathcal{N}(\gamma', e')\rho'(e') = \sum_{e \in E} \mathcal{N}(\gamma, e)\rho'(\tau(e)).$$

By definition, for $e \in E$, $\rho'(\tau(e)) = \rho(\varphi(\tau(e))) = \rho(e)$, so

$$\ell_{\rho'}(\gamma') = \sum_{e \in E} \mathcal{N}(\gamma, e)\rho(e) = \ell_\rho(\gamma) \geq 1.$$

So $\rho' \in \mathrm{Adm}(\Gamma')$ and by (2.5.1),

$$\begin{aligned}\mathrm{Mod}_p(\Gamma') \le \mathcal{E}_p(\rho') = \sum_{e'\in E'} \rho'(e')^p = \sum_{e'\in E'} \rho(\varphi(e'))^p = \sum_{e\in E}\sum_{e'\in\varphi^{-1}(e)} \rho(e)^p \\ = \sum_{e\in E} \sigma(e)\rho(e)^p = \mathcal{E}_{p,\sigma}(\rho) = \mathrm{Mod}_{p,\sigma}(\Gamma).\end{aligned}$$

This proves (2.5.6). Equation (2.5.7) follows from uniqueness of minimizers. $\square$

Theorem 2.5.9 allows one to solve the modulus problem on a weighted graph by replacing it with an unweighted graph. Additionally, it also establishes a connection between the minimum expected overlap problems on G and G' as follows.

Definition 2.5.10. Let $\mu' \in \mathcal{P}(\Gamma')$. The *push forward* $\varphi_*\mu' \in \mathcal{P}(\Gamma)$ is defined as

$$\varphi_*\mu'(\gamma) := \mu'(\varphi^\sharp\gamma) = \sum_{\gamma'\in\varphi^\sharp\gamma} \mu'(\gamma').$$

Theorem 2.5.11. *Let $\mu' \in \mathcal{P}(\Gamma)$ be optimal for MEO(Γ') and let $\eta' = \mathcal{N}^T\mu'$ be the associated expected edge usage. Then $\mu = \varphi_*\mu'$ is optimal for MEO$_{\sigma^{-1}}(\Gamma)$ and its expected edge usages $\eta = \mathcal{N}^T\mu$ satisfy*

$$\eta'(e') = \sigma(\varphi(e'))^{-1}\eta(\varphi(e')) \qquad \forall e' \in E'. \tag{2.5.8}$$

Proof. From the assumptions on μ' and η' and from (2.4.4), the density ρ'

$$\rho'(e') = \eta'(e')\,\mathrm{Mod}_2(\Gamma')$$

is optimal for the modulus problem $\mathrm{Mod}_2(\Gamma')$. Theorem 2.5.9 then implies that $\rho' = \rho\circ\varphi$, where $\rho \in \mathrm{Adm}(\Gamma)$ is the optimal density for $\mathrm{Mod}_{2,\sigma}(\Gamma)$.

Combining Theorem 2.5.9 with (2.4.4) and applying to the second modulus problem, shows that for all $e' \in E$

$$\eta'(e') = \mathrm{Mod}_2(\Gamma')^{-1}\rho'(e') = \mathrm{Mod}_{2,\sigma}(\Gamma)^{-1}\rho(\varphi(e')) = \sigma(\varphi(e'))^{-1}\eta(\varphi(e')),$$

therefore establishing (2.5.8). In order to show that $\mu = \varphi_* \mu'$ is optimal for $\mathrm{MEO}_{\sigma^{-1}}(\Gamma)$, then one only needs to show that $\eta = \mathcal{N}^T \mu$.

To that end, let $e \in E$ and apply Lemma 2.5.8 with $f \equiv 1$:

$$\begin{aligned}
(\mathcal{N}^T \mu)(e) &= \sum_{\gamma \in \Gamma} \mathcal{N}(\gamma, e)\mu(\gamma) = \sum_{\gamma \in \Gamma} \sum_{\gamma' \in \varphi^\sharp \gamma} \mathcal{N}(\gamma, e)\mu'(\gamma') \\
&= \sum_{\gamma \in \Gamma} \sum_{\gamma' \in \varphi^\sharp \gamma} \sum_{e' \in \varphi^{-1}(e)} \mathcal{N}(\gamma', e')\mu'(\gamma') \\
&= \sum_{e' \in \varphi^{-1}(e)} \sum_{\gamma' \in \Gamma} \mathcal{N}(\gamma', e')\mu'(\gamma') = \sum_{e' \in \varphi^{-1}(e)} \eta'(e').
\end{aligned}$$

Since ρ', and therefore η', is constant on $\varphi^{-1}(e)$ for each $e \in E$, it follows that

$$(\mathcal{N}^T \mu)(e) = \sigma(e)\eta'(e') = \eta(e)$$

for any $e \in E$ and any $e' \in \varphi^{-1}(e)$. □

In fact, there is a slightly stronger connection between the two MEO problems.

Theorem 2.5.12. *Let $\mu \in \mathcal{P}(\Gamma)$ be optimal for $MEO_{\sigma^{-1}}(\Gamma)$. Then there exists $\mu' \in \mathcal{P}(\Gamma')$ such that $\mu = \varphi_* \mu'$.*

In other words, not only does every optimal pmf $\mu' \in \mathcal{P}(\Gamma')$ give rise to an optimal pmf $\mu \in \mathcal{P}(\Gamma)$, but every optimal $\mu \in \mathcal{P}(\Gamma)$ can be obtained in this way.

For the remainder of this thesis, based on the theory provided in this section, assume that $G = (V, E)$ is an unweighted, connected, multigraph.

2.6 Feasible Partitions and the dual family

Recall from Section 2.2, that the set of admissible densities $\mathrm{Adm}(\Gamma)$ is the intersection of half-spaces. Hence it can be defined by its finite extreme points or vertices, denoted $\mathrm{ext}(\mathrm{Adm}(\Gamma))$. Using this idea, along with the work of Fulkerson [17], we can establish another related family of objects called the *blocker*.

Definition 2.6.1. Given a finite graph $G = (V, E)$ and the family of spanning trees Γ. We say that the family

$$\hat{\Gamma} := \mathrm{ext}(\mathrm{Adm}(\Gamma)) = \{\hat{\gamma}_1, \ldots, \hat{\gamma}_s\} \subset \mathbb{R}^E_{\geq 0}$$

consisting of the extreme points of $\mathrm{Adm}(\Gamma)$, is the *Fulkerson blocker* of Γ, and define the matrix $\hat{\mathcal{N}} \in \mathbb{R}^{\Gamma \times E}_{\geq 0}$ to be the matrix whose rows are the vectors $\hat{\gamma}^T$, for $\hat{\gamma} \in \hat{\Gamma}$

Recast in the preceding notation, the result of Chopra [10] is that when Γ is the set of spanning trees, the blocker $\hat{\Gamma}$ can be identified as a set of objects called feasible partitions. We review the definition and Chopra's results here.

Definition 2.6.2. A *feasible partition* P of a graph $G = (V, E)$ is a partition of the vertex set V into two or more nonempty subsets, $\{V_1, \ldots, V_{k_P}\}$, such that each of the induced subgraphs $G(V_i)$ is connected. The corresponding edge set, E_P, is defined to be the set of edges in G that connect vertices belonging to different V_i's. The family of all feasible partitions of G is denoted $\hat{\Gamma}$.

The results of [10] can be restated as follows.

Theorem 2.6.3. *Let $G = (V, E)$ be given and let Γ be the family of spanning trees on G. Then the blocker $\hat{\Gamma}$ is the set of vectors*

$$\hat{\Gamma} = \left\{ \frac{1}{k_P - 1} \mathbb{1}_{E_P} : P \text{ is a feasible partition of } G \right\}.$$

The weight of a partition is then defined as follows.

Definition 2.6.4. Let P be a feasible partition on G. The *weight* of P is defined as

$$w(P) = \frac{|E_P|}{k_P - 1}.$$

Example 2.6.5. Consider the slashed square of Example 2.2.1, the feasible partitions are given in figure 2.8 along with their corresponding weight. The darkened edges correspond to the edges in E_P. The weights are given as unreduced fractions, so that one can see where the numerator and denominator are coming from.

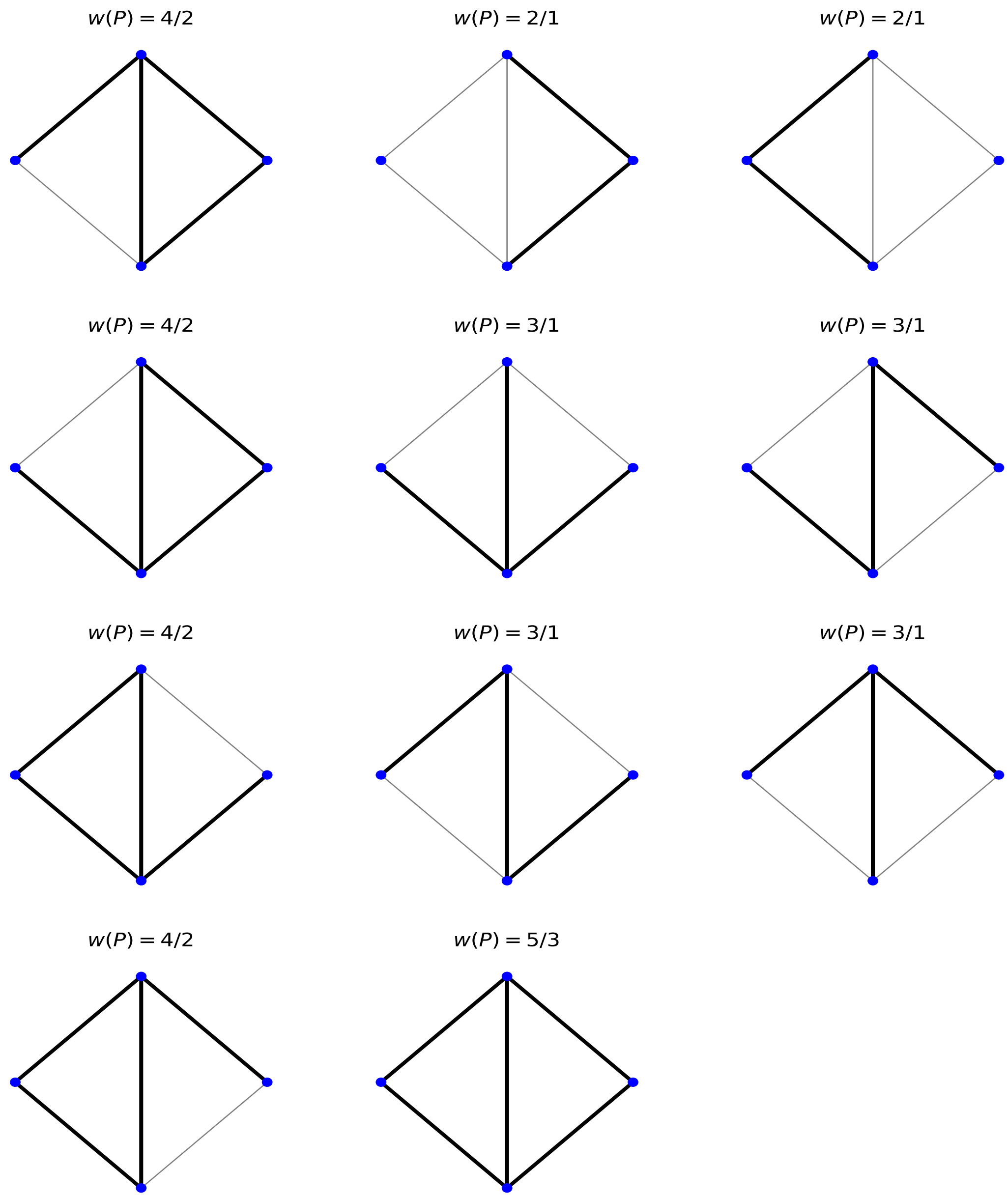


Figure 2.8: Feasible partitions of the slashed square

The usage matrix for $\hat{\Gamma}$ is denoted $\hat{\mathcal{N}}$, and note that unlike the usage matrix for spanning trees, this is not a (0/1) matrix. However,$\hat{\Gamma}$ is a nonempty family of objects on a graph G, and therefore defines another modulus problem, namely

$$\mathrm{Mod}_2(\hat{\Gamma})^{-1} = \min_{\nu \in \mathcal{P}(\hat{\Gamma})} \nu^T \hat{\mathcal{N}} \hat{\mathcal{N}}^T \nu$$

These two modulus problems can be connected through the following theorem from [2] (Although there it is written for the weighted case).

Theorem 2.6.6. *Let $G = (V, E)$ be a connected graph and let Γ be the family of spanning trees on G with Fulkerson blocker $\hat{\Gamma}$. Let $1 < p < \infty$ be given, with $q = \frac{p}{p-1}$ its Hölder conjugate exponent. Then*

$$\mathrm{Mod}_p(\Gamma)^{\frac{1}{p}} \,\mathrm{Mod}_q(\hat{\Gamma})^{\frac{1}{q}} = 1. \tag{2.6.1}$$

Moreover, the optimal $\rho^* \in \mathrm{Adm}(\Gamma)$ and $\eta^* \in \mathrm{Adm}(\hat{\Gamma})$ are unique and are related as follows:

$$\eta^*(e) = \frac{\rho^*(e)^{p-1}}{\mathrm{Mod}_p(\Gamma)} \qquad \forall e \in E. \tag{2.6.2}$$

Proof. For all $\rho \in \mathrm{Adm}(\Gamma)$ and $\eta \in \mathrm{Adm}(\hat{\Gamma})$, Hölder's inequality implies that

$$1 \le \sum_{e\in E} \rho(e)\eta(e) \le \left(\sum_{e\in E} \rho(e)^p\right)^{1/p} \left(\sum_{e\in E} \eta(e)^q\right)^{1/q}, \tag{2.6.3}$$

so

$$\mathrm{Mod}_p(\Gamma)^{1/p}\,\mathrm{Mod}_q(\hat{\Gamma})^{1/q} \ge 1. \tag{2.6.4}$$

Now, let $\alpha := \mathrm{Mod}_q(\hat{\Gamma})^{-1}$ and let $\eta^* \in \mathrm{Adm}(\hat{\Gamma})$ be the minimizer for $\mathrm{Mod}_q(\hat{\Gamma})$. Then (2.6.4) implies that

$$\mathrm{Mod}_p(\Gamma) \ge \alpha^{\frac{p}{q}} = \alpha^{\frac{1}{q-1}}. \tag{2.6.5}$$

Define

$$\rho^*(e) := \alpha \left(\eta^*(e)^q\right)^{1/p} = \alpha \eta^*(e)^{q/p}. \tag{2.6.6}$$

Note that

$$\mathcal{E}_p(\rho^*) = \sum_{e\in E} \rho^*(e)^p = \alpha^p \sum_{e\in E} \eta^*(e)^q = \alpha^{p-1} = \alpha^{\frac{1}{q-1}}.$$

Thus, all the is left to show is that $\rho^* \in \mathrm{Adm}(\Gamma)$. Moreover, if this is shown then (2.6.5) is attained and ρ^* must be extremal for $\mathrm{Mod}_p(\Gamma)$. In particular, (2.6.1) would follow. Moreover, (2.6.2) is another way of writing (2.6.6).

