UNIVERSITY OF CALIFORNIA, SAN DIEGO

Stochastic Processes Arising from Graph Manipulations

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Jacob Taylor Hughes

Committee in charge:

Professor Fan Chung Graham, Chair
Professor Sanjoy Dasgupta
Professor Ron Graham
Professor Jeff Remmel
Professor Jaques Verstraete

2013

The dissertation of Jacob Taylor Hughes is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2013

EPIGRAPH

It's wanting to know that makes us matter.
-Tom Stoppard, *Arcadia*

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and guidance of many teachers and friends for whom I am deeply grateful. First and foremost, I would like to thank Fan Chung for her encouragement, inspiration, and patience throughout these past few years. Her role as an advisor has been crucial for my formation as a researcher, and I am deeply thankful for her guidance. I would also like to thank the other professors in the department, in particular Jacques Verstrate, Jeff Remmel, and Bruce Driver.

Thanks is also due to my undergraduate advisor, Frank Connolly, who introduced me to research mathematics and set me on the path the graduate school. Also thanks to the research advisors I had during summer REUs, Michael Gehktman and Colin Bleak. A special thanks to John Gorman, who went above and beyond to inspire in me a love of mathematics in high school. Thanks also to Sr. Dolores Doohan, who at a young age taught me the importance of perseverance.

A special thanks to my friends and and officemates for their constant help and support. In particular, I would like to thank Franklin Kenter, Andy Parrish, and Mary Radcliffe for their help brainstorming, organizing reading groups, and proofreading early versions of my work. Also thanks to Andy, Bob, Chris, David, Eric, Helen, Hooman, Janine, Jimmy, Katie, Mark, Mike, Lyla, Scott, and everyone else in our department; you all helped make our department a community, and for that I am eternally grateful.

Thanks to all my family for their constant love and support. The support of my parents and sisters throughout this process has been invaluable. Mom, Dad, Kayla, Alyssa; thanks for being you. Finally, I cannot begin to express the profound gratitude for Suz. I cannot imagine these last years without her love, support, patience, and kindness. Despite the distance and difficulties she has been my constant font of support and inspiration. Thank you, this would not have been possible without you.

Chapter 2 is based on the paper "A note on an alternating upper bound for random walks on semigroups" written jointly with Fan Chung [24], which has been submitted for publication to Discrete Applied Mathematics. The dissertation

author is the primary author of this work.

Chapter 3 is based on the paper "Random Seidel Switching on Graphs" [37]. It has been accepted for publication in the Journal of Combinatorial Mathematics and Combinatorial Computing.

Chapter 4 is based on the paper "Random Lights Out Processes on Graphs" which has been published in Advances in Applied Mathematics [38].

Chapter 5 is based on the paper "Random Nearest Neighbor Interchanges on Phylogenetic Trees" which is currently in preparation.

Chapter 6 is based on the paper Multi-commodity allocation for dynamic demands using PageRank vectors which was written jointly with Fan Chung and Paul Horn. It appeared in the Proceedings of the 9th Workshop on Algorithms and Models for the Web Graph [22]. The dissertation author is the primary author of this work.

VITA

2008	B. S. in Mathematics, University of Notre Dame
2010	M. A. in Mathematics, University of California, San Diego
2013	Ph. D. in Mathematics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Stochastic Processes Arising from Graph Manipulations

by

Jacob Taylor Hughes
Doctor of Philosophy in Mathematics

University of California San Diego, 2013

Professor Fan Chung Graham, Chair

In this thesis, we study a collection of stochastic properties arising from graph manipulations. These manipulations may change the structure of the graph itself, recolor the vertices, or alter other assigned properties. Our goal with each process is to use a combination of combinatorial, spectral, and algebraic methods to analyze the the stationary distribution, mixing times, and hitting times of these stochastic processes. In particular, we do the following:

- We consider the random process on a class of semigroup known Left Regular Bands induced my multiplying by randomly selected generators. Such a process converges to a stationary distribution, and we present a bound the convergence time which slightly improves previous results.

- We consider the random process induced by random Seidel switching on a graph, an operation that transforms a graph by inverting the adjacency relations at a vertex. By computing the spectrum of the state graph of this process we show the process converges and bound the mixing time. We also consider several generalizations of Seidel switching.

- We study a stochastic process arising from randomly playing the Lights Out game on an arbitrary finite graph. In this game, the vertices are colored black or white, and a toggle changes state of that vertex and all of its neighbors.

We obtain a bounds on the mixing time and average hitting time of the all black coloring.

- We investigate the random process on phylogenetic trees induced by nearest neighbor interchanges, an operation that swaps two neighboring sub-trees. We use a comparison theorem to bound the spectral gap of the state graph and use this to bound the mixing time of the random nearest neighbor interchange process.

- We introduce the multi-commodity dynamic demand model on a graph, a variant of the standard contact process on graphs. Each vertex has a list of demands for various commodities, and the demands of a vertex influence the demands of its neighbors. We introduce a generalization of the PageRank vector, Kronecker PageRank, and use it to bound the amount of commodities needed to satisfy the demand at the vertices.

Chapter 1

Preliminaries and Overview

1.1 Notation and Types of Graphs

Throughout this discussion we will use standard notation and terminology for graphs. A *graph* $G = G(V, E)$ consists of a set V of vertices and a set E of pairs of vertices. If the edges are unordered pairs we say that G is an *undirected graph* and write $v \sim w$ to denote $\{v, w\} \in E$, and say that v and w are *adjacent*. A loop is an edge from a vertex to itself, and the term *simple graph* is used to refer to an undirected graph with no loops.

If instead the edges are ordered pairs, we say that G is a *directed graph* and write $v \to w$ to denote that $(v, w) \in E$. A graph G is a *multigraph* if the set of edges is allowed to be a multi-set, that is multiple edges are allowed between a pair of vertices. We will consider both directed and undirected multigraphs. A *weighted graph* $G = G(V, \omega)$ consists of a set V of vertices and a non-negative weight function $\omega \colon V \times V \to \mathbb{R}_{\geq 0}$. The case of an unweighted graph is then simply the special case where $\omega(v, w) = \mathbf{1}_{\{v \sim w\}}$, the $0 - 1$ indicator function testing if v and w are adjacent. Throughout this dissertation, all graphs will be finite.

The *degree* of a vertex v, denoted d_v, is the number of vertices that are adjacent to v. Each edge of a multigraph contributes 1 to the degree, and any loops contribute 1. For a weighted graph, the degree is defined to be the sum of

the weights of edges beginning at v. That is,

$$d_v = \sum_{w \in V} \omega(v, w)$$

A graph is *k-regular* if the degree of each vertex is k.

A *walk* of length k on a graph is a sequence of vertices $(v_0, v_1, ..., v_k)$ where $(v_i, v_{i+1}) \in E$. A cycle of length k, also called a $k - cycle$, is a walk of length k starting and ending at the same vertex. A graph G is *connected* if for any two vertices v, w there is a walk from v to w. A directed graph is *strongly connected* if for any two vertices v, w there is a walk from v to w and a walk from w to v.

For two vertices v, w the *distance* between v and w is the minimum length of a walk between them. The *diameter* of a graph is the maximum distance between two vertices in the graph.

For a subset $S \subseteq V$, the *volume* of S is defined by $\mathrm{vol\ S} = \sum_{\mathrm{v \in S}} \mathrm{d_v}$. We often will write $\mathrm{vol\ G}$ for $\mathrm{vol\ V(G)}$. Note that for a simple graph, $\mathrm{vol\ G} = 2|\mathrm{E}|$. The *Cheeger ratio* of a subset S is the ratio

$$h_G(S) := \frac{\mathrm{vol}\ \partial \mathrm{S}}{\min\{\mathrm{vol\ S}, \mathrm{vol}\ \bar{\mathrm{S}}\}}$$

where $\bar{S}$ is the complement of S, and $\partial S := \{v \in V \setminus S | \exists w \in S, w \sim v\}$. The *Cheeger constant* of a graph G is

$$h_G := \max_{S \subset V} h_G(S)$$

A common way to compare two graphs is to define a function between the sets vertices that preserves the adjacency relations. A *graph homomorphism* between two graphs G, G' is a function $f \colon V(G) \to V(G')$ such that for all vertices $v, w \in V(G)$, if $v \sim w$ then $f(v) \sim f(w)$. A *graph isomorphism* between two graphs G, G' is a bijection $f \colon V(G) \to V(G')$ such that for all vertices $v, w \in V(G)$, $v \sim w$ if and only if $f(v) \sim f(w)$. Two graphs are called *isomorphic* is there exists a graph isomorphism between them. Isomorphic graphs are identical except for labelings of the vertices. A *graph automorphism* of G is a graph isomorphism from G to itself, or in other words a relabeling of the vertices. A graph G is *vertex transitive* if for any two vertices v, w there is a graph automorphism $f \colon V(G) \to V(G)$ such

that $f(v) = w$. Informally, a vertex transitive graph is a graph where vertices are indistinguishable except for labels.

1.2 Graph Spectra

For a graph G with n vertices, there are several $n \times n$ matrices associated with G. In this section, we begin by defining several of the most commonly used matrices, and then illustrate several properties of their spectra.

Definition 1.2.1.
The adjacency matrix of a simple graph is the symmetric matrix A with entries

$$A(v,w) = \begin{cases} 1 & \text{if } v \sim w \\ 0 & \text{if } v \not\sim w \end{cases}$$

The adjacency matrix of a weighted graph $G(V, \omega)$ has entries

$$A(v,w) = \omega(v,w)$$

Note that the adjacency matrix of a directed graph is not necessarily symmetric.

Definition 1.2.2.
The *diagonal degree matrix* is the diagonal matrix with entries

$$D(v,v) = d_v$$

Definition 1.2.3.
The combinatorial Lapliacian, L, is defined by

$$L = D - A$$

Definition 1.2.4.
The *normalized Laplacian*, $\mathcal{L}$, is defined by

$$\mathcal{L} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$$

and thus has entries

$$\mathcal{L}(v,w) = \begin{cases} 1 & \text{if } v = w \\ -\frac{1}{\sqrt{d_v d_w}} & \text{if } v \sim w \\ 0 & \text{otherwise} \end{cases}$$

The advantage of working with these matrices is often graph properties can be understood by studying the spectra of these matrices. While many classical spectral graph theory results primarily use the adjacency matrix A, many of these results only hold for regular graphs. The advantage of using the normalized laplacian $\mathcal{L}$ is the ability to extend these results to all matrices, see [20] for a full discussion of $\mathcal{L}$ and its advantages.

1.3 Markov Processes and Random Walks

We assume basic knowledge of random variables and probability, and begin this section reviewing the class of random variables that we will use.

Definition 1.3.1.
A sequence of random variables $X_1, X_2, X_3...$ is a *Markov chain* if

$$\mathbb{P}(X_{t+1} = x | X_1 = x_1, X_2 = x_2 ... X_t = x_t) = \mathbb{P}(X_{t+1} = x | X_t = x_t)$$

for all t.

A Markov chain is *time homogeneous* if

$$\mathbb{P}(X_{t+1} = x | X_t = y) = \mathbb{P}(X_t = x | X_{t-1} = y)$$

for all t. All Markov chains in this discussion will be time homogenous.

Suppose $X_0, X_1, X_2, X_3, ...$ is a time homogeneous Markov chain with finite state space. Let f_t be the distribution at time t, that is

$$f_t(x) = \mathbb{P}(X_t = x).$$

The *transition matrix* is the matrix P with entries

$$P(x, y) = \mathbb{P}(X_{t+1} = y | X_t = x).$$

Note that since the Markov chain is time homogenous this quantity does not depend on t. It is easy to see that P has the property that for any t,

$$f_{t+1} = f_t P.$$

As a consequence of this definition, if f is the distribution of X_0, then X_t has distribution fP^t.

A Markov chain is *reversible* if there is a distribution π such that

$$\pi(x)P(x,y) = \pi(y)P(y,x)$$

for all vertices x, y.

π is called the *stationary distribution* of the Markov chain. A Markov chain is *ergodic* if there is a unique stationary distribution π such that for any vertex v,

$$\lim_{t\to\infty} fP^t(v) = \pi(v)$$

for any initial distribution f.

Definition 1.3.2.

Let G be a graph. A *random walk on* G is a sequence of random vertices $\{X_t\}_{t=0}^{\infty}$ where the starting vertex X_0 is chosen according to an initial distribution, and

$$\mathbb{P}(x_{i+1} = v|x_i) = \begin{cases} 1/d_{x_i} & \text{if } x_i \sim v \\ 0 & \text{if } x_i \nsim v \end{cases}.$$

More generally, if G is a weighted, directed graph then

$$\mathbb{P}(x_{i+1} = v|x_i) = \begin{cases} 1/d_{x_i} & \text{if } x_i \to v \\ 0 & \text{if } x_i \not\to v \end{cases}.$$

We note that a random walk on a graph is time homogeneous Markov process.

Proposition 1.3.3.

Let G be a connected, finite graph, and let $\{X_t\}_{t=0}^{\infty}$ be a random walk on G. Then this Markov process has transition matrix

$$P = D^{-1}A$$

Proposition 1.3.4.

Let G be a connected, finite graph, and let $\{X_t\}_{t=0}^{\infty}$ be a random walk on G with initial distribution f. Then the distribution at time t is given by

$$f_t = fP^t$$

where f is viewed as a row vector. That is,

$$\mathbb{P}(X_t = v) = fP^t(v),$$

the vth entry of the row vector fP^t.

There is a very clean characterization of ergodic random walks on graphs, see for example [20, Section 1.5].

Proposition 1.3.5.
A random walk on a weighted graph G is ergodic if and only if

1. *G is irreducible. That is, for any v, w there is a time t such that $P^t(v, w) > 0$*
2. *G is aperiodic. That is, the greatest common factor of the set $\{t \mid \exists v, w \text{ s.t. } P^t(v, w) > 0\}$ is 1.*

Aperiodicity can be artificially imposed by considering a *lazy random walk*, which has transition matrix $P' = \frac{1}{2}(P + I)$. This can be thought of as either the original random walk where half the time no action is taken (hence the name). Alternately, one can view it as a random walk on a modified graph where each vertex is given a self loop with weight equal to the degree of the vertex. The choice of using the constant $\frac{1}{2}$ ensures the eigenvalues of P' fall in the range $[0, 1]$, though in some contexts it is more natural to use other values.

Proposition 1.3.6.
Let G be a (possibly weighted) connected graph, and let P be the transition matrix for a random walk on G. If D is the diagonal degree matrix, then

$$\mathbb{1}DP = \mathbb{1}D$$

where $\mathbb{1}$ is the all 1's vector. Therefore a random walk on G has stationary distribution

$$\pi(v) = \frac{d_v}{\text{vol G}}.$$

Corollary 1.3.7.
Let G be a k-regular graph on n vertices. Then a random walk on G has stationary distribution

$$\pi(v) = \frac{1}{n}$$

When studying a random walk (or any convergent process), we care not only about the stationary distribution, but also about how quickly the process is converging to this distribution. However, in order to express the speed of convergence, we need a way of measuring the "distance" from the stationary distribution. There are several commonly used metrics, and here we define three we will use in this dissertation.

Definition 1.3.8.
Let P be the transition matrix of an ergodic random walk on a finite graph G. Let π be the stationary distribution.

- The relative pointwise distance after t steps, $\Delta(t)$, is defined as

$$\Delta(t) := \max_{x,y} \frac{|P^t(y,x) - \pi(x)|}{\pi(x)}$$

- The $\chi-$squared distance after t steps, $\Delta'(t)$, is defined as

$$\Delta'(t) := \max_{x \in V(G)} \left(\sum_{y \in V(G)} \frac{(P^t(x,y) - \pi(y))^2}{\pi(y)} \right)^{1/2}.$$

- The total variation distance after t steps, $\Delta_{\mathrm{TV}}(t)$, is defined as

$$\Delta_{\mathrm{TV}}(t) := \max_{A \subset V} \max_{y \in V} \left| \sum_{x \in A} P^t(y,x) - \pi(x) \right|$$

A term commonly used to discus the speed of convergence is the mixing time, given by

$$\tau_{\mathrm{mix}} = \min\{t | \Delta(t) \leq e^{-1}\}.$$

The choice of e^{-1} as the cut off constant, while traditional, is chosen primarily for algebraic convenience (see [1] for a discussion of the result of choosing a different constant). Our main results will actually provide bounds on the time before $\Delta(t)$ is less than an arbitrary constant (or even a function of the number of vertices).

For vertex transitive graphs, we can bound the Δ' distance using a simple expression, sometimes known as the Plancherel Formula. This theorem will be

one our main analytic tools throughout this discussion, and we provide a proof to illustrate some of the main techniques used for proving spectral results.

Theorem 1.3.9.

Let G be a vertex transitive graph on n vertices, and let W be the transition matrix for an ergodic random walk on G. Let λ_i be the eigenvalues of W, with $-1 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n = 1$. Then the $\chi-$squared distance after s steps is given by

$$\Delta'(t) = \left(\sum_{i \neq n} \lambda_i^{2t} \right)^{1/2}.$$

Proof. Let Ψ_x be the characteristic function of x, and let ϕ_i denote the i-th orthonormal eigenfunction of P. Let $J_n = \phi_n^* \phi_n$ be the projection matrix of of projecting onto the eigenfunction ϕ_n. Since G is vertex transitive,

$$\pi(x) = \frac{1}{|G|},$$

$$\phi_n(x) = \frac{1}{|G|^{1/2}}$$

for all x. Therefore

$$\Psi_x J_n \Psi_y^* = \Psi_x \phi_n^* \phi_n \Psi_y^* = \phi_n(x)\phi_n(y) = \pi(y).$$

Also note that both P and J_n are symmetric. Since G is vertex transitive, the quantity

$$\sum_{y \in V(G)} \frac{\left(P^t(x,y) - \pi(x)\right)^2}{\pi(y)}$$

is equal for all $x \in V(G)$. Summing over all vertices x we obtain that

$$\Delta'(t)^2 = \sum_{x \in V(G)} \pi(x) \sum_{y \in V(G)} \frac{\left(P^t(x,y) - \pi(x)\right)^2}{\pi(y)}.$$

Therefore

$$\begin{aligned}
\Delta'(t)^2 &= \sum_{x\in V(G)} \sum_{y\in V(G)} (P^t(x,y) - \pi(x))^2 \\
&= \sum_{x\in V(G)} \sum_{y\in V(G)} \left(\Psi_x(P^t - J_n)\Psi_y^*\right)^2 \\
&= \sum_{x\in V(G)} < \Psi_x(P^t - J_n), \Psi_x(P^t - J_n) > \\
&= \sum_{x\in V(G)} \Psi_x(P^t - J_n)^2\Psi_x^* \\
&= \sum_{x\in V(G)} \Psi_x(P^{2t} - J_n)\Psi_x^*
\end{aligned}$$

Since $\Psi_x = \sum_i \phi_i(x)\phi(i)$, we can replace that in the above sum to obtain

$$\begin{aligned}
\Delta'(t)^2 &= \sum_{x\in V(G)} \left(\sum_i (\phi_i(x)\phi_i)(P^{2t} - J_n) \sum_j \phi_j(x)\phi_j^* \right) \\
&= \sum_{x\in V(G)} \sum_i \sum_j \phi_i(x)\phi_j(x) \left(\phi_i(P^{2t} - J_n)\phi_j^*\right) \\
&= \sum_{x\in V(G)} \sum_i \sum_j \phi_i(x)\phi_j(x) \left(\lambda_i^{2t}\phi_i\phi_j^* - \phi_i J_n \phi_j^*\right) \\
&= \sum_{x\in V(G)} \sum_i \sum_j \phi_i(x)\phi_j(x) \left(\lambda_i^{2t}\mathbf{1}_{\{i=j\}} - \mathbf{1}_{\{i=j=0\}}\right) \\
&= \sum_{x\in V(G)} \sum_i \phi_i(x)^2 \left(\lambda_i^{2t} - \mathbf{1}_{\{i=0\}}\right) \\
&= \sum_{x\in V(G)} \sum_{i\neq 0} \phi_i(x)^2 \lambda_i^{2t} \\
&= \sum_{i\neq 0} \lambda_i^{2t} \sum_{x\in V(G)} \phi_i(x)^2 \\
&= \sum_{i\neq 0} \lambda_i^{2t}
\end{aligned}$$

□

The disadvantage of the above theorem is that it only applies to vertex

transitive graphs. While several of the graphs we consider throughout this discussion are veretex transitive, in Chapter 5 we will use the following more general theorem, that applies to any undirected graph. For a proof see [20, Section 1.5].

Theorem 1.3.10.

Let G be a weighted, undirected graph. Then we can choose a lazy random walk with transition matrix P so that

$$\Delta(t) < e^{-c}$$

if

$$t \geq \frac{1}{\lambda}\left(\log \frac{\text{vol G}}{\min_v d_v} + c\right)$$

where $\lambda = \lambda_1$ if $2 \geq \lambda_{n-1} + \lambda_1$ and $\lambda = 2\frac{\lambda_1}{\lambda_{n-1}+\lambda_1}$ otherwise, where $\lambda_0 \leq ... \leq \lambda_{n-1}$ are the eigenvalues of $\mathcal{L}$.

The other property of random walks that we will examine is the *hitting time.*

Definition 1.3.11.

Let $X_0, X_1, ...$ be a Markov chain on a finite state space. Then the first hitting time of a state x is defined by

$$T_x = \min\{t \geq 0 \mid X_t = x\}$$

Just as we used Theorem 1.3.9 to bound the mixing time of a random walk on a vertex transitive graph using eigenvalues, we can also bound the hitting times when the graph is vertex transitive using the results found in [1].

Theorem 1.3.12.

Let G be a vertex transitive graph on n verties and let v, w be two distinct vertices of G. Let $-1 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n = 1$ denote the eigenvalues of the transition matrix $P = D^{-1}A$. Consider a random walk starting at vertex v, and let T_w be the first hitting time of w. Then

$$\frac{n}{2} \leq \mathbb{E}[T_w] \leq \sum_{k=1}^{n-1} \frac{1}{1-\lambda_k}$$

Theorems 1.3.9, 1.3.10, 1.3.12 illustrate several ways to bound properties of random walks using spectral information about the graph. In order to use these theorems in a meaningful way, one must have a way to find the eigenvalues of a graph. In section 1.4 we will outline methods for explicitly computing eigenvalues of Cayley graphs. However, often it is sufficient to bound the eigenvalues using a "comparison theorem". Here we use the comparison theorem as presented in [20, Section 4.5], though we note that there is a long history of similar results. In 1989, Jerrum and Sinclair used it to approximate the permanent of a matrix. In 1991, Diaconis and Strook proved several "Poincaré inequalities" using this method, and showed how to obtain bounds on relaxation times in several examples. Further work used this method to compare methods of shuffling cards [26, 27]. We will use the following theorem in Chapter 5.

Theorem 1.3.13.
Let G and $\tilde{G}$ be two connected regular graphs with the same vertex set, with second smallest eigenvalues of the normalized laplacian λ_1 and $\tilde{\lambda}_1$ and degrees k and $\tilde{k}$ respectively. Suppose for each edge $\{x, y\}$ in $\tilde{G}$, there is a path $\gamma(x, y)$ in G joining x and y of length at most ℓ. Furthermore, suppose that every edge in G is contained in at most m paths $\gamma(x, y)$. Then

$$\lambda_1 \geq \frac{k\tilde{\lambda}_1}{\tilde{k}\ell m}$$

A broad strategy we will employ several times is to interpret a process as a random walk on a state graph, and then use several algebraic tools to understand the spectral properties of the state graph. In particular, these state graphs often have high degrees of symmetry, and in some cases are isomorphic to Cayley graphs of finite groups. This is a great boon for us, as there are many useful and elegant techniques for working with Cayley graphs, and in the following section we outline several that we will employ throughout this discussion.

1.4 Cayley Graphs and Finite Group Representations

Let Γ be a group and $\omega\colon \Gamma \to \mathbb{R}_{\geq 0}$ a non-negative function. We define the *Cayley graph* with respect to ω, $\mathrm{Cay}(\Gamma, \omega)$ to be the weighted, directed graph with vertex set Γ and weighted adjacency matrix $A(g, h) = \omega(g^{-1}h)$. We say that ω is symmetric if $\omega(h) = \omega(h^{-1})$ for all h.

For $S \subseteq \Gamma$, $\mathrm{Cay}(\Gamma, \mathrm{S})$ denotes $\mathrm{Cay}(\Gamma, \mathbf{1}_{\{\mathrm{S}\}})$. With this indicator function as the weight, there is an edge between g and h if and only if there is an element s in s such that $gs = h$. Typically, the set S is a generating subset, which ensures that $\mathrm{Cay}(\Gamma, \mathrm{S})$ is connected. We say that S is symmetric in G if $h \in S \iff h^{-1} \in S$. If S is symmetric, then $Cay(\Gamma, S)$ is undirected.

If S is instead a multi-set, then instead we use a counting function instead of an indicator function for the number of edges, and $\mathrm{Cay}(\Gamma, \mathrm{S})$ is a multigraph with number of edges between g and h equal to number of s in S such that $gs = h$.

In fact, this definition of a Cayley graph extends to the case where Γ is a semigroup. We will use this interpretation in Chapter 2, where we study a class of semigroups known as left regular bands. In that case, viewing the state graph as the Cayley graph of a semigroup gives us a nice way to visualize the problem, but is not crucial for our understanding. However, in Chapters 3 and 4, finding an isomorphism between the state graph and a Cayley graph is one of the key steps in our analysis, as it allows us to compute the eigenvalues.

Before we discus how to compute the spectrum of Cayley graphs, we begin with two elementary facts about that we will use several times throughout this discussion.

Proposition 1.4.1.

Let Γ be a finite group, and $\omega\colon \Gamma \to \mathbb{R}_{\geq 0}$ a non-negative function. Then $\mathrm{Cay}(\Gamma, \omega)$ is vertex transitive.

Proposition 1.4.2.

Let Γ_1, Γ_2 be two finite groups, $\omega\colon \Gamma_1 \to \mathbb{R}_{\geq 0}$ a non-negative function, and ϕ :

$\Gamma_1 \to \Gamma_2$ *be a group isomorphism. Then* ϕ *acts as a graph isomorphism from* $\mathrm{Cay}(\mathrm{G}_1, \omega)$ *to* $\mathrm{Cay}(\mathrm{G}_2, \omega \circ \phi^{-1})$.

These propositions follow from the symmetry in Cayley graphs. In the case where Γ is an abelian group, we can find the spectrum of the adjacency matrix of a Cayley graph in terms of the irreducible characters of Γ.

Theorem 1.4.3.

Let Γ *be a finite abelian group,* $\omega \colon \Gamma \to \mathbb{R}_{\geq 0}$ *a non-negative function on* Γ*, and* $\rho \colon \Gamma \to \mathbb{C}$ *a one dimensional representation of* Γ*. If* ρ *is viewed as a row vector in* $\mathbb{C}^{|\Gamma|}$*, then* ρ *is an eigenvector of the adjacency matrix of* $\mathrm{Cay}(\Gamma, \mathrm{S}, \omega)$ *corresponding to eigenvalue*

$$\lambda_\rho = \sum_{g \in \Gamma} \omega(g)\rho(-g).$$

Proof. Consider the s entry of the vector ρA.

$$\begin{aligned}
(\rho A)(s) &= \sum_{h \in \Gamma} \rho(h) A(h, s) \\
&= \sum_{h \in \Gamma} \omega(s - h)\rho(h) \\
&= \sum_{g \in \Gamma} \omega(g)\rho(s - g) \\
&= \sum_{g \in \Gamma} \omega(g)\rho(s)\rho(-g) \\
&= \rho(s) \left(\sum_{g \in \Gamma} \omega(g)\rho(-g) \right)
\end{aligned}$$

□

Since Cayley graphs are regular, the spectrum of the transition matrix of a (lazy) random walk follows immediately. Let $|\omega| = \sum_{g \in \Gamma} \omega(g)$. Then $\mathrm{Cay}(\Gamma, \omega)$ is a $|\omega|$-regular, weighted graph. Thus if $W = \frac{1}{2}(I + D^{-1}A)$ is the transition matrix of a lazy random walk on $\mathrm{Cay}(\Gamma, \mathrm{S})$ then W has eigenvalues given by

$$\lambda_\rho = \frac{1}{2} \left(1 + \frac{1}{|\omega|} \sum_{g \in \Gamma} \rho(-g)\omega(g) \right)$$

where ρ is a one-dimensional representation of Γ.

