Dartmouth College

Dartmouth Digital Commons

Dartmouth College Ph.D Dissertations

Theses and Dissertations

5-1-2016

Algorithmic Problems Arising in Posets and Permutations

Shahrzad Haddadan
Dartmouth College

Follow this and additional works at: https://digitalcommons.dartmouth.edu/dissertations

Part of the Computer Sciences Commons

Recommended Citation

Algorithmic Problems Arising in Posets and Permutations

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

in

Computer Science

by

Shahrzad Haddadan

DARTMOUTH COLLEGE

Hanover, New Hampshire

Examining Committee:

_____

Peter Winkler (chair)

_____

Amit Chakrabarti

_____

Scot Drysdle

_____

James Propp

_____

F. Jon Kull, Ph.D.

Dean of Graduate Studies

i

Abstract

Partially ordered sets and permutations are combinatorial structures having vast applications in theoretical computer science. In this thesis, we study various computational and algorithmic problems related to these structures.

The first chapter of the thesis contains discussion about randomized fully polynomial approximation schemes obtained by employing Markov chain Monte Carlo. In this chapter we study various Markov chains that we call: the gladiator chain, the interval chain, and cube shuffling. Our objective is to identify some conditions that assure rapid mixing; and we obtain partial results. The gladiator chain is a biased random walk on the set of permutations. This chain is related to self organizing lists, and various versions of it have been studied. The interval chain is a random walk on the set of points in $\mathbb{R}^n$ whose coordinates respect a partial order. Since the sample space of the interval chain is continuous, many mixing techniques for discrete chains are not applicable to it. The cube shuffle chain is a generalization of Håstad's square shuffle. The importance of this chain is that it mixes in constant number of steps.

In the second chapter, we are interested in calculating expected value of real valued function $f : S \to \mathbb{R}$ on a set of combinatorial structures S, given a probability distribution on it. We first suggest a Markov chain Monte Carlo approach to this problem. We identify the conditions under which our proposed solution will be efficient, and present examples where it fails. Then, we study homomesy. Homomesy is a phenomenon introduced by Jim Propp and Tom Roby. We say the triple $\langle S, \tau, f \rangle$ (τ is a permutation mapping S to itself) exhibits homomesy, if the average of f along all τ-orbits of S is a constant only depending on f and S. We study homomesy and obtain some results when S is the set of ideals in a class of simply described lattices.

Dedication

To my parents.

Acknowledgment

Contents

List of Figures

Introduction

Partially ordered sets, permutations, and ideals are among the combinatorial structures that have been interesting to both combinatorists and computer scientists. Here we study some computational and algorithmic problems concerning these structures. These problems vary from Markov Chain mixing problems to computing the expected value of different statistics of the above mentioned structures by employing randomized (Markov chains) or deterministic (homomesy) methods. The work can be divided into the following sections:

Markov Chain Mixing Problems. Suppose that samples are needed from a probability distribution π on a set Ω whose size is exponential in some parameter n. In most cases, no polynomial-time algorithm is known by which we can sample elements of S exactly from a given distribution; in such a case, computer scientists typically employ approximate sampling using Markov chains.

Let P be the transition matrix of an ergodic Markov chain $X_0, X_1, \ldots$

with state space S and stationary distribution π. Thus

$$
\begin{aligned}
& Pr(X_i = x_i \mid X_{i-1} = x_{i-1}, \ X_{i-2} = x_{i-2}, \ldots, X_0 = x_0) \\
&= Pr(X_i = x_i \mid X_{i-1} = x_{i-1}) \\
&= P(x_i, x_{i-1}).
\end{aligned}
$$

The idea of Markov chain Monte Carlo is to run such a chain for, say, t steps and use X_t as the sample. Since the distribution of X_t approaches π, this will work if t is large enough. The difficulty is in showing that one can get close to π with a value of t that is only polynomial in n.

The least value of t for which the total variation distance between π and the distribution of X_t is less than some constant (say, $1/4$) is known as *mixing time*. We say that a Markov chain has *polynomial mixing* if its mixing time is polynomial in the size parameter n. In this thesis, we will discuss two mixing problems.

In the first problem, the sample space Ω is the symmetric group S_n, that is, the set of all permutations of n objects. Let f be a positive real valued "strength" function on the set $\{1, \ldots, n\}$ and define a Markov chain on S_n by choosing i uniformly at random in $\{1, \ldots, n-1\}$, then swapping $\sigma(i)$ and $\sigma(i+1)$ with probability $f(\sigma(i))/(f(\sigma(i)) + f(\sigma(i+1)))$. We call this chain *the gladiator chain*. The gladiator chain has a simply-described stationary distribution, and it seems to mix rapidly. In fact, I showed that the chain mixes rapidly in some cases where the gladiators fall into only a few strength classes [20]. It is worth mentioning that Jim Fill (see [23]) has conjectured

that all cases of a more general swapping chain mix rapidly, but only certain special cases have so far been proved.