To see that $\rho^* \in \mathrm{Adm}(\Gamma)$, one need only show that $\sum_{e\in E} \rho^*(e)\eta(e) \geq 1$ for all $\eta \in \mathrm{Adm}(\hat{\Gamma})$. To see this, first consider the case when $\eta = \eta^*$. Then

$$\sum_{e\in E} \rho^*(e)\eta^*(e) = \alpha \sum_{e\in E} \eta^*(e)^q = 1.$$

Now let $\eta \in \mathrm{Adm}(\hat{\Gamma})$ be arbitrary. Since $\mathrm{Adm}(\hat{\Gamma})$ is convex, we have that $(1-\theta)\eta^* + \theta\eta \in \mathrm{Adm}(\hat{\Gamma})$ for all $\theta \in [0,1]$. Using Taylor's theorem, then

$$\begin{aligned}
\alpha^{-1} = \mathcal{E}_q(\eta^*) \leq \mathcal{E}_q((1-\theta)\eta^* + \theta\eta) &= \sum_{e\in E} [(1-\theta)\eta^*(e) + \theta\eta(e)]^q \\
&= \alpha^{-1} + q\theta \sum_{e\in E} \eta^*(e)^{q-1} \left(\eta(e) - \eta^*(e)\right) + O(\theta^2) \\
&= \alpha^{-1} + \alpha^{-1} q\theta \sum_{e\in E} \rho^*(e) \left(\eta(e) - \eta^*(e)\right) + O(\theta^2).
\end{aligned}$$

Since this inequality must hold for arbitrarily small $\theta > 0$, it follows that

$$\sum_{e\in E} \rho^*(e)\eta(e) \geq \sum_{e\in E} \rho^*(e)\eta^*(e) = 1,$$

and the proof is complete. $\square$

Lemma 2.6.7. *Let $\mu \in \mathcal{P}(\Gamma)$ and let $\eta = \mathcal{N}^T\mu$. Then μ is optimal for* (2.4.3) *if and only if $\rho := \frac{\eta}{\eta^T\eta}$ is admissible.*

Proof. First, suppose $\rho = \frac{\eta}{\eta^T\eta}$ is admissible, then

$$\mathrm{Mod}_2(\Gamma) \leq \rho^T\rho = \frac{\eta^T\eta}{(\eta^T\eta)^2} = \frac{1}{\eta^T\eta} \leq \mathrm{Mod}_2(\hat{\Gamma})^{-1} = \mathrm{Mod}_2(\Gamma)$$

On the other hand, if μ is optimal then (2.4.4) shows that ρ is optimal, hence it must be admissible. $\square$

Feasible partitions are important objects for understanding optimal pmfs for the minimum expected overlap (MEO) problem in the sense that if P is a feasible partition for G with k_P parts, then every spanning tree of G must use at least $k_P - 1$ edges in E_P. This fact gives rise to the following lemma.

Lemma 2.6.8. *Let P be a feasible partition, let $\mu \in \mathcal{P}(\Gamma)$, and let $\eta = \mathcal{N}^T \mu$. Then*

$$\frac{1}{|E_P|} \sum_{e \in E_P} \eta(e) \geq w(P)^{-1}.$$

Moreover, there exists an edge $e \in E_P$ such that

$$\eta(e) \geq w(P)^{-1}$$

Proof. For the first part, consider the fact that for any spanning tree $\gamma \in \Gamma$, $|\gamma \cap E_P| \geq k_P - 1$. So

$$\begin{aligned}\sum_{e \in E_P} \eta(e) = \sum_{e \in E_P} \sum_{\gamma \in \Gamma} \mathcal{N}(\gamma, e)\mu(\gamma) &= \sum_{\gamma \in \Gamma} \mu(\gamma) \sum_{e \in E_P} \mathcal{N}(\gamma, e) \\ &\geq (k_P - 1) \sum_{\gamma \in \Gamma} \mu(\gamma) = k_P - 1.\end{aligned}$$

Dividing by $|E_P|$ yields the desired result.

For the second part, notice that when dividing by $|E_P|$, this is just the average over all the values of η on the partition edges. Hence there must be at least one edge where the value of η is greater than or equal to this average. $\square$

This is a more general statement than given in Lemmas 4.7 and 4.8 of [3]. There, it is assumed that μ is optimal, but this is unnecessary as the result only relies on the fact that every spanning tree uses at least $k_P - 1$ edges of E_P.

Note that the definition of a feasible partition with $k_P = 2$ coincides with the usual definition of a (minimal) graph cut and that, in this case, $w(P)$ is the standard measure of “weight” or “value” of a cut. So feasible partitions can be thought of as generalized cuts of a graph. However, if $k_P \geq 2$ then the weight of the partition is, in a sense, normalized to take into account how many pieces it splits the graph into.

In light of section 2.5, if G' is the unweighted representation of a weighted graph G, then the feasible partition P' on G', with edge set $E_{P'}$, corresponds to a feasible partition (for similar reasons that Theorem 2.5.7 is true) P on G, with edge set E_P. Moreover, the edge sets are related through $E_P = \varphi(E_{P'})$ and the weight of the partition on G is computed by

$$w(P) := \frac{\sum\limits_{e \in \varphi(E_{P'})} \sigma(e)}{k_P - 1}$$

Another important partition is the *trivial partition* and is defined as follows.

Definition 2.6.9. Let $G = (V, E)$ be a finite undirected graph, the *trivial partition*, denoted P_0, is defined as the partition that isolates every node, in other words $E_{P_0} = E$. Moreover, the weight is given by

$$w(P_0) = \frac{|E|}{|V| - 1} \tag{2.6.7}$$

In [3], this is defined as the *sparseness* of a graph.

To complete the picture, using the analogue of η^* representing the expected edge usage for spanning trees when selected with pmf $\mu \in \mathcal{P}(\Gamma)$, one can introduce a pmf $\nu \in \mathcal{P}(\hat{\Gamma})$ so that ρ^* represents the expected edge usage of a feasible partition.

Feasible partitions will be revisited in Section 3.3.2 when describing their use in determining graph structure.

2.7 Homogeneous Graphs

Definition 2.7.1. A graph $G(V, E)$ is said to be *homogeneous* with respect to spanning tree modulus, if there exists a *pmf* $\mu \in \mathcal{P}(\Gamma)$ and $c \in \mathbb{R}$ such that $\eta^*(e) \equiv c$ for every edge $e \in E$.

If this happens, it can be shown that the value of c must be

$$c = \frac{|V| - 1}{|E|}$$

Notice that with the previous definition, along with definition 2.6.9 gives another interpretation of homogeneity.

Theorem 2.7.2. *[3, Theorem 4.11] A graph $G(V, E)$ is homogeneous if and only if for every feasible partition P of G with k_P parts, we have*

$$w(P) = \frac{|E_P|}{k_P - 1} \geq \frac{|E|}{|V| - 1} \tag{2.7.1}$$

Proof. Let η^* be the optimal density for $\mathrm{Mod}_2(\widehat{\Gamma})$. Suppose G is homogeneous, then by definition 2.7.1

$$\eta^*(e) = \frac{|V| - 1}{|E|} \qquad \forall e \in E.$$

Then by Lemma 2.6.8, (2.7.1) holds for every feasible partition P.

On the other hand, suppose 2.7.1 holds for every feasible partition P. The usage matrix $\widehat{\mathcal{N}}$ for $\widehat{\Gamma}$ is given by

$$\widehat{\mathcal{N}}(P, e) = \frac{1}{k_P - 1} \mathbb{1}_{e \in E_P}$$

So if P is a feasible partition and $\eta(e) \equiv \frac{|V|-1}{|E|}$, then

$$\sum_{e \in E} \widehat{\mathcal{N}}(P, e)\eta(e) = \frac{|V| - 1}{|E|} \sum_{e \in E_P} \frac{1}{k_P - 1} = \frac{|V| - 1}{|E|} \frac{|E_P|}{k_P - 1} \geq 1$$

where the last inequality holds by the assumption. This shows that η is admissible, and hence optimal. □

This means that if one wanted to find a lower bound for the edge probabilities η^*, the trivial partition gives that bound.

Section 3.3.1 will give a more thorough understanding of the importance of homogeneous graphs. Some common homogeneous graphs are:

- Simple Trees
- Cycle graph on n vertices (C_n)
- Complete graphs on n vertices(K_n)

If $G = (V, E)$ is a simple connected graph, then the family of spanning trees is nonempty. Moreover, every spanning tree of G is a spanning tree of the complete graph $K_{|V|}$. Combining these observations with the monotonicity of modulus (2.3.1), one can bound the 2-modulus problem as follows.

Theorem 2.7.3. *Given any simple connected graph $G = (V, E)$ and the family of spanning trees Γ, then*

$$\frac{1}{|V|-1} \leq \mathrm{Mod}_2(\Gamma) \leq \frac{|V|}{2(|V|-1)} \tag{2.7.2}$$

Proof. Let $\gamma \in \Gamma$ be a spanning tree of G, and consider the complete graph $K_{|V|}$, then

$$\gamma \subset \Gamma \subset \Gamma_{K_{|V|}}$$

hence by Lemma 2.3.1

$$\frac{1}{|V|-1} = \mathrm{Mod}_2(\gamma) \leq \mathrm{Mod}_2(\Gamma) \leq \mathrm{Mod}_2(\Gamma_{K_{|V|}}) = \frac{|V|}{2(|V|-1)}$$

□

2.7.1 Reducible Graphs

Definition 2.7.4. A homogeneous graph $G = (V, E)$ is called *reducible* if there exists a strict vertex-induced subgraph $G' = (V', E')$ with the property that every optimal pmf for G is supported on trees that restrict to trees on G'.

Theorem 2.7.5. *A homogeneous graph $G = (V, E)$ is reducible if and only if there exists a strict* *vertex-induced subgraph $G' = (V', E')$ with the property that*

$$\frac{|V|-1}{|E|} = \frac{|V'|-1}{|E'|}$$

Proof. First, suppose that G is reducible and let G' be the subgraph guaranteed by the definition. If μ^* is optimal for the spanning tree modulus of G, then it gives every edge in E a probability of $(|V|-1)/|E|$ of appearing in a random tree. Moreover, the restriction of μ^* to G' is optimal for the spanning tree modulus on the subgraph and thus gives every edge in E' a probability of $(|V'|-1)/|E'|$ of appearing. Since $E' \subset E$, the ratios are equal.

Conversely, suppose that G contains a strict edge-induced subgraph G' with the same ratio of "nodes-minus-1 to edges". Let μ^* be any optimal pmf for G. Then

$$\sum_{e\in E} \mathbb{P}_{\mu^*}[e \in \underline{\gamma}] = |E'|\frac{V|-1}{|E|} = |E'|\frac{|V'|-1}{|E'|} = |V'|-1.$$

But, also

$$\sum_{e\in E'} \mathbb{P}_{\mu^*}[e \in \underline{\gamma}] = \sum_{e\in E'}\sum_{\gamma\in\Gamma} \mu^*(\gamma)\mathcal{N}(\gamma,e) = \sum_{\gamma\in\Gamma} \mu^*(\gamma)\sum_{e\in E'}\mathcal{N}(\gamma,e) = \sum_{\gamma\in\Gamma}\mu^*(\gamma)|\gamma\cap E'|.$$

Since no spanning tree of G can use more than $|V'|-1$ of the edges in E', every $|\gamma\cap E'|$ is less than or equal to $|V'|-1$. A convex combination of these numbers, then, is also no larger than $|V'|-1$ and equality can only be attained if μ^* is positive only for trees that attain this upper bound. Thus, every spanning tree in the support of μ^* uses $|V'|-1$ edges in E' and, therefore, restricts to a spanning tree on G'. □

This theorem also lends itself to a similar statement using feasible partitions, that is

Corollary 2.7.6. *A homogeneous graph $G = (V, E)$ is reducible if and only if there is a nontrivial feasible partition P such that*

$$w(P) = w(P_0) = \frac{|V|-1}{|E|}$$

Note that this ratio implies that the number of nodes and edges in an irreducible graph are then necessarily relatively prime. This is made formal through the following corollary.

Corollary 2.7.7. *Given a homogeneous graph $G(V, E)$, if $\gcd(|V| - 1, |E|) = 1$, then G is irreducible.*

Notice that $\gcd(|V| - 1, |E|) \neq 1$ is a necessary condition for reducibility, however it is not sufficient. For example, complete graphs are irreducible, and $|E| = \frac{|V|(|V|-1)}{2}$.

2.7.2 Recurrence relation for number of spannning trees

This section is based on the work established by Raff in [34]. Based on the previous section, it can be seen that the house graph is a reducible graph, and is strikingly similar to the graph that Raff considered. The synopsis of his work is as follows, given the $2 \times n$ grid $G_2(n)$, the number of spanning trees is given by the recurrence relation

$$T_n = 4T_{n-1} - T_{n-2} \qquad \text{with initial conditions } T_0 = 1,\ T_1 = 4$$

In his construction, he uses the fact that when extending from $G_2(n)$ to $G_2(n + 1)$ the number of trees added is based on two factors:

1. Given a tree on $G_2(n)$, there are three ways to extend this to a tree on $G_2(n + 1)$.

2. Given a forest on $G_2(n)$ with the property that the two right most endpoints are in separate trees, and there are exactly two components in the forest, then there is one way to extend this to a tree on $G_2(n + 1)$.