Thus the problem of determining the eigenvalues of a random walk on the Cayley graph of an abelian group comes down to understanding the characters. Thankfully, these are simple to compute due to the following classic result in representation theory, see for example [39]

Theorem 1.4.4.
Let $\Gamma = \mathbb{Z}_{n_1} \times ... \times \mathbb{Z}_{n_k}$, and let $\theta_q = e^{\frac{2\pi i}{q}}$ for any positive integer q. For each $\overrightarrow{x} \in \Gamma$ define $\rho_{\overrightarrow{x}} \colon \Gamma \to \mathbb{C}$ be the homomorphism where $\rho_{\overrightarrow{x}}(e_i) = \theta_{n_i}$ where e_i is the cartesian product of the the additive generator 1 in the i-th group, and the identity 0 in all other groups. Then $\rho_{\overrightarrow{x}}$ is an irreducible character of Γ, and moreover every irreducible character of Γ is $\rho_{\overrightarrow{x}}$ for some $\overrightarrow{x} \in \Gamma$.

1.5 Overview

This remainder of this thesis is organized as follows. In Chapter 2 we consider the random process on a class of semigroup known Left Regular Bands induced my multiplying by randomly selected generators. Such a process converges to a stationary distribution, and we present a bound the convergence time which slightly improves previous results.

In Chapter 3, we consider the random process induced by random Seidel switching on a graph, an operation that transforms a graph by inverting the adjacency relations at a vertex. By computing the spectrum of the state graph of this process we show the process converges and bound the mixing time. We also consider several generalizations of Seidel switching.

In Chapter 4, we study a stochastic process arising from randomly playing the Lights Out game on an arbitrary finite graph. In this game, the vertices are colored black or white, and a toggle changes state of that vertex and all of its neighbors. We obtain a bounds on the mixing time and average hitting time of the all black coloring.

In Chapter 5, we study the random process on phylogenetic trees induced by nearest neighbor interchanges, an operation that swaps two neighboring sub-

trees. We use a comparison theorem to bound the spectral gap of the state graph and use this to bound the mixing time of the random nearest neighbor interchange process.

Finally, in Chapter 6, we introduce the multi-commodity dynamic demand model on a graph, a variant of the standard contact process on graphs. Each vertex has a list of demands for various commodities, and the demands of a vertex influence the demands of its neighbors. We introduce a generalization of the PageRank vector, Kronecker PageRank, and use it to bound the amount of commodities needed to satisfy the demand at the vertices.

Chapter 2

Random Walks on Left Regular Bands

2.1 Introduction

In this chapter we consider random walks on a class of semigroups known as Left Regular Bands (or LRB's, for short) which are idempotent with the additional relation $xyx = xy$ (see [41, 50, 52]). Many problems can be interpreted as random walks on LRB's, such as the move-to-front self-organizing schemes [14,30], hyperplane arrangements [10] and graph coloring games [21]. It is known that the random walks on Left Regular Bands have many amazing properties, including having real eigenvalues which can be expressed in elegant formula [14, 15]. In addition, Diaconis and Brown [16] gave a variation of the Plancherel formula for bounding the total variation distance $\Delta_{\mathrm{TV}}(t)$ of a LRB random walk after t steps:

$$\Delta_{\mathrm{TV}}(t) \leq \sum_{\{l \in L^* \mid l \text{ is co-maximal}\}} \lambda_l^t \tag{2.1}$$

where the eigenvalues are indexed by the co-maximal elements in the semilattice L associated with S and the random walk under consideration is on the ideal of chambers in S (definitions of these terms will be presented in the following section). This is in contrast with the Plancherel formula for random walks on groups (or,

on vertex transitive graphs), which states that

$$\Delta_{\mathrm{TV}}(t) \leq \frac{1}{2}\left(\sum_{\lambda_i<1} \lambda_i^{2s}\right)^{1/2}$$

where λ_i's are eigenvalues of the transition probability matrix.

2.2 Semilattices and Left Regular Bands

The formula that we will use for for the eigenvalues of the random walk on an LRB will use an associated semilattice and Möbius inverstion, so begin with the required combinatorial definitions.

Definition 2.2.1.
A partially ordered set is a *semilattice* if each pair of elements has a unique least upper bound.

Definition 2.2.2.
For a semilattice L, the *Möbius function* of L is the function $\mu : L \times L \to \mathbb{Z}$ defined by

$$\mu(x,y) = \begin{cases} 1 & \text{if } x = y \\ -\sum\limits_{x \leq z < y} \mu(x,z) & \text{if } x > \hat{0} \\ 0 & x \not\leq y \end{cases}$$

The main application of the Möbius function is in the following theorem.

Theorem 2.2.3.
Let L be a semilattice with Möbius function μ. Let $f : L \to \mathbb{C}$. If $g(x) = \sum\limits_{y \leq x} f(y)$, then

$$f(x) = \sum_{y \leq x} \mu(y,x) g(y)$$

By definition, a semigroup S is a left regular band if for all $x, y \in S$ it satisfies the relations

$$x^2 = x \tag{2.2}$$

$$xyx = xy \tag{2.3}$$

This definition was introduced by Klein-Barman in 1940 [41] and independently by Schtzenburger in 1947 [52]).

However, in 2000, Ken Brown gave an alternate characterization which we will use [14].

Lemma 2.2.4.
A semigroup S is a LRB if and only if there is a semilattice L and a surjection $\text{supp} : S \twoheadrightarrow L$ *such that*

$$\text{supp } xy = \text{supp } x \vee \text{supp } y \tag{2.4}$$

and

$$xy = x \quad \textit{if} \quad \text{supp } y \leq \text{supp } x \tag{2.5}$$

We briefly illustrate one direction of the lemma, how to create such an L and support map. For a semigroup S, there is a natural partial order $<$, defined by:

$$x \leq y \Leftrightarrow xy = y.$$

For x in S, we define

$$S_{\geq x} = \{y \in S \mid y \geq x\}.$$

The semilattice L can be derived from S as follows. First we define a relation $\preceq$ on S.

$$y \preceq x \Leftrightarrow xy = x.$$

The equivalence class under $\preceq$ containing x is defined to be the support of x, denoted by supp x. For x, y in S, we have

$$\text{supp } xy = \text{supp } x \vee \text{supp } y$$

and

$$x \leq y \Rightarrow \text{supp } x \preceq \text{supp } y.$$

Elements in L are called *flats* (following the terminology for semigroups associated with matroids [14]). A flat l is *co-maximal* if

$$x \succ l \Rightarrow x = \hat{1},$$

where $\hat{1}$ denotes the maximal element of L.

An element $c \in S$ is said to be a *chamber* if $cx = c$ for all $x \in S$. Therefore supp c is maximal in the semilattice L. The set of all chambers forms an ideal of S.

Below are several examples of LRBs that we will use throughout this chapter.

2.2.1 The Tsetlin Library

Let T be the set of permutations of the elements of $C = \{c_1, \cdots, c_m\}$. We define the action of C on T as follows. For c_j in C, and $t = c_{i_1}c_{i_2}\cdots c_{i_m}$, then $c_j t = c_j c_{i_1} c_{i_2} \cdots \widehat{c_j} \cdots c_{i_m}$, where the $\widehat{c_j}$ means to delete c_j where it appears later in the string. It is not hard to check that T is the ideal of chambers of the free semigroup generated by actions of C. This is known as the Tsetlin library, a well studied LRB, which has many applications including the "move-to-front" self-organizing search.

2.2.2 Strings of Colored Numbers

For given positive integers m, k, let $C = \{(i,j)\}_{i=1,j=1}^{m,k}$. It is convenient to think of (i,j) as being the number i with color j. We define the strings of colored numbers, $SCN(m,k)$, to be the LRB consisting of strings of elements of C with respect to the relation

$$(i,j_0)x(i,j_1) = (i,j_0)x$$

for all $x \in SCN(m,k)$. The support lattice L is the boolean lattice of the numbers $1, ..., m$ with supp $(i,j) = i \in L$. Thus the maximal ideal I consists of strings of length m, with each number represented exactly once.

2.2.3 Edge Flipping in Graphs

Let G be a graph with with no isolated vertices and edges $\{e_1, ..., e_m\}$. For edge e_i, define the function r_i that colors the endpoints of e_i red and b_i that colors them blue. Let $C = \{r_1, r_2, ..., r_m, b_1, ..., b_m\}$, and let S be the set compositions

of elements of C. Then S is a left regular band with respect to composition, and its chambers are colorings of the graph which have all vertices colored by a composition of elements of C. This operation is known as *edge flipping* and was studied extensively in [21].

2.3 Random Walks on LRBs

We will consider random walks on LRBs. Let S be a LRB, $C \subseteq S$ a set that generates S, p_C a probability distribution on C, and L the semilattice associated with S. The random on walk on S is a Markov chain where at each step we move from s to cs where c is a randomly chosen element of C. Formally,

$$\mathbb{P}(X_{t+1} = y \mid X_t = x) = \sum_{\{c \in C \mid cx = y\}} p_c$$

By definition, if s is a chamber, then cs is a chamber. Thus once the random walk on S enters the ideal of chambers, it remains there. Thus, we consider the random walk only on the chambers of S. to In [14, 16], Brown and Diaconis showed the following about this random walk

Theorem 2.3.1.
Let S be a semigroup, C a subset that generates S, p_c a probability distribution on C, and L the semilattice associated with S. Consider a random walk on the chambers of S where at each step we move from s to cs where c is a random element of C. Then:

1. *The random walk is ergodic*
2. *The transition matrix P is diagonalizable*
3. *For each $X \in L$ there is an eigenvalue $\lambda_X = \sum_{\text{supp } c \leq X} p_c$.*
4. *The multiplicity of λ_X, m_X, satisfies*

$$\sum_{Y \geq X} m_Y = c_X$$

where c_X is the cardinality of $S_{\succeq X} = S_{\geq x} = \{z \in S \mid z \geq x\}$ where x is any element with support X.

Note that c_X is independent of the choice of x. Additionally, we can calculate the multiplicities directly using Möbius inversion to get the expression

$$m_X = \sum_{Y \succeq X} \mu(X, Y) c_Y$$

where μ is the Möbius function of the semilattice L.

We can view a random walk on an LRB as a random walk on a weighted, directed graph. Define the function $w\colon S \times S \to \mathbb{R}$ by

$$w(u, v) = \sum_{\{c \mid cu = v\}} p_c.$$

Then a random walk on the weighted graph with vertex set the chambers of S and weight function given by w has the transition probabilities as the random walk on LRBs described above. For an illustration of why only the chambers are used, see Figure **??**. Any random walk will eventually reach and not leave the strongly connected core consisting of the chambers of the Tsetlin Library, those strings using all 3 numbers.

2.4 An Alternating Bound for the Mixing time of a Random Walk on a Left Regular Bound

We will give a slightly improved formula of (2.1). The proof combines the techniques of Diaconis-Brown in [16] and the methods of Bidigare, Hanlon, and Rockmore [10] for random walks on chambers of hyperplane arrangements. A similar result was independently obtained by Benjamin Steinberg [58].

Theorem 2.4.1.

For a random walk on chambers of an LRB semigroup, the total deviation distance

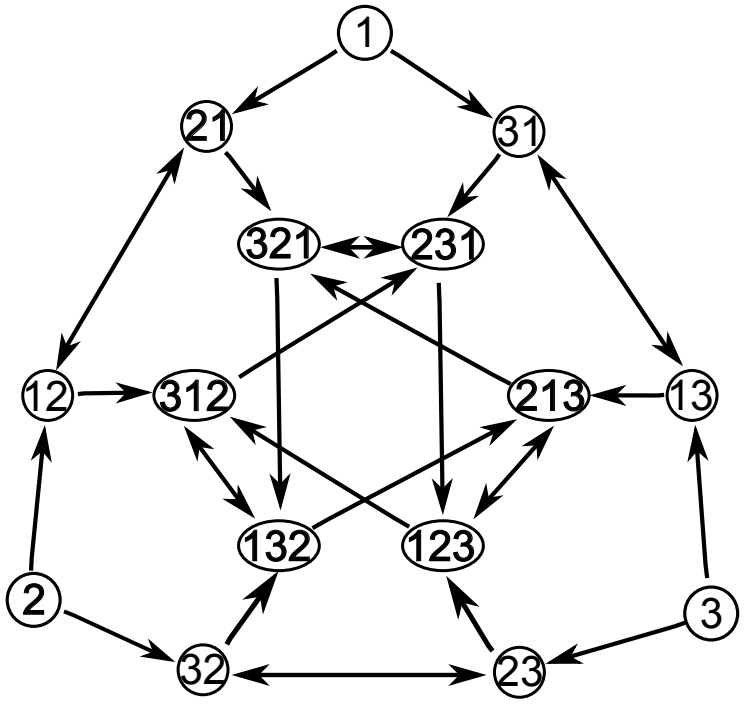


Figure 2.1: The Cayley Graph for the Tsetlin library on the set {1,2,3}. Self loops have been omitted for clarity.

after t steps is bounded by:

$$\Delta_{TV}(t) \leq \sum_{l \in L^*} -\mu(l, \hat{1})\lambda_l^t,$$

where $\mu(l, \hat{1})$ is the Möbius function on the support lattice L, and L^* denotes the lattice L with its maximal element, $\hat{1}$, removed.

To show that Theorem 2.4.1 is an improvement of (2.1), we will show the following.

Corollary 2.4.2.

$$\sum_{l \in L^*} -\mu(l, \hat{1})\lambda_l^t \leq \sum_{\{l \in L^* \mid l \text{ is co-maximal}\}} \lambda_l^t.$$

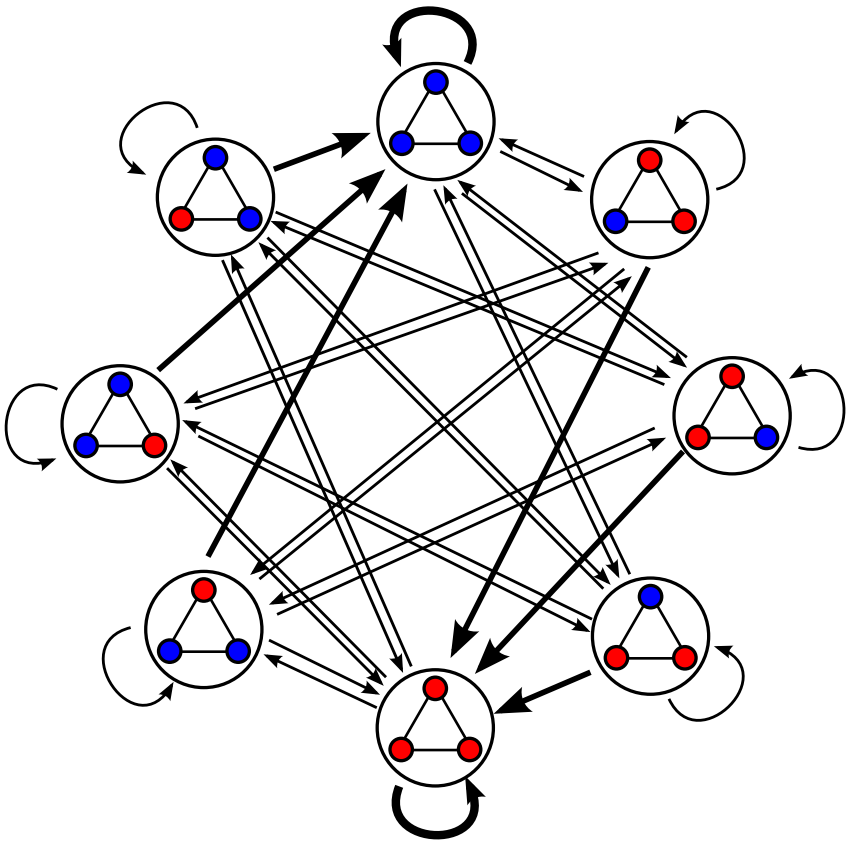


Figure 2.2: The Cayley Graph of the chambers of the LRB corresponding to edge flipping of a 3-cycle. The thickness of the edge corresponds to its weight.

When L is a Boolean lattice of a set V, then we can index elements of the lattice by subsets of V. For a subset $Y \subset V$, the Möbius function $\mu(Y, \hat{1}) = (-1)^{|V \setminus Y|}$.

Corollary 2.4.3.

If L is a Boolean lattice of a set V, then for a subset $Y \subset V$ $\mu(Y, \hat{1}) = (-1)^{|V \setminus Y|}$, and therefore

$$\Delta_{TV}(t) \leq \sum_{Y \subset V} (-1)^{|V \setminus Y|+1} \lambda_Y^t$$

Proof of Theorem 2.4.1:

Let $\{p_s\}$ be a probability distribution on S, so that $p_s \geq 0$ and $\sum_s p_s = 1$. The

transition probability matrix of the associated random walk is denoted by

$$P(u,v) = \sum_{\substack{s \\ su=v}} p_s.$$

Let $x_1, x_2, ...$ be an independent, identically distributed sequence of random elements of S. We consider $x^{(t)} = x_1 x_2 ... x_t$ and $x_{t,s} = x^{(t)}s$, which is the location of a random walk after t steps starting at s.

Note that if $x^{(t)}$ is a chamber then $x^{(t)}s_1 = x^{(t)}s_2$ for any $s_1, s_2 \in S$. We define $\pi_{t,s}$ to be the distribution of $x_{t,s}$, that is $\pi_{t,s}(u) = P(x_{t,s} = u)$, for any $u \in S$. Let π denote the stationary distribution of the random walk. Note that if $x^{(t_0)}$ is a chamber for a fixed time t_0, then $x^{(N)} = x^{(t)}$ for all $N \geq t_0$. Thus $\pi(u) = P(Cs = u)$ where C is a the random chamber first reached by $x^{(t)}$ as t increases.

We consider the total variation distance

$$\Delta_{TV}(t) := \max_{s \in S} \max_{A \subset S} \mid \pi_{t,s}(A) - \pi(A)|.$$

where $\pi_{t,s}(A) = P(x_{t,s} \in A) = \sum_{u \in A} \pi_{t,s}(u)$ and $\pi(A) = P(Cs \in A) = \sum_{u \in A} \pi(u)$. We split up both events according to whether or not $x^{(t)}$ is a chamber. Let B_t be the event that $x^{(t)}$ is a chamber. Then we have

$$\pi_{t,s}(A) = P(B_t \text{ and } x^{(t)}s \in A) + P(\neg B_t \text{ and } x^{(t)}s \in A)$$

$$\pi(A) = P(B_t \text{ and } Cs \in A) + P(\neg B_t \text{ and } Cs \in A)$$

If B_t occurs then $x^{(t)}$ is a chamber, and thus the first term of each expression is the same. This follows from the known fact that the stationary distribution in the original (unbounded) process (with replacement) is the same as the stationary distribution for the process without replacement. For detailed discussions, the reader is referred to Section 4 in [21].

Therefore we have

$$|\pi_{t,s}(A) - \pi(A)| = |P(\neg B_t \text{ and } x^{(t)}s \in A) - P(\neg B_t \text{ and } Cs \in A)|$$

which is at most $P(\neg B_t)$, as both terms in the difference are between 0 and $P(\neg B_t)$. Thus we have

$$\Delta_{TV}(t) \leq P(\neg B_t) = P(x^{(t)} \text{ is not chamber}).$$

By definition, the only way for $x^{(t)}$ not be a chamber is for some $m \in L, m \neq \hat{1}$, supp $x^{(t)} = m$. Therefore,

$$\Delta_{TV}(t) \leq \sum_{m \in L^*} P(\text{supp } x^{(t)} = m). \tag{2.6}$$

Let us denote $P(\text{supp } x^{(t)} = m)$ by $\beta_{t,m}$, then equation (2.6) becomes

$$\Delta_{TV}(t) \leq \sum_{m \in L^*} \beta_{t,m}. \tag{2.7}$$

We will evaluate $\beta_{t,m}$ using a Möbius inversion on the lattice L. From [14], we have

$$\lambda_m = \sum_{\text{supp } c \ \subseteq m} p_c = P(\text{supp } x \preceq m)$$

where $P(x = c) = p_c$. Since $x_1, ..., x_t$ are chosen independently, we have that

$$\lambda_m^t = P(\text{supp } x_1 \preceq m)P(\text{supp } x_2 \preceq m)...P(\text{supp } x_t \preceq m) = P(\text{supp } x^{(t)} \preceq m). \tag{2.8}$$

Note that supp $x^{(t)} \preceq m$ if and only if supp $x^{(t)} = l$ for some $l \preceq m$. Thus

$$P(\text{supp } x^{(t)} \preceq m) = \sum_{l \preceq m} \beta_{t,l}.$$

By (2.8) we have

$$\lambda_m^t = \sum_{l \preceq m} \beta_{t,l}.$$

Therefore we can use a Möbius inversion to derive

$$\beta_{t,m} = \sum_{l \preceq m} \mu(l, m) \lambda_l^t.$$

Plugging this into (2.7) we obtain

$$\begin{aligned}
\Delta_{TV}(t) &\leq \sum_{m \in L^*} \sum_{\{l \mid l \preceq m\}} \mu(l,m)\lambda_l^t \\
&= \sum_{l \in L^*} \sum_{\{m \mid l \preceq m, m \neq \hat{1}\}} \mu(l,m)\lambda_l^t \\
&= \sum_{l \in L^*} \left(\sum_{\{m \mid l \preceq m, , m \neq \hat{1}\}} \mu(l,m) \right) \lambda_l^t \\
&= \sum_{l \in L^*} \left(-\mu(l,\hat{1}) \right) \lambda_l^t
\end{aligned}$$

as desired. □

Proof of Corollary 2.4.2:
We note that the bound given in (2.1) can be compared to the bound of Theorem 2.4.1 as follows:

$$\begin{aligned}
\sum_{\{l \in L^* \mid l \text{ is co-maximal}\}} \lambda_l^t &= \sum_{\{l \in L^* \mid l \text{ is co-maximal}\}} P(\text{supp } x^{(t)} \leq l) \\
&= \sum_{m \in L^*} |\{l \mid m \preceq l \text{ and } l \text{ is co-maximal}\}| P(\text{supp } x^{(t)} = m) \\
&\geq \sum_{m \in L^*} P(\text{supp } x^{(t)} = m)
\end{aligned}$$

This last term is the right hand side of (2.6) above, which is equal to the bound given in Theorem 2.4.1. Thus Theorem 2.4.1 is an improvement over Equation 2.1. □

2.5 Examples Using the Alternating Bound

In this section we calculate and compare the bounds given by Equation 2.1 and Theorem 2.4.1 for the LRBs described earlier: The Tsetlin Library, Strings of Colored Numbers, and Edge Flipping.

2.5.1 The Tsetlin Library

Let T be the Tsetlin Library on a finite set C, and let us consider the random walk on T with respect to the uniform distribution . The support lattice L is the boolean lattice $\{0,1\}^C$. The eigenvalues for the random walk on T were first determined by Phatarfod [50] in 1991. For each subset $X \subset C$, there is an eigenvalue $\lambda_X = |X|/m$, with multiplicity equal to the so-called derangement number d_k where $k = m - |X|$. It is known that

$$d_k = k! \sum_{j=0}^{k} \frac{(-1)^j}{j!} = \left\lfloor \frac{k!}{e} + \frac{1}{2} \right\rfloor .$$

Note that the derangement numbers d_k satisfy

$$\sum_{Y \supseteq X} d_{m-|Y|} = c_X = (m - |X|)!.$$

However, one of the advantages of both the bounds in (1) and Theorem 2.4.1 is that they do not depend on the multiplicity of each eigenvalue. Equation (1) yields the bound

$$\Delta_{\mathrm{TV}}(t) \leq m \left(1 - \frac{1}{m}\right)^t .$$

Theorem 2.4.1 improves this to

$$\Delta_{\mathrm{TV}}(t) \leq \sum_{k=1}^{m-1} (-1)^{k+1} \binom{m}{k} \left(\frac{m-k}{m}\right)^t .$$

Since the old bound is simply the first term of this alternating series, it has been improved by

$$\sum_{k=2}^{m-1} (-1)^{k} \binom{m}{k} \left(\frac{m-k}{m}\right)^t .$$

To illustrate the improvement, Figure **??** is a plot of the two bounds for the case when $m = 10$.

2.5.2 Strings of Colored Numbers

Consider a random walk on the LRB $SCN(m,k)$ where each color from $1, ..., k$ is assigned a probability p_j of being chosen, where $\sum_j p_j = 1$, and each

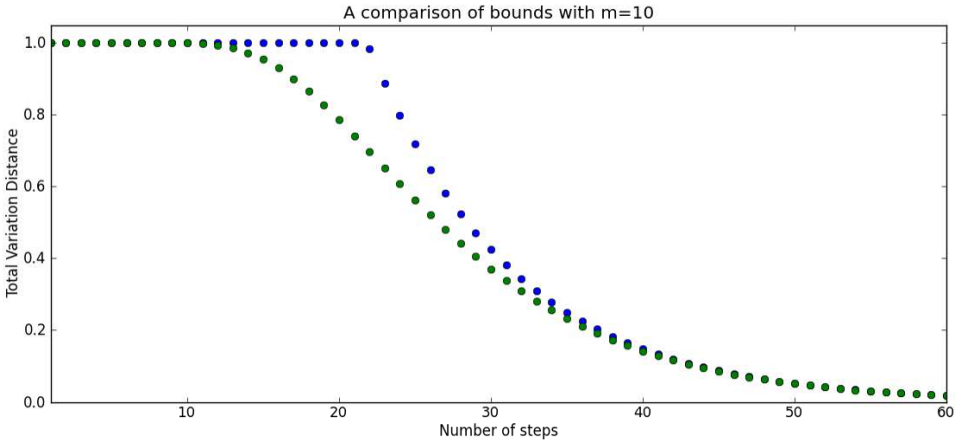


Figure 2.3: The bounds given by Equation 2.1 and Theorem 2.4.1 for the Tsetlin Library on 10 elements

number has equal probability of being chosen. More explicitly, for $(i,j) \in C$, $p_{(i,j)} = \frac{p_j}{m}$. Note that if $k = 1$, then we have the Tsetlin Library considered above. The support lattice L is the boolean lattice $\{0,1\}^m$, and thus elements are indexed by subsets $Y \subset \{1,...,m\}$. The eigenvalues are thus

$$\lambda_Y = \sum_{\text{supp } c \subset Y} p_c = \sum_{\substack{(i,j) \\ i \in Y}} \frac{p_j}{m} = \sum_{i \in Y} \sum_j \frac{p_j}{m} = \sum_{i \in Y} \frac{1}{m} \sum_j p_j = \frac{|Y|}{m}.$$

Thus the eigenvalues and support lattice are the same as in the Tsetlin library, and so we have the same bound of

$$\Delta_{\text{TV}}(t) \leq \sum_{k=1}^{m-1} (-1)^{k+1} \binom{m}{k} \left(\frac{m-k}{m}\right)^t.$$

It is interesting to note that the bound does not depend at all on the choice of probabilities for a color p_j, so long as each number has equal probability of being chosen.

2.5.3 Edge Flipping

In [21], the following random process was studied: Initially each vertex of a graph G is colored red or blue. At each step in the process, we select a random

edge of G and (re-)color both its endpoints blue with probability p, or red with probability $q = 1 - p$. This process is then repeated some large number of times. The color configuration of G changes at each step.

The re-coloring actions (each of which is associated with picking an edge e and changing the colors of its endpoints) form a LRB with respect to composition, and the chambers are simply the colorings of the graph G which are created by this process.

It was shown in [21] that, for example, for the uniform case of $p = 1/2$, the random walk on the state graph has, for each subset T of the vertex set V of G, the eigenvalue λ_T (with multiplicity 1) being the ratio of the number of edges in the induced subgraph of T divided by the total number of edges in G.