The second mixing problem is motivated by the hard-core gas model in dimension 1. Given n points in the unit interval, a natural way to mix them is to choose one uniformly at random and move it to a random point between its two neighbors, or between its neighbor and an endpoint; it has been shown that $n^3 \log n$ such steps are necessary and sufficient to mix [36]. What if the points are constrained not by linear order but by a partial order? If they are all incomparable, then mixing in time $n \log n$ is automatic (via "coupon-collection"). We study this chain in Section 1.4, and present experimental and theoretical evidences that it mixes rapidly.

Mixing Faster by Taking Larger Steps. Randomized algorithms allow us to find approximate solutions to some NP-hard problems in polynomial time. However, Markov chains that make only local changes are limited in speed, since it takes $n \log n$ steps (coupon collector problem) just to hit all the sites of a space of configurations based on an n-element set. This bound can sometimes be overcome by the use of chains with big ("global") steps.

In the following we will study two problems in this topic. The problems are Håstad's square lattice shuffle (see [38]) and the top to random shuffle.

Suppose an $m \times m$ array of distinct objects is given. In order to perform one step of Håstad's square lattice shuffle, we pick m independent permutations $(\sigma_i)_{i=1}^m$ of $\{1, \ldots, m\}$ uniformly at random. At odd time steps, we apply each σ_i to the elements in row i. At even steps another such set of

permutations is applied to elements of each column.

This is a huge step, and it is shown in Håstad's paper [38] that only a constant number of these steps are required to mix. Håstad's theorem sticks to the two dimensional case. Here, by adding a few lemmas we generalize Håstad's work to cubes in arbitrary dimension. In particular, we prove that for any arbitrary k, a cube of size n^k mixes in time t, where t is only dependent on k but not on n.

Calculating the Expected Value of Statistics and Homomesy. Consider a sample space S of combinatorial structures and statistic $f : S \to \mathbb{R}$. For instance, we take S to be the set of all ideals in a poset and f to be the size of an ideal. We are interested in the expected value of f; $\hat{f} = \sum_{s \in S} f(s)\pi(s)$, when π is an arbitrary distribution on S. We study two different approaches to this problem: running a Markov chain (randomized method) and verifying *homomesy* (deterministic method).

In the first approach, the goal is to run a Markov chain for τ steps to take m different samples $s_1 \ldots s_m$ and calculate the average $1/m \sum_{i=1}^{m} f(s_i)$. There are examples of S and f where despite the existence of a rapidly mixing Markov chain to sample from S, the average of f cannot be approximated by the aforementioned method in polynomial time. For instance, take S to be the set of all Eulerian orientations of an Eulerian graph, $f : S \to \mathbb{R}$ the number of Eulerian circuits corresponding to each orientation, and π the uniform distribution on S. Mihail and Winkler showed that there is a rapidly mixing Markov chain for sampling Eulerian orientations of an

4

Eulerian graph uniformly [71]. However, by running Mihail-Winkler chain we can not obtain a good approximation of the average number of Eulerian circuits of an Eulerian graph. Despite the existence of the examples such as aforesaid, we succeed to show that by employing the above method we can achieve an ϵ approximation of $\hat{f}$ by taking $M \geq 8v/\tilde{f}^2\epsilon^3$ samples and running the chain for $\tau = 2t_{mix}\log(\frac{f_{\max}}{\tilde{f}}\epsilon^{-1/2})$ number of steps where t_{mix} is the mixing time of the chain, v is the variance of f and f_{max} the maximum value of f in S.

In the other approach we look for a phenomenon called homomesy. Homomesy was studied and named by Jim Propp and Tom Roby [7]. The word is derived from the Greek for "same average."

Formally, let S be a set of combinatorial objects and $\tau : S \to S$ a one-to-one operation on elements of S. Let $f : S \to \mathbb{R}$ be a statistic mapping S to the reals (or any field of characteristic 0).

We say the triple $\langle S, \tau, f \rangle$ exhibits homomesy if there is a constant c such that for any orbit $\mathcal{O}$ of τ,

$$\frac{1}{|\mathcal{O}|} \sum_{s \in \mathcal{O}} f(s) = c.$$

Thus, in the presence of homomesy, the average value of the statistic (considering uniform distribution) can be determined from any orbit, without needing to compute f at every member of S.