So it is also important to keep track of these forests, denoted $\mathcal{F}_n$ with $|\mathcal{F}_n| = F_n$. As it turns out, as Raff shows, F_n satisfies the same recurrence relation.

$$F_n = 4F_{n-1} - F_{n-2} \qquad \text{with intial conditions } F_0 = 1,\ F_1 = 3$$

Example 2.7.8. Consider Figure 2.9, $\mathcal{H}_r$, or the r-story house graph, is similar to $G_2(n)$ with the exception of the "roof", or top triangle. The figure is annotated and the labels will be referred to later.

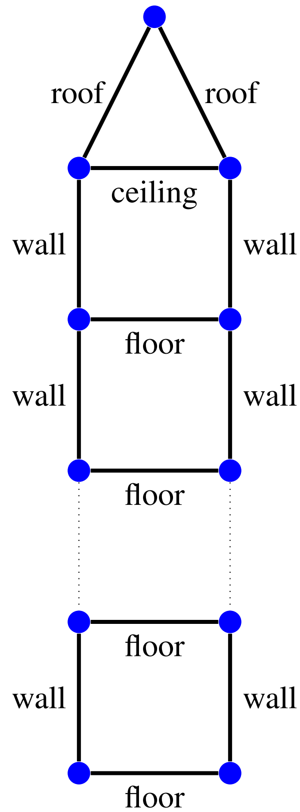


Figure 2.9: $\mathcal{H}_r$

Theorem 2.7.9. *The number of spanning trees on the r-story house, H_r, $r \geq 0$, is given by the recurrence relation*

$$H_r = 2T_r + F_r.$$

Rationale: Notice that using only one edge from the roof will not produce any cycles with trees in $\mathcal{T}_r$, since there are two roof edges to choose from this yields the first part. Also, notice that using both roof edges will certainly create a cycle with any tree in $\mathcal{T}_n$, and therefore are not trees on the r-story house. However, if we consider $\mathcal{F}_n$, then no cycle will be created because of the condition that both endpoints that adjoin the roof are in seperate trees, hence every forest is admissable, and there is only one way to use both edges.

Check with known: We have computed the number of spanning trees for the 1,2 and 3 story house.

- House - 11 spanning trees, r=1:

$$2T_r + F_r = 2(4) + 3 = 11$$

- 2-House - 41 spanning trees, r=2:

$$2T_r + F_r = 2(4(4) - 1) + (4(3) - 1) = 2(15) + 11 = 41$$

- 3-House - 153 spanning trees, r=3:

$$2T_r + F_r = 2(4(15) - 4) + (4(11) - 3) = 2(56) + 41 = 153$$

Notice that this recurrence relation could be rewritten as the following:

$$H_r = 2T_r + H_{r-1} \qquad \text{with initial conditions } T_0 = 1, T_1 = 4, H_0 = 3$$

More importantly, notice that F_r can be thought of as having its own recurrence relation. Notice that for any tree in $G_2(r-1)$ there are two ways of creating a forest in $G_2(r)$, and for any forest, there is only one way of extending it to a forest on $G_2(r)$, hence

$$F_r = 2T_{r-1} + F_{r-1}$$

Which gives that H_r satisfies the same recurrence relation, that is,

$$H_r = 2T_r + F_r = F_{r+1} = 4F_r - F_{r-1} = 4H_{r-1} - H_{r-2}$$

2.7.3 Number of forbidden trees

Since we know that the number of fair trees on $\mathcal{H}_r$ is given by

$$3^{r+1}$$

this gives a usefull way of calculating the number of "forbidden" trees, $\tilde{F}_r$.

$$\tilde{F}_r = H_r - 3^{r+1}$$

2.7.4 Calculating Effective Resistance

We know that every edge will show up in exactly $\frac{2}{3}$ of the "good" trees. However, knowing how many of the forbidden trees each edge shows up in is the question. Since we know the number of total spanning trees of $\mathcal{H}_r$, then the effective resistance of each edge is given by

$$\mathcal{R}_{\text{eff}}(e) = \frac{2 \cdot 3^r + |\{f \in \tilde{\mathcal{F}}_r : e \in f\}|}{H_r}$$

In order to calculate the effective resistance we make note of a few facts:

1. The r-story house, $\mathcal{H}_r$, will be considered to be upright, with the bottom floor being floor number 1, and the top floor being floor r.
2. There are four main concerns, the roof, the walls, the floors, and the ceiling of the floor r.

We can find the effective resistances for the walls and floors using the diagram in Figure 2.10. Where α and β are equivalent resistors given by collapsing the levels below and above respectively, using Kirchoff's laws.

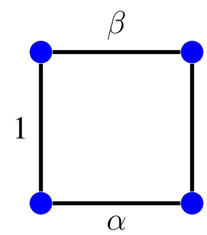


Figure 2.10: Resistance Diagram

Where if this is the k^{th} level (counting from the bottom up), then

$$\alpha = \frac{H_{k-2}}{T_{k-1}} \qquad \text{and} \qquad \beta = \frac{2T_{r-k}}{H_{r-k}}$$

Both of these formulas can be easily verified, if one considers the trees that each edge is not in (for the top, the r-k house, and the bottom, $G_2(k-1)$).

This allows us to establish a formula for the effective resistances of the floors and walls of any level.

$$\begin{aligned}
\mathcal{R}_{\text{eff}}(k^{th}\text{ floor}) &= \frac{\alpha(2+\beta)}{2+\alpha+\beta} \\
&= \frac{2H_{k-2}(H_{r-k}+T_{r-k})}{2H_{r-k}T_{k-1}+H_{k-2}H_{r-k}+2T_{r-k}T_{k-1}} \\
\mathcal{R}_{\text{eff}}(k^{th}\text{ level walls}) &= \frac{1+\alpha+\beta}{2+\alpha+\beta} \\
&= \frac{T_{k-1}H_{r-k}+H_{k-2}H_{r-k}+2T_{r-k}T_{k-1}}{2T_{k-1}H_{r-k}+H_{k-2}H_{r-k}+2T_{r-k}T_{k-1}}
\end{aligned}$$

Notice that for $k = 1$ this gives the same values, that is

$$\mathcal{R}_{\text{eff}}(\text{bottom floor and walls}) = \frac{2T_r}{H_r}$$

It is also worth noting that for $k = r$ this becomes

$$\mathcal{R}_{\text{eff}}(\text{top floor}) = \frac{8H_{r-2}}{H_r}$$

To find the effective resistances for the two roof edges, notice that the diagram only needs slight modification, see figure 2.11.

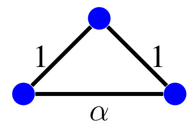


Figure 2.11: Resistance Diagram

Where here $\alpha = \frac{H_{r-1}}{T_r}$. This gives

$$\mathcal{R}_{\text{eff}}(\text{roof}) = \frac{T_r + H_{r-1}}{H_r}$$

Finally, for the effective resistance of the ceiling, again the diagram in figure 2.11 shows that

$$\begin{aligned}\mathcal{R}_{\text{eff}}(\text{ceiling}) &= \frac{2\alpha}{\alpha + 2} \\ &= \frac{2H_{r-1}}{H_r}\end{aligned}$$

Conjecture 2.7.10. *Given the r-story house $\mathcal{H}_r$ the following equality holds for any k.*

$$2T_{k-1}H_{r-k} + H_{k-2}H_{r-k} + 2T_{r-k}T_{k-1} = H_r$$

2.8 Greedy Algorithm for Computing Spanning Tree Modulus

2.8.1 Previous approach

Before going into the main results of this section, it should be pointed out how modulus has been computed up until now. The previous algorithm, laid out in [4, 6], when $p = 2$ used a primal-dual active set method based on an algorithm developed by Goldfarb and Idnani [19], where in each step a subproblem is optimized and then a new constraint is added and a re-optimization is done. The pseudocode for this algorithm, given in [6], is given for the readers edification in Algorithm 1.

Algorithm 1 Previous approach

$\rho \leftarrow 0$
$\Gamma' \leftarrow \emptyset$
loop
    $\gamma \leftarrow \mathrm{mnspt}(\rho)$
    if $\ell_\rho(\gamma)^p \geq 1 - \epsilon_{\text{tol}}$ **then**
        stop
    end if
    $\Gamma' \leftarrow \Gamma' \cup \{\gamma\}$
    $\rho \leftarrow \text{argmin}\{\mathcal{E}_p(\rho) : \rho \in \mathrm{Adm}(\Gamma')\}$
end loop

Where $\mathrm{mnspt}(\rho)$ denotes the minimum spanning tree with respect to edge weights $\rho(e)$. This can be found, for example, using Kruskal's algorithm. The algorithm is demonstrated in Example 2.8.1.

Example 2.8.1. The previous algorithm utilized finding the most violated constraint, in this case the minimum spanning tree. Note that if $\ell_\rho(\Gamma') > 0$ in any iteration, then $\rho/\ell_\rho(\Gamma') \in \mathrm{Adm}(\Gamma')$, providing an upper bound on $\mathrm{Mod}_p(\Gamma')$, hence the stopping condition. Consider the graph in Figure 2.12 for visualization of this process.

2.8.2 Greedy Approach

Going back to the minimum spanning tree problem, it would seem like a natural extension of the 2-modulus problem minimizing overlap representation (2.4.3), that if one chooses a minimum spanning tree and the subsequently chooses another, taking into account the previous choice(s) that this should lead to a sequence of trees that can estimate the spanning tree modulus problem.

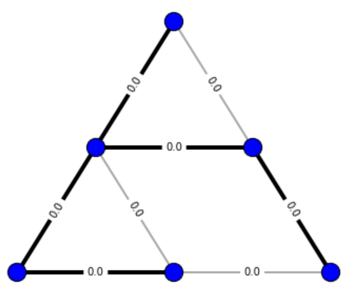
This idea is the basis of the greedy algorithm developed for spanning tree modulus, that is, if $G(V,E)$ is a connected graph, let $\{\gamma_1, \gamma_2, \ldots\}$ be a the sequence of spanning trees with the property that, for $k \geq 2$, γ_{k+1} is the minimum spanning tree with respect to the edge weights

$$\omega_k(e) = \sum_{\ell=1}^{k} \mathcal{N}(\gamma_\ell, e) \qquad \text{for } k = 1, 2, \ldots.$$

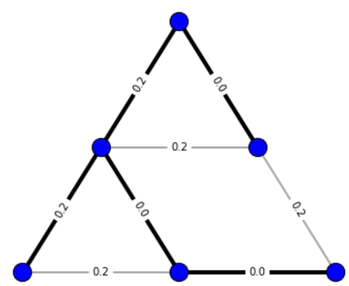
When $k = 1$, γ_1 is an arbitrary spanning tree, or can be thought of as the minimum spanning tree with respect to edge weights identically zero, essentially taken uniformly at random (although in practice this is not the case). This gives rise to the first set of weights ω_1 which is 1 for any edge in γ_1 and zero for all other edges, in other words $\omega_1 = \mathbb{1}_{\gamma_1}$, γ_2 is then chosen to be minimal with respect to these edge weights. The new edge weights are then given by $\omega_2 = \mathbb{1}_{\gamma_1} + \mathbb{1}_{\gamma_2}$, and the process is repeated. For this reason, the algorithm is named the *Plus-1 algorithm*,

Example 2.8.2. The Plus-1 algorithm is illustrated in Figure 2.13 below.

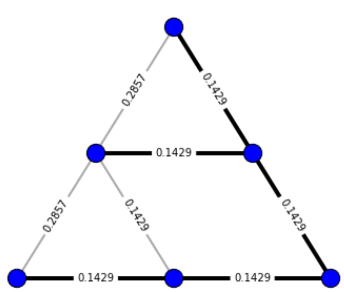
With this general example, we can expand this to a random geometric graph to see if the algorithm can detect the different components of the graph (see next chapter).