To bound the total variation distance, we can use Theorem 1 to improve previous bounds since the negative terms in the Möbius function lead to cancelations. For the example of the path P_5, the bound given in [21] yields

$$\Delta_{\mathrm{TV}}(t) \leq 2\left(\frac{3}{4}\right)^t + 6\left(\frac{1}{2}\right)^t + 10\left(\frac{1}{4}\right)^t,$$

whereas Theorem 2.4.1 and Corollary 2.4.3 give the improved bound

$$\Delta_{\mathrm{TV}}(t) \leq 2\left(\frac{3}{4}\right)^t - 2\left(\frac{1}{4}\right)^t.$$

2.6 Remarks

There is room for improvement in Theorem 2.4.1. In the proof, we use the bound

$$|\pi_{t,s}(A) - \pi(A)| = |P(\neg B_t \text{ and } x^{(t)}s \in A) - P(\neg B_t \text{ and } Cs \in A)| \leq P(\neg B_t)$$

This is a crude estimate, and a more delicate handling of the term $|\pi_{t,s}(A) - \pi(A)|$ may lead to a more precise bound.

Acknowledgement:
This chapter is based on the paper "A note on an alternating upper bound for random walks on semigroups" written jointly with Fan Chung [24]. It has been submitted for publication to Discrete Applied Mathematics.

Chapter 3

Random Siedel Switching

3.1 Introduction

Let $G = G(V, E)$ be a finite, simple graph. For a vertex $v \in V$, the operation of *switching at* v transforms G to a new graph G_v by deleting all edges adjacent to v, and adding all potential edges from v to vertices not previously connected. This operation is also known as *vertex switching*, *node switching*, or *Seidel Switching*. It was originally introduced by J.H. van Lint and J.J. Seidel [61] as tool to study equilateral point sets in elliptic spaces.

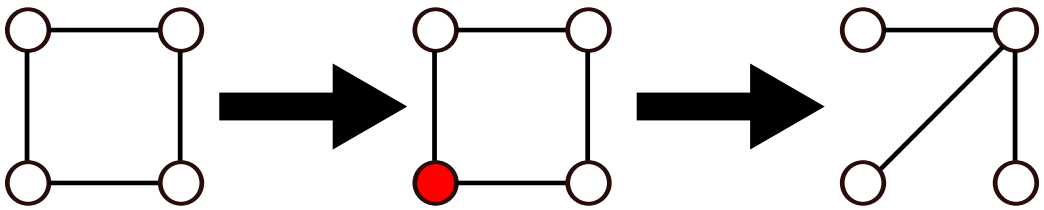


Figure 3.1: Seidel switching on 4 vertices. We begin with a graph on 4 vertices, select one vertex, then switch the adjacency relations for that vertex.

Two graphs G_0 and G_k are said to be *switching equivalent* if there is a sequence of vertices $v_1, v_2, ..., v_k$ such that G_i is obtained by switching v_i in G_{i-1} for $i = 1, ..., k$. It is easy to see that performing the sequence of operations in

reverse order will transform G_k back to G_0, and so the relation is both reflexive and symmetric. Also note that this demonstrates each switching operation is invertible, a fact that will be needed later. Transitivity follows immediately from the definition, and so this is an equivalence relation among graphs on a fixed number of vertices. The equivalence classes are known as *switching classes*. The number of switching classes on n vertices is equal to the number of two-graphs with n vertices [55] as well as Euler graphs [44]. Seidel Switching has applications to spectral graph theory as illustrated by the following theorem, also due to Seidel [54].

Theorem 3.1.1.
Let G and G' be two regular graphs of degree d in the same switching class. Then G and G' are cospectral.

In this chapter, we consider a randomized switching process. At each step we randomly select a vertex, and apply a switching operation at that vertex. We analyze this random process, and obtain results about the times of convergence to the stationary distribution. Our method can be outlined as follows:

1. The set of compositions of switching operations is an abelian group, which we denote $\Gamma_{SS}(n)$.

2. There is an isomorphism $\nu\colon \Gamma_{SS}(n) \to \mathbb{Z}_2^{n-1}$

3. The random process can be understood as a lazy random walk on a Cayley graph of $\Gamma_{SS}(n)$.

4. The eigenvalues of the transition matrix of this random walk can be found using the irreducible representations of $\Gamma_{SS}(n)$, and can be used to bound the convergence time to the stationary distribution.

This chapter is organized as follows. In Section 3.2 we formally introduce randomized Seidel Switching, and show that the stationary distribution is uniform and give bounds on the convergence time. We omit proofs in this section as results are special cases of more general results stated in the following section. In Section 3.3 we introduce two generalizations of Seidel switching. We consider

both switching with multiple colors and restricted to a fixed host graph. We then study a randomized version of this generalized switching, again showing how it can be viewed as a random walk on a graph and use spectral methods to bound the convergence times.

3.2 Randomized Seidel Switching

We wish to consider the behavior of the random process arising from a random sequence of switchings on graphs with n vertices. It will be useful to think of the switching actions as functions on the state space of all graphs on n vertices. Let $\mathfrak{G}(n)$ denote the set of all graphs on n labeled vertices. Let $s_v \colon \mathfrak{G}(n) \to \mathfrak{G}(n)$ be the action of switching at vertex v, so that $s_v(G) = G_v$ for any graph G. We let $s_\emptyset$ denote the identity function, and let $\overline{K_n} \in \mathfrak{G}(n)$ denote the empty graph on n vertices.

Let $\{x_t\}_1^\infty$ be a sequence of vertices, independently chosen uniformly at random among the n vertices. Consider the random process $\hat{X}(t)$ where

$$\hat{X}(0) = \overline{K_n},$$

and for each $t \geq 1$,

$$\hat{X}(t+1) = s_{x_t}(\hat{X}(t)).$$

As we will see later, this sequence may be periodic and so will not converge to a stationary distribution. To eliminate this concern, we consider a "lazy version" of this process where half the time no action is taken. We define the sequence $X(t)$ where

$$X(0) = \overline{K_n},$$

$$X(t+1) = s_{y_t}(X(t))$$

where

$$\mathbb{P}(y_t = \emptyset) = \frac{1}{2}$$

and

$$\mathbb{P}(y_t = x_t) = \frac{1}{2}.$$

Our goal is to analyze $X(t)$. In particular, we wish to consider the following questions:

1. How can we understand or interpret $X(t)$?
2. What is the stationary distribution of $X(t)$?
3. How fast does $X(t)$ converge to its stationary distribution?

We begin by examining the algebraic structure of compositions of switching functions s_v. Again, proofs in this section are omitted, as the results stated are special cases of results that appear in Section 3.3

Theorem 3.2.1.
Let $\Gamma_{SS}(n)$ be the group of all compositions of the switching operators $\{s_{v_i}\}_{i=1}^{n}$ with respect to composition. Then $\Gamma_{SS}(n) \cong \mathbb{Z}_2^{n-1}$.

For our analysis, we will need to explicitly understand the isomorphism between the groups, and in particular the image of the switching functions s_v.

For the groups $\mathbb{Z}_2 \times ... \times \mathbb{Z}_2$, let e_i correspond to the cartesian product of the the additive generator 1 in the i-th group, and the identity 0 in all other groups.

Corollary 3.2.2.
There exists an isomorphsim $\nu \colon \Gamma_{SS}(n) \to \mathbb{Z}_2^{n-1}$ with

$$\nu(s_{v_i}) = \begin{cases} e_i & \text{if } i = 1, ..., n-1 \\ \sum_{i=1}^{n-1} e_i & \text{if } i = n \end{cases}$$

The key idea in our analysis is recognizing that $X(t)$ can be viewed a lazy random walk on a state graph. Moreover, we will show that this state graph is isomorphic to a Cayley graph of $\Gamma_{SS}(n)$.

We begin by defining the *Switching State Graph* as follows:

Definition 3.2.3.
The *Switching State Graph* of n vertices, denoted $C_{SS}(n)$, is a graph with vertex set the set of all graphs on n (labeled) vertices in the same switching class as the empty graph. For two graphs G and H, $G \sim H$ if there is a vertex v such that $s_v(G) = H$.

The notion of the Switching Graph is useful because it allows us to view our random process $X(t)$ as a lazy random walk on a graph. At each step in a random walk on $C_{SS}(n)$ one moves to a neighbor with probability $\frac{1}{2}$, and stays at the same vertex with probability $\frac{1}{2}$. Since moving to a random neighbor is equivalent to picking a random vertex v and applying the switching function s_v, we have the following proposition.

Proposition 3.2.4.
The random process $X(t)$ described above is identically distributed to a lazy random walk on the switching graph starting from the empty graph.

Thus we have reduced the problem of analyzing $X(t)$ to that of understanding the lazy random walk on $C_{SS}(n)$. We begin by recognizing that we can fully understand the structure of $C_{SS}(n)$ in terms of the switching group $\Gamma_{SS}(n)$.

Proposition 3.2.5.
The Switching State Graph $C_{SS}(n)$ is isomorphic to $\mathrm{Cay}(\mathbb{Z}_2^{n-1}, S)$, where $S = \{e_i\}_{i=1}^{n-1} \cup \left\{ \sum_{i=1}^{n-1} e_i \right\}$ and e_i is the standard basis element of $\mathbb{Z}_2^{n-1}$. This is the hypercube of dimension $n-1$ with diagonal chords added.

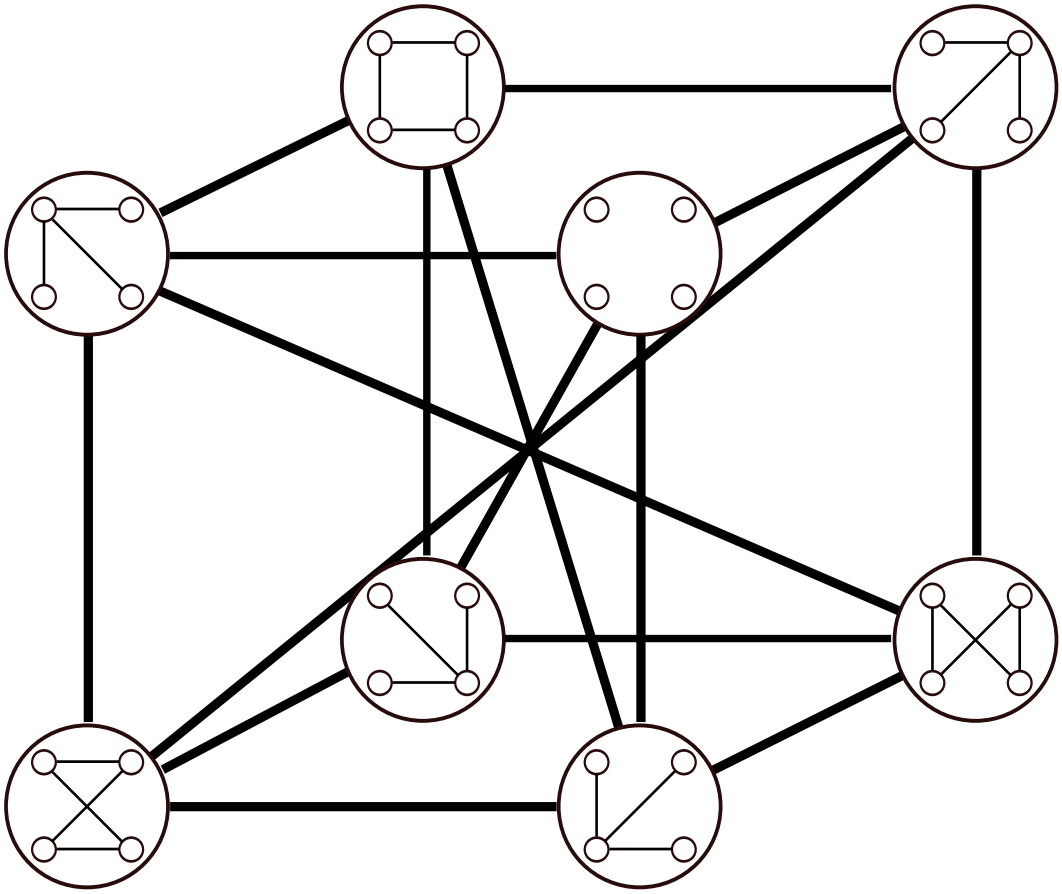


Figure 3.2: The Switching State Graph for 4 vertices. The "vertices" of the state graph are the graphs on 4 vertices that are in the same switching class as the empty graph. Two graphs are adjacent if they differ by a single switching operation. Note that the graph is bipartite, which demonstrates why we consider a lazy random walk to guarantee convergence.

Using Propositions 3.2.4 and 3.2.5, along with the tools presented in Chapter 1, we can answer the questions posed at the beginning of this section.

Theorem 3.2.6.
Let $X(t)$ be the random graph at time t obtained from the randomized Seidel Switching switching process described above. Then

1. *$X(t)$ converges to a uniform distribution on all graphs in the switching class of $\overline{K_n}$.*

2. *The χ-squared distance from the stationary distribution after t steps is bounded by*

$$\Delta'(t) \leq \left(\sum_{j=1}^{n-1} \binom{n-1}{j} \left(1 - \frac{j}{n}\right)^{2t} \right)^{\frac{1}{2}}$$

3. $\Delta'(t) \leq e^{-c}$ *if* $t > \frac{1}{2} n \log n + c$ *for any* $c > 0$

3.3 Restricted and Multi-colored Switching

One can generalize the switching operation to colorings of graphs in a number of ways. Consider if instead of two states (off or on) for any of the edges, there are q states for some fixed integer q. It is natural to think of these as different colors an edge. The switching operation is then some permutation π on the set of potential colors of an edge. Brewster and Graves [12] considered the action of an arbitrary, fixed permutation π and studied homomorphisms between colorings of graphs. Cameron and Tarzi [18] considered the action under all transpositions, as well as restricted cases where not all transpositions were allowed. We will consider the case where π is the cyclic operation "$+1 \mod q$", though many of the techniques generalize to arbitrary permutations.

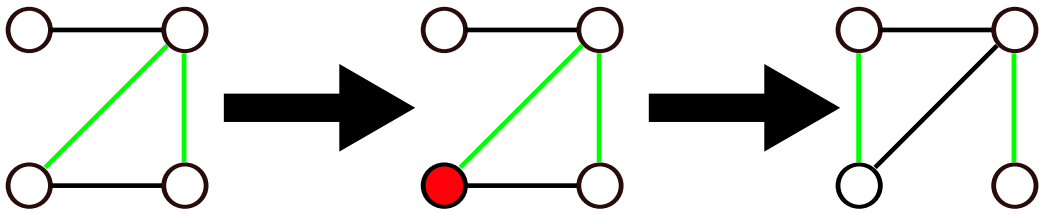


Figure 3.3: An example of multi-color switching with $q = 3$. No edge corresponds to state 0, a grey edge to state 1, and a black edge to state 2. After selecting bottom left vertex, we increment the state of each adjacent edge by 1 mod 3

We consider a further generalization simultaneously. Previously, any non-edge could become an edge. We consider instead when only certain edges can

be created. For a fixed *host graph* H, we can define switching actions that are restricted to colorings of the edges of H.

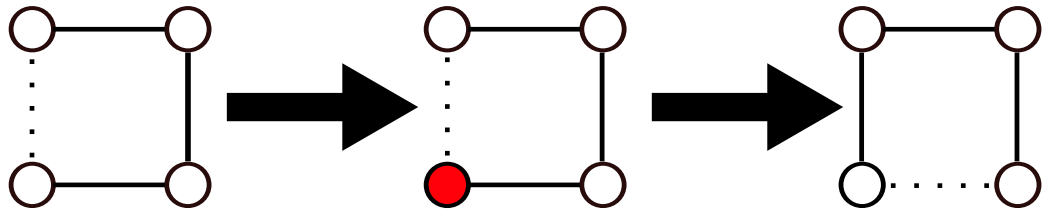


Figure 3.4: Switching restricted to the host graph $H = C_4$. Note that switching at the bottom left vertex only creates new edges that are in the host graph

We let $\mathfrak{G}_q(H)$ be the set of all edge-colorings of H using q colors, that is all functions $\tau\colon E(H) \to \{0, 1, ..., q-1\}$. We call τ a q-coloring of the edges of H.

Definition 3.3.1.
We define the q-H-switching of a vertex v to be the operator $s_v\colon \mathfrak{G}_q(H) \to \mathfrak{G}_q(H)$ where

$$s_v(\tau)(e) = \tau(e) + 1 \mod q.$$

$s_\emptyset$ will refer to the identity map on $\mathfrak{G}_q(H)$. The operators s_v for $v \in V$ will be called the elementary q-H-switching operators.

Note that the Seidel Switching as explored in Section 3.2 is q-H-switching with $q = 2$ and $H = K_n$.

3.3.1 The q-H Switching Group

It follows from definition 3.3.1 that the elementary switching operators commute and that

$$s_v^q = s_\emptyset$$

for every v. Thus each elementary operator s_v is invertible with

$$s_v^{-1} = s_v^{q-1}.$$

Letting $\Gamma_{qSS}(H)$ be the set of all compositions of elementary q-H-switching operators, the following proposition follows.

Proposition 3.3.2.
$\Gamma_{qSS}(H)$ is an abelian group under composition.

We call $\Gamma_{qSS}(H)$ the *q-H-switching group*, or just *switching group* if q and H are clear from context. We begin with a somewhat surprising result. We can view $\Gamma_{qSS}(H)$ as the vector space over $\mathbb{Z}_q$ spanned by $\{s_v\}_{v\in V}$, where here s_v is viewed as a vector in $\mathbb{Z}_q^E$ with 1 for all edges with endpoint v. Thus, $\Gamma_{qSS}(H)$ is the column space over $\mathbb{Z}_q$ of the edge-vertex adjacency matrix. It seems that such an object should vary depending on the structure of the graph H, but it turns out only to depend on whether H is bipartite or not, and on the parity of q.

Theorem 3.3.3.
Let H be a connected graph on n vertices.

$$\Gamma_{qSS}(H) \cong \begin{cases} \mathbb{Z}_q^{n-1} & \text{if } H \text{ is bipartite} \\ \mathbb{Z}_q^{n} & \text{if } H \text{ is not bipartite and } q \text{ is odd} \\ \mathbb{Z}_q^{n-1} \times Z_r & H \text{ is not bipartite and } q = 2r \end{cases}$$

Proof. Index the n generators of $\mathbb{Z}_q^n$ by the vertices of H, and denote them by unit vectors $\{f_v\}_{v\in V}$, where f_v corresponds to a 1 in the v coordinate and 0 elsewhere. We define the map $\phi\colon \mathbb{Z}_q^n \to \Gamma_{qSS}(H)$ by $\phi(f_v) = s_v$, extended linearly so that ϕ is a homomorphism.

We first note that ϕ is surjective, as any element in $\Gamma_{qSS}(H)$ occurs from switching vertices of H some number of times less than q. By the First Isomorphism Theorem, $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^n / \ker\phi$.

It remains to analyze $\ker\phi$. We will view $\Gamma_{qSS}(H)$ as a module over $\mathbb{Z}_q$. Suppose that $g = \sum_v \alpha_v f_v \in \ker\phi$, where $\alpha_i \in \mathbb{Z}_q$. The value on $e = \{u, w\}$ is only influenced by s_u and s_w, so if $(u, w) \in E(H)$, then

$$\alpha_u = -\alpha_w. \tag{3.1}$$

So if $v_{i_1}, v_{i_2}, ..., v_{i_k}$ is a walk on the vertices of H, then

$$\alpha_{i_l} = (-1)^{l+1}\alpha_{i_1} \tag{3.2}$$

That is, values alternate between c and $q-c \in \mathbb{Z}_q$ for a fixed c as one moves along the path. We now split into cases based on the structure of H.

Case 1. H is bipartite

Let A and B be the independent sets with respect to which H is bipartite, and without loss of generality suppose that $v_1 \in A$. Let $c = \alpha_1$. Since H is connected, for any vertex v_j, there exists a walk $v_1 = v_{i_1}, v_{i_2}, ..., v_{i_{k_j}} = v_j$ for some $k_j \in \mathbb{N}$. Thus

$$\alpha_j = (-1)^{k_j - 1}\alpha_1.$$

Because H is bipartite, k_j is even precisely when $v_j \in A$. Therefore

$$\alpha_v = \begin{cases} c & \text{if } v \in A \\ -c & \text{if } v \in B \end{cases}.$$

It follows that

$$\ker \phi = \{\sum_{v \in A} c f_v + \sum_{w \in B} (q-c) f_w\}_{c=0}^{q-1},$$

and thus $\ker \phi \cong \mathbb{Z}_q$ and

$$\Gamma_{qSS}(H) \cong \mathbb{Z}_q^n / \mathbb{Z}_q \cong \mathbb{Z}_q^{n-1}.$$

Case 2. H contains an odd cycle

Suppose H contains a cycle of length j for some odd integer j. Let v_1 be a vertex in that cycle, and let $v_1 = v_{i_1}, v_{i_2}, ..., v_{i_{j+1}} = v_1$ be a walk around the cycle, starting and ending at v_1. Then

$$\alpha_1 = (-1)^j \alpha_k = -\alpha_1.$$

Therefore $2\alpha_1 = 0$.

Subcase 1. q is odd

If q is odd, $2\alpha_1 = 0$ implies $\alpha_1 = 0$. Since H is connected, for every vertex v_i there is a walk from v_1 to v_i. Thus by Equation 3.2,

$$\alpha_i = 0$$

for all i. Thus $\ker \phi = 0$, so $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^n$.

Subcase 2. q is even

If q is even, $q = 2r$, and $2\alpha_1 = 0$ implies $\alpha_1 = 0$, or $\alpha_1 = r$. As above, if $\alpha_1 = 0$, then $\alpha_i = 0$ for all i. Similarly, if $\alpha_1 = r$, $\alpha_i = \pm r$ for all i. But since $q = 2r$, $-r = r$ in $\mathbb{Z}_q$, so $\alpha_i = r$ for all i. Thus

$$\ker \phi = \left\{ 0, \sum_v r f_v \right\} \cong \mathbb{Z}_2.$$

Thus $\ker \phi \cong \mathbb{Z}_2$, and therefore

$$\Gamma_{qSS}(H) \cong \mathbb{Z}_q^n / \mathbb{Z}_2 \cong \mathbb{Z}_q^{n-1} \times \mathbb{Z}_r.$$

□

For calculations later, it will be necessary to explicitly construct the isomorphism. In particular, we will need to determine the image of the elementary switching operators s_v.

Corollary 3.3.4.

1. *When H is bipartite with respect to two disjoint subsets $A, B \subset V$, there exists an isomorphsim $\nu \colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1}$ with $\nu(s_{v_i}) = e_i$ for $i \leq n-1$, and $\nu(s_{v_n}) = \sum_{v_i \in A, i \neq n} e_i - \sum_{v_j \in B, j \neq n} e_j$.*

2. *When H is not bipartite, and q is odd, there exists an isomorphism $\nu \colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^n$ with $\nu(s_{v_i}) = e_i$ for all i.*

3. *When H is not bipartite, and q is even, there exists an isomorphism $\nu \colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1} \times \mathbb{Z}_r$ with $\nu(s_{v_i}) = e_i$ for all i. Note the final generator $e_n = (0, ...0, 1)$ is of order r, whereas all the other e_i are order $q = 2r$.*

Proof. As in the proof of Theorem 3.3.3, we let $\{f_v\}_{v \in V}$ denote the standard basis of $\mathbb{Z}_q^n$, indexed by the vertices of H. We let $\{e_i\}$ denote the standard generator of $\mathbb{Z}_{n_1} \times ... \mathbb{Z}_{n_k}$ with a 1 in the i-th spot and 0 elsewhere.

To construct ν, we simply follow the standard proof of the First Isomorphism Theorem.

Case 1. H is bipartite

The use of the First Isomorphism Theorem in the proof of Theorem 3.3.3 above yields the isomorphism $\widetilde{\phi}\colon Z_q^n/\ker\phi \to \Gamma_{n,q}$ given by

$$\widetilde{\phi}(a+\ker\phi) = \phi(a).$$

Thus $\Gamma_{qSS}(H)$ corresponds to the cosets of

$$\ker\phi = \left\{\sum_{v\in A} cf_v + \sum_{w\in B}(q-c)f_v\right\}_{c=0}^{q-1}.$$

Without loss of generality, suppose that $v_n \in B$. There are q elements in each coset, and we define the map $\psi\colon Z_q^n/\ker\phi \to Z_q^{n-1}$ to be the map sending $a+\ker\phi$ to its representative with the coefficient of $e_n = 0$. This is an isomorphism, and so we define $\nu\colon \Gamma_{n,q} \to \mathbb{Z}_q^{n-1} \times Z_r$ to be the composition of $\widetilde{\phi}^{-1}$ and ψ. Tracing back through the composition of maps we see that $\nu(s_{v_i}) = e_i$ for $i \leq n-1$, and

$$\nu(s_{v_n}) = \sum_{v_i\in A, i\neq n} e_i - \sum_{v_j\in B, j\neq n} e_j.$$

Case 2. H is not bipartite, q is odd

Let $\nu\colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^n$ be the inverse of ϕ, that is $\nu(s_{v_i}) = e_i$ for all i.

Case 3. H is not bipartite, q is even, $q > 2$

As in the first case, the use of the First Isomorphism Theorem yields the isomorphism $\widetilde{\phi}\colon Z_q^n/\ker\phi \to \Gamma_{qSS}(H)$ given by

$$\widetilde{\phi}(a+\ker\phi) = \phi(a).$$

Thus $\Gamma_{qSS}(H)$ corresponds to the cosets of

$$\ker\phi = \left\{0, \sum_i r f_{v_i}\right\}.$$

There are two elements in each coset, and we define the map $\tau\colon Z_q^n/\ker\phi \to Z_q^{n-1} \times Z_r$ to be the map sending $a+\ker\phi$ to its unique representative with the coefficient of e_n lying in the set $\{0,1,...,r-1\}$. This is an isomorphism, and so we define $\nu\colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1} \times Z_r$ to be the composition of $\widetilde{\phi}^{-1}$ and τ. Then $\nu(s_{v_i}) = e_i$ for all i as required.

Case 4. H is not bipartite, $q = 2$

As above, the use of the First Isomorphism Theorem yields the isomorphism $\widetilde{\phi}\colon Z_q^n/\ker\phi \to \Gamma_{qSS}(H)$ given by

$$\widetilde{\phi}(a + \ker\phi) = \phi(a).$$

Thus, $\Gamma_{qSS}(H)$ corresponds to the cosets of

$$\ker\phi = \left\{0, \sum_v f_v\right\}.$$

There are two elements in each coset, and we define the map $\tau\colon Z_q^n/\ker\phi \to Z_q^{n-1}$ as follows. For each coset $a + \ker\phi$ there is a unique representative with the coefficient of e_n equal to 0. We define $\tau(a + \ker\phi)$ to be the projection of this element to $\mathbb{Z}_2^{n-1}$. This is an isomorphism, and so we define $\nu\colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1} \times Z_r$ to be the composition of $\widetilde{\phi}^{-1}$ and τ. Then $\nu(s_{v_i}) = e_i$ for all $i \leq n-1$, and $\nu(s_{v_n}) = \sum_i e_i$.

□

We have identified $\Gamma_{qSS}(H)$ for all connected H, but it remains to consider the case when H is disconnected. Fortunately, the switching group decomposes as a product of the switching groups of the connected components in the most natural way possible. Suppose that H is the disjoint union of H_1 and H_2. Then for $v \in H_i$, s_v only changes the color of edges in H_i. Therefore there is no interaction between the switching functions on H_1 and H_2. In other words, we have the following;

Proposition 3.3.5.

If H is the disjoint union of two subgraphs H_1 and H_2, then

$$\Gamma_{qSS}(H) \cong \Gamma_q(H_1) \times \Gamma_q(H_2).$$

Proof. We note that if $v \in H_1$, $w \in H_2$, then $s_v \perp s_w$, when viewed as vectors in $\mathbb{Z}_q^{|E|}$. Thus $\text{span}\{s_v\}_{v\in H_1} \perp \text{span}\{s_v\}_{v\in H_2}$, and so

$$\Gamma_q(H) \cong \text{span}\{s_u\}_{u\in H} \cong \text{span}\{s_v\}_{v\in H_1} \oplus \text{span}\{s_v\}_{v\in H_2} \cong \Gamma_q(H_1) \times \Gamma_q(H_2)$$

□

We can now give a classification of the structure of the abelian groups that are isomorphic to $\Gamma_{qSS}(H)$ for any q and H.