Let $J([a] \times [b])$ be the set of all ideals in the lattice $[a] \times [b]$, where $[n]$ is the

5

ordered set $\{1 < 2 < \cdots < n\}$. Propp and Roby studied homomesy when $S = J([a] \times [b])$ and τ is the action of rowmotion or promotion, both bijective operations of special combinatorial interest. They showed in particular that homomesy holds when the statistic is the size of an ideal.

Homomesy turns out to be both ubiquitous and, often, surprisingly non-trivial to prove. In this thesis, we study homomesy when $S = J([a] \times [b])$, as well as the set of ideals in some other simply described posets. By extending the definition of rowmotion and promotion, we define a class of permutations on $J([a] \times [b])$, show that they produce the same orbit structure, and find a family of homomesic statistics for them [25].

In another paper that grew out of a workshop held by the American Institute of Mathematics (AIM), we study the expected value of a statistic that we call *jaggedness* of ordered ideals in a poset P under certain distributions that we call *toggle symmetric*. We derived a formula for this expected value when the poset P is a skew Young diagram [21].

Chapter 1

Markov Chain Monte Carlo

"Ah, fill the cup, what boots it to repeat,

How time is slipping underneath our feet,

Unborn tomorrow, and dead yesterday,

Why fret about them if today be sweet!"

–Omar Khayyám.

1.1 Introduction and Preliminaries

Consider Ω a set of combinatorial objects. A **Markov chain** or a **random walk** on Ω is a sequence of random variables $X_1, X_2, \ldots, X_n$ satisfying the condition:

$\forall s_1, \ldots, s_i \in \Omega,$ and $i \in \mathbb{N},$

$Pr(X_i{=}s_i|X_{i-1}{=}s_{i-1}, X_{i-2}{=}s_{i-2}, \ldots, X_1{=}s_1) = Pr(X_i{=}s_i|X_{i-1}{=}s_{i-1}),$

The set Ω is called the **state space** of the chain. Note that $Pr(X_i{=}s|X_{i-1}{=}s')$ only depends on s and s'. This property is known as **memorylessness**. We call the probability of going from s to s' the **transition probability** from state s to state s'.

The following terminology is used in the literature:

Definition 1. The matrix P defined by $P(s, s') = Pr(X_i = s|X_{i-1} = s')$ is called the **transition matrix**, and it is an $|\Omega| \times |\Omega|$ matrix.

Throughout, we denote the probability of an event $\mathcal{E} \subseteq \Omega$ given the probability distribution γ on Ω by $Pr_\gamma(\mathcal{E})$, and we denote the transition probability of a Markov chain $\mathcal{M}$ going from state s to state s' by $P_\mathcal{M}(s, s')$. We drop the subscripts when they are understood.

Definition 2. The underlying graph of a Markov chain $\mathcal{M}$ is a graph $G = \langle S, E \rangle$, where E is the set containing all the edges e satisfying: $e = (s, s') \in E$ if and only if $P(s, s') > 0$. We denote this graph by $G(\mathcal{M})$.

Markov chains were introduced by Andrey Markov, a Russian mathematician (14 June 1856 – 20 July 1922). Today, Markov chains have applications in different fields of science such as computer science, genetics, economics, finance, physics, etc.

In computer science, **ergodic** Markov chains has received special attention since they can be employed for designing approximation algorithms to solve counting or samplings problems. We define an ergodic Markov chain as follows:

Definition 3. A Markov chain $\mathcal{M}$ is called irreducible if in $G(\mathcal{M})$ any two vertices u and v are connected via at least one path. The chain is called aperiodic if in $G(\mathcal{M})$ for any vertex v the least common divisor of lengths of circuits connecting v to itself is 1. An irreducible and aperiodic Markov chain is called ergodic.

Theorem 1. *Let $\mathcal{M} = (X_i)_{i=1}$ be an ergodic Markov chain on Ω, we have $\lim_{i \to \infty} X_i = \pi$, where π is a distribution on Ω that does not depend on X_1.*

Definition 4. We call a Markov chain $\mathcal{M}$ on state space Ω with transition probability P and stationary distribution π **reversible** if for any $x, y \in \Omega$ we have the following equation:

$$\pi(x)P(x,y) = \pi(y)P(y,x). \tag{1.1}$$

Equation 1.1 is known as the **detailed balanced** equation. If a distribution γ on Ω satisfies the detailed balanced equation, then $\pi = \gamma$.

Consider the following problem:

Problem 1. *Consider a set Ω whose size is exponential in some parameter n and a probability distribution π on it. Assume that sampling from S with respect to distribution π is not in the complexity class P. Can we design an approximation algorithm for this problem by which we can take samples from S with respect to another distribution π' where π is very close to π?*

Example 1. Let Ω be the set of permutations of the numbers $1, 2, \ldots, n$.

Then size of S will be $n!$. Do we have an efficient algorithm for taking samples from S according to some arbitrary distribution π?