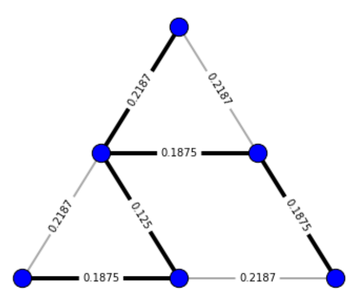
(a) Initialization with $\rho \equiv 0$ and darkened edges denote spanning tree chosen.
$0.0 \leq \text{Mod} \leq \infty$



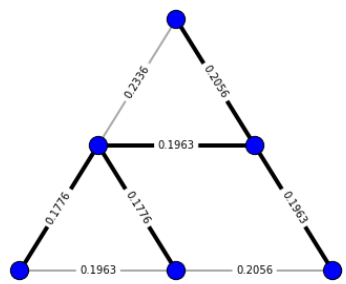
(b) Lightest tree has weight .4,
$.200000004629 \leq \text{Mod} \leq 1.25$



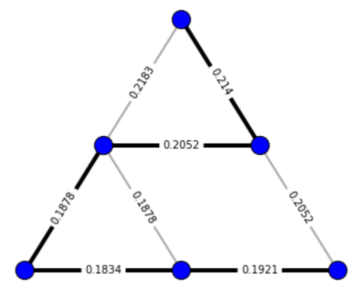
(c) Lightest tree has weight .714285693614,
$0.285714269177 \leq \text{Mod} \leq 0.56$



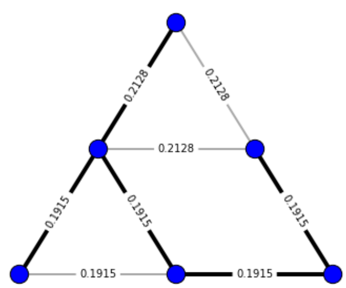
(d) Lightest tree has weight 0.906249955951
$0.312499967084 \leq \text{Mod} \leq 0.38049940238$



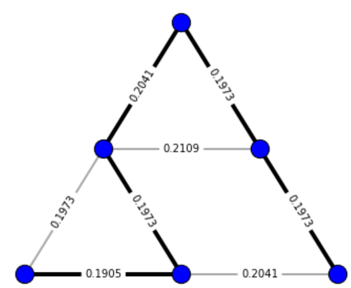
(e) Lightest tree has weight 0.953271027473
$0.31775700897 \leq \mathrm{Mod} \leq 0.349673202614$



(f) Lightest tree has weight 0.965065505777
$0.318777296087 \leq \mathrm{Mod} \leq 0.342273910436$



(g) Lightest tree has weight 0.98639453468
$0.31972789446 \leq \mathrm{Mod} \leq 0.328608817864$



(h) Lightest tree has weight 0.99999990553
$0.320000003367 \leq \mathrm{Mod} \leq 0.320000063828$

Figure 2.12: Basic algorithm approach to spanning tree modulus

Algorithm 2 Plus one algorithm

procedure PLUS-1(G, ϵ_{tol})
    $\omega_0(E) \leftarrow 0$
    $\gamma_1 \leftarrow \text{mnspt}(G, \omega_0)$
    $\omega_1(e \in \gamma_1) += 1$
    $\gamma_2 \leftarrow \text{mnspt}(G, \omega_1)$
    $\omega_2(e \in \gamma_2) += 1$
    $k = 2$
    while $k \geq 2$ **and** $\|\frac{\omega_{k-1}}{k-1} - \frac{\omega_k}{k}\| > \epsilon_{\text{tol}}$ **do**
        $k += 1$
        $\gamma_k \leftarrow \text{mnspt}(G, \omega_{k-1})$
        $\omega_k(e \in \gamma_k) += 1$
    end while
    return $\frac{w_k^T w_k}{k^2}$
end procedure

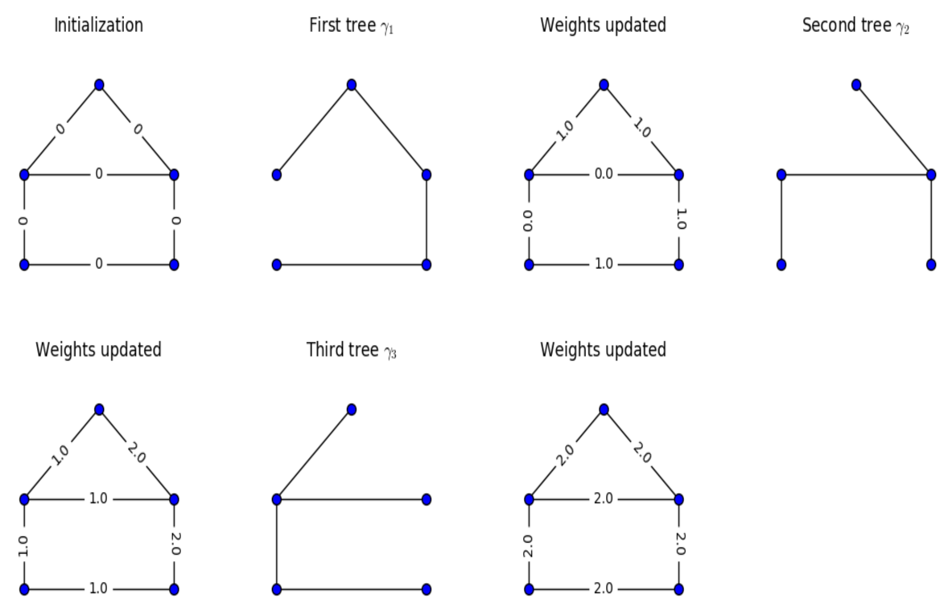


Figure 2.13: Plus-1 algorithm illustration

Example 2.8.3. Running the Plus-1 algorithm on an Erdós-Rényi random geometric graph, $G(n, p)$, shows that as the number of iterations grows, $\frac{w_k(e)}{k}$ converges, presumably to $\eta^*(e)$.

Figure 2.14: Erdós-Rényi $G(n, p)$ graph with $n = 50$ and $p = .25$

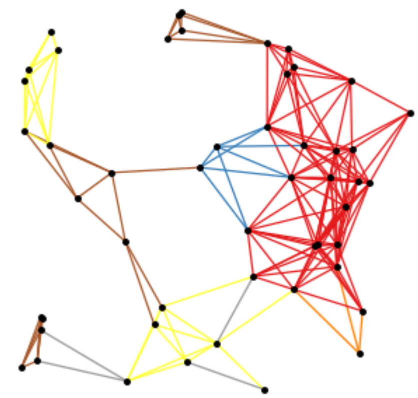


Figure 2.15: Edges colored based on η^* on each edge

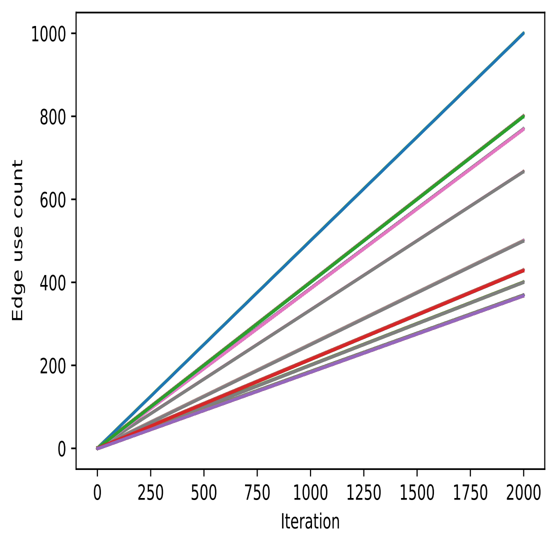


Figure 2.16: Edge usage counts ($w_k(e)$) over 10,000 iterations.

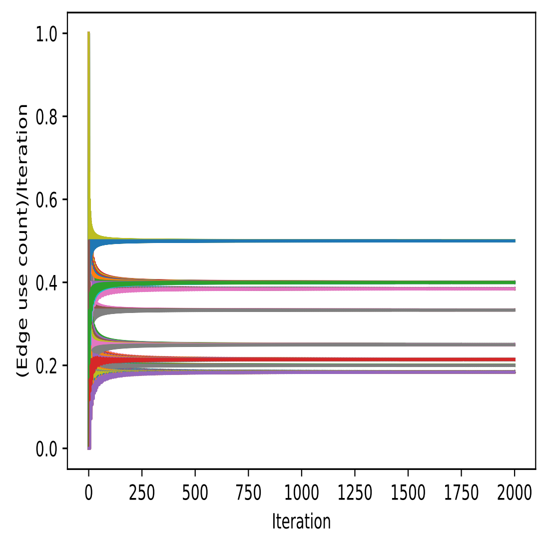


Figure 2.17: $\frac{w_k(e)}{k}$ over 10,000 iterations

Example 2.8.4. Running the algorithm on the uncapicitated Zachary Karate club graph [43] (Figure 2.18), Figure 2.19 shows that after about 2500 iterations, the values of η_k on each edge are fairly distinct.

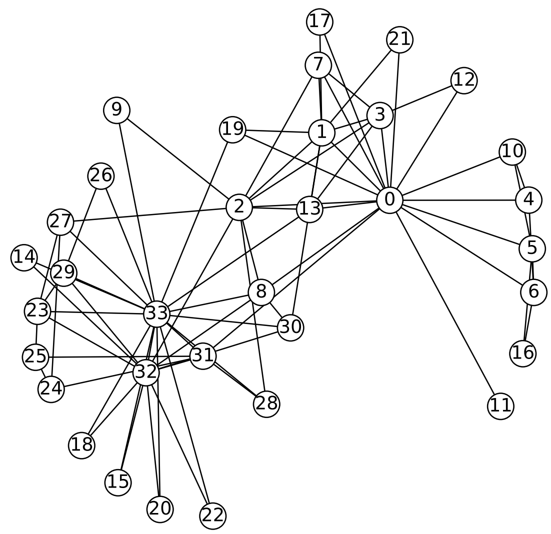


Figure 2.18: Zachary Karate club graph

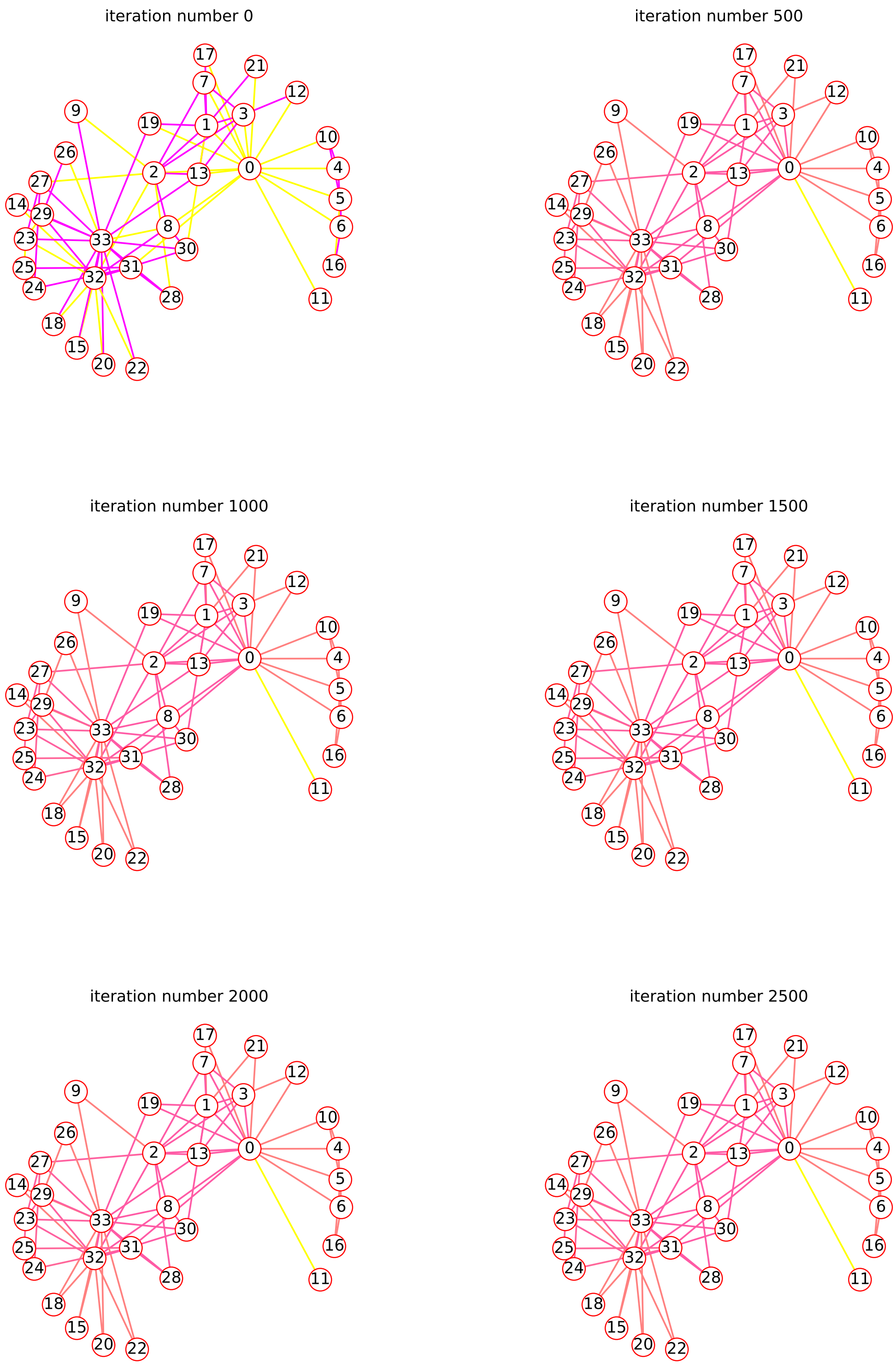


Figure 2.19: Plus-1 on the Zachary Karate Club graph.

To show the convergence of this algorithm in solving the spanning tree modulus problem, first consider the sequence, $\{\mu_k\}$, of pmfs and corresponding sequence of edge probabilities, $\{\eta_k\}$, formed by this process:

$$\mu_k = \frac{1}{k}\sum_{\ell=1}^{k}\delta_{\gamma_\ell}, \qquad \eta_k = \mathcal{N}^T\mu_k = \frac{w_k}{k}.$$

Since each γ_k for $k \geq 2$ is chosen to be minimal with respect to the edge weights w_{k-1}, it is also minimal with respect to the weights η_{k-1}. The relationship between successive iterations is given by

$$\mu_k = \frac{1}{k}\delta_{\gamma_k} + \frac{k-1}{k}\mu_{k-1}, \qquad \eta_k = \frac{1}{k}\mathbb{1}_{\gamma_k} + \frac{k-1}{k}\eta_{k-1}. \tag{2.8.1}$$

This recurrence relation will prove to be very useful in showing convergence.

Theorem 2.8.5. *Given a connected graph $G(V,E)$, let $\eta_k = \frac{\omega_k}{k}$ be given by the Plus-1 algorithm and let η^* be the minimizer of the optimization problem (2.4.3), then*

$$\|\eta_k - \eta^*\|^2 \to 0 \ \ \textit{as}\ k \to \infty$$

moreover, the rate of convergence is $O\left(\frac{1}{k}\right)$.