Theorem 3.3.6.

1. *Let H be a graph on n vertices. Then there exist $b, c \geq 0$ such that $2b+3c \leq n$ and*
$$\Gamma_{qSS}(H) \cong \begin{cases} \mathbb{Z}_q^{n-b-c} \times \mathbb{Z}_r^j & \text{if } q = 2r \\ \mathbb{Z}_q^{n-c} & \text{if } q \text{ is odd} \end{cases}$$
2. *If $q = 2r$, and $b, c \geq 0$ such that $2b + 3c \leq n$, then there exists a graph H on n vertices such that $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{n-b-c} \times \mathbb{Z}_r^j$. If q is odd and $2b \leq n$, then there exists a graph H on n vertices such that $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{n-b}$*

Proof. Let $H_1 \times ... \times H_k$ be the connected components of H, and let $n_1, ..., n_k$ be the number of vertices in each component. Then by Proposition 3.3.5, $\Gamma_{qSS}(H) \cong \Gamma_{H_1,q} \times ... \times \Gamma_{H_k,q}$. We first consider the case when q is even. By Theorem 3.3.3,

$$\Gamma_{H_i,q} \cong \begin{cases} Z_q^{n_i-1} \times \mathbb{Z}_r & \text{if } H_i \text{ is not bipartite} \\ \mathbb{Z}_q^{n_i-1} & \text{if } H \text{ is bipartite} \end{cases}.$$

Let b be the number of non-bipartite connected components and c the number of bipartite connected components. Then $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{n-b-c} \times \mathbb{Z}_r^j$. For a component H_i to be bipartite, $n_i \geq 2$, and to be non-bipartite, $n_i \geq 3$. Thus $2b + 3c \leq n$. The case when q is odd is simpler; let $b = 0$ and let c be the number of bipartite components. Then since

$$\Gamma_{H_i,q} \cong \begin{cases} Z_q^{n_i} & \text{if } H_i \text{ is not bipartite} \\ \mathbb{Z}_q^{n_i-1} & \text{if } H \text{ is bipartite} \end{cases},$$

$\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{n-c}$.

For the other half of the proof, suppose $q = 2r$ and $2b+3c \leq n$. Let H be the disjoint union of c 3-cycles, $b-1$ edges, and one path of length $n+2-2b-3c$. Then H is a graph on n vertices and $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{2j} \times \mathbb{Z}_r^j \times \mathbb{Z}_q^{i-1} \times \mathbb{Z}_q^{n+1-2b-3c} \cong \mathbb{Z}_q^{n-b-c} \times \mathbb{Z}_r^j$.

If q is odd, and $2b \leq n$, let H be a disjoint union of $b-1$ edges, and one path of length $n+2-2b$. Then H is a graph on n vertices with switching group $\Gamma_{qSS}(H) \cong \mathbb{Z}_q^{b-1} \times \mathbb{Z}_q^{n+1-2b} \cong \mathbb{Z}_q^{n-b}$ □

3.3.2 Random q-H-Switching

We wish to consider the process generated by a sequence of randomly chosen q-H-switching operators. We introduce one further generalization. Let H be a connected graph with vertex set V, and $q \geq 2$ a positive integer. While in Section 3.2 the operators were chosen uniformly at random, we now allow an arbitrary probability distribution on the vertices. Formally, let $\omega \colon V \to (0,1)$ be a probability distribution on V.

Let $\tau_\emptyset$ refer to the 0 coloring of the edges; that is, $\tau_\emptyset(e) = 0$ for all edges $e \in E$. For an elementary switching operation s_v, recall that its inverse $s_v^{-1} = s_v^{q-1}$. Also, recall that $s_\emptyset$ refers to the identity operator.

Let $\{x_t\}_1^\infty$ be a sequence of independent identically distributed random vertices, where $\mathbb{P}(x_t = v) = \omega(v)$. We consider the following random process. Starting from the empty coloring, we pick a vertex at random and apply either the q-H-switching operator or its inverse (with equal probability) at that vertex. To guarantee convergence, we consider a lazy random process where half the time no action is taken. Formally, this is the sequence $X(t)$ where

$$X(0) = \tau_\emptyset,$$

$$X(t+1) = \sigma_t(X(t))$$

where $\{\sigma_t\}$ is a sequence of random independently chosen switching operators where for $q > 2$,

$$\mathbb{P}(\sigma_t = s_\emptyset) = \frac{1}{2},$$

$$\mathbb{P}(\sigma_t = s_v) = \frac{\omega(v)}{4},$$

and

$$\mathbb{P}(\sigma_t = s_v^{-1}) = \frac{\omega(v)}{4}.$$

When $q = 2$, then $s_v^{-1} = s_v$, and so

$$\mathbb{P}(\sigma_t = s_\emptyset) = \frac{1}{2},$$

$$\mathbb{P}(\sigma_t = s_v) = \frac{\omega(v)}{2}.$$

Our immediate goal is to interpret $X(t)$ as a random walk on a weighted graph, then use the myriad of tools available for studying such processes. For an arbitrary probability distribution ω on V, we define the symmetric distribution $\hat{\omega}$ on $\Gamma_{qSS}(H)$ by

$$\hat{\omega}(\gamma) = \begin{cases} \frac{\omega(v)}{2} & \text{if } \gamma = s_v, v \in V \\ \frac{\omega(v)}{2} & \text{if } \gamma = s_v^{q-1}, v \in V \\ 0 & \text{otherwise} \end{cases}.$$

We define the *switching state graph* to be the directed, weighted graph

$$C_{qSS}(H) = \mathrm{Cay}(\Gamma_{\mathrm{qSS}}(\mathrm{H}), \hat{\omega}).$$

Thus $C_{qSS}(H)$ has weighted adjacency matrix given by

$$A(\tau_i, \tau_j) = \hat{\omega}(\tau_i \circ \tau_j^{-1}) = \begin{cases} \frac{\omega(v)}{2} & \text{if } \tau_j = s_v(\tau_i), v \in V \\ \frac{\omega(v)}{2} & \text{if } \tau_j = s_v^{-1}(\tau_i), v \in V \\ 0 & \text{otherwise} \end{cases}$$

That is, the weight on the edge from τ_i to τ_j corresponds to the elementary operator or its inverse that sends τ_i to τ_j.

We will show that $X(t)$ is equal in distribution to a lazy random walk on the weighed graph $C_{qSS}(H)$ starting at the empty coloring. In order to guarantee that $C_{qSS}(H)$ is connected we only consider distributions ω whose support on G is a generating set.

Let W be the transition matrix of the lazy random walk on $C_{qSS}(H)$.

Proposition 3.3.7.

For any coloring τ_i in the q-H switching class of $\tau_\emptyset$, $\mathbb{P}(X(t) = \tau_i) = \mathbf{1}_{\{\tau_\emptyset\}} W^t(\tau_i)$, where $\mathbf{1}_{\{\tau_\emptyset\}}$ a row vector with a 1 in the entry for $\tau_\emptyset$ and 0 elsewhere.

Proof. We must show that

$$P(X(t+1) = \tau_i \mid X(t) = \tau_j) = W(\tau_j, \tau_i)$$

for any t. W is the transition matrix of a lazy random walk on $C_{qSS}(H)$, so $W = \frac{1}{2}(I + D^{-1}A)$, where I is the identity matrix, D the weighted diagonal degree matrix, and A the weighted adjacency matrix of $C_{qSS}(H) = \mathrm{Cay}(\Gamma_{\mathrm{qSS}}(\mathrm{H}), \omega)$. Since ω is a probability distribution, the weighted out-degree of each vertex is 1, and so $W = \frac{1}{2}(I + A)$. If $\tau_i = \tau_j$, then $W(\tau_i, \tau_i) = \frac{1}{2} = \mathbb{P}(X(t+1) = \tau_i \mid X(t) = \tau_i)$. For $\tau_i \neq \tau_j$,

$$W(\tau_i, \tau_j) = \frac{1}{2}A(\tau_i, \tau_j) = \begin{cases} \frac{\omega(v)}{4} & \text{if } s_v(\tau_i) = \tau_j \text{ for some } v \in V \\ \frac{\omega(v)}{4} & \text{if } s_v^{-1}(\tau_i) = \tau_j \text{ for some } v \in V \\ 0 & \text{otherwise} \end{cases}.$$

But this is precisely equal to $\mathbb{P}(X(t+1) = \tau_j | X(t) = \tau_i)$. We have demonstrated that $X(t)$ is a Markov chain with transition matrix W, proving the proposition. □

Since we can now understand $X(t)$ as the lazy random walk on a connected graph, it remains to analyze that graph. Since $C_{qSS}(H)$ is a Cayley graph with respect to the symmetric distribution ω, it is a connected, vertex transitive, undirected, weighted graph. Thus a lazy random walk converges to the uniform distribution π. It remains to understand how quickly the walk converges.

We remark that in the case where q is odd and H is not bipartite, this reduces to the case of a geometric random walk on $\mathbb{Z}_q^n$, see for example [49].

By Theorem 1.3.9, the $\chi-$squared distance can be calculated as

$$\Delta'(t) = \left(\sum_{i \neq n} \lambda_i^{2t} \right)^{1/2},$$

where $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n = 1$ are the eigenvalues of the transition matrix. Thus to understand the rate of convergence of $X(t)$ to its stationary distribution it remains only to understand the eigenvalues of W.

Theorem 3.3.8.

Let H be a connected graph on n vertices, ω a probability distribution on V, and let $\theta_q = e^{\frac{2\pi i}{q}}$ denote a qth root of unity. Let W be the transition matrix of the random walk on the state graph $C_{qSS}(H)$. Then the spectrum of W depends on the parity of q and the structure of H as follows:

1. *If $q = 2$, there is one eigenvalue of W corresponding to each vector $\overrightarrow{x} \in \mathbb{Z}_2^{n-1}$ with*
$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \omega(v_n)\prod_{i=1}^{n-1}(-1)^{x_i} + \sum_{i=1}^{n-1}\omega(v_i)(-1)^{x_i}\right)$$

2. *If H is bipartite with respect to subsets $A, B \subset V$, then there is one eigenvalue of W corresponding to each vector $\overrightarrow{x} \in \mathbb{Z}_q^{n-1}$ with*
$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \omega(v_n)\Re\left(\frac{\prod_{v_i \in A, i \neq n}\theta_q^{x_i}}{\prod_{v_j \in B, j \neq n}\theta_q^{x_i}}\right) + \sum_{i=1}^{n-1}\omega(v_i)\Re(\theta_q^{x_i})\right)$$
where $\Re(\cdot)$ denotes the real part.

3. *If H is not bipartite and q is odd, then there is one eigenvalue of W corresponding to each vector $\overrightarrow{x} \in \mathbb{Z}_q^{n-1}$ with*
$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \sum_{i=1}^{n}\omega(v_i)\Re(\theta_q^{x_i})\right).$$

4. *If H is not bipartite and $q \geq 4$ is even, $q = 2r$, then there is one eigenvalue of W corresponding to each vector $\overrightarrow{x} \in \mathbb{Z}_q^{n-1} \times \mathbb{Z}_r$ with*
$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \sum_{i=1}^{n-1}\omega(v_i)\Re(\theta_q^{x_i}) + \omega(v_n)\Re(\theta_r^{x_n})\right).$$

Proof. By Theorem 1.4.3, the eigenvalues of W are
$$\lambda_\rho = \frac{1}{2} + \sum_{g \in \Gamma}\hat{\omega}(g)\rho(-g)$$
where ρ is a one dimensional irreducible representation of $\Gamma_{qSS}(H)$. Theorem 1.4.4 states the irreducible representations of $\mathbb{Z}_q^n$ are the functions $\rho_{\overrightarrow{x}}$ for $\overrightarrow{x} \in \mathbb{Z}_q^{n-1}$

where

$$\rho_{\overrightarrow{x}}(\sum_{i=1}^{n-1}\alpha_i e_i) = \prod_{i=1}^{n-1}\theta_q^{x_i\alpha_i}.$$

We will handle each of the cases separately, though the techniques used for each case are identical.

Case 1. $q = 2$

The irreducible representations of $\Gamma_{qSS}(H)$ are $\rho_{\overrightarrow{x}} \circ \nu$ where $\nu\colon \Gamma_{qSS}(H) \to \mathbb{Z}_2^{n-1}$ is the isomorphism defined by Corollary 3.3.4 with

$$\nu(s_{v_i}) = e_i \text{ for } i \leq n-1$$
$$\nu(s_{v_n}) = \sum_{i=1}^{n-1} e_i.$$

Let $\tilde{\rho}_{\overrightarrow{x}}\colon \Gamma_{qSS}(H) \to \mathbb{C}$ be the composition for $\rho_{\overrightarrow{x}}$ and ν. Since the $\tilde{\rho}_{\overrightarrow{x}}$ are the irreducible characters of $\Gamma_{qSS}(H)$, we have that for each $\overrightarrow{x} \in \mathbb{Z}_2^{n-1}$ there is an eigenvalue

$$\begin{aligned}\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2}\sum_{g\in\Gamma_{qSS}(H)} \hat{\omega}(g)\tilde{\rho}_{\overrightarrow{x}}(-g) &= \frac{1}{2} + \frac{1}{2}\sum_{v} \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v) + \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v^{-1}) \\ &= \frac{1}{2} + \frac{1}{4}\sum_{i=1}^{n} \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1})) \\ &= \frac{1}{2} + \frac{1}{2}\sum_{i=1}^{n} \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}))\end{aligned}$$

Let $\zeta_{\overrightarrow{x}}(i) = \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}))$. Then

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2}\sum_{i=1}^{n} \zeta_{\overrightarrow{x}}(i).$$

For $i \leq n-1$,

$$\begin{aligned}\zeta_{\overrightarrow{x}}(i) &= \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i})) \\ &= \omega(v_i)\rho_{\overrightarrow{x}}(e_i) \\ &= \omega(v_i)(-1)^{x_i}\end{aligned}$$

It remains to calculate $\zeta_{\overrightarrow{x}}(n)$.

$$\begin{aligned}\zeta_{\overrightarrow{x}}(n) &= \omega(v_n)\rho_{\overrightarrow{x}}(\nu(s_{v_n})) \\ &= \omega(v_n)\rho_{\overrightarrow{x}}\left(\sum_{i=1}^{n-1} e_i\right) \\ &= \omega(v_n)\prod_{i=1}^{n-1}(-1)^{x_i}\end{aligned}$$

Thus

$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \omega(v_n)\prod_{i=1}^{n-1}(-1)^{x_i} + \sum_{i=1}^{n-1}\omega(v_i)(-1)^{x_i}\right)$$

Case 2. H is bipartite

Let $A, B \subset V$ be the subsets of vertices with respect to which H is bipartite. Thus the irreducible representations of $\Gamma_{qSS}(H)$ are $\rho_{\overrightarrow{x}} \circ \nu$ where $\nu\colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1}$ is the isomorphism defined in Corollary 3.3.4 with

$$\begin{aligned}\nu(s_{v_i}) &= e_i \text{ for } i \leq n-1 \\ \nu(s_{v_n}) &= \sum_{v_i \in A, i \neq n} e_i - \sum_{v_j \in B, j \neq n} e_j.\end{aligned}$$

Let $\tilde{\rho}_{\overrightarrow{x}}\colon \Gamma_{qSS}(H) \to \mathbb{C}$ be the composition for $\rho_{\overrightarrow{x}}$ and ν. Since the $\tilde{\rho}_{\overrightarrow{x}}$ are the irreducible characters of $\Gamma_{qSS}(H)$, we have that for each $\overrightarrow{x} \in \mathbb{Z}_q^n$ there is an eigenvalue

$$\begin{aligned}\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2}\sum_{g \in \Gamma_{qSS}(H)} \hat{\omega}(g)\tilde{\rho}_{\overrightarrow{x}}(-g) &= \frac{1}{2} + \frac{1}{2}\sum_{v} \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v) + \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v^{-1}) \\ &= \frac{1}{2} + \frac{1}{4}\sum_{i=1}^{n} \omega(v_i)\rho_{\overrightarrow{x}}\left(\nu(s_{v_i})\right) + \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\end{aligned}$$

Let $\zeta_{\overrightarrow{x}}(i) = \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right)$. Then

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{4}\sum_{i=1}^{n} \zeta_{\overrightarrow{x}}(i).$$

For $i \leq n-1$,

$$\begin{aligned}
\zeta_{\overrightarrow{x}}(i) &= \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right) \\
&= \omega(v_i)\left(\rho_{\overrightarrow{x}}(e_i) + \rho_{\overrightarrow{x}}(-e_i)\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \frac{1}{\theta_q^{x_i}}\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \overline{\theta_q^{x_i}}\right) \\
&= \omega(v_i) 2\Re(\theta_q^{x_i})
\end{aligned}$$

It remains to calculate $\zeta_{\overrightarrow{x}}(n)$.

$$\begin{aligned}
\zeta_{\overrightarrow{x}}(n) &= \omega(v_n)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_n})) + \rho_{\overrightarrow{x}}(\nu(s_{v_n}^{-1}))\right) \\
&= \omega(v_n)\left(\rho_{\overrightarrow{x}}(\sum_{v_i \in A, i \neq n} e_i - \sum_{v_j \in B, j \neq n} e_j) + \rho_{\overrightarrow{x}}(\sum_{v_i \in A, i \neq n} -e_i + \sum_{v_j \in B, j \neq n} e_j)\right) \\
&= \omega(v_n)\left(\frac{\prod\limits_{v_i \in A, i \neq n} \theta_q^{x_i}}{\prod\limits_{v_j \in B, j \neq n} \theta_q^{x_j}} + \frac{\prod\limits_{v_j \in B, j \neq n} \theta_q^{x_j}}{\prod\limits_{v_i \in A, i \neq n} \theta_q^{x_i}}\right) \\
&= \omega(v_n) 2\Re\left(\frac{\prod\limits_{v_i \in A, i \neq n} \theta_q^{x_i}}{\prod\limits_{v_j \in B, j \neq n} \theta_q^{x_j}}\right)
\end{aligned}$$

Thus

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2}\left(\omega(v_n)\Re\left(\frac{\prod\limits_{v_i \in A, i \neq n} \theta_q^{x_i}}{\prod\limits_{v_j \in B, j \neq n} \theta_q^{x_j}}\right) + \sum_{i=1}^{n-1} \omega(v_i)\Re(\theta_q^{x_i})\right)$$

Case 3. H not bipartite, q odd

The irreducible representations of $\Gamma_{qSS}(H)$ are $\rho_{\overrightarrow{x}} \circ \nu$ where $\nu \colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^n$ is the isomorphism defined in Corollary 3.3.4 with

$$\nu(s_{v_i}) = e_i \text{ for all } n.$$

Let $\tilde{\rho}_{\overrightarrow{x}} \colon \Gamma_{qSS}(H) \to \mathbb{C}$ be the composition for $\rho_{\overrightarrow{x}}$ and ν. Since the $\tilde{\rho}_{\overrightarrow{x}}$ are the irreducible characters of $\Gamma_{qSS}(H)$, we have that for each $\overrightarrow{x} \in \mathbb{Z}_q^n$ there is an

eigenvalue

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2} \sum_{g \in \Gamma_{qSS}(H)} \hat{\omega}(g)\tilde{\rho}_{\overrightarrow{x}}(-g) = \frac{1}{2} + \frac{1}{2} \sum_{v} \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v) + \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v^{-1})$$
$$= \frac{1}{2} + \frac{1}{4} \sum_{i=1}^{n} \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))$$

Let $\zeta_{\overrightarrow{x}}(i) = \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right)$. Then

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{4} \sum_{i=1}^{n} \zeta_{\overrightarrow{x}}(i).$$

In this case,

$$\begin{aligned}
\zeta_{\overrightarrow{x}}(i) &= \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right) \\
&= \omega(v_i)\left(\rho_{\overrightarrow{x}}(e_i) + \rho_{\overrightarrow{x}}(-e_i)\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \frac{1}{\theta_q^{x_i}}\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \overline{\theta_q^{x_i}}\right) \\
&= \omega(v_i) 2\Re(\theta_q^{x_i}).
\end{aligned}$$

Thus,

$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \sum_{i=1}^{n} \omega(v_i)\Re(\theta_q^{x_i})\right).$$

Case 4. H not bipartite, $q \geq 4$ is even

Let $q = 2r$. The irreducible representations of $\Gamma_{qSS}(H)$ are $\rho_{\overrightarrow{x}} \circ \nu$ where $\nu \colon \Gamma_{qSS}(H) \to \mathbb{Z}_q^{n-1} \times \mathbb{Z}_r$ is the isomorphism defined in Corollary 3.3.4 with

$$\nu(s_{v_i}) = e_i \text{ for } i \leq n-1$$
$$\nu(s_{v_n}) = e_n.$$

Recall that e_n here is the generator of $\mathbb{Z}_r$ and is of order r, whereas the other e_i are of order q.

Let $\tilde{\rho}_{\overrightarrow{x}} \colon \Gamma_{qSS}(H) \to \mathbb{C}$ be the composition for $\rho_{\overrightarrow{x}}$ and ν. Since the $\tilde{\rho}_{\overrightarrow{x}}$ are the irreducible characters of $\Gamma_{qSS}(H)$, we have that for each $\overrightarrow{x} \in \mathbb{Z}_q^{n-1} \times \mathbb{Z}_r$ there

is an eigenvalue

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{2} \sum_{g \in \Gamma_{qSS}(H)} \hat{\omega}(g)\tilde{\rho}_{\overrightarrow{x}}(-g) = \frac{1}{2} + \frac{1}{2} \sum_{v} \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v) + \hat{\omega}(v)\tilde{\rho}_{\overrightarrow{x}}(s_v^{-1})$$
$$= \frac{1}{2} + \frac{1}{4} \sum_{i=1}^{n} \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \omega(v_i)\rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))$$

Let $\zeta_{\overrightarrow{x}}(i) = \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right)$. Then

$$\lambda_{\overrightarrow{x}} = \frac{1}{2} + \frac{1}{4} \sum_{i=1}^{n} \zeta_{\overrightarrow{x}}(i).$$

For $i \leq n-1$,

$$\begin{aligned}
\zeta_{\overrightarrow{x}}(i) &= \omega(v_i)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_i}^{-1}))\right) \\
&= \omega(v_i)\left(\rho_{\overrightarrow{x}}(e_i) + \rho_{\overrightarrow{x}}(-e_i)\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \frac{1}{\theta_q^{x_i}}\right) \\
&= \omega(v_i)\left(\theta_q^{x_i} + \overline{\theta_q^{x_i}}\right) \\
&= \omega(v_i) 2\Re(\theta_q^{x_i})
\end{aligned}$$

We calculate $\zeta_{\overrightarrow{x}}(n)$ in the same way.

$$\begin{aligned}
\zeta_{\overrightarrow{x}}(n) &= \omega(v_n)\left(\rho_{\overrightarrow{x}}(\nu(s_{v_i})) + \rho_{\overrightarrow{x}}(\nu(s_{v_n}^{-1}))\right) \\
&= \omega(v_n)\left(\rho_{\overrightarrow{x}}(e_i) + \rho_{\overrightarrow{x}}(-e_i)\right) \\
&= \omega(v_n)\left(\theta_r^{x_i} + \frac{1}{\theta_r^{x_i}}\right) \\
&= \omega(v_n)\left(\theta_r^{x_i} + \overline{\theta_r^{x_i}}\right) \\
&= \omega(v_n) 2\Re(\theta_r^{x_i})
\end{aligned}$$

Thus

$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \sum_{i=1}^{n-1} \omega(v_i)\Re(\theta_q^{x_i}) + \omega(v_n)\Re(\theta_r^{x_n})\right).$$

□

Combining Theorem 1.3.9 and Theorem 3.3.8 yields the following bounds on convergence times of the random generalized switching process.

Theorem 3.3.9.
Let H be a connected graph on n vertices, and let ω be a probability distribution on the vertices. Let $\theta_q = e^{\frac{2\pi i}{q}}$ denote a qth root of unity. The random walk on $C_{qSS}(H)$, and hence the random q-H-switching process after t steps has χ^2 distance from the stationary distribution bounded as follows:

1. *If $q = 2$,*

$$\Delta'(t) \leq \frac{1}{2^t}\left[\sum_{\overrightarrow{x}\in\mathbb{Z}_2^{n-1}}\left(1+\omega(v_n)\prod_{i=1}^{n-1}(-1)^{x_i}+\sum_{i=1}^{n-1}\omega(v_i)(-1)^{x_i}\right)^{2t}\right]^{\frac{1}{2}}$$

2. *If H is bipartite with respect to subsets $A, B \subset V$,*

$$\Delta'(s) \leq \frac{1}{2^t}\left[\sum_{\overrightarrow{x}\in\mathbb{Z}_q^{n-1}}\left(1+\omega(v_n)\Re\left(\frac{\prod_{v_i\in A, i\neq n}\theta_q^{x_i}}{\prod_{v_j\in B, j\neq n}\theta_q^{x_j}}\right)+\sum_{i=1}^{n-1}\omega(v_i)\Re(\theta_q^{x_i})\right)^{2t}\right]^{\frac{1}{2}}$$

3. *If H is not bipartite and q is odd,*

$$\Delta'(t) \leq \frac{1}{2^t}\left[\sum_{\overrightarrow{x}\in\mathbb{Z}_q^{n}}\left(1+\sum_{i=1}^{n}\omega(v_i)\Re(\theta_q^{x_i})\right)^{2t}\right]^{\frac{1}{2}}$$

4. *If H is not bipartite and $q \geq 4$ is even, $q = 2r$,*

$$\Delta'(t) \leq \frac{1}{2^t}\left[\sum_{\overrightarrow{x}\in\mathbb{Z}_q^{n-1}\times\mathbb{Z}_r}\left(1+\sum_{i=1}^{n-1}\omega(v_i)\Re(\theta_q^{x_i})+\Re(\theta_r^{x_n})\right)^{2t}\right]^{\frac{1}{2}}$$

Corollary 3.3.10.

When $q = 2$ and $\omega(v_i) = \frac{1}{n}$, and $H = K_n$ then we are considering randomized Seidel Switching as discussed in Section 3.2. In this case, we obtain the bounds

1. $\Delta'(t) \leq \left(\sum_{j=1}^{n-1} \binom{n-1}{j} (1 - \frac{j}{n})^{2t} \right)^{\frac{1}{2}}$
2. $\Delta'(t) \leq e^{-c}$ *if* $t > \frac{1}{2} n \log n + c$.

Proof. There will be $\binom{n-1}{j}$ vectors $\overrightarrow{x} \in \mathbb{Z}_2^{n-1}$ with j 1's, and $(n-1-j)$ 0's. For each of these, $\lambda_{\overrightarrow{x}} = \frac{1}{2}(1 + \frac{1}{n}(n-1-2j) + \frac{(-1)^j}{n} \leq (1 - \frac{j}{n})$.

For the second fact, we note that

$$\begin{aligned}
\left(\sum_{j=1}^{n-1} \binom{n-1}{j} \left(1 - \frac{j}{n}\right)^{2t} \right)^{\frac{1}{2}} &\leq \left(\sum_{j=1}^{n-1} e^{j \log(n-1) - \frac{2jt}{n}} \right)^{\frac{1}{2}} \\
&\leq \left((n-1) e^{(n-1)\log(n-1) - \frac{2(n-1)t}{n}} \right)^{\frac{1}{2}} \\
&= e^{\frac{(n-1)\log^2(n-1)}{2} - \frac{(n-1)\log(n-1)t}{n}} \\
&\leq e^{-c}
\end{aligned}$$

if

$$\frac{(n-1)\log^2(n-1)}{2} - \frac{(n-1)\log(n-1)t}{n} \leq -c. \tag{3.3}$$

Solving for t and simplifying shows that (3.3) is satisfied when

$$t > \frac{1}{2} n \log n + c.$$

□

Acknowledgement:

This chapter is based on the paper "Random Seidel Switching on Graphs" [37]. It has been accepted for publication in the Journal of Combinatorial Mathematics and Combinatorial Computing.