Example 2. Let Ω be the set of all perfect matchings in an arbitrary graph. Is there an efficient algorithm to take samples from this set according to an arbitrary distribution π?

Example 3. Let Ω be the set of all indecent sets in an arbitrary graph. Is there an efficient algorithm to take samples from this set according to an arbitrary distribution π?

To formalize the notion of "closeness" we use the following metric on distributions:

Definition 5. The total variation distance of two probability distributions γ and ζ on Ω is denoted by $\| \gamma - \zeta \|_{TV}$ and defined as:

$$\| \gamma - \zeta \|_{TV} = 1/2 \sum_{s \in \Omega} |Pr_\gamma(s) - Pr_\zeta(s)|.$$

It is known that for any event $\mathcal{E} \in 2^\Omega$, $|Pr_\gamma(\mathcal{E}) - Pr_\zeta(\mathcal{E})| \leq \| \gamma - \zeta \|_{TV}$.

Definition 6. The **mixing time** of a Markov chain $\mathcal{M}$ having stationary distribution π and convergence factor ϵ is the minimum natural number t satisfying the following condition:

$$\| X_t - \pi \|_{TV} \leq \epsilon.$$

We denote the mixing time of $\mathcal{M}$ by $t_\epsilon(\mathcal{M})$. We often take $\epsilon = 1/4$ and drop the subscript. It is known that $t_\epsilon(\mathcal{M}) = t(\mathcal{M})\log(1/\epsilon)$ for arbitrary $\epsilon > 0$.

We say a Markov chain $\mathcal{M}$ is **rapidly mixing** or we say it mixes in polynomial time, if $t(\mathcal{M})$ is a polynomial function of the problem's parameter. The problem of bounding the mixing time of a Markov chain is called a **mixing problem**.

Markov Chain Monte Carlo. Markov chain Monte Carlo is a technique used for designing approximation algorithms when the exact solution is computationally hard. Consider Problem 1, having a rapidly mixing Markov chain $\mathcal{M}$, converging to the stationary distribution π, we can run $\mathcal{M}$ for $t(\mathcal{M})$ number of steps, and take samples that are close to π.

The idea of Markov chain Monte Carlo has been applied in approximation algorithms for many instances of Problem 1. In 1989 Jerrum and Sinclair [70] proved that for self reducible[1] structures, sampling uniformly from a set is computationally equivalent to counting the set. As a result, the following problem can also be approximated by employing Markov chain Monte Carlo:

Problem 2. *Consider a set Ω. Assume we know $|\Omega|$ is an exponential function of some parameter n, but calculating the exact value of $|\Omega|$ is computationally hard. Can we design an approximation algorithm for this problem?*

---

[1] All of the structures that are discussed in this thesis are self reducible. We will not present the definition of self reducibly here. The interest reader can refer to [70].

Example 4. A partially ordered set consists of a set $\mathcal{P}$ and a reflexive, transitive, antisymmetric relation $\leq_{\mathcal{P}}$. Let $|\mathcal{P}| = n$. A linear extension of $\mathcal{P}$ is a linear ordering $\leq_{\mathcal{L}}$ of elements of $\mathcal{P}$ satisfying: $x \leq_{\mathcal{P}} y \implies x \leq_{\mathcal{L}} y$. Is there any algorithm for finding the number of linear extensions of an arbitrary poset $\mathcal{P}$?

Example 5. Consider a bipartite graph of $2n$ vertices where each part of graph contains n vertices. The number of perfect matching in this graph is equal to the permanent of its adjacency matrix. It is known that calculating permanent of a 0 and 1 $n \times n$ matrix is a #P-complete problem. Is there any way to sample a perfect matching of a graph, or equivalently, approximate the permanent of a 0 and 1 matrix?

We conclude this section by presenting two widely studied classes of Markov chains: the Gibbs sampler and shuffling.

Gibbs Sampler

A Gibbs sampler or Glauber dynamics is a Markov chain for which the sample space Ω is contained in a set of the form S^V, where V is the vertex set of a graph and S is a set of values for labeling vertices of V. We call each labeling of vertices a configuration. Assume, we want to sample from Ω with respect to distribution π. The Gibbs sampler moves as follows when being at configuration $c \in \Omega$:

- Pick $v \in V$ uniformly at random (if we pick one $v \in V$ at a time step,

we say the Gibbs sampler is performing **single site updates**. We will see a discussion on single site updates in Section 2.1).

- Update the label of v according to the following distribution:

  $Pr(v \text{ is labeled by } \omega) = Pr_\pi(v \text{ is labeled by } \omega | \text{All of the other vertices } u \neq v \text{ have the same label as in } c); \omega \in S.$

Problem 3. (The hard-core model.)