Proof. Let μ_k and η_k be defined as in (2.8.1) Then, for $k = 2, 3, \ldots$,

$$\begin{aligned}\|\eta_k - \eta^*\|^2 &= \|\eta_k\|^2 - 2\eta_k^T\eta^* + \|\eta^*\|^2 \\ &= \left(\frac{1}{k}\right)^2(|V|-1) + \frac{2}{k}\left(1-\frac{1}{k}\right)\mathbb{1}_{\gamma_k}^T\eta_{k-1} + \left(1-\frac{1}{k}\right)^2\|\eta_{k-1}\|^2 \\ &\quad - \frac{2}{k}\mathbb{1}_{\gamma_k}^T\eta^* - 2\left(1-\frac{1}{k}\right)\eta_{k-1}^T\eta^* + \|\eta^*\|^2.\end{aligned} \tag{2.8.2}$$

Note that the convex combination

$$\begin{aligned}\sum_{\gamma\in\Gamma}\mathbb{1}_\gamma^T\eta_{k-1}\mu^*(\gamma) &= \sum_{\gamma\in\Gamma}\sum_{e\in E}\mathbb{1}_\gamma(e)\eta_{k-1}(e)\mu^*(\gamma) \\ &= \sum_{e\in E}\eta_{k-1}(e)\sum_{\gamma\in\Gamma}\mathbb{1}_\gamma(e)\mu^*(\gamma) \\ &= \sum_{e\in E}\eta_{k-1}(e)\eta^*(e) = \eta_{k-1}^T\eta^*.\end{aligned}$$

Since any convex combination of a set of numbers is bounded below by the minimum of the set, and since γ_k is chosen so that

$$\mathbb{1}_{\gamma_k}^T \eta_{k-1} = \min_{\gamma \in \Gamma} \mathbb{1}_{\gamma}^T \eta_{k-1},$$

it follows that

$$\mathbb{1}_{\gamma_k}^T \eta_{k-1} \leq \eta_{k-1}^T \eta^*.$$

Substituting this into (2.8.2) and rearranging the terms shows that

$$\begin{aligned}
\|\eta_k - \eta^*\|^2 \leq \left(1 - \frac{1}{k}\right)^2 \|\eta_{k-1}\|^2 - 2\left(1 - \frac{1}{k}\right)^2 \eta_{k-1}^T \eta^* + \left(1 - \frac{1}{k}\right)^2 \|\eta^*\|^2 \\
+ \left(\frac{1}{k}\right)^2 (|V| - 1) - \frac{2}{k} \mathbb{1}_{\gamma_k}^T \eta^* + \left(\frac{2}{k} - \left(\frac{1}{k}\right)^2\right) \|\eta^*\|^2 \\
= \left(1 - \frac{1}{k}\right)^2 \|\eta_{k-1} - \eta^*\|^2 + \frac{2}{k}(\|\eta^*\|^2 - \mathbb{1}_{\gamma_k}^T \eta^*) + \left(\frac{1}{k}\right)^2 (|V| - 1 - \|\eta^*\|^2).
\end{aligned}$$

Now, since $\rho^* = \mathrm{Mod}(\Gamma)\eta^* = \|\eta^*\|^{-2}\eta^*$ is admissible for the family of spanning trees, it follows that $\mathbb{1}_{\gamma_k}^T \rho^* = w_{\rho^*}(\gamma_k) \geq 1$. So,

$$\mathbb{1}_{\gamma_k}^T \eta^* = \|\eta^*\|^2 \mathbb{1}_{\gamma_k}^T \rho^* \geq \|\eta^*\|^2,$$

and so

$$\|\eta_k - \eta^*\|^2 \leq \left(1 - \frac{1}{k}\right)^2 \|\eta_{k-1} - \eta^*\|^2 + \left(\frac{1}{k}\right)^2 (|V| - 1 - \|\eta^*\|^2).$$

By induction, it can be shown that for $k \geq 2$

$$\|\eta_k - \eta^*\|^2 \leq \prod_{\ell=2}^{k} \left(1 - \frac{1}{\ell}\right)^2 \|\eta_1 - \eta^*\|^2 + \sum_{j=2}^{k} \left(\frac{1}{j}\right)^2 \prod_{\ell=j+1}^{k} \left(1 - \frac{1}{\ell}\right)^2 (|V| - 1 - \|\eta^*\|^2). \quad (2.8.3)$$

Also by induction, it can be shown that for $k \geq 2$

$$\prod_{\ell=2}^{k} \left(\frac{\ell - 1}{\ell}\right)^2 = \frac{1}{k^2}$$

and

$$\sum_{j=2}^{k} \frac{1}{j^2} \prod_{\ell=j+1}^{k} \left(\frac{\ell-1}{\ell}\right)^2 = \sum_{j=2}^{k} \frac{1}{k^2} = \frac{k-1}{k^2}.$$

Therefore,

$$\left\| \frac{w_k}{k} - \eta^* \right\|^2 = \|\eta_k - \eta^*\|^2 \leq \frac{1}{k^2}\|\eta_1 - \eta^*\|^2 + \frac{k-1}{k^2}(|V| - 1 - \|\eta^*\|^2) = O\left(\frac{1}{k}\right).$$

□

This establishes an efficient way of computing spanning tree modulus on any graph $G = (V, E)$, and also establishes a convergence rate that is not currently known for the basic algorithm, Algorithm 1. The next chapter will show what spanning tree modulus can say about the global and local structure of a graph.

Chapter 3

Graph Structure

With recent advancements in network science, the development and understanding of the structure of a graph has been improved through the concepts of community structure, or clustering. Such clusters or communities can be thought of as fairly independent compartments, relying heavily on the information passed through the community, and less from other communities. Community detection plays a large role in sociology, biology, and computer science, where systems are generally modeled by graphs.

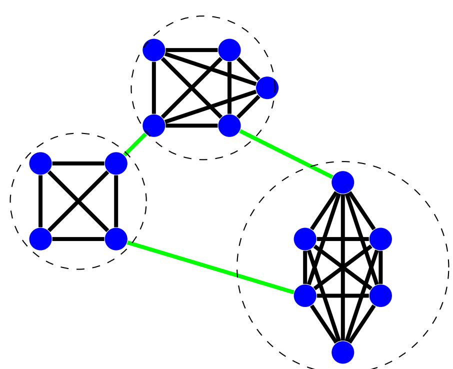


Figure 3.1: Simple graph with three communities

3.1 Community Detection

Historically there have been several methods for detecting community structure in a graph, however, there is still no standard definition of what constitutes a community. In [14], Fortunato provides a thorough review of known techniques (at the time). Here, I will give a brief overview of the main methods used historically, and how the processes relate to what is obtained through spanning tree modulus.

In [29], an approximation algorithm is used to determine the optimal size of the communities based on the use of a conductance measure. The best possible community size is then determined by using a profile plot the ranges over a wide range of community sizes. It is also shown in this paper, that many community detection algorithms may work on small scale graphs but then give a completely different result on large scale graphs such as social networks.

3.1.1 Graph Partitioning

The graph partitioning problem consists of dividing the vertices into k groups (often with constraints on the their sizes), such that the number of edges lying between the groups is minimal in some sense. However, this is an unusual assumption that one would have *a priori* knowledge of the size of the communities at the start. A common saying in biology is that structure determines function, not the other way around.

A common problem in this area is called the *minimum k-way cut problem* or the *k-terminal cut problem*, in which one looks to minimize the number of edges $E' \in E$ such that the removal of the edges in E' would disconnect the graph G into at least k connected components. A deterministic algorithm developed in [23], based on a divide and conquer scheme, and is shown to run in $O(n^{4k/(1-1.71/\sqrt{k})-16})$.

In [26], it is shown that if the number of edges allowed to be removed is bounded by s is fixed parameter tractable, where as the original k-way cut problem is shown to be NP-hard [12]. A similar problem, is that of the network reliability problem, in which each edge of a graph G is assumed to fail independently with some probability, and to determine the probability that the surviving network is connected. In [24], a fully polynomial randomized approximation scheme is given to approximate the failure probability.

3.1.2 Hierarchical Clustering

In many real-world graph, it may not be clear a priori the number of clusters into which the graph should be split. Moreover, community inclusion may not be exclusive; a node may belong to multiple communities. In these cases, methods like graph partitioning are of little use. However, the graph may have a clear hierarchical structure, that is, highly connected groups of nodes within more loosely connected, larger clusters. Hierarchical clustering is very common in social networks, biology, engineering, etc.

Similar to other methods, hierarchical clustering algorithms start with a similarity measure between vertices, once the measure is chosen, one can compute the similarity matrix, an $n \times n$ matrix where the (i, j) entry is the similarity measure between node i and node j. Hierarchical clustering techniques group vertices with high similarity measure.

As with any method, there are certain advantages and disadvantages. One advantage of hierarchical clustering is it does not require any assumptions on the number or size of clusters. However, it does not provide a way to discern a difference between the partitions obtained, and whether the community structure is a better representation of the graph. Another problem observed with these methods, is that vertices are often mis-classified putting periphery nodes into a community when in reality they may not be an important factor for information spread, particularly in the case where a vertex has just one neighbor.

Hierarchical clustering techniques are classified into two categories:

- *Agglomerative algorithms*
- *Divisive algorithms*

Agglomerative algorithms merge clusters if they have a high similarity(low dissimilarity) measure, where as divisive algorithms split clusters by removing edges connecting clusters with low similarity (high dissimilarity).

As will be shown in the following sections, through the use of spanning tree modulus, and the optimal value of η^* on each edge, one can use this as a similarity measure to group nodes into communities based on having identical expected edge usage on the edge connecting them.

This will however put nodes into multiple communities when not every adjacent edge has the same expected usage. In the light of this, spanning tree modulus is closer to solving the problems of dealing with this fact that there may be important nodes used for communicating information between communities.

This type of graph structure is notably different than another type of graph structure, named core-periphery.

3.2 Core-Periphery

Graphs are said to exhibit a *core-periphery* structure if there is a densely connected "core" of vertices and sparsely connected "peripheral" vertices. Core vertices tend to be connected to peripheral vertices, where as peripheral vertices tend not to be connected to other peripheral vertices.

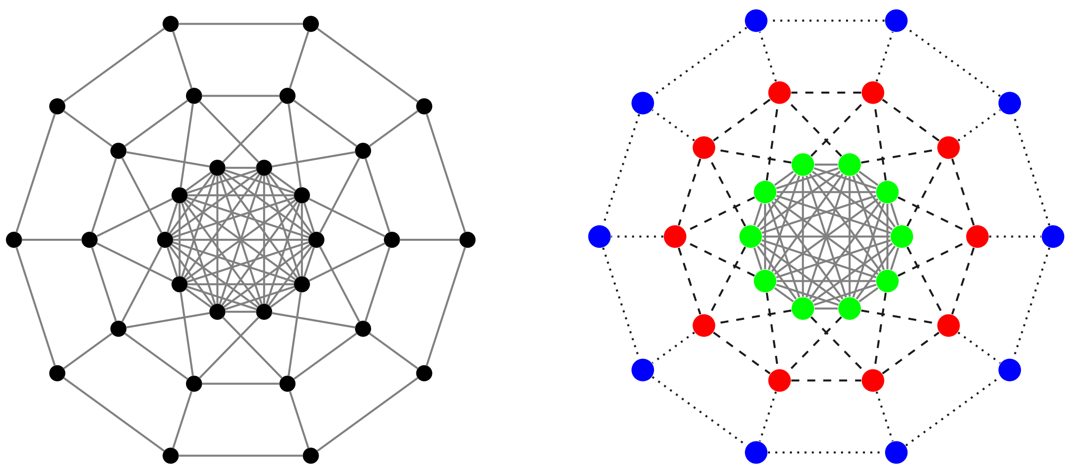


Figure 3.2: Core-Periphery structure

This definition has evolved over time. Originally introduced in [8], it was assumed that the periphery had no internal edges. However, with the development of onion networks [36, 41] the periphery nodes do connect to each other. Defining what the core of a graph is has also been the subject of recent research. These can vary from cliques (complete graphs), k-cliques, k-component, LS-set, etc. A nice table defining all the different types of core definitions used is given in [11]. Also included in [11] is a well detailed history on the core-periphery problem.

The concept of network core and periphery are not simply a mathematical construct created for ease of defining certain graphs, but emerged from several different fields, for example social networks, scientific citation networks, or economic models. [35, 42, 44]

3.3 Spanning Tree Modulus as a Tool for Community Detection/Graph Structure

To use spanning tree modulus as a tool for determining the structure of a graph, just as with previous methods mentioned, one must define a similarity measure, or in this case, a dissimilarity measure, to determine how to group nodes into clusters.

Definition 3.3.1. A *community* or *core* is a maximal, vertex-induced, connected subgraph G' of G such that η^* is constant on all edges in G'.

Notice that this alleviates the need to distinguish between reducible and irreducible homogeneous graphs. If a core happens to be reducible, that is, a smaller core inside, the community does not see this. In other words, with this definition, spanning tree modulus can be seen as a hierarchical method as there may be underlying communities that are being passed over. With this definition in place, one can create an iterative method to successfully detect, and remove, successively weaker communities or cores of a graph until all have been utilized. This leads to an ability to completely determine the set of optimal pmfs on the spanning trees of a graph, completely in terms of the optimal pmfs obtained on each core.

3.3.1 Deflation

The process of deflation was described in detail in [3]. This will provide more of a heuristic guide of the process and visualization through an example. It is important, however, to note the important definitions and theorems that make deflation a technique for community detection. The reader is encouraged to return to Section 2.7 and review the definitions of homogeneity, as deflation relies heavily on the idea of homogeneous components or *homogeneous cores*.

The first step to understanding deflation and the homogeneous core is to define *fair trees* and *forbidden trees*. A *fair tree* is defined as a spanning tree $\gamma \in \Gamma$ such that there exists at least one optimal pmf $\mu^* \in \mathcal{P}(\Gamma)$ with the property that $\mu^*(\gamma) > 0$. In other words, γ is in the support of at least one optimal pmf μ^*. A tree that is not in the support of any optimal pmf μ^* is then said to be a *forbidden tree*.

The existence of forbidden trees is a consequence of the restriction property. A subgraph H of G is said to have the *restriction property* if every tree in the support of any optimal μ^* is required to restrict to a spanning tree of H. In particular, G itself has the restriction property. The following theorem shows that, if G is not homogeneous, then there is a strict, nontrivial subgraph H of G that has the restriction property.

Theorem 3.3.2. *[3, Theorem 5.3] Let $G = (V_G, E_G)$ be a connected multigraph containing at least one edge. Let η^* be the optimal density for $\mathrm{Mod}_2(\widehat{\Gamma_G})$. Let $H_{min} \subset E$ be the subgraph where η^* attains its minimum. Let H be a connected component of $H_{\min}$. Then, the following hold:*

1. *H has at least one edge.*

2. *H has the restriction property.*

3. *H is a vertex-induced subgraph of G.*

4. *H is itself a homogeneous graph.*

Corollary 3.3.3. *If G is nonhomogeneous and biconnected, then the set of forbidden trees of G is nonempty.*

Proof. Consider the fact that since, by assumption, G is nonhomogeneous, then necessarily $\eta^*_{\max} > \eta^*_{\min}$. Consider the set of edges $E' \subset E$ that attain this minimum. Let $\gamma \in \Gamma$ be any spanning tree that restricts to a spanning tree on the connected, homogeneous component H guaranteed by Theorem 3.3.2. Consider two nodes, x and y, on the boundary of H. Since γ is spanning tree on H there is a unique path along the spanning tree, strictly inside H, connecting x and y. Also, consider a node (that must exist because of nonhomogeneity) z not in H, then because γ is also a spanning tree on G, there is a path from x and y to z along the spanning tree, that must use edges outside of H.