Chapter 4

Random Lights Out Processes

4.1 Introduction

Lights Out is a single player game played on graph. The game begins with a two-coloring of a graph, with vertices assigned a color of "0" or "1." The goal is to get all the vertices colored 0 via sequence of vertex toggles. Each vertex toggle changes state of that vertex and all of its neighbors. This name comes from an electronic toy introduced 1995 by Tiger Electronics, though the game had been studied previously due to its connection to linear cellular autometa [59]. It is also known as Sutner's σ^+ game [60].

The game can be played on any finite graph G. Let $G = G(V, E)$ be a graph on n vertices. Let $\mathfrak{F}(G)$ denote the set of functions $\tau\colon V \to \mathbb{Z}_2$, which we think of as colorings of the vertices of G. The "all 0 coloring" $\tau_\emptyset$ will refer to the function with $\tau_\emptyset(v) = 0$ for all $v \in V$. For each vertex v, we define the toggling operator $s_v\colon \mathfrak{F}(G) \to \mathfrak{F}(G)$ to be the operator that changes the color of v and its neighbors. We call the action of applying s_v "toggling vertex v", which switches the color of v and all its neighbors. Formally,

$$s_v(\tau(w)) = \begin{cases} \tau(w) + 1 & \text{if } \{w, v\} \in E \text{ or } w = v \\ \tau(w) & \text{if } \{w, v\} \notin E \end{cases}$$

where the addition occurs in $\mathbb{Z}_2$. We let $s_\emptyset$ denote the identity operator. A coloring τ is *winnable* if there are a sequence of vertices such that toggling them will

transform τ to the all 0 coloring $\tau_\emptyset$. A graph G is called *always winnable* if all colorings $\tau \in \mathfrak{F}(G)$ are winnable.

Let $r_2 = r_2(G) = \text{Rank}_{\mathbb{Z}_2}(A + I)$, where A is the adjacency matrix of G. This is equal to $n - PD(G)$, where $PD(G)$ is the parity dimension of G investigated in [4, 5]. $r_2(G)$ will be a key parameter in investigating the Lights Out game. For example, the following lemma follows from [4].

Lemma 4.1.1.
A graph G on n vertices is always winnable if and only if $r_2(G) = n$

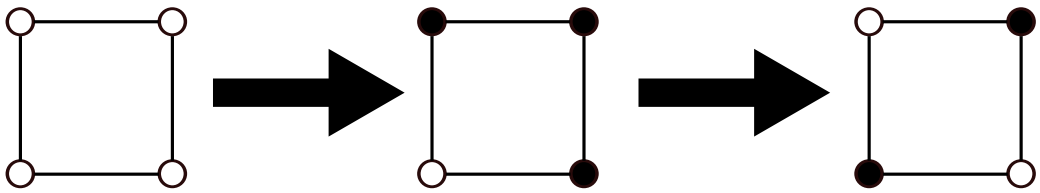


Figure 4.1: An example demonstrating how the colorings change under two switching operations. Beginning with the "all 1" coloring, the top right vertex is toggled, followed by the bottom left.

Many authors have investigated what graphs families of graphs are always winnable or which colorings are winnable for particular families of graphs. Sutner [59] showed that for any finite graph, the "all ones" coloring is winnable. The Lights Out electronic toy consisted of a 5×5 grid of buttons, and this version on a 5×5 grid was studied by Anderson and Feil [7] who used linear algebra over $\mathbb{Z}_2$ to study configurations where it is possible to win. Goldwasser, Klostermeyer, and Trapp [34] studied the game on $n \times m$ grids. Using Fibonacci Polynomials, they classified which values of m and n yield an always winnable grid. Amin, Clark, and Slater [4] studied this game on several families of graphs including trees, paths, cycles, wheels, and fans. Their study was graph theoretic and focused on the parity dimension of a graph, which is the rank of the nullspace of the matrix $A+I$, where A is the adjacency matrix of the graph. In a follow up paper [5], Amin, Slater,

and Zhang used an linear algebraic approach to expand upon their results, finding the parity dimension for products of graphs. Giffen and Parker [33] considered a generalization of the puzzle with multiple colors possible for each vertex. Edwards et. al [29] generalized the techniques of both Amin, Clark, and Slater to the multicolor generalization of Giffen and Parker to classified winnable configurations for several families of graphs. Gervacio and Maehara [32] classified all the *parity state graphs*, for which there solvability of a coloring τ depends only on the parity of the number of 1's assigned. Goldwasser, Wang, and Wu [35, 62] considered a variant where one is only allow to toggle vertices colored 1, which we will discuss briefly in the conclusion.

This previous work on the Lights Out Puzzle involved studying the parity dimension of graphs and classifying graphs that are always winnable. In this chapter, we consider a different problem. Given a fixed graph G, we will consider the stochastic process arising from a sequence of random vertex toggles. This is the first treatment of this random Lights Out process. We investigate the stationary distribution of this process, the speed of convergence to the stationary distribution, and the hitting times.

Our approach can be summarized as follows:

1. The random lights out process on a finite graph G, $X(t)$, has the same distribution as a random walk on the state graph of the configuration space, denoted $C_{LO}(G)$ (Lemma 4.2.1)

2. The state graph $C_{LO}(G)$ is isomorphic to a Cayley graph of $\mathbb{Z}_2^{r_2}$ (Corollary 4.2.6)

3. The eigenvalues of the transition matrix can be determined using the irreducible representations of $\mathbb{Z}_2^{r_2}$. (Lemma 4.3.2, Corollary 4.3.3)

4. A lazy version of the process converges to the uniform distribution. The $\chi-$squared distance between the distribution of $X(t)$ and the uniform distribution satisfies $\Delta'(t) \leq e^{-c}$ for all $t > \frac{n}{2}(1 + \frac{1}{r_2}) \log(r_2) + c$ (Lemma 4.3.1, Theorem 4.3.4)

5. Let T_0 be the hitting time of the empty coloring, that is $\min\{t|X_\tau(t) = \tau_\emptyset\}$. Then $2^{r_2(G)-1} \leq \mathbb{E}[T_0] \leq 2n\frac{r_2}{r_2-2-\sqrt{2}}\binom{r_2}{\lfloor \frac{r_2}{2} \rfloor}$ (Theorem 4.3.5)

4.2 Characterizing the State Graph

Let $G = G(V, E)$ be a finite graph on n vertices. We wish to consider the random process arising from applying a sequence of lights out toggles to the vertices of G. Let $\{x_t\}_1^\infty$ be a sequence of independent random vertices, each chosen uniformly among all vertices. Consider initially the random process $\hat{X}(t)$ where

$$\hat{X}(0) = \tau_\emptyset,$$
$$\hat{X}(t+1) = s_{x_t}\left(\hat{X}(t)\right).$$

We will analyze this process by viewing it as a random walk on a multigraph. Let $C_{LO}(G)$ be the multigraph with vertex set consisting of winnable colorings of G and edge set with an edge between colorings τ_1 and τ_2 for each vertex $v \in G$ such that toggling v changes between τ_1 and τ_2.

The following lemma is the key to our analysis of this process, reducing our study of this process to that of a random walk on a graph.

Lemma 4.2.1.
The random lights out process $\hat{X}(t)$ on a finite graph G has the same distribution as a random walk on $C_{LO}(G)$ starting at $\tau_\emptyset$.

Proof. Let $Y(t)$ the location of a random walk on $C_{LO}(G)$ after t steps. Note that $\hat{X}(0) = Y(0)$, and that for any $t > 0$ and colorings $\tau_1 \neq \tau_2$,

$$\begin{aligned}\mathbb{P}(\hat{X}(t+1) = \tau_1|\hat{X}(t) = \tau_2) &= \frac{|\{v \in V|s_v(\tau_2) = \tau_1\}|}{n} \\ &= \mathbb{P}(Y(t+1) = \tau_1|Y(t) = \tau_2).\end{aligned}$$

□

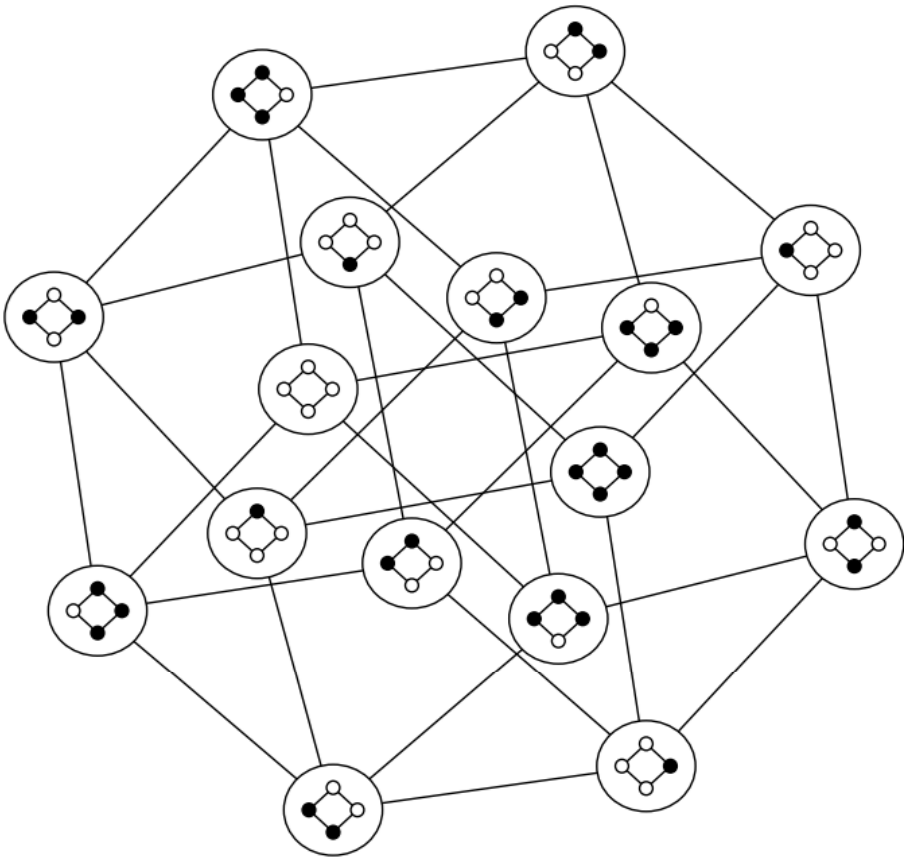


Figure 4.2: The Lights Out state graph $C_{LO}(G)$ for the 4-cycle. The vertices of the state graph are the winnable colorings of the 4-cycle. Two graphs are adjacent if they differ by the toggling of a single vertex. Note that the graph is bipartite, which demonstrates why we consider a lazy random walk to guarantee convergence.

In general, this process may fail to be ergodic. For example, consider the state graph for the lights out process of the cycle on 4 vertices shown in Figure **??**. Since the graph is bipartite, a random walk will be periodic and hence not ergodic and therefore may not converge a stationary distribution. To eliminate this concern we consider a "lazy version" of this process where half the time no action is taken. We define the sequence $X(t)$ where

$$X(0) = \tau_{\emptyset},$$

$$X(t+1) = s_{y_t}(X(t))$$

where

$$\mathbb{P}(y_t = \emptyset) = \frac{1}{2},$$
$$\mathbb{P}(y_t = x_t) = \frac{1}{2}.$$

Lemma 4.2.2.

A lazy random walk on $C_{LO}(G)$ is ergodic.

Proof. The vertices of $C_{LO}(G)$ are winnable colorings, so by definition there is a path from the all off coloring $\tau_\emptyset$ to every vertex τ, so the graph is irreducible. Aperiodicity comes from the fact that a lazy random walk has $W(v, v) = \frac{1}{2} > 0$ for all vertices. □

The choice to consider a lazy random walk as opposed to a random walk is to ensure aperiodicity.

Since we have reduced the study of the random lights out process $X(t)$ to the study of a lazy random walk on $C_{LO}(G)$, we now characterize the structure of $C_{LO}(G)$. We begin with the following observation about the toggling operations s_v.

Lemma 4.2.3.

Let $\Gamma_{LO}(G)$ denote the set of all compositions of the switching operators s_v on a graph $G = G(V, E)$ along with identity $s_\emptyset$. Then $\Gamma_{LO}(G)$ is an abelian group with respect to composition.

Before proving this lemma, we note that this is the main technique of Anderson and Feil in [7], that of turning the Lights Out game on a 5×5 grid into a problem in linear algebra.

Proof. If we view a coloring $\tau \in \mathfrak{F}(G)$ as a vector in $\mathbb{Z}_2^V$, then the toggling operator s_v acts as addition of the vector with 1's for v and all its neighbors, and 0's otherwise. Thus the s_v commute, and $s_v^2 = s_\emptyset$, so $s_v^{-1} = s_v$. Since the s_v commute, we see that each composition of s_v is also its own inverse, and also commute, proving the lemma. □

The reason we introduce the group $\Gamma_{LO}(G)$ is so we can interpret $C_{LO}(G)$ as a Cayley graph of $\Gamma_{LO}(G)$.

Lemma 4.2.4.
$C_{LO}(G)$ is isomorphic to the Cayley graph $\mathrm{Cay}(\Gamma_{\mathrm{LO}}(\mathrm{G}), \mathrm{S})$*, where* $S = \{s_v\}_{v \in V}$

Proof. Let $\phi \colon \Gamma_{LO}(G) \to V(C_{LO}(G))$ be the map sending a composition of toggling operations to the coloring of G obtained by applying those toggles to the all off coloring $\tau_\emptyset$, i.e.

$$\phi(s) = s(\tau_\emptyset).$$

Our goal is to show that this map between the vertex sets of $C_{LO}(G)$ and $\mathrm{Cay}(\Gamma_{\mathrm{LO}}(\mathrm{G}), \mathrm{S})$ is a graph isomorphism. Since $V(C_{LO}(G))$ is by definition all winnable colorings of G, ϕ is surjective. Suppose $\phi(s_1) = \phi(s_2)$. Then

$$s_1(\tau_\emptyset) = s_2(\tau_\emptyset),$$

which implies that

$$s_1 s_1(\tau_\emptyset) = s_1 s_2(\tau_\emptyset),$$

and so

$$\tau_\emptyset = s_1 s_2(\tau_\emptyset).$$

Thus $s_1 s_2$ does not toggle any vertices, so

$$s_1 s_2 = s_\emptyset,$$

which implies

$$s_1 = s_2^{-1} = s_2.$$

Thus ϕ is a bijection. It remains to show that the number of edges between s_1 and s_2 in $\mathrm{Cay}(\Gamma_{\mathrm{LO}}(\mathrm{G}), \mathrm{S})$ is the same as those between $\phi(s_1)$ and $\phi(s_2)$ in $C_{LO}(G)$.

For two vertices v, w in a multigraph, let $|E(v, w)|$ denote the number of edges between v and w. We must then show that

$$E(s_1, s_2) = E(\phi(s_1), \phi(s_2)).$$

$E(s_1, s_2)$ is the number elements $s \in S$ such that $s\ s_1 = s_2$. Since $S = \{s_v\}_{v \in V}$, this is the number of $v \in V$ such that $s_v s_1 = s_2$. $E(\phi(s_1), \phi(s_2))$ is the number of v such that toggling v will transform the coloring $s_1(\tau_\emptyset)$ to $s_2(\tau_\emptyset)$.

Suppose

$$s_v s_1 = s_2.$$

Then $s_v s_1(\tau_\emptyset) = s_2(\tau_\emptyset)$, so

$$E(s_1, s_2) \leq E(\phi(s_1), \phi(s_2)).$$

Conversely, suppose that $s_v s_1(\tau_\emptyset) = s_2(\tau_\emptyset)$. We must show that

$$s_v s_1(\tau) = s_2(\tau) \text{ for all } \tau \in V(C_{LO}(G)).$$

Fix a $\tau \in V(C_{LO}(G))$. Since τ is winnable, there exists a sequence of vertices $v_1 ... v_k$ such that

$$\tau = s_{v_1} ... s_{v_k}(\tau_\emptyset).$$

Since the s_v commute, we see that

$$\begin{aligned} s_v s_1(\tau) &= s_v s_1 s_{v_1} ... s_{v_k}(\tau_\emptyset) \\ &= s_{v_1} ... s_{v_k} s_v s_1(\tau_\emptyset) = s_{v_1} ... s_{v_k} s_2(\tau_\emptyset) \\ &= s_2 s_{v_1} ... s_{v_k}(\tau_\emptyset) \\ &= s_2(\tau). \end{aligned}$$

Thus $E(s_1, s_2) \geq E(\phi(s_1), \phi(s_2))$ proving the result. □

The previous lemma suggests that a way to proceed would be to characterize $\Gamma_{LO}(G)$. The technique used in Lemma 4.2.3 of viewing the Lights Out process in terms of linear algebra over $\mathbb{Z}_2$ can be used to give a much stronger result.

Theorem 4.2.5.
Let r_2 denote the rank over $\mathbb{Z}_2$ of $A + I$, where A is the adjacency matrix of $C_{LO}(G)$. There is an group isomorphism $\nu \colon \Gamma_{LO}(G) \to \mathbb{Z}_2^{r_2}$ such that for each generator $e_i \in \mathbb{Z}_2^{r_2}$ there is a vertex w such that $\nu(s_w) = e_i$.

Proof. Let $\nu_1 \colon \Gamma_{LO}(G) \to \mathbb{Z}_2^V$ be the homomorphism described above sending s_w to the vector with 1's for w and all its neighbors, and 0's otherwise. We note that $\nu_1(s_w)$ is the wth column of $A + I$, where A is the adjacency matrix of G, and I is the identity matrix. Since composition of the s_v corresponds to vector addition of the $\nu_1(s_w)$, we see that the range of ν_1 is simply the column space of $A + I$. Let $\nu_2 \colon \mathrm{Col}(A + I) \to \mathbb{Z}_2^{r_2}$ be the canonical isomorphism between these two vector spaces, sending the r_2 linearly independent columns of $A + I$ to the generators of $\mathbb{Z}_2^{r_2}$. Letting ν to be the composition of ν_1 and ν_2 completes the proof. □

Since group isomorphism induce graph isomorphisms between cayley graphs (see Proposition 1.4.2), we have the following corollary.

Corollary 4.2.6.
$C_{LO}(G) \cong \mathrm{Cay}(\mathbb{Z}_2^{r_2}, \nu(S))$, where $S = \{s_w \mid w \in V\}$ and ν is the isomorphism in Theorem 4.2.5

4.3 Eigenvalues of the Lights Out Graph, Rates of Convergence, and Hitting Times

Since we have reduced the study of the random Lights Out process $X(t)$ to a random walk on a graph, we can use tools from graph theory to investigate $X(t)$. We begin with a simple result.

Lemma 4.3.1.
For each coloring τ equivalent to $\tau_\emptyset$,

$$\lim_{t\to\infty} \mathbb{P}(X(t) = \tau) = \frac{1}{2^{r_2}}$$

Proof. The graph $C_{LO}(G)$ is undirected, connected, and regular (since it is isomorphic to a Cayley graph with a symmetric generating set), and so a lazy random walk will converge to the uniform distribution. There are 2^{r_2} vertices since $C_{LO}(G)$ is isomorphic to a Cayley graph of $\mathbb{Z}_2^{r_2}$ □

We wish to investigate two further properties of the random Lights Out process $X(t)$: the mixing time and the hitting time. Informally, these are the number of steps until the process converges to the stationary distribution, and the number of steps until the process reaches the all 0 coloring, respectively. We will begin with the mixing time.

In order to express the speed of convergence, we need a way of measuring the "distance" from the stationary distribution. There are several commonly used metrics (see section 1.3), in this chapter we will use the χ-squared distance.

Combining the results of Corollary 4.2.6, Theorem 1.4.3, and Theorem 1.4.4 we obtain the following key fact.

Theorem 4.3.2.
Let $\Gamma = \mathbb{Z}_{n_1} \times ... \times \mathbb{Z}_{n_k}$ be an abelian group, and let e_i correspond to the element of Γ defined by the product of the generator 1 of $\mathbb{Z}_{n_i}$ and the identity 0 in the others. Let $\theta_q = e^{\frac{2\pi i}{q}}$ for any positive integer q. For each $\overrightarrow{x} \in \Gamma$ define $\rho_{\overrightarrow{x}} \colon \Gamma \to \mathbb{C}$ be the homomorphism where $\rho_{\overrightarrow{x}}(e_i) = \theta_{n_i}$.

Then the eigenvalues of the transition matrix for a lazy random walk on $C_{LO}(G)$ can be indexed by the vectors $\overrightarrow{x} \in \{-1, 1\}^{r_2}$, with

$$\lambda_{\overrightarrow{x}} = \frac{1}{2}\left(1 + \frac{1}{n}\sum_{w \in V} \rho_{\overrightarrow{x}}(\nu(s_w))\right)$$

Proof. Since $C_{LO}(G) \cong \mathrm{Cay}(\mathbb{Z}_2^{\mathrm{r}_2}, \nu(\mathrm{S}))$, we can find the eigenvalues of the transition matrix for a random walk on $\mathrm{Cay}(\mathbb{Z}_2^{\mathrm{r}_2}, \nu(\mathrm{S}))$. We will do this by exploiting two facts about Cayley graphs of abelian groups.

We know that $C_{LO}(G) \cong \mathrm{Cay}(\mathbb{Z}_2^{\mathrm{r}_2}, \nu(\mathrm{S}))$ where $S = \{s_w\}_{w \in V}$. By Claim 1.4.4, the irreducible representations of $\mathbb{Z}_2^{r_2}$ are $\rho_{\overrightarrow{x}} \colon \mathbb{Z}_2^{r_2} \to \mathbb{C}$ where $\rho_{\overrightarrow{x}}(e_i) = x_i$, for each $\overrightarrow{x} \in \{-1, 1\}^{r_2}$. Thus we can index the eigenvalues according to the vectors

$\overrightarrow{x}$ and get

$$\begin{aligned}
\lambda_{\overrightarrow{x}} &= \frac{1}{2}\left(1 + \sum_{g \in \mathbb{Z}_2^{r_2}} \frac{|g \cap \nu(S)|}{|\nu(S)|} \rho_{\overrightarrow{x}}(-g)\right) \\
&= \frac{1}{2}\left(1 + \sum_{s \in \Gamma_{LO}(G)} \frac{|s \cap S|}{|S|} \rho_{\overrightarrow{x}}(-\nu(s))\right) \\
&= \frac{1}{2}\left(1 + \sum_{s \in \Gamma_{LO}(G)} \frac{|s \cap S|}{|S|} \rho_{\overrightarrow{x}}(\nu(s))\right) \\
&= \frac{1}{2}\left(1 + \sum_{w \in V} \frac{1}{n} \rho_{\overrightarrow{x}}(\nu(s_w))\right)
\end{aligned}$$

□

Corollary 4.3.3.
Let $\overrightarrow{x} \in \{-1, 1\}^{r_2}$ be a vector with j entries equal to -1, and $r_2 - j$ entries equal to 1. Then

$$\lambda_{\overrightarrow{x}} \leq 1 - \frac{j}{n}$$

Proof. By Theorem 4.2.5, we know for $e_i \in \mathbb{Z}_2^{r_2}$ there is a vertex v such that

$\nu(s_v) = e_i$. Let the set of these vertices be V_1, and let $V_2 = V \setminus V_1$. Then

$$\begin{aligned}
\lambda_{\vec{x}} &= \frac{1}{2}\left(1 + \frac{1}{n}\sum_{v\in V}\rho_{\vec{x}}(\nu(s_v))\right) \\
&= \frac{1}{2}\left(1 + \sum_{v\in V_1}\frac{1}{n}\rho_{\vec{x}}(\nu(s_v)) + \sum_{v\in V_2}\frac{1}{n}\rho_{\vec{x}}(\nu(s_v))\right) \\
&= \frac{1}{2}\left(1 + \frac{1}{n}\sum_{i=1}^{r_2}\rho_{\vec{x}}(e_i) + \frac{1}{n}\sum_{v\in V_2}\rho_{\vec{x}}(\nu(s_v))\right) \\
&= \frac{1}{2}\left(1 + \frac{1}{n}(r_2 - 2j) + \frac{1}{n}\sum_{v\in V_2}\rho_{\vec{x}}(\nu(s_v))\right) \qquad (1) \\
&\leq \frac{1}{2}\left(1 + \frac{1}{n}(r_2 - 2j) + \frac{1}{n}\sum_{v\in V_2}1\right) \\
&= \frac{1}{2}\left(1 + \frac{r_2 - 2j}{n} + \frac{n - r_2}{n}\right) \\
&= 1 - \frac{j}{n}
\end{aligned}$$

□

Putting all the pieces together, we can now state the main result concerning convergence times.

Theorem 4.3.4.
Let $G = G(V, E)$ be a connected graph on n vertices. Let $X(t)$ be the random lights out process on G. Let r_2 denote the rank of $A + I$ over $\mathbb{Z}_2$. Let $\Delta'(t)$ be the $\chi-$squared distance between $X(t)$ and the uniform distribution. Then

1. $\Delta'(t) = \left(\sum_{\vec{x}\neq\mathbb{1}}\left(\frac{1}{2} + \frac{1}{2n}\sum_{v\in V}\rho_{\vec{x}}(\nu(s_v))\right)^{2t}\right)^{1/2}$
2. $\Delta'(t) \leq \left(\sum_{j=1}^{r_2}\binom{r_2}{j}\left(1 - \frac{j}{n}\right)^{2t}\right)^{1/2}$
3. $\Delta'(t) \leq e^{-c}$ *for all* $t > \frac{n}{2}(1 + \frac{1}{r_2})\log(r_2) + c$

Proof. By Lemma 4.2.1, $X(t)$ has the same distribution as a lazy random walk on $C_{LO}(G)$, which is vertex transitive. Thus by Theorem 1.3.9 and Lemma 4.3.2

$$\begin{aligned}\Delta'(t) &= \left(\sum_{\overrightarrow{x}\neq\mathbb{1}} \lambda_{\overrightarrow{x}}^{2t}\right)^{1/2} \\ &= \left(\sum_{\overrightarrow{x}\neq\mathbb{1}} \left(\frac{1}{2}+\frac{1}{n}\sum_{v\in V}\rho_{\overrightarrow{x}}(\nu(s_v))\right)^{2t}\right)^{1/2}\end{aligned}$$

Using Corollary 4.3.3 and the fact that there will be $\binom{r_2}{j}$ vectors $\overrightarrow{x}$ with j '-1' entries yields the second fact. Finally, since

$$\begin{aligned}\Delta'(t)^2 &\leq \sum_{j=1}^{r_2}\binom{r_2}{j}\left(1-\frac{j}{n}\right)^{2t} \\ &\leq \sum_{j=1}^{r_2} e^{j\log(r_2)-\frac{2jt}{n}} \\ &\leq r_2 e^{r_2\log(r_2)-\frac{2r_2t}{n}} \\ &= e^{\log(r_2)+r_2\log(r_2)-\frac{2r_2t}{n}},\end{aligned}$$

and so

$$\Delta'(t) < e^{-c} \text{ if } \frac{2r_2t}{n} > \log(r_2)+r_2\log(r_2)+c.$$

Solving for t shows this holds for all $t > \frac{n}{2}(1+\frac{1}{r_2})\log(r_2)+c$. □

The other stochastic quantity we will study is the *hitting time*. Informally, we want to answer the following question; if we start the random lights out process at a winnable coloring τ_1, how many steps does it take until we reach the all 0 coloring $\tau_\emptyset$? Formally, let $X_{\tau_1}(t)$ denote the lights out process starting from τ_1 coloring after t steps. Then the define hitting time $T_{\tau_\emptyset} = \min\{t|X_{\tau_1}(t) = \tau_\emptyset\}$. The expectation of the hitting time can bound in terms of the spectrum of the transition matrix.

Theorem 4.3.5.

Let G be a connected graph on n vertices, and let r_2 denote the rank of $A+I$ over $\mathbb{Z}_2$.