In the hard-core model, we have a graph with vertex set V, and to each vertex we assign a number from the set $S = \{0, 1\}$, such that no adjacent vertices of v can both have label 1 to constitute Ω. The hard-core model has two variations, the first one being a special case of the second one:

- **Uniform case.** Consider the uniform distribution on Ω. The Gibbs sampler picks $v \in V$ uniformly at random at each step. If v has label 1 with probability 1/2 it updates the label from 1 to 0, and keeps the other labels unchanged. If v has label 0, with probability 1/2 it updates the label from 0 to 1 under the condition that non of v's neighbors are labeled 1, and keeps the other labels intact.

  Remark 1. If the uniform hard core is rapidly mixing we will have a randomized fully polynomial approximation scheme to find the number of independent sets in an arbitrary graph.

- **The hard-core model with fugacity λ.** The hard-core model with fugacity λ is the following Gibbs sampler: Pick $v \in V$ uniformly at

13

random at each step. If v has label 0 and all of its adjacent vertices also have label 0 then change its label to 1 with probability $\lambda/(1+\lambda)$. If v has label 1 change it to zero with probability $1/\lambda$. This chain has stationary distribution π, where $\pi(I) = \lambda^{|I|}/Z$ where Z is a normalizing factor often called the partition function, and I is any independent set.

Hard core model is a simple Markov chain, and it has vast applications in statistical physics to model the behavior of gas molecules.. However, if the maximum degree Δ in a graph is greater than or equal to 6 there are examples in which the uniform hard core model does not become close to stationarity in fewer than exponential number of steps (n being number of vertices.). (See [89]) The only mixing result for general graphs was proved by Luby and Vigoda [90], where they demonstrated that the hard-core model is rapidly mixing when $\lambda \leq 2/(\Delta - 2)$. This proof was initially presented for triangle free graphs, and then extended to general graphs [90, 91]. To summarize, for a general graph we know that the uniform hard core model is rapidly mixing when $n < 6$, and it is not rapidly mixing when $n \geq 6$. We say that at $n = 6$ **phase transition** is occurring. The following question arises:

Question 1. We know that if the maximum degree of a general graph is greater than or equal to 6, the uniform hard core model takes exponential time to converge. Does this mean the hard-core model is not a good model to approximately sample (or equivalently approximately count) the independent sets of a graph; or the problem itself is hard to approximate?

14

In 2010 Sly proved the following theorem which addresses this question ([86]).

Theorem 2. *Unless NP=RP for every $\Delta \geq 6$ there does not exist a fully polynomial approximation scheme for counting independent sets on graphs of maximum degree at most Δ.*

Sly also proves that unless NP=RP there is no polynomial time approximation scheme for the partition function on graphs of maximum degree d for fugacity $\lambda_c(d) \leq \lambda \leq \lambda_c(d) + \epsilon(d)$, where $\lambda_c(d)$ is the point where the hard-core model experiences a phase transition if the graph is a $d-$regular tree [87]. The relationship between the phase transition and computational complexity of sampling (i.e. counting) problems has received considerable attention recently and has been demonstrated for other sampling problems [88]. These results indicate the strength of the Markov chain Monte Carlo technique.

Problem 4. (Graph coloring.) Consider a graph with G with vertex set V, and the set of colors $Q = \{1, 2, \ldots, q\}$. A coloring of G is an assignment of a color from the set Q to each vertex $v \in V$, such that no adjacent vertices have the same color. Consider the uniform distribution on this set. The Gibbs sampler picks $v \in V$ uniformly at random at each step. Let $\tilde{Q}_v \subseteq Q$ be the set of colors of neighbors of v. Pick color c from $Q \setminus \tilde{Q}_v$ uniformly at random, and change color of v to c.

We will get back to this problem and the papers tacking it in Section

1.2.1.

Shuffling

Let S_n be the set containing all the permutations of numbers $1, 2, \ldots, n$. A Markov chain converging to the uniform distribution on S_n is a shuffling chain. When $n = 52$ we can think of each number as a playing card. The riffle shuffle which is the most common way for shuffling a deck of cards was studied by Bayer and Diaconis in 1992 [81].

Example 6. (The Riffle Shuffle) Consider a deck of size n. Take M from the distribution $Binomial(n, 1/2)$, and split the deck into piles of size M and $n - M$. Cards are dropped from one of the piles one at a time . When the first pile has size a and the other one has size b, a card will drop from the first pile with probability $a/a + b$ and from the other one with probability $b/a + b$.

Bayer and Diaconis showed that after $\log n$ number of steps, the total variation distance drops drastically. They call this phenomenon the **cutoff** phenomenon: for $n = 52$, the cutoff happens after 7 steps.