The biconnectedness of G guarantees there is at least two edge disjoint paths from any two distinct pairs of nodes. Consider the edge $e \in E'$ that creates a cycle between the nodes x, y, and z (with possibly other nodes involved in this cycle). Removing an edge from inside H that is along the path between x and y on γ then creates a forest on H. However, this is still a tree on G, and therefore must be a forbidden tree. □

Corollary 3.3.4. *If G is nonhomogeneous and biconnected, then G is not uniform.*

Proof. This follows directly from Corollary 3.3.3, since the set of forbidden trees is non-empty. □

Definition 3.3.5. Let H be a connected subgraph of G, satisfying properties 1–4 in Theorem 3.3.2. Such a subgraph is called a *homogeneous core* of G.

As shown in the following example, homogeneous cores are not necessarily unique.

Example 3.3.6. Consider the graph in figure 3.3a, colloquially referred to as the "house graph". This graph satisfies properties 1-4 in Theorem 3.3.2, but so does the "roof", illustrated in figure 3.3b. This graph has 11 spanning trees, and it can be shown that 9 of the trees are fair trees. The 2 forbidden trees, figure 3.4, do not restrict to spanning trees of the "roof" as they are only using one edge from it.

(a) House Graph

(b) Roof of house

Figure 3.3: Illustrating non-uniqueness of homogeneous cores

Figure 3.4: Forbidden trees on the house

The heuristic process of deflation is as follows: Given a graph G, let H be a homogeneous core, then collapse all the nodes of H into one node and remove all edges of H. The remaining graph, G/H, has $|V| - |V_H| + 1$ nodes and $|E| - |E_H|$ edges. Moreover, [3, Theorem 5.8] shows that the 2-modulus problem on G can be decoupled into two 2-modulus problems, that is

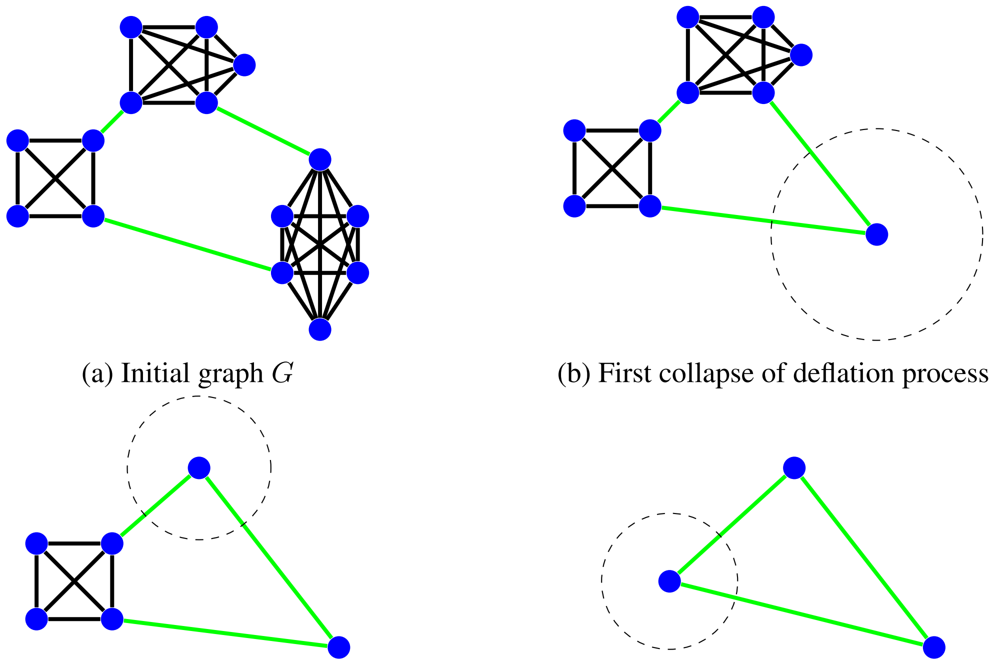
$$\begin{aligned} \mathrm{Mod}_2(\Gamma_G) &= \mathrm{Mod}_2(\Gamma_H) + \mathrm{Mod}_2(\Gamma_{G/H}) \\ &= \frac{(|V_H| - 1)^2}{|E_H|} + \mathrm{Mod}_2(\Gamma_{G/H}) \end{aligned}$$

where the second equality comes from the fact that H is homogeneous. The second is just another modulus problem and the process can be repeated until there is only a singleton node left.

Since each subgraph satisfies the restriction property at each step, one can work backwards creating to create an optimal pmf for $\mathrm{Mod}_2(\Gamma)$ by simply using the optimal pmfs for each subgraph. The deflation process is illustrated on the graph in figure 3.1 in the following example.

Example 3.3.7. Referring to figure 3.1, one can see the three communities present. Close examination shows that these are $K_4, K_5,$ and K_6, complete graphs on 4, 5, and 6 vertices resp. By equation, these are homogeneous graphs, and when viewed as subgraphs of the entire graph, are homogeneous cores. The order in which to proceed with deflation doesn't matter, however, in practice it seems to be natural to start with the homogeneous core with smallest expected edge usage. Shrinking this first core, gives the graph in figure 3.5b. Notice the other two cores are unaffected by this collapse, and hence are still homogeneous cores of the remaining graph. The process is repeated two more times, figures 3.5c and 3.5d, until finally the shrunk graph is K_3 which is the last step since it is itself homogeneous.

To visualize an optimal pmf for G, in this case μ_0 will work for every subgraph since complete graphs are also uniform. An alternative view point is that since each community is homogeneous one can build a spanning tree on each community and simply "glue" the trees together to create fair trees on G



(a) Initial graph G (b) First collapse of deflation process

(c) Second collapse of deflation (d) Third collapse of deflation

Figure 3.5: Deflation process

Typically real world graphs will not have the clear cut community structure as the graph in the previous example. Often times in practice, one will use a Erdős Rényi random graph to emulate a real world example. It is important to note here, unlike historic community detection models, no *a priori* knowledge of the number or size of the communities is needed. This would suggest, that possibly spanning tree modulus may be an efficient community detection method for graphs.

A typical graph used for comparison and validity of community detection algorithms is that of the Zachary Karate club, in figure 3.6, the graph is colored based on the optimal expected edge usage η^* computed in Section 2.8.2 with 3000 iterations.

Example 3.3.8. Recall from Example 2.8.4, the Zachary Karate club graph takes on 5 distinct values of η^*. The edges in figure 3.6 are colored based on these values of η^* which come from the set $\{.353, .375, .4, .5, 1\}$.