Let $X_\tau(t)$ be the random lights out process starting at a winnable coloring τ for any $\tau \neq \tau_\emptyset$. Let T_0 be the hitting time of the all 0 coloring, that is $\min\{t|X_\tau(t) = \tau_\emptyset\}$. *Suppose that $r_2 > 5$. Then*

$$2^{r_2(G)-1} \leq \mathbb{E}[T_0] \leq 2n \sum_{j=1}^{r_2} \binom{r_2}{j} \frac{1}{j} \leq 2n \frac{r_2}{r_2 - 2 - \sqrt{2}} \binom{r_2}{\lfloor \frac{r_2}{2} \rfloor}.$$

Proof. Since the random Lights Out process is equivalent to a lazy random walk on the (vertex transitive) state graph $C_{LO}(G)$, we can use Theorem 1.3.12 which states that

$$2^{|V|-1} \leq \mathbb{E}[T_0] \leq 2 \sum_{\lambda \neq 1} \frac{1}{1-\lambda}.$$

The lower bound then follows the fact that $\mathrm{Cay}(\Gamma_{\mathrm{LO}}(\mathrm{G}), \mathrm{S})$ is isomorphic to $\mathrm{Cay}(\mathbb{Z}_2^{\mathrm{r_2}}, \nu(\mathrm{S}))$, and hence has 2^{r_2} vertices. Using Corollary 4.3.3 to the upper bound the eigenvalues yields

$$\mathbb{E}[T_0] \leq 2n \sum_{j=1}^{r_2} \binom{r_2}{j} \frac{1}{j}. \tag{4.1}$$

The upper bounds for $r_2 \leq 5$ can be computed directly from equation 4.1. For values of $r_2 > 5$, we use the following estimate

Claim 1.
For all $r_2 > 5$, $\sum_{j=1}^{r_2} \binom{r_2}{j} \frac{1}{j} \leq r_2 \binom{r_2}{\lfloor \frac{r_2}{2} \rfloor} \frac{2}{r_2 - 2 - \sqrt{2}}$

Proof of Claim. Let $f(j) = \binom{r_2}{j} \frac{1}{j}$. Consider $f(j+1) - f(j) = \frac{r_2!}{(j+1)!(r_2-j-1)!(j+1)} - \frac{r_2!}{(j)!(r_2-j)!(j)} = \frac{r_2!}{(j)!(r_2-j-1)!} \left(\frac{1}{(j+1)^2} - \frac{1}{(r_2-j)(j)} \right)$. Thus

$$\begin{aligned} f(j+1) > f(j) &\Leftrightarrow \frac{1}{(j+1)^2} > \frac{1}{(r_2-j)(j)} \\ &\Leftrightarrow r_2 j - j^2 > j^2 + 2j + 1 \\ &\Leftrightarrow j^2 + (1 - \frac{r_2}{2})j + \frac{1}{2} < 0 \\ &\Leftrightarrow \frac{r_2 - 2 - \sqrt{(r_2-2)^2 - 8}}{4} < j < \frac{r_2 - 2 + \sqrt{(r_2-2)^2 - 8}}{4} \end{aligned}$$

Let $b(r_2) = \frac{r_2-2+\sqrt{(r_2-2)^2-8}}{4}$. If $r_2 > 5$, then $\frac{r_2-2-\sqrt{(r_2-2)^2-8}}{4} < 1$. Therefore, $f(j)$ is increasing for all $j < b_{r_2}$ and decreasing for $j \geq b_{r_2}$, and thus $f(j)$ is maximized at $f(\lfloor b_{r_2} \rfloor + 1)$.

$$\begin{aligned} f(\lfloor b_{r_2} \rfloor + 1) &= \binom{r_2}{\lfloor b_{r_2}+1 \rfloor} \frac{1}{\lfloor b_{r_2}+1 \rfloor} \\ &\leq \binom{r_2}{\lfloor b_{r_2}+1 \rfloor} \frac{1}{b_{r_2}} \\ &= \binom{r_2}{\lfloor b_{r_2}+1 \rfloor} \frac{4}{r_2-2+\sqrt{(r_2-2)^2-8}} \\ &\leq \binom{r_2}{\lfloor b_{r_2}+1 \rfloor} \frac{2}{r_2-2-\sqrt{2}} \end{aligned}$$

□

To conclude, we note that $b_{r_2} + 1 < \frac{r_2}{2}$, and thus $\binom{r_2}{\lfloor b_{r_2}+1 \rfloor} < \binom{r_2}{\lfloor \frac{r_2}{2} \rfloor}$. Thus $f(j) \leq \binom{r_2}{\lfloor \frac{r_2}{2} \rfloor} \frac{2}{r_2-2-\sqrt{2}}$ for all j, giving the desired bound. □

Corollary 4.3.6.

The following table summarizes the results of Theorems 4.3.4 and 4.3.5 for certain graphs where $r_2(G)$ is known.

	r_2	$\Delta'(t) \le e^{-c}$ *for* $t >$	$\le \mathbb{E}[T_0] \le$	
Path, P_n				
$n \equiv 2 \bmod 3$	$n-1$	$\frac{n^2}{2n-2}\log(n-1)+c$	2^{n-2}	$\frac{2n^2-2n}{n-3-\sqrt{2}}\binom{n-1}{\lfloor\frac{n-1}{2}\rfloor}$
$n \not\equiv 2 \bmod 3$	n	$\frac{n+1}{2}\log(n)+c$	2^{n-1}	$\frac{2n^2}{n-2-\sqrt{2}}\binom{n}{\lfloor\frac{n}{2}\rfloor}$
Cycle, C_n				
$n \equiv 0 \bmod 3$	$n-2$	$\frac{n^2-n}{2n-4}\log(n-2)+c$	2^{n-3}	$\frac{2n^2-4n}{n-4-\sqrt{2}}\binom{n-2}{\lfloor\frac{n-2}{2}\rfloor}$
$n \not\equiv 0 \bmod 3$	n	$\frac{n+1}{2}\log(n)+c$	2^{n-1}	$\frac{2n^2}{n-2-\sqrt{2}}\binom{n}{\lfloor\frac{n}{2}\rfloor}$
Wheel, $C_{n-1}+K_1$				
$n \not\equiv 1 \bmod 3$, *n odd*	n	$\frac{n+1}{2}\log(n)+c$	2^{n-1}	$\frac{2n^2}{n-2-\sqrt{2}}\binom{n}{\frac{n-1}{2}}$
$n \not\equiv 1 \bmod 3$, *n even*	$n-1$	$\frac{n^2}{2n-2}\log(n-1)+c$	2^{n-2}	$\frac{2n^2-2n}{n-3-\sqrt{2}}\binom{n-1}{\frac{n-2}{2}}$
$n \equiv 1 \bmod 3$, *n odd*	$n-2$	$\frac{n^2-n}{2n-4}\log(n-2)+c$	2^{n-3}	$\frac{2n^2-4n}{n-4-\sqrt{2}}\binom{n-2}{\frac{n-3}{2}}$
$n \equiv 1 \bmod 3$, *n even*	$n-3$	$\frac{n^2-2n}{2n-6}\log(n-3)+c$	2^{n-4}	$\frac{2n^2-6n}{n-5-\sqrt{2}}\binom{n-3}{\frac{n-4}{2}}$
Complete bipartite, K_{n_1,n_2}				
$n_1 n_2$ *even*	n	$\frac{n+1}{2}\log(n)+c$	2^{n-1}	$\frac{2n^2}{n-2-\sqrt{2}}\binom{n}{\lfloor\frac{n}{2}\rfloor}$
$n_1 n_2$ *odd*	$n-1$	$\frac{n^2}{2n-2}\log(n-1)+c$	2^{n-2}	$\frac{2n^2-2n}{n-3-\sqrt{2}}\binom{n-1}{\lfloor\frac{n-1}{2}\rfloor}$
k-hypercube, $Q_k, n=2^k$				
k even	n	$\frac{n+1}{2}\log(n)+c$	2^{n-1}	$\frac{2n^2}{n-2-\sqrt{2}}\binom{n}{\frac{n}{2}}$
k odd	$\frac{n}{2}$	$\frac{n+2}{2}\log(\frac{n}{2})+c$	$2^{\frac{n-2}{2}}$	$\frac{2n^2}{n-2\sqrt{2}}\binom{\frac{n}{2}}{\frac{n}{4}}$
5 by 5 grid	23	$48+c$	4×10^6	8×10^7

Proof. The parity dimensions of paths, cycles, wheels, and complete bipartite graphs were investigated in [4]. The hypercube was considered in [5], and the 5×5 grid (the graph used in the original electronic toy) was analyzed in [7]. The final columns follow from Theorems 4.3.4 and 4.3.5. □

4.4 Remarks

There are two interesting generalizations that may be worth further consideration. In this study, we considered colorings of graphs by elements of $\mathbb{Z}_2$. While this was the original context of the Lights Out game, several papers [29, 33] consider a multi-color variation, where the colorings are instead over $\mathbb{Z}_q$ for some integer q. The random $\mathbb{Z}_q$ lights out process on graphs could be interesting to investigate. Most of the proofs in this chapter generalize in the case where q is a prime (and so we can still do linear algebra over the finite field with q elements), but the case where q is composite would require more care and possiblely new techniques.

Another variant involves the "lit-only" restriction. In this version of Lights Out, one can only toggle vertices that are currently colored 1. While initially one would expect this would limit the number of winnable colorings, Goldwasser, Wang, and Wu [35] proved the following surprising result;

Theorem 4.4.1.
Let G be a connected graph, and let $\tau_1, \tau_2 \in \mathfrak{F}(G)$ with τ_1 not equal to the all 0 coloring and τ_2 not the all 1 coloring. Then it is possible to go from τ_1 to τ_2 in the lights out game if and only if it is possible to do so in the lit-only lights out game.

One could consider the random lit-only lights out process, where at each step the next vertex to toggle is chosen uniformly at random from all vertices currently labeled 1. In this version, the state graph would be directed, but Theorem 4.4.1 implies that the state graph with the all 0 and all 1's colorings removed is strongly connected.

The lit-only process appears much more difficult to analyze, for several reasons. For example, it is not regular, and so we cannot view the state graph as the Cayley graph of a group or semi-group. However, experimental results on several small graphs gives rise to the following conjecture.

Conjecture 1.

Let G be a connected graph on n vertices. Let $X(t)$ denote the location at time t of the random lights out process starting at a random winnable coloring. Let T_0 be the hitting time of the all 0 coloring, that is $\min\{t|X_\tau(t) = \tau_\emptyset\}$. Similarly, let $\tilde{X}(t)$ and $\tilde{T}_0$ be the location and hitting time of the random lit-only process. Then

$$\mathbb{E}[T_0] = (1 + o(1))\mathbb{E}[\tilde{T}_0]$$

Acknowledgement:

This chapter is based on the paper "Random Lights Out Processes on Graphs" which has been accepted for publication in Advances in Applied Mathematics [38].

Chapter 5

Random Nearest Neighbor Interchanges on Phylogenetic Trees

5.1 Introduction

A phylogenetic tree is a tree structure that represents the evolutionary relationships within a group of organisms. Reconstructing phylogenetic trees from genetic data an area of active biological research, and several approaches use Markov Chain Monte Carlo (MCMC) methods, see for example [28, 43, 45, 57]. MCMC methods depend on an underlying Markov chain against which one samples from a distribution based on real data. While direct analysis of the data-dependent chain is often intractable, an understanding of the properties of the base chain is a natural first step.

In this chapter, we analyze a random walk on labeled unrooted bifurcating trees. For the purposes of this discussion, the term *phylogenetic tree* will refer to an unrooted tree with $n \geq 4$ vertices of degree 1 labelled $\{1, 2, ..., n\}$, and $n - 2$ unlabelled internal vertices of degree 3. The edges are unweighted. The n labeled vertices are called *leaves*, and the $n - 2$ internal nodes are called *branch points*. The term *cladogram* is also used to describe such trees. An *internal edge* is an edge

that is connects two branch points. Each internal edge is adjacent to 4 subtrees, and a *nearest neighbor interchange* (NNI) interchanges two of these. Note that interchanging two subtrees connected to the same branch point does not change the tree. The (non-trivial) NNIs are called "crossovers" of the tree, and in general the term nearest neighbor interchange refers to the non-trivial manipulations.

We consider the random process induced by random NNIs, where at each step an internal edge and two adjacent subtrees are selected uniformly at random and swapped. A variety of other chains on the space of phylogenetic trees have been studied. Diaconis and Holmes use study a Markov chain on (rooted) phylogenetic trees using a bijection between perfect matchings of the complete graph on $2n$ vertices and (rooted) phylogenetic trees. Aldous studied a chain based on moving a random leaf to a random position [2]. This chain was further studied by Schweinsberg, who obtained sharper bounds [53]. These chains all consider trees with unweighed edges, though in practice most biological applications use weighted trees where edges correspond to the degree of genetic similarity between species. Štefankovič and Vigoda analyzed a chain based removing and reattaching subtrees for a specific class of phylogenetic subtrees and showed it is rapidly mixing.

In Section 5.2 we discuss basic facts about phylogenetic trees and illustrate several common tree manipulations. In Section 5.3 we formally define the Markov process mentioned above and the linear algebraic tools we use to study its convergence. In Section 5.4 we use a comparison theorem and the results of Schweinsberg to bound the spectral gap of the configuration graph for nearest neighbor interchanges. In particular, we show that the spectral gap is $\Omega(\frac{1}{n^4})$ and thus the relaxation time is $O(n^4)$. We use this bound on the spectral gap to show that the mixing time of the random NNI process for phylogenetic trees with n leaves is $O(n^5 \log n)$.

5.2 Phylogenetic Trees and Tree Manipulations

Let $\mathcal{T}_n$ denote the set all of all phylogenetic trees on n leaves. We note that $|\mathcal{T}_n| = (2n-5)!! = \frac{(2n-5)!}{2^{n-3}(n-3)}$ and that $(2n-5)!! \sim \frac{1}{2\sqrt{2}} \left(\frac{2}{e}\right)^n n^{n-2}$ as $n \to \infty$ [56].

Each phylogenetic tree has $2n - 3$ edges, and since n of these are connected to leaves there are $n - 3$ internal edges. Each internal edge partitions the tree into 4 subtrees.

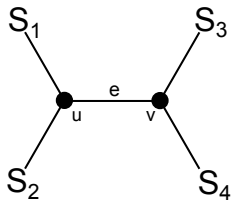


Figure 5.1: An internal edge $e = \{u, v\}$ and its four adjacent subtrees.

Let e be an internal edge joining branch points u and v. Let S_1 and S_2 denote the subtrees adjacent to u, S_3 and S_4 the subtrees adjacent to v. Then we say that S_1 and S_2 are *neighboring subtrees*, and similarly for S_3 and S_4. Subtrees adjacent to the same internal edge but not neighboring (e.g. S_1 and S_3) are said to be *separated* by e. For a subtree S, we let $\bar{S}$ denote the rooted subtree obtained by adding the nearest edge and vertex, which becomes the root. For example, $\bar{S_1}$ has the vertex u as its root.

5.2.1 Tree Manipulations

There are many operations used to manipulate phylogenetic tree structures, see for example [17]. Here we will define three operations: *Nearest Neighbor Interchange* (NNI), *Subtree Prune and Regraft* (SPR), and *Leaf Pluck and Reattach* (LPR).

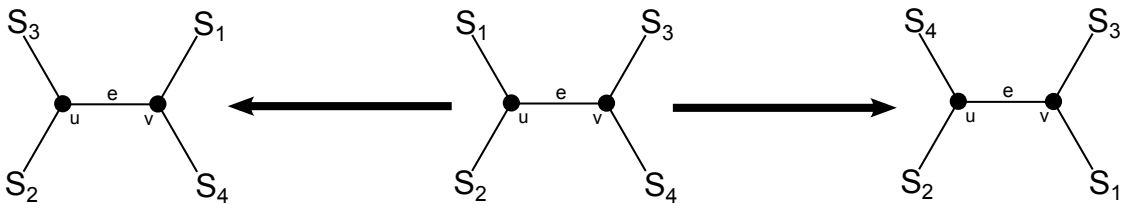


Figure 5.2: The two possible NNIs across an internal edge e

Definition 5.2.1.
Let e be an internal edge of a phylogenetic tree T. A *nearest neighbor interchange* (NNI) swaps two subtrees separated by e, as in Figure 5.2

Nearest Neighbor Interchanges were introduced in 1969 by D.F. Robinson [51] and independently by Moore, Goodman, and Barnabas in 1973 [47]. The *NNI distance* between two trees is the number of NNIs needed to transform one tree into the other. This is a commonly used metric that is difficult to compute, see for example [42] for a history of the problem and [3] for current results.

Robinson defined a graph with vertex set consisting of all phylogenetic trees with n leaves, and an edge between two trees if there is an NNI that transforms one to the other. This graph has become known as the *Robinson Graph*, which we denote in this manuscript by R_n. Note that since every NNI is reversible, R_n is undirected. Robinson proved showed that R_n is connected for all n. Because there are two NNIs that can occur at each of the $n-3$ internal edges, R_n is $(2n-6)$-regular. Although R_n is vertex transitive for $n = 4, 5$, it is not vertex transitive for $n \geq 6$ [8]. This graph has been extensively researched, see [8, 9] for a thorough study.

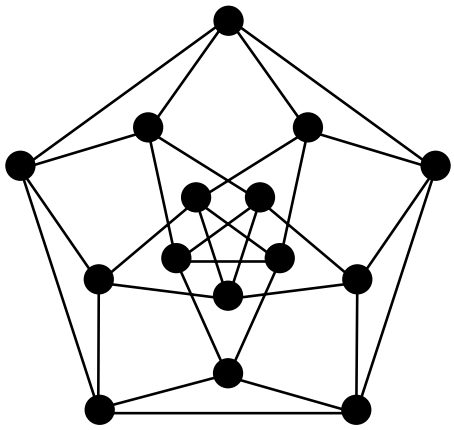


Figure 5.3: The Robinson graph for phylogenetic trees with 5 leaves.

We now define a more general tree operation, the Subtree Prune and Regraft.

Definition 5.2.2.

A *subtree prune and regraft* (SPR) on a phylogenetic tree T is an operation consisting of three parts. See Figure 5.4 for an illustration.

1. A subtree S is selected, and it is removed along with the edge from S to the branch point v in to the rest of T
2. The two edges separated by v are joined into a single edge.
3. A new branch point u is created in the middle of an edge and S is connected to that branch point

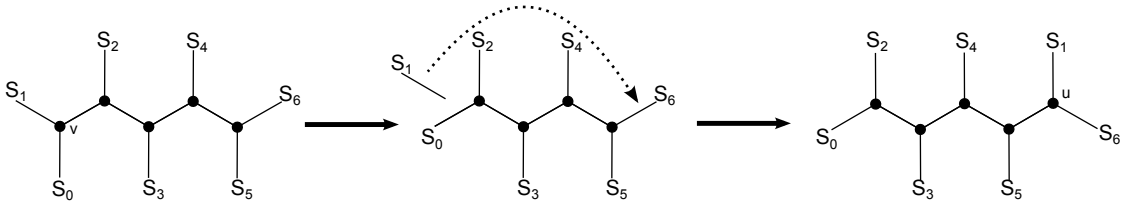


Figure 5.4: An example of an SPR move. The subtree S_1 is removed, the edge split by v is combined into a single edge, and S_1 is attached to the edge connecting S_6 where the new branchpoint u is created.

Note that branch points are actually unlabeled, and are included in the description above for clarity and to coincide with Figure 5.4

As noted in [3], the SPR operation is of particular note to biologists as it is used to model biological processes such as horizontal gene transfer and recombination, which are especially common among bacteria.

Definition 5.2.3.

A leaf pluck and regraft (LPR) is a SPR where the removed subtree consists of a single leaf.

Let $\tilde{G}_n$ denote the configuration space with respect to LPR moves. That is, $\tilde{G}_n$ has vertex set $\mathcal{T}_n$ and two trees are adjacent if there is an LPR move that transforms one to the other. Since one can remove any of the n leaves and attach it to any of $2n-5$ edges, $\tilde{G}_n$ is $n(2n-5)$-regular .

5.3 The Random Nearest Neighbor Interchange Process on Phylogenetic Trees

We define a Markov chain $\{X_t\}_{t=0}^{\infty}$ on the space of trees with n leaves. At each step, an internal edge e of X_t is chosen uniformly at random. Two of the four subtrees adjacent to e are selected uniformly at random and swapped to create X_{t+1}. Note that if the two subtrees are neighboring then the swap does not alter the tree and $X_{t+1} = X_t$. For two trees $T, T' \in \mathcal{T}_n$, we write $T \sim T'$ if there exists a NNI that transforms T to T' (equivalently, if {T,T'} is an edge of the Robinson Graph R_n). One third of the time the two selected subtrees will be neighbors, resulting in no change so $X_{t+1} = X_t$. The other two thirds of the time, X_{t+1} is one of X_t's $2n-6$ neighbors in $\mathcal{T}_n$. Therefore we see that the transition probabilities for this Markov chain are given by

$$P(X_{t+1} = T' | X_t = T) = \begin{cases} \frac{1}{3n-9} & \text{if } T' \sim T \\ \frac{1}{3} & \text{if } T' = T \end{cases} \tag{5.1}$$

That is, the random Nearest Neighbor Interchange Process can be viewed as a lazy random walk on R_n. Letting P be the matrix with entries indexed by elements of $\mathcal{T}_n$ and $P(T, T') = P(X_{t+1} = T' | X_t = T)$, we see that

$$\begin{aligned} P &= \frac{1}{3} \cdot I + \frac{2}{3} \cdot D^{-1} \cdot A \\ &= \frac{1}{3} \cdot I + \frac{2}{3} \cdot \frac{1}{2n-6} \cdot A \end{aligned}$$

where I is the identity matrix and A and D are the adjacency and diagonal degree matrices of the the Robinson Graph R_n, respectively. Another matrix associated with the graph R_n is *normalized Laplacian* $\mathcal{L} = I - D^{-1/2} A D^{-1/2}$. Since R_n is $(2n-6)$-regular,

$$\mathcal{L} = I - \frac{1}{2n-6} A,$$

and so

$$P = I - \frac{2}{3}\mathcal{L}.$$

We let $0 = \lambda_0 \leq \lambda_1 \leq ... \leq \lambda_{(2n-5)!!-1} \leq 2$ denote the eigenvalues of $\mathcal{L}$. The second smallest eigenvalue λ_1 is known as the *spectral gap*, and is a crucial parameter when considering the convergence of the process.

Since R_n is undirected, non-bipartite, and connected, the process is ergodic and converges to the uniform stationary distribution as $n \to \infty$. We wish to quantify the time it takes for the chain to converge to its stationary distribution, commonly referred to as the "mixing time" of the process. In order to ask the question "How long until the process' distribution is close to the uniform distribution?", we must first establish what metric we are using to measure the "distance" from the stationary distribution. Here we will use the *relative pointwise distance*, $\Delta(t)$.

5.4 Bounding the Spectral Gap of the Robinson Graph

We begin with one of our two main results, which bounds the size of the spectral gap for the normalized laplacian of R_n. The following theorem will allow us to upper bound the mixing time in Theorem 5.4.5

Theorem 5.4.1.
Let R_n denote the Robinson graph, the NNI configuration graph. Let A denote the adjacency matrix of the Robinson graph, D the diagonal degree matrix, and $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{(2n-5)!!-1}$ denote the eigenvalues of the Normalized Laplacian $\mathcal{L} = I - D^{-1/2}AD^{-1/2}$. Similarly define $\tilde{A}, \tilde{\mathcal{L}}, \left\{\tilde{\lambda}\right\}_{i=0}^{(2n-5)!!-1}$ for the LPR configuration graph $\tilde{G}_n$. Then

1. $\lambda_1 \geq \frac{n(2n-5)\tilde{\lambda}_1}{(2n-6)(n-3)2(n-3)^2}$
2. $\lambda_1 = \Omega(\frac{1}{n^4})$

The proof of the first statement uses on the so called "method of distinguished paths" or "comparison theorems" discussed in the introduction. We note that it has been used specifically in the context of Phylogenetic Trees. A modification of this method is the main tool used in [53] to bound the *relaxation time* of a random walk on the LPR configuration graph, given by $\tilde{\lambda}_1^{-1}$.

The goal is to use Theorem 1.3.13 to compare the Robinson Graph and the LPR configuration graph. In order to do this, we must show two things:

1. For every edge $\{T_1, T_2\}$ in $\tilde{G}_n$, bound the length of a path $\gamma(T_1, T_2)$ in R_n connecting T_1 and T_2

2. For every edge E in R_n, bound the number of $\gamma(T_1, T_2)$ that contain E, where T_1 and T_2 differ by a single LPR move.

Lemmas 5.4.2 and 5.4.4 address these items. We remark that the diameter of R_n is bounded above by $n \log n + O(n)$ [42], but the result of Corollary 5.4.3 will give a sharper bound in the case of two trees differing by a LPR. We begin by bounding the NNI distance of two trees differing by a single SPR move. We note that the proof of Lemma 5.4.2 uses a similar technique to that used by Robinson to show that G_n is connected [51, Theorems 4 & 5].

Lemma 5.4.2.
Let T_1 and T_2 be two phylogenetic trees with n leaves. Suppose that T_1 and T_2 differ by a single SPR move, and the subtree moved contains k of the n labeled leaves. Then the NNI distance is less than $n - k - 2$.

Proof. Let S_1 denote the subtree of T_1 that is pruned and regrafted. Let v_1 denote the root of S, where it attaches to T_1, and note that v_1 is an internal node of T_1. S_1 has $k + 1$ leaves (the k labeled leaves, plus v_1, and thus since it itself is a phylogenetic tree it has $2(k + 1) - 2 = 2k$ vertices. Let e_ℓ denote the internal edge of T_1 where S_1 is attached by the SPR move that transforms T_1 to T_2. Let v_ℓ denote the endpoint of e_ℓ that is closest to v_1, and label all the intermediate vertices so that $v_1, v_2, ..., v_\ell$ is the unique shortest path from v_1 to v_ℓ. Let e_i denote the edge between v_i and v_{i+1}, and let S_i denote the subtree attached to v_i not containing e_{i-1} or e_i. We first bound ℓ by a simple counting argument. Note first that each of $v_1, ..., v_\ell$ is an internal vertex of T_1. Recall that a phylogenetic tree on n leaves has $n - 2$ internal vertices. Since S has $k + 1$ leaves (including v_1), there are $k - 1$ internal vertices of T_1 that are also internal vertices of S_1. Thus

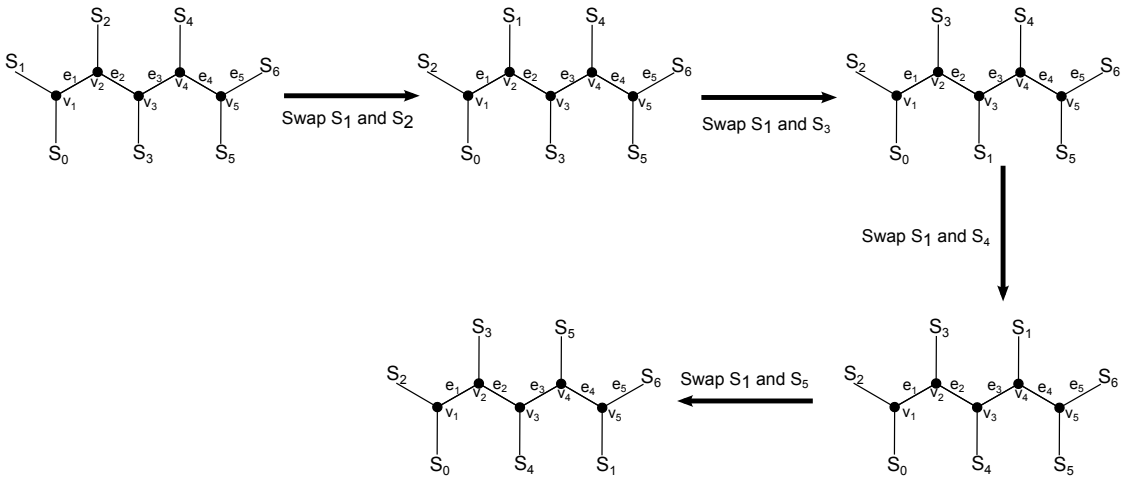


Figure 5.5: An example of a sequence of NNIs corresponding to the SPR move in Figure 5.4

there are $n-k-1$ internal vertices of T_1 that are not internal vertices of S, and so $\ell \leq n-k-1$.