Example 7. (The Top to Random Shuffle) In the top to random shuffle, the top card of the deck is taken and put back uniformly in any position inside the deck. It can be shown by a stopping time argument that a deck of n cards needs $n \log n$ number of steps in expectation to become uniform. The top to random shuffle and a generalization of it are explained in detail in

Section 1.6.

Example 8. (The Adjacent Transposition Shuffle) In the adjacent transposition shuffle, we take two adjacent cards from the deck, and we swap their position with some probability. Depending on whether or not the probabilities are equal we can have a shuffling of cards or we can converge to distribution very different from the uniform distribution. A complete discussion about various versions of this chain is presented in Section 1.3.

We will discuss another method of shuffling in Section 1.5.

1.2 Techniques

1.2.1 Coupling

Consider Markov chain $\mathcal{M}$ and run two copies of it $X : X_0, X_1, \ldots$ and $Y :$ $Y_0, Y_1, \ldots$. We say (X, Y) is a coupling if both of the marginal distributions are faithful to $\mathcal{M}$. Meanwhile, the coupling should be designed so that X and Y are not running independently, and they will collide with high probability at some t_{coup}. The time of collision is called the **coupling time**.

Theorem 3. *Let (X, Y) be a coupling of two copies of Markov chain $\mathcal{M}$. We define the coupling time to be t_{coup} such that $Pr(X_{t_{coup}} \neq Y_{t_{coup}}) \leq \epsilon$. We have, $t_\epsilon(\mathcal{M}) \leq t_{coup}$.*

Coupling two arbitrary copies of a Markov chain is usually a difficult task. Thus, **path coupling** is often employed. In path coupling, instead of coupling two copies of the Markov chain, we couple any arbitrary pair of states in the sample space. The following theorem was proved by Bubley and Dyer, and it appeared in [64].

Theorem 4. *Let δ be an integer values metric with maximum value D on the sample space Ω of a Markov chain $\mathcal{M}$ satisfying: for any $x, y \in \Omega$: there is a path $x = z_0, z_1, z_2, \ldots z_{r-1}, y = z_r$ between x, y, and $\delta(x, y) = \sum_{i=0}^{r-1} \delta(z_i, z_{i+1})$. (We can take δ to be the shortest path between two states in $G(\mathcal{M})$.) Assume that we have a coupling such that there is a $\beta < 1$ and for any pair of states $(x, y) \in \Omega \times \Omega$, after one step of coupling $\mathbb{E}(\delta(x, y)) \leq \beta \delta(x, y)$. Then,*

$$t_\epsilon(\mathcal{M}) \leq \frac{D\epsilon^{-1}}{1-\beta}.{}^2$$

Coupling was introduced and applied to some mixing problems by Aldous [92]. In this thesis we use it in Section 1.5 to show that the cube shuffling is rapidly mixing. Some classical examples where coupling was the key technique to bound the mixing time are the following:

Recall the Gibbs sampler the for graph coloring problem (Problem 4.) Jerrum [31] applied coupling to this chain and demonstrated that it is rapidly mixing when $q > 2\Delta$, where q is the number of colors and Δ is the maximum degree in the graph. Later, Vigoda [32] found a path coupling solution and showed the chain mixed rapidly when $q > 11/6\Delta$.

Another way of employing coupling is when Ω, the state space is ordered by a partial order, and we have a coupling respecting this order. In that case, we couple the maximal elements of Ω to the minimals. Having them coupled, all the other pairs will also be sandwiched and coupled.

This way of coupling has appeared in a few places such as: Wilson's paper bounding mixing time of the uniform adjacent transposition chain [52] (also explained in Section 1.3), and Randall and Winkler paper [36] to bound the mixing time of points moving in the unit interval whose coordinates respect an order (more details in Section 1.4.)

---

[2]There is a more complicated version of path coupling for which $\beta \leq 1$ which we do not state here.

1.2.2 Conductance

Conductance is a quantity to measure the geometric structure of a Markov chain; and it produces lower bounds and upper bounds of the mixing time of a chain:

Theorem 5. *(Conductance)[79]*

For any Markov chain, define conductance Φ by:

$$\Phi = \min_{\substack{S \in \Omega \\ \pi(S) \leq 1/2}} \left(\sum_{s \in S s' \in S^c} \pi(s) Pr(s, s') \right) / \pi(S).$$

Then, for arbitrary $\epsilon > 0$

$$\left(\frac{1}{\Phi} - 1/2 \right) \log \left(\frac{1}{2\epsilon} \right) \leq \tau_\epsilon(\mathcal{M}) \leq \frac{1}{2\Phi} \log \left(\frac{1}{\pi_{\min}\epsilon} \right). \qquad (1.2)$$

One famous application of conductance is its application by Dyer and Frieze to the mixing problem on a Markov chain for sampling from convex bodies. [33].