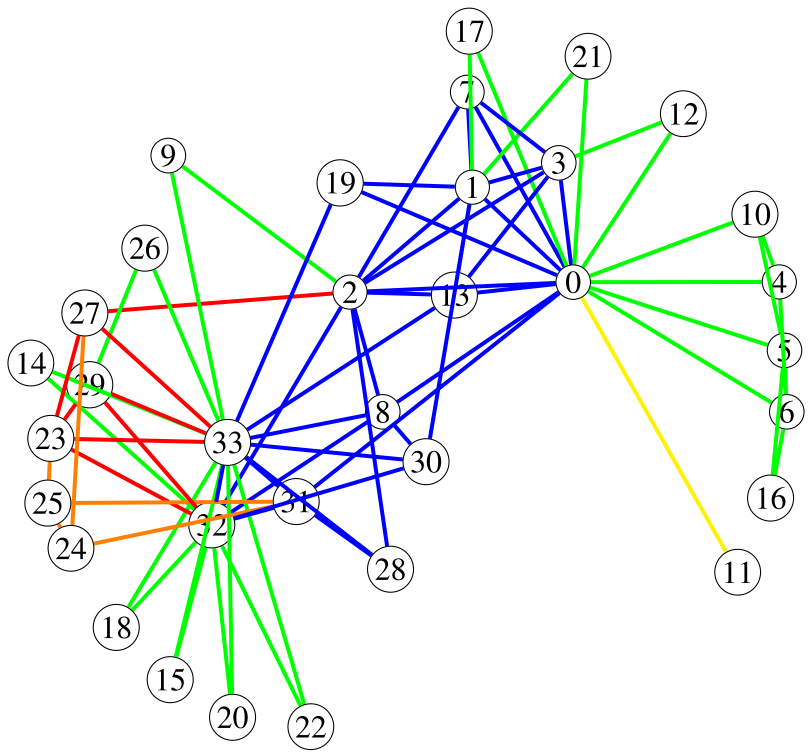


Figure 3.6: Zachary Karate club

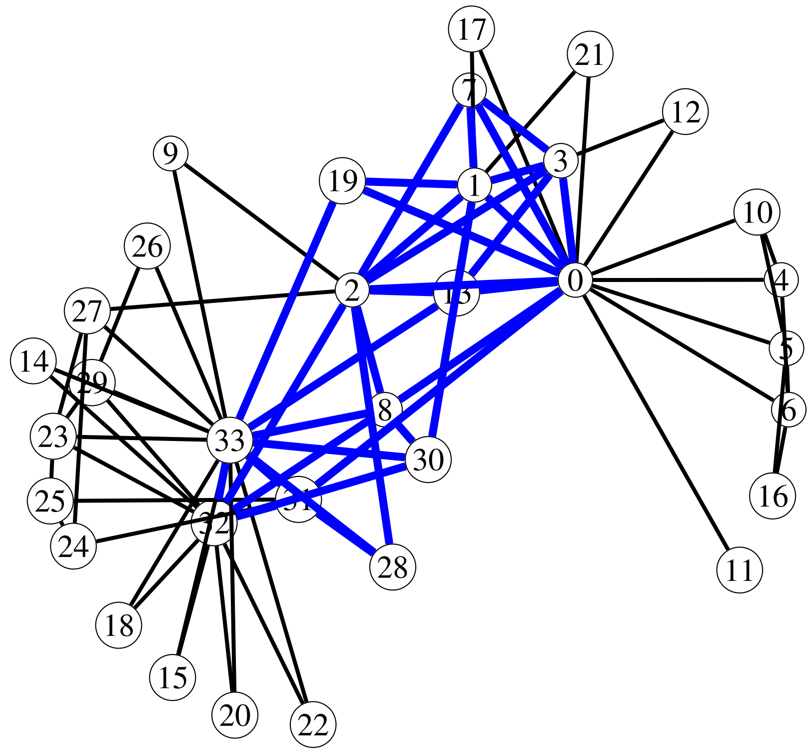


Figure 3.7: First component in deflation process for Zachary Karate club graph

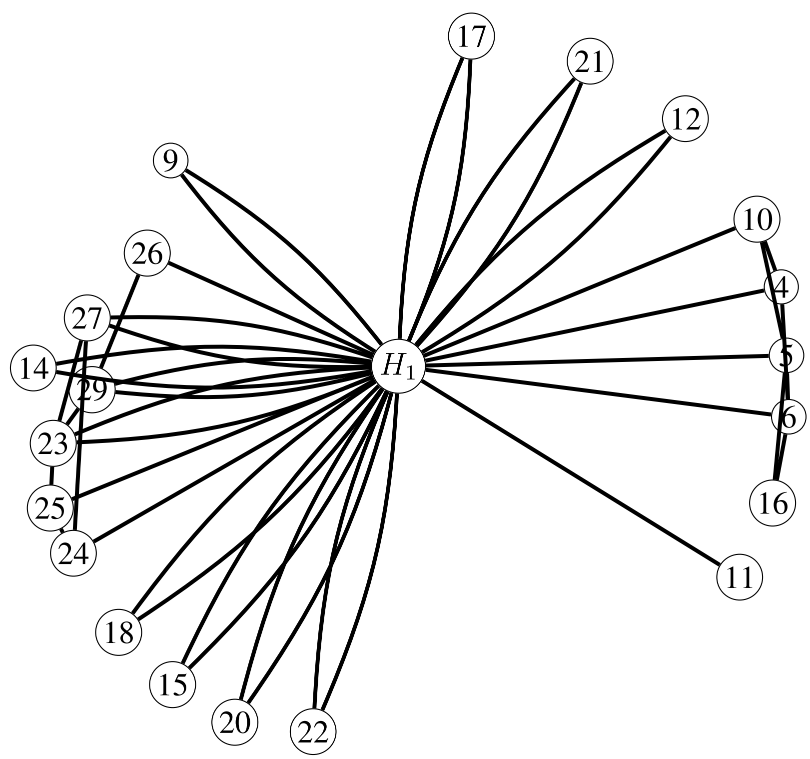


Figure 3.8: Resulting shrunk graph after first component deflated

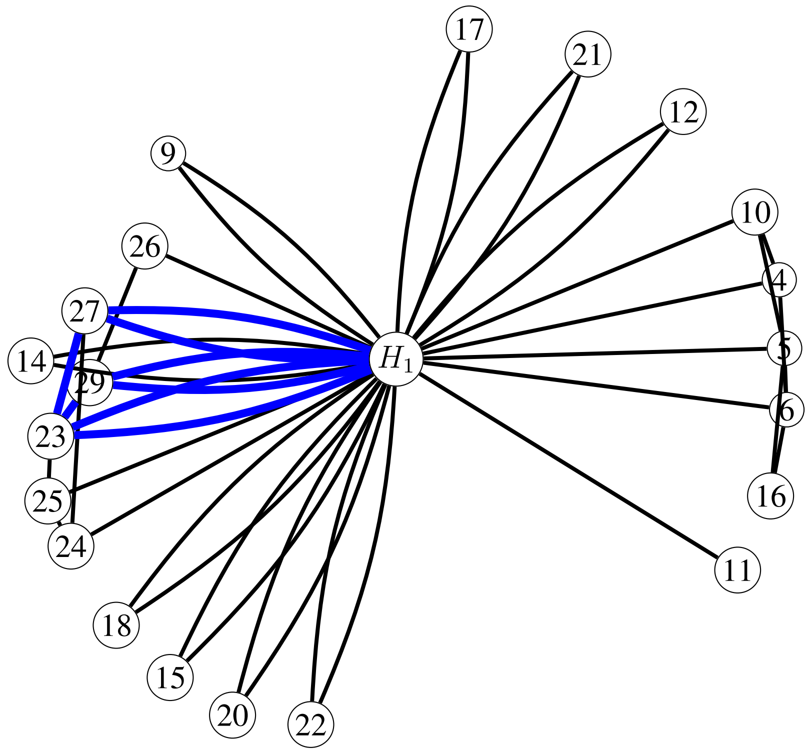


Figure 3.9: Second component in deflation

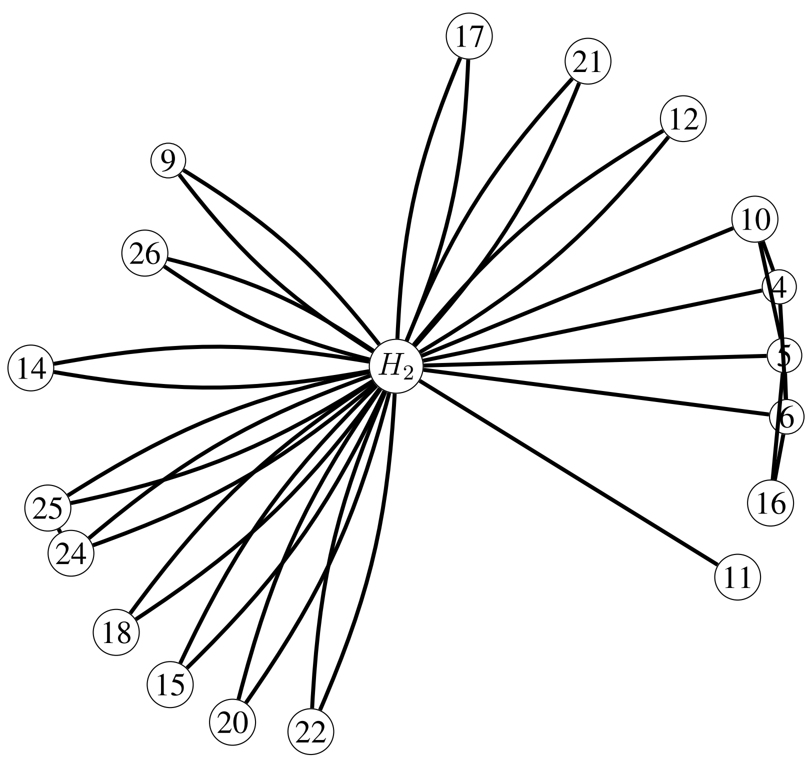


Figure 3.10: Resulting Shrunk Graph

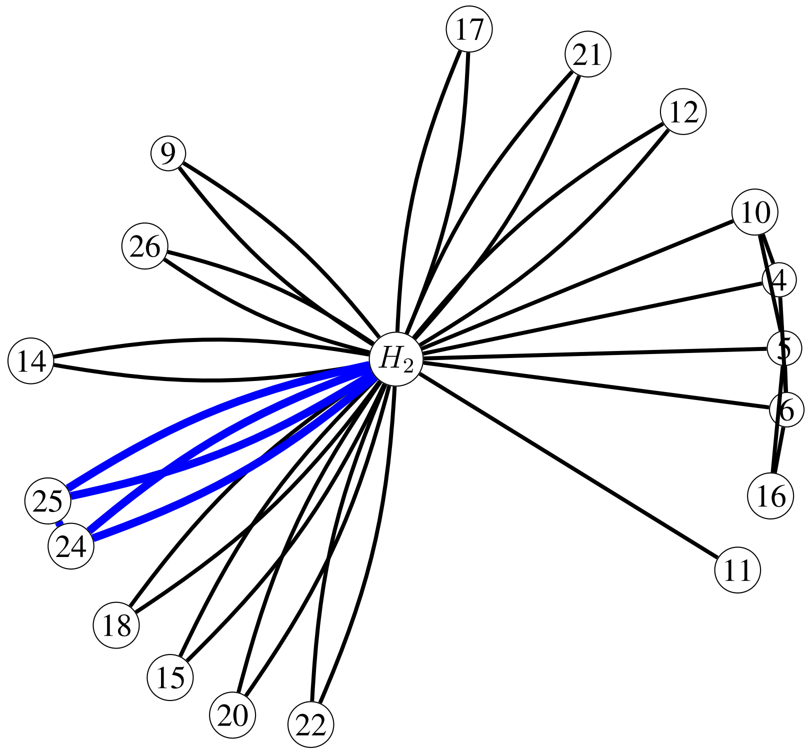


Figure 3.11: Third component in deflation

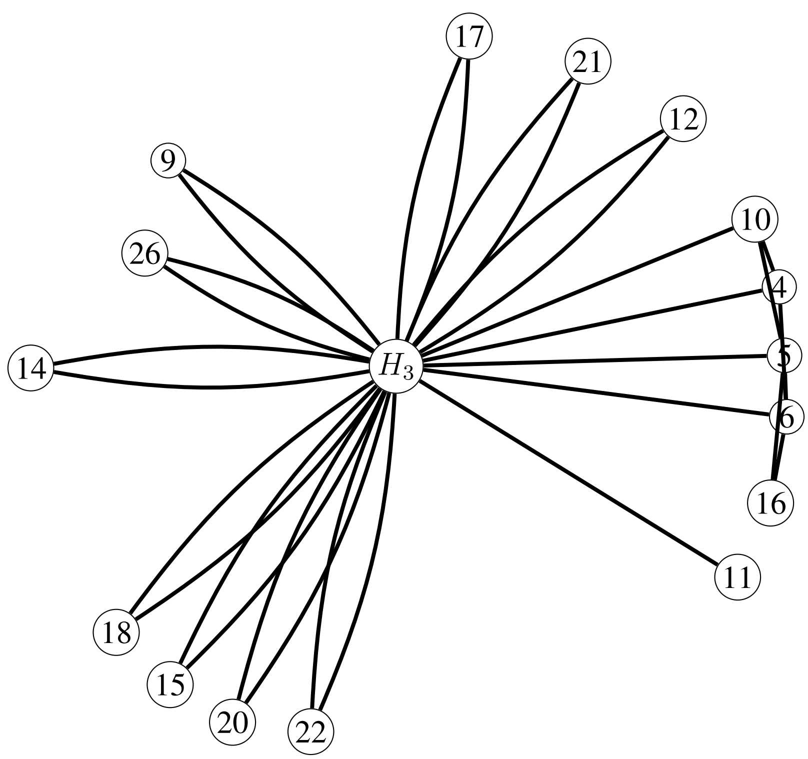


Figure 3.12: Resulting shrunk graph

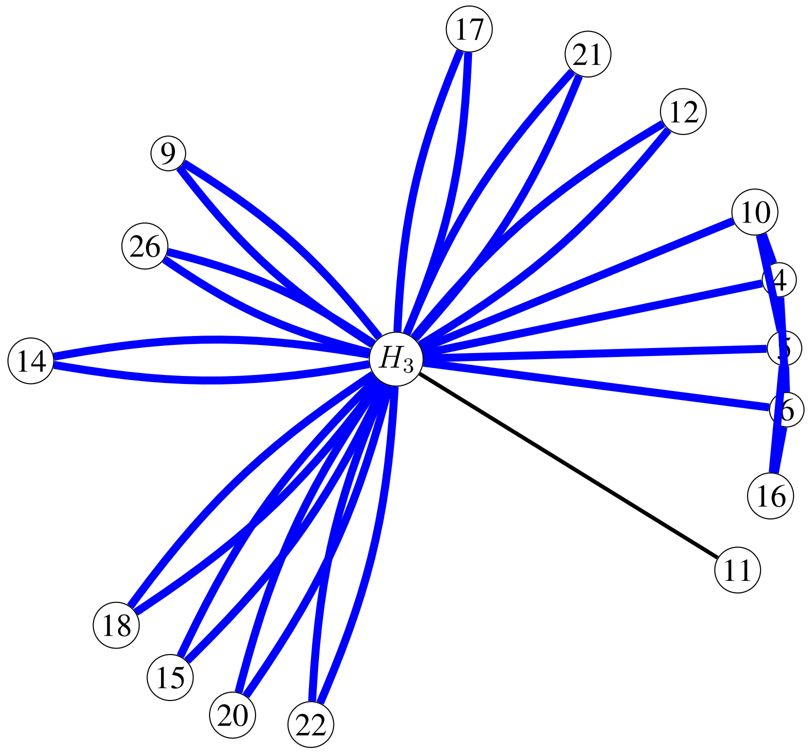


Figure 3.13: Fourth component in deflation

Remark: Note that this is not saying that although two communities may share the same value of η^*, they are in fact two separate communities if there is other edges between them with different values of η^*. All this means is that locally, the communities "share information" in a similar manner.

Looking at the process of deflation, one can see that this graph has a core-periphery structure, in the sense that every component is being shrunk into the original core, and there are no other isolated homogeneous cores. When comparing this to core-periphery algorithms previously used, for example the method in [13], the results are strikingly similar. That is, in [13] they obtain 14 core nodes and 20 periphery nodes, where as with deflation, 13 core nodes are obtained (figure 3.7).

From the previous examples, it is easy to see that deflation should be classified as an agglomerative algorithm, as it iteratively merges components based on the value of η^* on the edges.

Example 3.3.9. In [43], an additional measure, called the capacity, was used to quantify the strength of each connection. Based on Section 2.5, one could consider this a weighted graph, and use the unweighted representation (Definition 2.5.2) to compare to the community structure approach. This graph is given in Figure 3.14. Again, each edge is colored based on the expected edge usage η^*, and it can be seen that the core differs slightly than with the uncapacitated graph in the previous example.

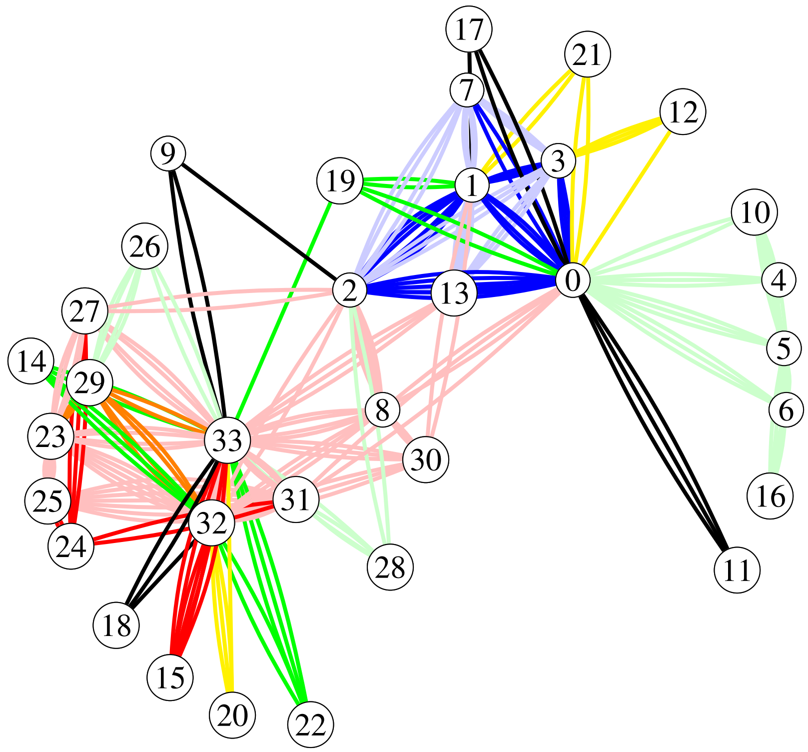


Figure 3.14: Capacitated Zachary Karate Club graph colored based on optimal values of η^*.

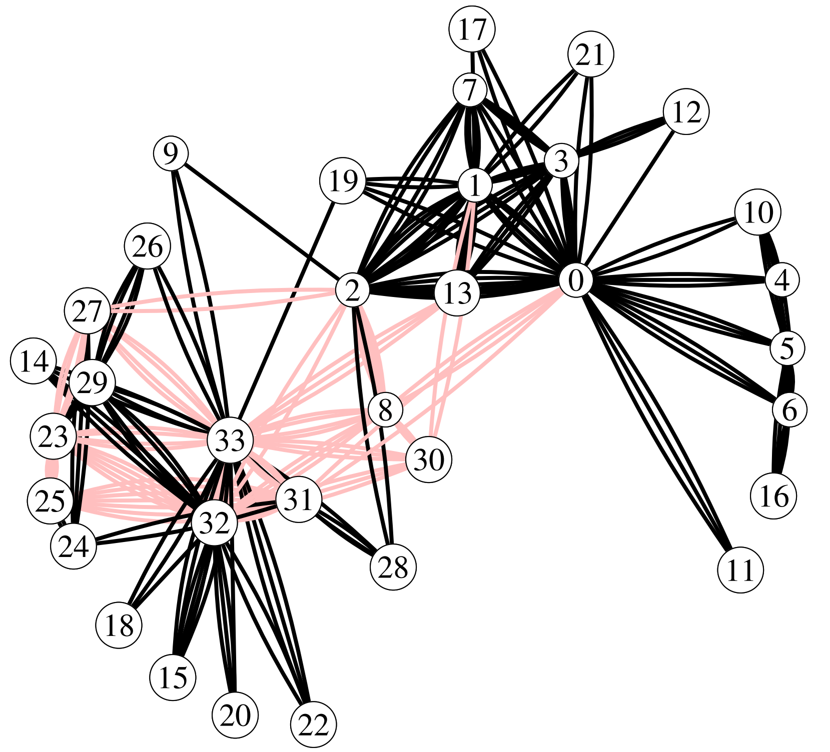


Figure 3.15: Capacitated Zachary Karate Club first component in deflation.