We now show that a sequence of NNI moves, across each edge $e_1, ..., e_{\ell-1}$, will transform T_1 to T_2. Begin with the NNI move across e_1 that swaps S_1 and S_2. Continue to swap S_1 "down the line" so that at each step, S_1 is swapped with S_{i+1} across the edge e_i. This preserves the relative order of $S_2, ..., S_\ell$, and moves S_1 so that it is attached to the same vertex as $S_{\ell+1}$. Also S_0 and S_2 are now connected to the same vertex. This is exactly the result of the SPR move removing $\bar{S}_1$ and regrafting at e_ℓ, so we have successfully transformed T_1 to T_2 in $\ell - 1 \leq n-k-2$ NNI moves. □

Corollary 5.4.3.

Let T_1 and T_2 be two phylogenetic trees with n leaves. Suppose that T_1 and T_2 differ by a single LPR move. Then the NNI distance is less than $n-3$.

Lemma 5.4.4.

For any LPR move φ transforming T to T', let $\gamma(\varphi)$ be the path in R_n corresponding to the sequence of NNI moves between them constructed in the proof of Lemma 5.4.2. Let E be any edge of R_n. Then E is in contained in at most $2(n-3)^2$ paths $\gamma(\varphi)$.

Proof. E corresponds to a single NNI move between two phylogenetic trees T_1, T_2. Let e denote the edge in T_1 across which the NNI move is occurring, and let $\{S_1, S_2, S_3, S_4\}$ denote the four subtrees adjacent to e. Without loss of generality, assume that S_1 and S_2 are adjacent, S_3 and S_4 are adjacent, and E corresponds to swapping S_1 and S_4 (or, equivalently, swapping S_2 and S_3). This is the NNI illustrated in Figure 5.2

Suppose there are j_1, j_2, j_3, j_4 nodes (including the root attached to e) in $\bar{S}_1$, $\bar{S}_2$, $\bar{S}_3$, $\bar{S}_4$ respectively, where $j_1 + j_2 + j_3 + j_4 - 4 = n$. In every NNI in the sequence $\gamma(\varphi)$, (at least one) of the subtrees swapped is a leaf. We split into cases based on the number of S_1, S_2, S_3, S_4 that are leaves.

Case 1. 1 Leaf

Without loss of generality, let S_1 consist of a single node, l, and let S_2, S_3, S_4 all be subtrees with 2 or more leaves. Then the node l is the leaf being pushed "down the line", and the original LPR move φ must have moved l. We will bound the number of LPR moves that could use E by first considering the number of starting positions of l, and the number of ending positions. Note that $j_1 = 2$ and $j_i \geq 3$ for $i = 2, 3, 4$. Either l must originally have been attached to an edge in S_2, or E is the first move in $\gamma(\varphi)$. Since $\bar{S}_2$ contains j_2 nodes, there are $2j_2 - 3$ edges where l could have originally been attached, and note that attaching l to the root edge of S_2 is the same as starting with T_1.

We also see that ϕ moves l to an edge of $\bar{S}_3$. There are $2j_3 - 3 \leq 2(n - 1 - j_2) - 3 = 2n - 2j_2 - 5$ edges in $\bar{S}_3$ where l could be attached by φ.

So there are $2j_2 - 3$ possibilities for T, and $2n - 2j_2 - 5$ possibilities for T', and thus E is contained in at most $(2n - 2j_2 - 5)(2j_2 - 3) \leq (n-4)^2$ paths. Therefore if e has only one leaf attached, then there are at most $(n-4)^2$ paths $\gamma(\varphi)$ that use E.

Case 2. 2 Leaves

Now we consider the case where there are two leaves. We must consider the fact that the original LPR move could have plucked and reattached either leaf. Again we assume that S_1 is a leaf l_1, but now there are three cases corresponding to

which other subtree is a leaf. Suppose first that the two leaves are on the same side, that is that S_2 consists of a single leaf and so $j_2 = 2$. Then φ moves l_1 to an edge in S_3 or l_2 to an edge in S_4, and the NNI move E is the first move in $\gamma(\varphi)$. Therefore E is involved in at most $2j_3 - 3 + 2j_4 - 3 = 2n - 6$ paths $\gamma(\varphi)$. The case where $j_3 = 2$ follows in the same manner. It remains to consider the case where $j_1 = j_4 = 2$ (corresponding to single leaves l_1, l_4) and $j_2, j_3 \geq 3$. In this case, the leaf moved by φ could be either l_1 or l_4. Suppose first it is l_1. Proceeding in a manner similar to case 1, there are $2j_2 - 3$ places where l_1 could have originally been attached, and $2(j_3) - 3 = 2(n - j_2) - 3$ edges of S_3 where it could be moved. If instead l_4 is the leaf moved by φ, then there are $2(j_3) - 3 = 2(n - j_2) - 3$ edges in S_3 where it could have started, and $2j_2 - 3$ edges in S_2 where it could be placed. Thus at most $2(2(n-j_2)-3)(2j_2-3) \leq 2(n-3)^2$ paths $\gamma(\varphi)$ use the edge E.

Case 3. 3 Leaves

Without loss of generality, assume that S_1, S_2, S_3 be the subtrees corresponding to single leaves l_1, l_2, l_3, and S_4 the non-trivial subtree. We must consider each of the three cases that φ moves l_1, l_2, l_3. First suppose φ moves l_1. l_1 is adjacent to another leaf both before and after E, so if φ moves l_1 then $\gamma(\varphi) = \{E\}$. Suppose that φ moves l_2. Since l_2 is initially adjacent to the leaf l_1, E must be the first move in $\gamma(\varphi)$, and there are $2n - 7$ edges of S_4 where l_2 could be placed. Similarly, if l_3 is the leaf moved by φ, then E must be the last move in $\gamma(\varphi)$. In this case, there are $2n - 7$ edges of S_4 where it could have started. Therefore we find that in the case of 3 leaves, there are at most $4n - 13$ paths $\gamma(\varphi)$ that use E.
Therefore we see that every edge E is contained in at most $2(n-3)^2$ paths $\gamma(\varphi)$.

□

Proof of Theorem 5.4.1. The first statement follows immediately from Theorem 1.3.13, and Corollary 5.4.3, and Lemma 5.4.4. The second statement follows from the first, and the fact that $\tilde{\lambda}_1 = \Omega(\frac{1}{n^2})$ [53]. The third statement follows immediately from the second. □

Theorem 5.4.5.

Let R_n denote the Robinson Graph of phylogenetic trees with n leaves. Let $\Delta(t)$ denote the relative pointwise distance between the location of a lazy random walk and the stationary distribution after t steps. Let $t_f = \min\{t \mid \Delta(t) \leq e^{-f(n)}\}$. Then

$$t_f = O(n^5 \log n + n^4 f(n))$$

In particular, for any constant c, $\Delta(t) \leq e^{-c}$ after $O(n^5 \log n)$ steps and thus

$$\tau_{mix} = O(n^5 \log n).$$

Proof. From Theorem 1.3.10, we know that

$$\Delta(t) \leq e^{-c}$$

if

$$t \geq \frac{\log(2n-5)!! + c}{\lambda}$$

where $\lambda = \frac{2}{3}\lambda_1$. Since

$$(2n-5)!! \sim \frac{1}{2\sqrt{2}} \left(\frac{2}{e}\right)^n n^{n-2}$$

$$\log\left((2n-5)!!\right) = O(n \log n).$$

Therefore $t_c = O(n^5 \log n + cn^4)$. □

Acknowledgement:

This chapter is based on the paper "Random Nearest Neighbor Interchanges on Phylogenetic Trees" which is currently in preparation. The author would like to thank Courtney Hoffman of the University of Maryland for the introduction to phylogenetic trees and NNIs.

Chapter 6

Resource Allocation Using pageRank Vectors

6.1 Introduction

Efficient allocation of resources to meet changing demands is a task arising in numerous applications. For example, institutions such as governments or corporations respond to the needs of a populace, and wish to meet the demands within allowed expenditure of resources. In some cases where demand spreads, one has to be able to act before demand becomes unmanageable. In the case of an epidemic, for instance, it is desirable to find a way to distribute medicine so that the disease can be contained. Such problems have been studied in several contexts using the contact process model [40] [31], [11], [48], [23]. In [23], it was demonstrated how to use PageRank vectors to both restrict the number of nodes inoculated and to provide certain containment guarantees.

In this chapter, we study a variant of the classical contact process, a continuous time Markov process on a contact graph. This model was previously used for modeling the spread of disease. In our scenario, vertices in the graph each have varying levels of demand for multiple commodities. Demand at a vertex propagates to its neighbors at a rate depending on the current demand. Our model allows for rich interactions between different commodities; for instance, demand for one

commodity can influence demand for another. This fits many scenarios that arise, for instance demand for iPhones may accelerate the demand for iPads. As another example, demand at a node can sometimes be viewed as a measure of discontent with the current supply of a resource. It is natural for an unhappy node to create unrest in its neighbors. As the contact process continues, the demands at a vertex are increased based on the demands at neighboring vertices and are decreased at a satisfaction rate, which can be thought of as a frequency of shipments. The rates at which demand spreads will be a linear combination of demands from neighboring vertices. These rates will be encapsulated as a spread matrix, B, roughly analogous to the infectivity parameter in the classical contact process. The goal of this chapter is to find satisfaction rates, dependent on the spread matrix B and the geometry of the contact graph which ensure that all vertices have no demand and the process dies out. This process will be defined, in detail, in Section 6.2.

To satisfy the demands which evolve according to our model as defined in Section 6.2, the goal is to ship commodities and supply to vertices with unsatisfied demands in an efficient way. The model here differs somewhat from typical resource allocation problems in the sense that we do not specify the location of the "warehouses" for the supply. We will not be concerned with either the sources of the supply or the detailed incremental costs of shipping supply. Instead, our goal is to identify *how often* to ship each commodity to a particular vertex, in order to contain and satisfy demands, given an initial seed set. The reader is referred to [19] for the usual resource allocation problem.

After we describe the demand model in Section 6.2, we proceed to analyze our supply scheme. First, we introduce the Kronecker PageRank in Section 6.3 based on the PageRank originally introduced by Brin and Page [13]. Our analysis is comprised of two parts. First, we give conditions which ensure that all demand is satisfied in $O(\log n)$ time, with high probability, regardless of the initial demand. This is a global solution, in this sense that it involves "scheduling shipments" to all vertices in the graph in a way that will be made precise once the model is formally introduced in Section 6.2 and Theorem 6.4.1. Next, in Section 6.5 we analyze a situation where shipments are scheduled to only a subset of vertices containing the

initial demand. In particular, when the contact graph has some clustering structure we are interested in subsets so that the demand within the subset is satisfied quickly (in $O(\log n)$ time) and demands reach a vertex not receiving shipments with low probability. Precise results to this end are given in Section 6.5.

Our analysis provides a tradeoff in the following sense: On one hand, if one would like to guarantee that the demand escapes a set with probability at most ϵ, our results will allow us to use PageRank (or a sharper Kronecker PageRank as defined in Section 6.3) to identify a subset of vertices where supplies will be shipped, and provide a guarantee that this will be sufficient. On the other hand, if one would like to send shipments to a particular set of vertices, then our analysis in Theorem 6.5.3 allows a guaranteed bound on the escape probability which depends on the clustering structure of the contact graph.

6.2 The Demand Model

We model the demand spreading along an undirected simple graph $G = (V, E)$. Let n be the number of vertices of G, A the adjacency matrix, D the diagonal degree matrix, and $W = D^{-1}A$ the transition matrix of a random walk on G.

If B is an $k \times k$ matrix, and A an $n \times n$ matrix, then the *Kronecker product* $A \otimes B$ is the $nk \times nk$ block matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{pmatrix}$$

An *exponential random variable* with parameter λ has probability density function given by $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$, and 0 for $x < 0$. This distribution will be denoted $\text{Exp}(\lambda)$. Exponential random variables are memoryless, that is to say, if X is an exponential random variable then for any constants $a, b > 0$,

$$\mathbb{P}(X > a + b | X > a) = \mathbb{P}(X > b).$$

If X and Y are independent and $X \sim \text{Exp}(\lambda_1)$, $Y \sim \text{Exp}(\lambda_2)$ then $\min\{X, Y\} \sim$

$\text{Exp}(\lambda_1 + \lambda_2)$. A Poisson point process at rate λ is a sequence of random variables $X_1, X_2, \ldots$ so that $X_i - X_{i-1}$ has distribution $\text{Exp}(\lambda)$.

Before we describe our model, let us briefly recall the contact process on a graph G, which we denote $CP(T, \beta, \boldsymbol{\sigma}, G)$. In the contact process (see for example [11] or [23]), a disease initially infects a set $T \subseteq V(G)$. The disease has an infectivity parameter, β, and each vertex has a certain amount of "medicine" σ_v. Each infected vertex independently infects its neighbors at times given by a Poisson point process at rate β, and each infected vertex is cured at times given by a Poisson point process at rate σ_v. In the most frequently studied case, σ is constant and the host graph is an infinite graph. The process ends when all vertices are cured, and the basic problem is to determine under which conditions on σ, and β the process ends almost surely. In the case of finite graphs, if $\sigma_v > 0$ for every vertex, it is easy to observe that the process ends a.s., so the problem becomes determining *how fast* the process ends.

The $k-$commodity dynamic demand model on a graph G is a variant of the contact process, $DD(\boldsymbol{\tau}(0), B, \boldsymbol{\sigma}, G)$. In this situation, the spread matrix B is an $k \times k$ real matrix (not assumed to be symmetric, or even non-negative), along with a supply function $\boldsymbol{\sigma} \colon V \to \mathbb{R}^k$, and the initial demand $\boldsymbol{\tau}(0) \colon V \to \mathbb{N}^k$. At time t, each node v has demand $\boldsymbol{\tau}_v(t) \in \mathbb{N}^k$, with each coordinate representing a different product in demand. A node v is said to be *satisfied* at time t if $\boldsymbol{\tau}_v(t) = \mathbf{0}$, and *unsatisfied* otherwise. The state of the process is described by the demand vector $\boldsymbol{\tau}$, where $\tau_v^j(t)$ is the demand for commodity j at node v at time t.

The spread matrix $B = [\beta_{ij}]$ describes how the demand for one commodity influences demands for other commodities. The i, j entry of B, β_{ij}, determines the spread rate of the demand for commodity j that is caused by demand for commodity i. In particular, we can describe the rate of spread events as follows. If v is a node that is unsatisfied at time t, and w an adjacent vertex, then there are spread events from v to w with rates $\max\{\boldsymbol{\tau}_v(t)B, 0\}$. That is the rate at which τ_w^j increases due to the demand at v is given by $\max\{\sum_i \tau_v^i(t)\beta_{ij}, 0\}$. Here, when we say an event occurs with rate λ, we mean that the elapsed time until that event takes place is distributed as $\text{Exp}(\lambda)$. Because the minimum of exponential

random variables is itself an exponential random variable, we can capture the total spreading rates in a condensed form. We define the rate function at time t, $\boldsymbol{\rho}(t)\colon V \to \mathbb{R}^k$, by

$$\boldsymbol{\rho}_v = \sum_{w \sim v} \boldsymbol{\tau}_w(t) B = \left(\boldsymbol{\tau}(t)(A \otimes B)\right)_v,$$

where $\boldsymbol{\tau}(t)$ is viewed as a vector with indices indexed by $V \times k$. $\rho_v^i(t)$ is the rate at which τ_v^i is increasing at time t.

Supply events occur with rates given by $\boldsymbol{\tau}(t)\mathrm{Diag}(\boldsymbol{\sigma})$, independently of any neighboring supply events. That is, the time until τ_v^i is decreased by 1 is distributed as $\mathrm{Exp}(\sigma_v^i \tau_v^i)$.

We briefly give a construction of the process, to show it is well-defined. Let $\overrightarrow{E}$ denote the set of ordered edges; that is, ordered pairs that are edges in the graph, so that uv and vu are distinct. We run independent Poisson point processes $\{X_e^{j,\boldsymbol{\rho}}\}_{e \in \overrightarrow{E}(G), j \in [k], \boldsymbol{\rho} \in \mathbb{N}^k}$ so that $X_e^{j,\boldsymbol{\rho}}$ is at rate $\max\{0, [\boldsymbol{\tau}_v B]_j\}$ and independent Poisson point processes $\{X_v^{i,n}\}_{v \in V(G), i \in [k], n \in \mathbb{N}}$ so that $X_v^{i,n}$ is at rate $n\sigma_v^i$. Then these countably many point processes can easily be seen to define the entire process; a spread event of type j from a vertex v to a vertex u which is currently in state $\boldsymbol{\rho}$ is controlled by the point process $X_{vu}^{j,\boldsymbol{\rho}}$ with satisfaction events handled similarly.

An advantage of such a formulation is that it gives an easy coupling between processes that shows that if $B' \leq B$ pointwise, the stochastic process $DD(\boldsymbol{\tau}(0), B, \boldsymbol{\sigma}, G)$ stochastically dominates $DD(\boldsymbol{\tau}(0), B', \boldsymbol{\sigma}, G)$ in the sense that in the coupling the demands in the B process are always at least those in the B' process. This is accomplished by noting that the rates $\boldsymbol{\rho}B \geq \boldsymbol{\rho}B'$ pointwise for all $\boldsymbol{\rho} \in \mathbb{N}^l$. We thus take point processes $Y_e^{j,\rho}$ at rate $[\rho B - \rho B']_i$. If the point processes $\{X_e^{j,\rho}\}$ and $\{X_v^{i,n}\}$ are used to determine $DD(\boldsymbol{\tau}(0), B, \boldsymbol{\sigma}, G)$, then the point processes $\{X_e^{j,\rho} \cup Y_e^{j,\rho}\}$ and $\{X_v^{i,n}\}$ are used to determine $DD(\boldsymbol{\tau}(0), B', \boldsymbol{\sigma}, G)$.

In particular, this allows us to replace B with B', where $B'_{ij} = \max\{B_{ij}, 0\}$, and conclusions about the extinction of the B' process still hold for B. Furthermore, this turns out not to be entirely unreasonable; one hopes that the negative entries in B would afford better bounds on the extinction time, but in many cases with negative entries in B extinctions of some demand types mean that the process is eventually run in a non-negative case. In light of this, we will assume for the

rest of this chapter that B is *non-negative* for convenience.

Given an initial demand $\boldsymbol{\tau}(0)$ and spread matrix B, our goal is to find a supply function $\boldsymbol{\sigma}$ such that demand is satisfied. Ideally we would like to do this with small supply rates. Furthermore, the supply rates should only depend on the contact graph G, the spread matrix B, and the initial demand $\boldsymbol{\tau}(0)$, but not on t or $\boldsymbol{\tau}(t)$.

6.3 The Kronecker PageRank

The notion of PageRank was first introduced by Brin and Page [13] in 1998 for Google's search algorithms. Although the PageRank was originally used for the Web graph, we can define the PageRank for any finite graph G. Here we will use a modified version of the PageRank, called personalized PageRank which has two parameters, a jumping constant $\alpha \in [0, 1]$ and a seed vector $\mathbf{s}$ which is some probability distribution on the vertex set V of G.

The *personalized PageRank* vector $\mathbf{pr}(\alpha, s)$ for jumping constant α and the seed distribution $\mathbf{s}$ on V is given by

$$\mathbf{pr}(\alpha, \mathbf{s}) = \alpha \sum_{l=0}^{\infty} (1-\alpha)^l \mathbf{s} W^l.$$

Note that here we view $\mathbf{s}$ as a row vector. We note that the PageRank vector is also the solution to the recurrence relation

$$\mathbf{pr}(\alpha, \mathbf{s}) = \alpha \mathbf{s} + (1-\alpha)\mathbf{pr}(\alpha, \mathbf{s})W.$$

The original formulation of PageRank [13] is the special case where $\mathbf{s}$ is the uniform distribution over all the vertices.

Recall that for a subset of vertices $H \subset V$, the volume of H is the sum of degrees of the vertices of H. The Cheeger Ratio of H, $h(H)$, measures the cut between H and $\bar{H}$ via the relationship

$$h(H) = \frac{e(H, \bar{H})}{\min\{\text{vol (H)}, \text{vol } (\bar{\text{H}})\}}.$$

The $\alpha-core$ of a subset H is the set of vertices

$$C_\alpha = \left\{ v \in H | \mathbf{pr}(\alpha, 1_v) 1_H \geq 1 - \frac{h}{\alpha} \right\}$$

A basic tool for analyzing PageRank is the fact (see [6]) that for a subset with Cheeger ratio h, we can choose α, say, $\alpha = h/2$, so that at least half of the vertices of H is in the α-core of H. Therefore, if the seed is in α-core of H, then we can use PageRank to identify a large part of H. Another advantage of using PageRank is the fact that there are very efficient algorithms for approximating PageRank vectors [6].

One tool that we will use to understand the $k-$commodity dynamic demand model will be a the Kronecker PageRank vector, which we define below. This is a generalization of the personalized PageRank vector.

Definition 6.3.1.

Let B be a square $k \times k$ matrix with spectral radius strictly less than 1, and W be the transition matrix for a random walk on a graph G. Let $\mathbf{s}$ be a non-negative vector in $\mathbb{R}^{k \times |V|}$. The Kronecker PageRank vector with parameters B and $\mathbf{s}$ is defined as

$$\mathbf{Kpr}(B, \mathbf{s}) = \sum_{l=0}^{\infty} \mathbf{s}(W \otimes B)^l = \sum_{l=0}^{\infty} \mathbf{s}(W^l \otimes B^l)$$

The condition that the spectral radius of B less than 1 is necessary to ensure convergence of the infinite sum, as the spectrum of $W \otimes B$ is the product of the spectra of W and B. Since the eigenvalues of W have absolute value at most 1, the sum will converge.

We note that in the case where B is a 1×1 matrix $B = \beta < 1$ and $\mathbf{s}$ is a probability distribution, then we have the relationship

$$\mathbf{Kpr}(B, \mathbf{s}) = \sum_{l=0}^{\infty} \mathbf{s}(W \otimes \beta)^l = \sum_{l=0}^{\infty} \mathbf{s}\beta^l W^l = \frac{1}{1-\beta} \mathbf{pr}(1-\beta, \mathbf{s}),$$

so the Kronecker PageRank is a natural extension of personalized PageRank. We will see in Theorem 6.5.3 that the Kronecker PageRank will arise naturally in our analysis in Section 6.3, and give better bounds than those that will be afforded by standard PageRank by incorporating the spread matrix. We remark that the

Kronecker PageRank vectors can be efficiently computed and approximated along the same line as that for the usual PageRank.

For a square matrix A, there are many different matrix norms that can be used (see, for example, [36]). We will use the following notation for the following norms:

1. $||A||_1 = \sum_{i,j} |a_{ij}|$ is the *ℓ_1 norm.*
2. $|||A|||_1 = \max_j \sum_i |a_{ij}|$ is the *maximum column sum norm.*
3. $|||A|||_\infty = \max_i \sum_j |a_{ij}|$ is the *maximum row sum norm.*
4. $|||A|||_2 = \max\{\sqrt{\lambda} \mid \lambda \text{ is an eigenvalue of } A^*A\}$ is the *spectral norm.*

6.4 A Solution that Supplies Every Vertex

Here we show that if supply rates are above a certain threshold, then with probability approaching 1 demands will be satisfied.

Theorem 6.4.1.
Consider the $k-$commodity demand model on a graph G with n vertices parameterized by spread matrix $B = [\beta_{ij}]$. If the supply rates to each vertex v satisfy

$$\sigma_v^i > d_v \left(\sum_j \frac{\beta_{ij} + \beta_{ji}}{2} \right) + \delta$$

for $\delta > 0$, then with probability $1 - \epsilon$ all vertices are satisfied at time t for all

$$t > \frac{1}{\delta} \left(\frac{1}{2} \log(nk) + \log(X(0)) + \log\left(\frac{1}{\epsilon}\right) \right).$$

Proof. We consider the expectation $\mathbb{E}[\boldsymbol{\tau}(t)]$. Let $X(t) = ||\boldsymbol{\tau}(t)||_1$, the total demand at time t. We will begin by considering the quantity $\frac{\partial}{\partial t}\mathbb{E}[\boldsymbol{\tau}(t)]$.

From the discussion in section 6.2, we know that demand is increasing with rates given by $\boldsymbol{\rho}(t) = \boldsymbol{\tau}(t)\,(A \otimes B)$, but also demand decreases according to the supply rates. Let $S = \text{diag}(\boldsymbol{\sigma})$, the diagonal $nk \times nk$ matrix with entries given

by the supply vector. Then we can see that demand decreases at each vertex according to rates given by the supply rate vector $\boldsymbol{\tau}(t)S$.

To proceed, we need the following two well known and simple facts concerning exponentially distributed random variables. We use the notation

$$f(h) = O_{h\to 0}(g(h))$$

to indicate that there exists a constant C such that $f(h) \leq C\dot{g}(h)$ for h sufficiently small.

Suppose X is an exponentially distributed waiting time with rate λ, then

$$\mathbb{P}(X < h) = \lambda h + O_{h\to 0}(h^2). \tag{6.1}$$

Equation 6.1 follows from the fact that the probability that an exponential waiting time is at most h is given by

$$\int_0^h \lambda e^{-\lambda h} = 1 - e^{-\lambda h} = \lambda h + O_{h\to 0}(h^2).$$

As an immediate consequence, we see that if X and Y are independent exponentially distributed waiting times with rates λ_1, λ_2, then

$$\mathbb{P}(X, Y < h) = O_{h\to 0}(h^2). \tag{6.2}$$

Using these two facts, we will show that we can neatly encode the behavior of the k-commodity demand model in a single differential equation.

Fix a vertex v and commodity i. We will show that

$$\frac{\partial}{\partial t}\mathbb{E}[\boldsymbol{\tau}_v^i(t)] = [\mathbb{E}[\boldsymbol{\tau}(t)](A \otimes B - S)]_v^i,$$

To do this, we compute the derivative by the definition, that is we compute

$$\lim_{h\to 0} \frac{\mathbb{E}[\boldsymbol{\tau}_v^i(t) - \boldsymbol{\tau}_v^i(t+h)]}{h}.$$

To do this, consider the conditional expectation, $\mathbb{E}[\boldsymbol{\tau}_v^i(t) - \boldsymbol{\tau}_v^i(t+h)|\boldsymbol{\tau}(t)]$. Note that by Equation 6.2, the probability that two independent events (either two spread events, or two satisfy events or a spread and a satisfy event) occur is $O_{h\to 0}(h^2)$. On the other hand, given a neighbor u of v, and a commodity j, the

probability of a spread event originating from this neighbor and commodity in time $(t, t+h)$ is exactly $B_{ji}\tau_u^j(t)h + O_{h\to 0}(h^2)$. Likewise, the probability of a satisfaction event in this time is $\tau_v^i(t)\sigma_v^i h + O_{h\to 0}(h^2)$. Linearity of expectation yields

$$\mathbb{E}[\boldsymbol{\tau}_v^i(t) - \boldsymbol{\tau}_v^i(t+h)|\boldsymbol{\tau}(t)] = \boldsymbol{\tau}(t)(A \otimes B - S)h + o(h^2).$$

Applying the tower property of conditional expectation yields the result for this vertex v and commodity i. Since this holds for all choices of v and i, we obtain the single equation

$$\frac{\partial}{\partial t}\mathbb{E}[\boldsymbol{\tau}(t)] = \mathbb{E}[\boldsymbol{\rho}(t) - \boldsymbol{\tau}(t)S] = \mathbb{E}[\boldsymbol{\tau}(t)](A \otimes B - S). \tag{6.3}$$

Solving the matrix differential equation with initial condition $\mathbb{E}[\boldsymbol{\tau}(0)] = \boldsymbol{\tau}(0)$ yields

$$\mathbb{E}[\boldsymbol{\tau}(t)] = \boldsymbol{\tau}(0)e^{t(A\otimes B - S)}. \tag{6.4}$$

Let $Q = A \otimes B - S$. Then by [25], $|||e^{tQ}|||_2 \leq e^{t\nu}$, where ν is the largest eigenvalue of $\frac{Q+Q^*}{2}$. We note that $\frac{Q+Q^*}{2} = A \otimes (\frac{B+B^*}{2}) - S$, which has diagonal terms $\beta_{ii} - \sigma_v^i$, ranging over all values of v and i. By the Gershgorin Circle Theorem, the eigenvalues of $\frac{Q+Q^*}{2}$ are contained in the intervals

$$[-(d_v - 2)\beta_{ii} - d_v\left(\sum_{j\neq i}\frac{\beta_{ij}+\beta_{ji}}{2}\right) - \sigma_v^i, d_v\left(\sum_{j}\frac{\beta_{ij}+\beta_{ji}}{2}\right) - \sigma_v^i].$$

Since $\sigma_v^i > d_v\left(\sum_j \frac{\beta_{ij}+\beta_{ji}}{2}\right) + \delta$ all the eigenvalues of $\frac{Q+Q^*}{2}$ are less than $-\delta$. Therefore

$$\begin{aligned}
\mathbb{E}[X(t)] &= ||\tau(0)e^{t(A\otimes B-S)}||_1 \leq \sqrt{nk}||\tau(0)e^{t(A\otimes B-S)}||_2 \\
&\leq \sqrt{nk}||\tau(0)||_2|||e^{t(A\otimes B-S)}|||_2 \leq \sqrt{nk}||\tau(0)||_1 e^{t\nu} \\
&\leq \sqrt{nk}X(0)e^{-t\delta}
\end{aligned}$$

Thus Markov's inequality gives that $\mathbb{P}(X(t) > 0) < \epsilon$ if
$t > \frac{1}{\delta}\left(\frac{1}{2}\log(nk) + \log(X(0)) + \log\left(\frac{1}{\epsilon}\right)\right)$. □

We note that this approach works for all initial distributions $\boldsymbol{\tau}(0)$. This indicates that in many situations this approach may be overkill and that we could have used smaller supply rates. In the next section, we analyze the process more carefully and give conditions that depend on the initial distribution of demand.