3.3.2 Minimal feasible partitions

Feasible partitions can also give an insight into the structure of a graph. Consider the fact that for homogeneous graphs, Theorem 2.7.2 says that every feasible partition must have weight greater than $\frac{1}{\eta^*}$. This gives minimum node and edge connectivity bounds stated in the next two lemmas.

Corollary 3.3.10. *Let $G = (V, E)$ be a homogeneous graph, Γ the family of spanning trees of G, and η^* be optimal for $\mathrm{Mod}_2(\Gamma)$, then for any node $v \in V$*

$$\deg(v) \geq \left\lceil \frac{1}{\eta^*} \right\rceil$$

Proof. This follows directly from Theorem 2.7.2, since removing all edges connected to v is a feasible partition, and the number of edges connected to v must be an integer. $\square$

More than just minimum node degree, feasible partitions say something about the edge connectivity of graph as well. The *edge-connectivity* of a graph G, denoted $\lambda(G)$, is size of the smallest edge cut.

Corollary 3.3.11. *Let $G = (V, E)$ be a homogeneous graph, Γ the family of spanning trees of G, and η^* be optimal for $\mathrm{Mod}_2(\Gamma)$, then*

$$\lambda(G) \geq \left\lceil \frac{1}{\eta^*} \right\rceil$$

The proof is straightforward again since all graph cuts are feasible partitions, and the number of edges in a cut is integer valued.

The question that comes up, is what happens in the case where G is not homogeneous? There is an analogous result to Corollary 3.3.11 that says that the communities or cores must be better connected than the graph as a whole.

To see this, it will be helpful to introduce the concept of a minimal feasible partition.

Definition 3.3.12. Given a graph G, a *minimum feasible partition*, P^*, is a feasible partition that satisfies

$$w(P^*) = \min\{w(P) : P \text{ is a feasible partition of } G\} \tag{3.3.1}$$

Similar to the non-uniqueness of homogeneous cores, minimum feasible partitions are also not unique. In other words, if a graph is reducible (Section 2.7.1) then the entire graph will be a minimum feasible partition (trivial partition P_0), as well as the trivial partition on the reduced graph when viewed as a partition on the original graph. Combining this with Theorem 2.7.5 gives another characterization for reducible graphs.

Corollary 3.3.13. *A homogeneous graph $G = (V, E)$ is reducible if and only if the minimum feasible partition is not unique.*

In general, just as with the deflation process, it does not matter which minimum feasible partition is used in each step, but typically a partition that maximizes $|E_P^*|$ is used. Since if not, the process will iterate until all these edges have been removed before finding a new feasible partition. The next theorem gives light on how to find such a minimum feasible partition based on the optimal solution η^* to the MEO problem.

Theorem 3.3.14 ([3], Theorem 4.10)**.** *Let G be a graph, and Γ the family of spanning trees on G. Let μ^* be optimal for $\mathrm{Mod}_2(\Gamma)$ and let $\eta^* = \mathcal{N}^T\mu^*$. Define*

$$E^* = \{e \in E : \eta^*(e) = \max_{e\in E} \eta^*(e)\}$$

Then there exists a minimum feasible partition P^* such that

1. $E^* = E_{P^*}$
2. $|\gamma^* \cap E_{P^*}| = k_{P^*} - 1 \qquad \forall \gamma^* \in$ *supp* μ^*
3. $\eta^*(e) = w(P^*)^{-1} \qquad \forall e \in E_{P^*}$

Theorem 3.3.15. *Let G be a graph, and let P^* be a minimum feasible partition, then for any partition P_i of the connected subgraphs $G_i(V_i)$ satisfies*

$$w(P_i) \geq w(P^*) \tag{3.3.2}$$

In particular, the weight of any minimum feasible partition, P_i^*, for any subgraph satisfies the inequality.

Proof. Let P^* be a minimum feasible partition of G, and let P_i be any feasible partition of G_i. Then $\tilde{P} := P^* \cup P_i$ is a feasible partition of G, with $|E_{\tilde{P}}| = |E_{P^*}| + |E_{P_i}|$ and $k_{\tilde{P}} = k_{P^*} + k_{P_i} - 1$. Moreover, since P^* is a minimum feasible partition of G, then

$$\frac{|E_{P^*}| + |E_{P_i}|}{(k_{P^*} - 1) + (k_{P_i} - 1)} \geq \frac{|E_{P^*}|}{k_{P^*} - 1}$$

which gives

$$\frac{|E_{P_i}|}{k_{P_i} - 1} \geq \frac{|E_{P^*}|}{k_{P^*} - 1}$$

□

Since graph cuts are feasible partitions of a graph, the following corollary gives a relationship to the weight of the minimum feasible partition and the connectivity of each subgraph.

Corollary 3.3.16. *Let G be a graph, and let P^* be a minimum feasible partition, then each of the connected subgraphs $G_i(V_i)$ satisfies*

$$\lambda(G_i) \geq w(P^*) \tag{3.3.3}$$

Proof. Let P^* be a minimum feasible partition of G. Without loss of generality consider $\lambda(G_1)$, and let P' be the feasible partition of G where $E_{P'} = E_{P^*} \cup E_{\lambda(G_1)}$. Then $|E_{P'}| = |E_{P^*}| + \lambda(G_i)$ and $k_{P'} = k_{P^*} + 1$, and since P^* is a minimum feasible partition of G

$$w(P') = \frac{|E_{P^*}| + \lambda(G_1)}{k_{P^*}} \geq \frac{|E_{P^*}|}{k_{P^*} - 1}.$$

Rearranging terms gives

$$(k_{P^*} - 1)\lambda(G_1) \geq k_{P^*}|E_{P^*}| - (k_{P^*} - 1)|E_{P^*}|,$$

a final rearrangement gives the desired result. □

Chapter 4

Conclusion

In this dissertation, it has been shown that for spanning tree modulus, one can treat weighted graphs as multi-graphs using an appropriate push-forward technique. The optimal solutions on the multi-graph representation can then be used to construct optimal solutions on the original graph.

In order to compute spanning tree modulus, a greedy algorithm was introduced and the convergence rate was given. However, to extend this result as a future interest, would be to determine the running time. That is, how many iterations would be needed to approximate the solution within some preset tolerance.

In Chapter 3, it was shown that spanning tree modulus is an effective tool for determining graph structure. Through either the use of the deflation process or minimum feasible partitions, one can construct optimal probability mass functions on the set of fair trees of smaller subgraphs to build a global solution on the graph as a whole.

The future uses of spanning tree modulus currently being explored are its applications to data analysis, in particular to aiding in the determination of key components of the data structure and what variables may be more significant than others.

Bibliography

[1] Lars Valerian Ahlfors. Collected papers. Vol. 1. Contemporary Mathematicians. Birkhäuser, Boston, Mass., 1982. 1929–1955, Edited with the assistance of Rae Michael Shortt.

[2] N. Albin, J. Clemens, N. Fernando, and P. Poggi-Corradini. Blocking duality for p-modulus on networks and applications. ArXiv e-prints, December 2016.

[3] N. Albin, J. Clemens, D. Hoare, P. Poggi-Corradini, B. Sit, and S. Tymochko. Fairest edge usage and minimum expected overlap for random spanning trees. 2018. Preprint.

[4] N. Albin and P. Poggi-Corradini. Minimal subfamilies and the probabilistic interpretation for modulus on graphs. Journal of Analysis, pages 1–26, 2016. arXiv:1605.08462.

[5] Nathan Albin, Megan Brunner, Roberto Perez, Pietro Poggi-Corradini, and Natalie Wiens. Modulus on graphs as a generalization of standard graph theoretic quantities. Conformal Geometry and Dynamics, 19(13):298–317, 2015. arXiv:1504.02418.

[6] Nathan Albin, Pietro Poggi-Corradini, Faryad Darabi Sahneh, and Max Goering. Modulus of families of walks on graphs. In Proceedings of Complex Analysis and Dynamical Systems VII, to appear. arXiv:1401.7640.

[7] Arne Beurling. The collected works of Arne Beurling. Vol. 1. Contemporary Mathematicians. Birkhäuser Boston, Inc., Boston, MA, 1989. Complex analysis, Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer.

[8] Stephen P Borgatti and Martin G Everett. Models of core/periphery structures. Social networks, 21(4):375–395, 2000.

[9] Bernard Chazelle. A minimum spanning tree algorithm with inverse-ackermann type complexity. Journal of the ACM (JACM), 47(6):1028–1047, 2000.

[10] Sunil Chopra. On the spanning tree polyhedron. Operations Research Letters, 8(1):25–29, February 1989.

[11] Peter Csermely, AndrÃ¡s London, Ling-Yun Wu, and Brian Uzzi. Structure and dynamics of core/periphery networks. Journal of Complex Networks, 1(2):93–123, 2013.

[12] Elias Dahlhaus, David S Johnson, Christos H Papadimitriou, Paul D Seymour, and Mihalis Yannakakis. The complexity of multiway cuts. In Proceedings of the twenty-fourth annual ACM symposium on Theory of computing, pages 241–251. ACM, 1992.

[13] Fabio Della Rossa, Fabio Dercole, and Carlo Piccardi. Profiling core-periphery network structure by random walkers. Scientific reports, 3:1467, 2013.

[14] Santo Fortunato. Community detection in graphs. Physics reports, 486(3):75–174, 2010.

[15] Michael L Fredman and Robert Endre Tarjan. Fibonacci heaps and their uses in improved network optimization algorithms. Journal of the ACM (JACM), 34(3):596–615, 1987.

[16] Michael L Fredman and Dan E Willard. Trans-dichotomous algorithms for minimum spanning trees and shortest paths. Journal of Computer and System Sciences, 48(3):533–551, 1994.

[17] DR Fulkerson. Blocking polyhedra. Technical report, RAND CORP SANTA MONICA CA, 1968.

[18] Harold N Gabow, Zvi Galil, and Thomas H Spencer. Efficient implementation of graph algorithms using contraction. In Foundations of Computer Science, 1984. 25th Annual Symposium on, pages 347–357. IEEE, 1984.

[19] Donald Goldfarb and Ashok Idnani. A numerically stable dual method for solving strictly convex quadratic programs. Mathematical programming, 27(1):1–33, 1983.

[20] John C Gower and Gavin JS Ross. Minimum spanning trees and single linkage cluster analysis. Applied statistics, pages 54–64, 1969.

[21] Ronald L Graham and Pavol Hell. On the history of the minimum spanning tree problem. Annals of the History of Computing, 7(1):43–57, 1985.

[22] Michael Held and Richard M Karp. The traveling-salesman problem and minimum spanning trees: Part ii. Mathematical programming, 1(1):6–25, 1971.

[23] Yoko Kamidoi, Noriyoshi Yoshida, and Hiroshi Nagamochi. A deterministic algorithm for finding all minimum k-way cuts. SIAM Journal on Computing, 36(5):1329–1341, 2006.

[24] David R Karger. A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem. SIAM review, 43(3):499–522, 2001.

[25] David R Karger, Philip N Klein, and Robert E Tarjan. A randomized linear-time algorithm to find minimum spanning trees. Journal of the ACM (JACM), 42(2):321–328, 1995.

[26] Ken-ichi Kawarabayashi and Mikkel Thorup. The minimum k-way cut of bounded size is fixed-parameter tractable. In Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on, pages 160–169. IEEE, 2011.

[27] Joseph B Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. Proceedings of the American Mathematical society, 7(1):48–50, 1956.

[28] Sukhamay Kundu. Bounds on the number of disjoint spanning trees. Journal of Combinatorial Theory, Series B, 17(2):199–203, 1974.

[29] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. Internet Mathematics, 6(1):29–123, 2009.

[30] CSJA Nash-Williams. Edge-disjoint spanning trees of finite graphs. Journal of the London Mathematical Society, 1(1):445–450, 1961.

[31] Jaroslav Nešetřil, Eva Milková, and Helena Nešetřilová. Otakar Borůvka on minimum spanning tree problem translation of both the 1926 papers, comments, history. Discrete mathematics, 233(1-3):3–36, 2001.

[32] Alex K Obruča. Spanning tree manipulation and the travelling salesman problem. The Computer Journal, 10(4):374–377, 1968.

[33] Robert Clay Prim. Shortest connection networks and some generalizations. Bell Labs Technical Journal, 36(6):1389–1401, 1957.

[34] Paul Raff. Spanning trees in grid graphs. arXiv preprint arXiv:0809.2551, 2008.

[35] Puck Rombach, Mason A Porter, James H Fowler, and Peter J Mucha. Core-periphery structure in networks (revisited). SIAM Review, 59(3):619–646, 2017.

[36] Christian M Schneider, André A Moreira, José S Andrade, Shlomo Havlin, and Hans J Herrmann. Mitigation of malicious attacks on networks. Proceedings of the National Academy of Sciences, 108(10):3838–3841, 2011.

[37] H. Shakeri, P. Poggi-Corradini, N. Albin, and C. Scoglio. Network clustering and community detection using modulus of families of loops. to appear in Physical Review E, 2016. https://arxiv.org/pdf/1609.00461v2.pdf.

[38] Heman Shakeri, Nathan Albin, Faryad Darabi Sahneh, Pietro Poggi-Corradini, and Caterina Scoglio. Maximizing algebraic connectivity in interconnected networks. Physical Review E, 93:030301, Mar 2016. arXiv:1510.06785.

[39] Heman Shakeri, Pietro Poggi-Corradini, Caterina Scoglio, and Nathan Albin. Generalized network measures based on modulus of families of walks. Journal of Computational and Applied Mathematics, 307:307–318, 2016.

[40] Robert Endre Tarjan. Efficiency of a good but not linear set union algorithm. Journal of the ACM (JACM), 22(2):215–225, 1975.

[41] Zhi-Xi Wu and Petter Holme. Onion structure and network robustness. Physical Review E, 84(2):026106, 2011.

[42] Jaewon Yang and Jure Leskovec. Overlapping communities explain core–periphery organization of networks. Proceedings of the IEEE, 102(12):1892–1902, 2014.

[43] Wayne W Zachary. An information flow model for conflict and fission in small groups. Journal of anthropological research, 33(4):452–473, 1977.

[44] Xiao Zhang, Travis Martin, and Mark EJ Newman. Identification of core-periphery structure in networks. Physical Review E, 91(3):032803, 2015.