6.5 A Solution that Supplies a Small Subset

For the remainder of the discussion, it is convenient to introduce reformulation of the model that takes advantage of the fact that demands take on integer values. Rather than view demands as a function $\tau\colon V \to \mathbb{N}^k$, we view demands as discrete objects sitting on each node. Borrowing language from chip-firing games on graphs (see, for example, [46]) we view units of the demand as chips located on vertices of the graph. For example, if $k = 7$ and for a vertex v $\boldsymbol{\tau}_v(t) = (0, 1, 2, 0, 2, 0, 3)$ then we would say that at time t there was 1 "2-chip", 2 "3-chips", 2 "5-chips", and 3 "7-chips" at vertex v, corresponding to 1 "unit of demand" for commodity 2, etc. Unlike in classic chip-firing games, the number of chips is not static, and the game is parameterized by continuous time. We restate the possible transitions in terms of demand chips. For an $i-$chip at vertex v, there are two types of transition events:

- For each vertex $w \sim v$ and each $j = 1, ..., l$, a $j-$chip is added at w with rate β_{ij}. When this occurs we say that the new $j-$chip is created by the $i-$chip.
- The $i-$chip itself is removed with rate σ_v^i.

It is important to note that due to the properties of exponential random variables, the rates add linearly and the model is equivalent to the original description discussed in Section 6.2. The main advantage of this reformulation is the ability to trace back the history of a chip. If there is a chip c_l at vertex v_l at time t, then either c_l existed at time $t = 0$, or there is a sequence of l chips $(c_0, ..., c_l)$ located at vertices along a path $\pi = (v_0, v_1, ..., v_l)$ where c_0 existed at $t = 0$, and c_r is created by c_{r-1} for $r = 1, ..., l$. We allow π to have repeated vertices to allow for the case where demand created more demand at the same vertex. If a chip c exists at time 0, we refer to it as an *initial chip*.

For a path $\pi = (v_0, v_1, ..., v_l)$ and a chip c_0 located at v_0, we define the event $\mathcal{S}_{\pi,c_0}$ to be the event there is a sequence of m chips $(c_0, ..., c_l)$ located respectively at $(v_0, v_1, ..., v_l)$ and c_r is created by c_{r-1} for $r = 1, ..., l$.

It is important to note that $\mathcal{S}_{\pi,c_0}$ occurring does not imply that there is any demand at v_l at time t because it could be satisfied sometime before t. However, if there is a demand at v_l at time t, then $\mathcal{S}_{\pi,c}$ must have occurred for some initial chip c at vertex v_0 and some walk π from v_0 to v_l.

We begin by showing that if supply rates are large enough, then the probability of demands spreading along a long walk is small. Inspired by Theorem 6.4.1 we make the assumption that supply rates are proportional to the degree of the vertices. That is, we assume that $\sigma_v^i > \mu_i(d_v)$ for all v for constants $\mu_i > 0$.

Lemma 6.5.1.

Let $M = diag(\mu_1, ..., \mu_k)$, $\hat{B} = M^{-1}B$ and $\zeta = \min\{|||\hat{B}|||_1, |||\hat{B}|||_\infty\}$

Then for any chip c_0 located at v_0 and any walk $\pi = (v_0, ..., v_l)$ of length l,

$$P(S_{\pi,c_0}) \leq k \prod_{j=0}^{l} \frac{1}{d_{v_j}} \zeta^l$$

Proof. Let S_r denote the event that a chip c_r at v_r creates a chip at v_{r+1}. If c_r is an $i-$chip, then for it to create any chip at v_{r+1} a spread event must occur before c_r is removed. The time until c_r creates a $j-$chip at v_{r+1} is an exponential random variable with rate β_{ij}. Since the time until c_r is removed is given by $\text{Exp}(\sigma_v^i)$, the probability of c_r creating a $j-$chip is $\frac{\beta_{ij}}{\beta_{ij}+\sigma_v^i} \leq \frac{\beta_{ij}}{\sigma_v^i} < \frac{\beta_{ij}}{\mu_i d_{v_r}}$. Thus $\mathbb{P}(S_r) < \sum_{i,j} \frac{\beta_{ij}}{\mu_i d_{v_r}} = \frac{1}{d_{v_r}} \mathbb{1}\hat{B}\mathbb{1}^*$.

For a walk π of length l, we want to consider the intermediate steps more carefully. Since there are l transitions that occur, we can use the same reasoning as above to obtain the bound

$$\mathbb{P}(S_{\pi,c}) < \prod_{r=0}^{l} \frac{1}{d_{v_r}} \mathbb{1}\hat{B}^l\mathbb{1}^* = \prod_{r=0}^{l} \frac{1}{d_{v_r}} ||\hat{B}^l||_1 \leq k \prod_{r=0}^{l} \frac{1}{d_{v_r}} |||\hat{B}^l|||_1 \leq k \prod_{r=0}^{l} \frac{1}{d_{v_r}} |||\hat{B}|||_1^l$$

The factor of k that appears in the final lines above is just a consequence of switching from the vector 1-norm $||\hat{B}^l||_1$ to maximum column sum norm $|||\hat{B}^l|||_1$ (see [36]).

We could have just as easily switched to the maximum row sum norm and obtained the term $k|||\hat{B}|||_\infty^l$, and so it follows that

$$\mathbb{P}(S_{\pi,c}) < \min\{k\prod_{r=0}^{l}\frac{1}{d_{v_r}}|||\hat{B}|||_1^l, k\prod_{r=0}^{l}\frac{1}{d_{v_r}}|||\hat{B}|||_\infty^l\} = k\prod_{j=0}^{l}\frac{1}{d_{v_j}}\zeta^l$$

□

We note that the decision to use $\zeta = \min\{|||\hat{B}|||_1, |||\hat{B}|||_\infty\}$ in Lemma 6.5.1 reflects the difficulty in working with arbitrary spread matrices B. For certain classes of spread matrices (e.g. if B is symmetric or diagonalizable) it is possible to obtain tighter bounds. While the previous lemma will be useful for obtaining a bound using PageRank, a more careful analysis is possible which will lead naturally to use of Kronecker PageRank which we explore in Theorem 6.5.3.

Theorem 6.5.2.

Suppose that initial demand is contained in $S \subset H \subset V$ with and each vertex $v \in H$ has supply rates $\sigma_v^i > \mu_i d_v$, and $\sigma_w^i = 0$ for $w \in \bar{H}$. Let $M = diag(\mu_1, ..., \mu_k)$, $\hat{B} = M^{-1}B$ and $\zeta = \min\{|||\hat{B}|||_1, |||\hat{B}|||_\infty\}$. Let $\boldsymbol{x}(t)$ be defined by $x_v(t) = \sum_i \tau_v^i(t)$, and $X(t) = ||\boldsymbol{x}(t)||_1 = ||\boldsymbol{\tau}(t)||_1$. Let E_H denote the event that demands spread outside the set H. Then

1. $\mathbb{P}(E_H) \leq \frac{X(0)}{\zeta} pr\left(1-\zeta, \frac{\boldsymbol{x}(0)}{X(0)}\right) 1_{\bar{H}}^*$
2. *If S in the $(1-\zeta)$ core of H, then $\mathbb{P}(E_H) \leq \frac{2X(0)h(H)}{\zeta(1-\zeta)}$, where $h(H)$ is the Cheeger ratio of H.*

Proof. Let P_l denote the set of all paths of length l from an initial chip in S to $\bar{H}$ such that the first $k-1$ steps are in H. Let $P = \bigcup_{l=1}^{\infty} P_l$. The key observation is that if $w \in \bar{H}$ ever has demand, then $S_{\pi,c}$ must have occurred for some initial chip c and path π from the location of c to w. Thus we can use the union bound to get

that

$$\begin{aligned}
\sum_{\pi\in P}\mathbb{P}(S_{\pi,c}) &\leq \sum_{l}\sum_{(\pi,c)\in P_l}\mathbb{P}(S_{\pi,c}) \\
&\leq \sum_{l}\sum_{v_0\in S}\sum_{c\ at\ v_0}\sum_{v_l\in\bar{H}}\sum_{\pi=(v_0,\ldots,v_l)\in P_l}\mathbb{P}(S_{\pi,c}) \\
&\leq \sum_{l}\sum_{v_0\in S}\sum_{c\ at\ v_0}\sum_{v_l\in\bar{H}}\sum_{\pi=(v_0,\ldots,v_l)\in P_l}\zeta^l\prod_{r=0}^{l}\frac{1}{d_{v_r}} \\
&= \sum_{l}\mathbf{x}(0)\zeta^l(D^{-1}A)^l 1^*_{\bar{H}} \\
&= \sum_{l}\mathbf{x}(0)\zeta^l W^l 1^*_{\bar{H}} \\
&= \frac{X(0)}{\zeta}\mathbf{pr}\left(1-\zeta,\frac{\mathbf{x}}{X(0)}\right)1^*_{\bar{H}},
\end{aligned}$$

proving the first statement. The second statement follows the same proof as Theorem 3.2 of [23]. □

Theorem 6.5.3.

Suppose that the initial demand is contained in $S \subset H \subset V$ with and each vertex $v \in H$ has supply rates $\sigma_v^i \geq \mu_i d_v$. Let $M = diag(\mu_1, ..., \mu_k)$, $\hat{B} = M^{-1}B$ and $\zeta = ||\hat{B}||_1$. Let $X(t) = ||\boldsymbol{\tau}(t)||_1$, the total amount of demands at time t. Let $\mathcal{E}_H$ denote the event that demands spread outside the set H. Then $\mathcal{E}_H$ can be bounded above using the Kronecker PageRank vector via the relationship:

$$\mathbb{P}(\mathcal{E}_H) \leq X(0)\boldsymbol{Kpr}\left(\hat{B},\frac{\tau(0)}{X(0)}\right)1_{\bar{H}}$$

Proof. Let f be a vector indicator function of commodity type on chips, that is $f(c) = \mathbf{e}_i$ if c is an $i-$chip, where $\mathbf{e}_i$ denotes the ith standard basis vector for $\mathbb{R}^k$. Let C_0 denote the set of initial chips. By the same methods that were used in the proof of Lemma 6.5.1, we can bound the probability that demand originating from c ever spreads along a path $\pi = (v_0, v_1, ..., v_l)$ by the sum

$$\mathbb{P}(\mathcal{S}_{\pi,c}) \leq f(c)\hat{B}^l\mathbb{1}^*\prod_{r=0}^{l}\frac{1}{d_{v_r}}$$

Therefore using the same technique as in the proof of Theorem 6.5.2 we obtain the bound

$$\begin{aligned}\sum_{\pi\in P}\mathbb{P}(\mathcal{S}_{\pi,c}) &\leq \sum_{l}\sum_{u\in S}\sum_{\pi\in B_l}\mathbb{P}(\mathcal{S}_{\pi,u}) \leq \sum_{l}\sum_{c\in C_0}\sum_{v_l\in\bar{H}}\sum_{\pi=(v_0,\ldots,v_l)\in P_l}\mathbb{P}(\mathcal{S}_{\pi,u})\\ &\leq \sum_{l}\sum_{c\in C_0}\sum_{v_l\in\bar{H}}\sum_{\pi=(v_0,\ldots,v_l)\in P_l} f(c)\hat{B}^l\mathbb{1}^*\prod_{r=0}^{l}\frac{1}{d_{v_r}} = \sum_{l}\tau(0)(D^{-1}A\otimes\hat{B})^l\mathbf{1}_{\bar{H}}\\ &= \sum_{l}\tau(0)(W\otimes\hat{B})^l\mathbf{1}_{\bar{H}} = X(0)\mathbf{Kpr}(\hat{B},\frac{\tau(0)}{X(0)})\mathbf{1}_{\bar{H}}\end{aligned}$$

□

Let $\mathcal{S}_t$ denote the event that all of the vertices are satisfied at time t. In order to complete the analysis of the local case, we would like to bound $\mathbb{P}(\mathcal{S}_t|\overline{\mathcal{E}}_H)$, where $\mathcal{E}_H$ is as in Theorems 6.5.2 and 6.5.3. Such a bound is not immediately given by Theorem 6.4.1, but this is not difficult. To derive a bound on $\mathbb{P}(\mathcal{S}_t|\overline{\mathcal{E}}_H)$, consider running a modified Dirichlet version which is identical to the standard process with the same supply rates, except demand leaving H is ignored. Let $\mathcal{S}'_t$ denote the event that in Dirichlet process, all of the events are satisfied at time t then $\mathbb{P}(\mathcal{S}'_t)$ can be bounded directly by Theorem 6.4.1 as this Dirichlet process restricted to vertices in H is the standard process on H. Furthermore $\mathbb{P}(\mathcal{S}_t\cap\mathcal{E}_H)\leq\mathbb{P}(\mathcal{S}'_t)$. Therefore

$$\mathbb{P}(\mathcal{S}_t|\overline{\mathcal{E}}_H) = \frac{\mathbb{P}(\mathcal{S}_t\cap\overline{\mathcal{E}}_H)}{\mathbb{P}(\overline{\mathcal{E}}_H)} \leq \frac{\mathbb{P}(\mathcal{S}'_t)}{\mathbb{P}(\overline{\mathcal{E}}_H)}.$$

Combining this observation along with Theorems 6.5.2 and 6.5.3, yields that the probability of escape from H is bounded and if the process does not escape from H it dies quickly.

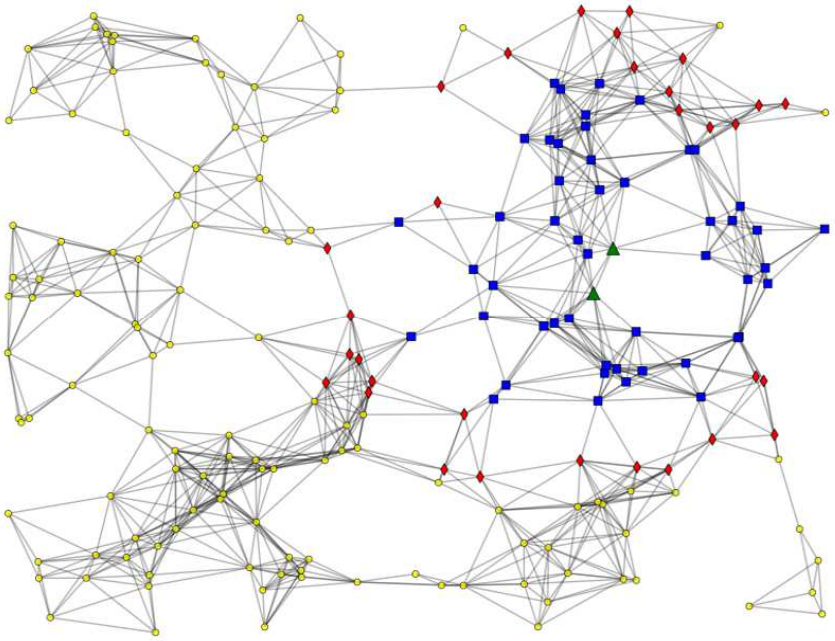
Theorems 6.5.2 and 6.5.3 can be used in two different ways. As stated, they provide a way to bound the probability demands escape from a given subset. However, they can be also used to construct such a bounding subset. For example, given initial demand $\boldsymbol{\tau}(0)$ contained in an initial set of vertices $S\subset V$, we can algorithmically construct H such that demand stays in H with probability $1-\epsilon$ as follows. We do this by constructing a increasing family of subsets $\{H_r\}$. We begin

by setting $H_0 = S$ and $\mathbf{pr} = \frac{X(0)}{\zeta}\mathbf{pr}\left(1-\zeta, \frac{\boldsymbol{\tau}(0)}{X(0)}\right)$. Then we follow the following procedure:

1. Compute $f(r) = \langle \mathbf{pr}, \mathbf{1}^*_{\bar{H}_r} \rangle$
2. If $f(r) < \epsilon$, set $H = H_r$ and end the process.
3. If $f(r) \geq \epsilon$, let $H_{r+1} = H_r \cup v$ for some $v \notin H_r$, and return to step 1.

This process will eventually terminate, since $|H_{r+1}| = |H_r| + 1$, and $f(r) = 0$ once $H_r = V$.

6.6 Resource Allocation on a Random Geometric Graph



We conclude with an example calculation on synthetic data. Our graph G is an instance of a random geometric graph. 200 vertices were placed uniformly at random in a unit square, and two vertices are adjacent if the distance between

them is less than .13 units. We let $k = 3$, and

$$B = \begin{pmatrix} .4 & .2 & .3 \\ .2 & .3 & .4 \\ .3 & .4 & .5 \end{pmatrix}.$$

The initial demand is given by $\tau_v^j(0) = 1$ for all commodities for the vertices marked by triangles, and $\tau_v^j(0) = 0$ for all other vertices. In addition we set $\mu_i = 3$ for all i.

We demonstrate the difference between Theorems 6.5.2 and 6.5.3 in the following way. The figure above shows the graph G. The demands start in the green triangular vertices and spread outward from there. Theorem 6.5.3 states that with 95% probability, demands stay in the blue square vertices. Theorem 6.5.2 states that with 95% probability, demands stay in the red diamond and blue square vertices. This small example illustrates how the Kronecker PageRank can be used to obtain improved results.

Acknowledgement:
This chapter is based on the paper Multi-commodity allocation for dynamic demands using PageRank vectors which was written jointly with Fan Chung and Paul Horn. It appeared in the Proceedings of the 9th Workshop on Algorithms and Models for the Web Graph [22].

Index

Bibliography

[1] D. Aldous. Hitting times for random walks on vertex-transitive graphs. *Math. Proc. Cambridge Philos. Soc*, 106(1):179–191, 1989.

[2] D. Aldous. Mixing time for a markov chain on cladograms. *Combinatorics Probability and Computing*, 9(03):191–204, 2000.

[3] B. L. Allen and M. Steel. Subtree transfer operations and their induced metrics on evolutionary trees. *Annals of Combinatorics*, 5(1):1–15, 2001.

[4] A. T. Amin, L. H. Clark, and P. J. Slater. Parity dimension for graphs. *Discrete mathematics*, 187(1):1–17, 1998.

[5] A. T. Amin, P. J. Slater, and G. Zhang. Parity dimension for graphs - a linear algebraic approach. *Linear and Multilinear Algebra*, 50(4):327–342, 2002.

[6] R. Andersen, F. Chung, and K. Lang. Local partitioning for directed graphs using PageRank. *Proceedings of the 5th international conference on Algorithms and models for the web-graph*, pages 166–178, 2007.

[7] M. Anderson and T. Feil. Turning lights out with linear algebra. *Mathematics Magazine*, 71(4):300–303, 1998.

[8] O. Bastert, D. Rockmore, P. F. Stadler, and G. Tinhofer. Some properties of Robinson graphs. Technical report, 2000.

[9] O. Bastert, D. Rockmore, P. F. Stadler, and G. Tinhofer. Landscapes on spaces of trees. *Applied Mathematics and Computation*, 131(2-3):439–459, 2002.

[10] P. Bidigare, P. Hanlon, and D. Rockmore. A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Mathematical Journal*, 99:135–174, 1999.

[11] C. Borgs, J. Chayes, A. Ganesh, and A. Saberi. How to distribute antidote to control epidemics. *Random Structures & Algorithms*, 37(2):204–222, 2010.

[12] R. Brewster and T. Graves. Edge-switching homomorphisms of edge-coloured graphs. *Discrete Mathematics*, 309(18):5540–5546, 2009.

[13] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107–117, 1998.

[14] K. S. Brown. Semigroups, rings, and markov chains. *Journal of Theoretical Probability*, 13(3):871–938, 2000.

[15] K. S. Brown. *Semigroup and ring theoretical methods in probability*. 2004.

[16] K. S. Brown and P. Diaconis. Random walks and hyperplane arrangements. *The Annals of Probability*, 26(4):1813–1854, 1998.

[17] A. J. J. Caceres, S. Daley, J. DeJesus, M. Hintze, D. Moore, and K. St. John. Walks in phylogenetic treespace. *Information Processing Letters*, 111(12):600–604, 2011.

[18] P. J. Cameron and S. Tarzi. Switching with more than two colours. *European J. Combinatorics*, 25(2):169–177, 2009.

[19] Y. Chevaleyre, P. E. Dunne, U. Endriss, J. Lang, M. Lemaitre, N. Maudet, J. Padget, S. Phelps, J. A. Rodriguez-Aguilar, and P. Sousa. Issues in multi-agent resource allocation. *Informatica*, 2006.

[20] F. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1997.

[21] F. Chung and R. L. Graham. Edge flipping in graphs. *Advances in Applied Mathematics*, 48(1):37–63, 2012.

[22] F. Chung, P. Horn, and J. Hughes. Multi-commodity allocation for dynamic demands using PageRank vectors. *Lecture Notes in Computer Science*, Proceedings of the 9th Workshop on Algorithms and Models for the Web Graph(7323), 2012.

[23] F. Chung, P. Horn, and A. Tsiatas. Distributing antidote using PageRank vectors. *Internet Mathematics*, 6(2):237–254, 2009.

[24] F. Chung and J. Hughes. A note on an alternating upper bound for random walks on semigroups. 2012. Submitted to Discrete Applied Mathematics.

[25] G. Dahlquist. Stability and error bounds in the numerical integration of ordinary differential equations. *Kungl. Tekn. Hgsk. Handl. Stockholm. No.*, 130:87, 1959.

[26] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walk on finite groups. *The Annals of Probability*, pages 2131–2156, 1993.

[27] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible markov chains. *The Annals of Applied Probability*, pages 696–730, 1993.

[28] A. J. Drummond and A. Rambaut. BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7(1):214, 2007.

[29] S. Edwards, V. Elandt, N. James, K. Johnson, Z. Mitchell, and D. Stephenson. Lights out on finite graphs. *Involve*, 3(1):17–32, 2010.

[30] J. Fill. An exact formula for the move-to-front rule for self-organizing lists. *Journal of Theoretical Probability*, 9(1):113–160, 1996.

[31] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 2, pages 1455–1466 vol. 2, 2005.

[32] S. V. Gervacio and H. Maehara. A note on lights-out-puzzle: Parity-state graphs. *Graphs and Combinatorics*, 27:109–119, 2010.

[33] A. Giffen and D. Parker. On generalizing the Lights outgame and a generalization of parity domination. *preprint*, 2009.

[34] J. Goldwasser, W. Klostermeyer, and G. Trapp. Characterizing switch-setting problems. *Linear and Multilinear Algebra*, 43(1-3):121–135, 1997.

[35] J. Goldwasser, X. Wang, and Y. Wu. Does the lit-only restriction make any difference for the sigma game and the sigma+-game? *European Journal of Combinatorics*, 30(4):774–787, 2009.

[36] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[37] J. Hughes. Random seidel switching on graphs. *The Journal of Combinatorial Mathematics and Combinatorial Computing*, 2012. To Appear.

[38] J. Hughes. Random lights out processes on graphs. *Advances in Applied Mathematics*, 2013.

[39] G. James and M. Liebeck. *Representations and Characters of Groups*. Cambridge University Press, 2nd edition, 2001.

[40] I. Z Kiss, D. M Green, and R. R Kao. Infectious disease control using contact tracing in random and scale-free networks. *Journal of The Royal Society Interface*, 3(6):55–62, 2006.

[41] F. Klein-Barmen. Uber eine weitere verallgemeinerung des verbandsbegriffes. *Mathematische Zeitschrift*, 46:472–480, 1940.

[42] M. Li, J. Tromp, and L. Zhang. On the nearest neighbour interchange distance between evolutionary trees. *Journal of Theoretical Biology*, 182(4):463–467, 1996.

[43] S. Li, D. K. Pearl, and H. Doss. Phylogenetic tree construction using Markov chain monte carlo. *Journal of the American Statistical Association*, 95(450):493–508, 2000.

[44] C. L. Mallows and N. J. A. Sloane. Two-graphs, switching classes and euler graphs are equal in number. *SIAM Journal on Applied Mathematics*, 28(4):876–880, 1975.

[45] B. Mau, M. A. Newton, and B. Larget. Bayesian phylogenetic inference via markov chain monte carlo methods. *Biometrics*, 55(1):1–12, 1999.

[46] C. Merino. The chip-firing game. *Discrete Mathematics*, 302(1-3):188–210, 2005.

[47] G. W. Moore, M. Goodman, and J. Barnabas. An iterative approach from the standpoint of the additive hypothesis to the dendrogram problem posed by molecular data sets. *Journal of Theoretical Biology*, 38(3):423–457, 1973.

[48] M. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66(1), 2002.

[49] I. Pak and V. H. Vu. On mixing of certain random walks, cutoff phenomenon and sharp threshold of random matroid processes. *Discrete Applied Mathematics*, 110(2-3):251–272, 2001.

[50] R. M. Phatarfod. On the matrix occurring in a linear search problem. *Journal of Applied Probability*, 28(2):336–346, 1991.

[51] D. F. Robinson. Comparison of labeled trees with valency three. *Journal of Combinatorial Theory, Series B*, 11(2):105–119, 1971.

[52] M. Schtzenberger. Sur certains treillis gauches. *C. R. Acad. Sci. Paris*, 224:776–778, 1947.

[53] J. Schweinsberg. An o(n2) bound for the relaxation time of a markov chain on cladograms. *Random Structures and Algorithms*, 20(1):59–70, 2002.

[54] J. J. Seidel. Strongly regular graphs with (-1,1,0) adjacency matrix having eigenvalue 3. *Linear Algebra and Appl.*, 1:281–298, 1968.

[55] J. J. Seidel. A survey of two-graphs. In *Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo I*, pages 481–511. Accad. Naz. Lincei, Rome, 1976.

[56] C. Semple and M. A Steel. *Phylogenetics*. Oxford University Press, Oxford; New York, 2003.

[57] D. Stefankovic and E. Vigoda. Fast convergence of markov chain monte carlo algorithms for phylogenetic reconstruction with homogeneous data on closely related species. *SIAM Journal on Discrete Mathematics*, 25(3):1194–1211, 2011.

[58] B. Steinberg. An algebraic approach to the brown-diaconis upper bound lemma for left regular band walks. 2009. preprint.

[59] K. Sutner. Linear cellular automata and the garden-of-eden. *The Mathematical Intelligencer*, 11(2):49–53, 1989.

[60] K. Sutner. Sigma - automata and chebyshev-polynomials. *Theoretical Computer Science*, 230(1-2):49–73, 2000.

[61] J. H. van Lint and J. J. Seidel. Equilateral point sets in elliptic geometry. *Nederl. Akad. Wetensch. Proc. Ser. A*, 28(69):335–348, 1966.

[62] Y. Wu. Lit-only sigma game on a line graph. *European Journal of Combinatorics*, 30(1):84–95, 2009.