Conductance is often employed to find lower bounds for mixing time of Markov chains. We use it for this purpose in Section 1.3. Lemma 10.

In addition we compare conductance of a chain to another chain for which the mixing time in known to derive upper bounds for the mixing time of the initial one. We use this technique in Section 1.3. Theorem 14 and Section 1.4. Theorem 19.

1.2.3 Canonical Paths

The following theorem appeared in [64], and further explained in [60].

Theorem 6. *Let $\mathcal{M}$ be a Markov chain with stationary distribution π and E the set of the edges in its underlying graph. For any two states σ and τ in the state space Ω we define a path $\gamma_{\sigma,\tau}$. The congestion factor for any edge $e \in E$ is denoted by Φ_e and is defined by*

$$\Phi_e = \frac{1}{Q(e)} \sum_{\substack{x,y \\ e \in \gamma_{x,y}}} \pi(x)\pi(y). \tag{1.3}$$

We can bound the mixing time of $\mathcal{M}$ using the congestion factor:

$$t_\epsilon(\mathcal{M}) \le 8\Phi^2(\ln \pi_{min}^{-1} + \ln \epsilon), \tag{1.4}$$

Where $\Phi = \max_{e \in E} \phi_e$, $\pi_{min} = \min_{x \in \Omega} \pi(x)$ and ϵ is the convergence factor.

The canonical path theorem has been used to solve some of the most difficult mixing problems such as: sampling perfect matchings which is computationally equivalent to approximating the permanent of a matrix ([56]). We use this technique to prove Theorem 15 of Section 1.3.

1.2.4 The Comparison Method

The comparison method was introduced by Diaconis and Saloff-Coste [53] and then Randall and Tetali extended it and employed it for analysis of Gibbs samples [54].

The comparison method is appealing when analyzing Markov chains that make local moves at each step (for instance single site updates in Gibbs samplers). In those cases, a variation of that chain is studied which has the same sample space and stationary distribution, while making a set of moves simultaneously. Then, the mixing times of the two chains are compared:

Theorem 7. *(Comparison method [54])*

Let $\mathcal{M}$ and $\mathcal{M}'$ be two reversible Markov chains on state space Ω and with stationary distribution π. Let P be the transition matrix of $\mathcal{M}$ and P' the transition matrix of P'. Consider the two underlining graphs of $\mathcal{M}$ and $\mathcal{M}'$, and let them be $G(\mathcal{M})$ and $G(\mathcal{M}')$ respectively. i.e. $G(\mathcal{M}) = \langle \Omega, E(\mathcal{M}) \rangle$ where $E(\mathcal{M}) = \{(\sigma, \tau) | \sigma, \tau \text{ in } \Omega \text{ and } P(\sigma, \tau) > 0\}$, for each $e = (\alpha, \beta) \in E(\mathcal{M})$, we define the capacity of e by $\mathcal{C}(e) = \pi(\alpha) P(\alpha, \beta)$. Similarly, we define $E(\mathcal{M}')$, $G(\mathcal{M}')$ and capacity for $e \in E(\mathcal{M}')$.

For any edge (σ, τ) in $G(\mathcal{M}')$, we find a path in $G(\mathcal{M})$ from σ to τ. Let this path be $\gamma_{\sigma, \tau}$. For an arbitrary edge $e \in E(\mathcal{M})$, let $\Gamma(e) = \{\gamma_{\sigma, \tau} | \gamma_{\sigma, \tau}$ traverses through $e\}$ and $\pi_{min} = \min_{\zeta \in \Omega} \{\pi(\zeta)\}$. We define,

$$\mathcal{A}_e = \frac{\sum_{\Gamma(e)} |\gamma_{\sigma, \tau}| \mathcal{C}((\sigma, \tau))}{\mathcal{C}(e)}. \tag{1.5}$$

We have,

$$t_\epsilon(\mathcal{M}) \leq \frac{4 \log(1/\epsilon \pi_{min})}{\log(\epsilon/2)} t_\epsilon(\mathcal{M}') \max_{e \in E(\mathcal{M})} \mathcal{A}_e. \tag{1.6}$$

1.2.5 The Decomposition Technique

The decomposition technique is an effective method, when the sample space can be divided to smaller sets such that the restriction of the Markov chain $\mathcal{M}$ to these smaller sets becomes easier to analyze. There are two variations of the decomposition technique appearing in [82] and [77]. Here we state the one which partitions the sample space to non intersecting subsets:

Theorem 8. *(Decomposition Theorem) [77]*

Let $\mathcal{M}$ be a Markov chain on state space Ω partitioned into $\Omega_1, \Omega_2, \ldots, \Omega_k$. For each i, let $\mathcal{M}_i$ be the restriction of $\mathcal{M}$ to Ω_i that rejects moves that go outside of Ω. Let $\pi_i(A) = \pi(A \cap \Omega_i)/\pi(\Omega_i)$ for $A \subseteq \Omega$. We define the Markov chain $\bar{\mathcal{M}}$ on state space $\{1, \ldots k\}$ as follows:

$Pr_{\bar{\mathcal{M}}}(i,j) = \sum_{x \in \Omega_i, y \in \Omega_j} \pi_i(x) Pr_{\mathcal{M}}(x,y)/\pi(\Omega_i)$, where $Pr_{\mathcal{M}}$ and $Pr_{\bar{\mathcal{M}}}$ are transition probabilities of $\mathcal{M}$ and $\bar{\mathcal{M}}$ respectively.

Then,

$$t(\mathcal{M}) \leq 2t(\bar{\mathcal{M}}) \max_i \{t(\mathcal{M}_i)\}. \tag{1.7}$$

We use the decomposition technique in the proof of Theorem 14. Section 1.3, and the proof of Theorem 19. Section 1.4.

1.2.6 Summary

We conclude this section by presenting the following table which summarizes our use of the aforementioned techniques together with some of the important problems in the literature employing these techniques.

Håstad's square shuffle	[38]	Coupling
The cube shuffle	Section 1.5	Coupling
The interval chain of Randall and Winkler	[36] (Section 1.4.)	Coupling
The Gibbs sampler for $k-$colorings of a graph	[31], [32]	Coupling
Three constant ratio gladiators	Section 1.3	Decomposition Comparison path congestion
Lower bound for gladiators		Conductance
Linear extensions of a poset	[75]	Coupling
Linear extensions of a poset (Karzanov-Khachiyan Chain)	[74]	Conductance
Approximating the permanent	[56]	Path congestion

1.3 The gladiator Chain

Remember that for any arbitrary natural number $n \in \mathbb{N}$, we define S_n to be the set that contains all the permutations of numbers $1, 2, \ldots n$. A natural Markov chain on S_n is the chain which picks a number $1 \leq i \leq n-1$ uniformly at random and operating on $\sigma \in S_n$, puts $\sigma(i+1)$ ahead of $\sigma(i)$ w.p. $p_{\sigma(i),\sigma(i+1)}$. We call this chain the *adjacent transposition* Markov chain. This Markov chain has been studied widely for various choices of $p_{i,j}$ [52, 51, 55, 23].

In this paper, we consider the total variation mixing time, which is defined as the time it takes until the total variation distance between the distribution of the current state and stationarity is less than ϵ (where ϵ is some fixed convergence factor). For Markov chain $\mathcal{M}$ we denote this time by $t_\epsilon(\mathcal{M})$, or if $\epsilon = 1/4$, simply by $t(\mathcal{M})$.

A special case of the adjacent transposition chain which we call the *gladiator chain* has been introduced by Jim Fill ([51]). Fill was interested in probabilistic analysis of algorithms for *self-organizing lists* (SOLs). Self-organizing lists are data structures that facilitate linear searching in a list of records; the objective of a self-organizing list is to sort the records in non-decreasing order of their access frequencies [68]. Since these frequencies are not known in advance, an SOL algorithm aims to move a particular record ahead in the list when access on that record is requested. There are two widely used SOL algorithms: the *move ahead one* algorithm (MA1) and the

move to front algorithm (MTF). In MA1, if the current state of the list is $(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$ and the ith record is requested for access, it will go ahead in the list only one position and the list will be modified to $(x_1, x_2, \ldots, x_i, x_{i-1}, x_{i+1}, \ldots, x_n)$. In MTF it will go to the front and the list will be modified to $(x_i, x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. It appears that MA1 should perform better than MTF when the list is almost sorted and worse when the low frequency records are standing in front; however, this has not been analytically studied [69]. Considering the adjacent transposition Markov chain corresponding to MA1, Fill shows ([51]) that there are cases in which the chain is not rapidly mixing. Hence, he poses the question of sampling from the stationary distribution of MA1, and he introduces the gladiator chain which has the same stationarity as MA1 and seems to be rapidly mixing for arbitrary choice of parameters. He makes the following conjecture for mixing time of the adjacent transposition chain in general:

Fill's conjecture ([51]).[3] If the adjacent transposition Markov chain is monotone, then it is rapidly mixing. Monotonicity in this context means: for all i, j satisfying $1 \leq i < j \leq n$: $p_{i,j} \geq 1/2$, and $p_{i,j-1} \leq p_{i,j} \leq p_{i,j+1}$.

Here we provide a brief history of the results on the adjacent transposition Markov chain. All of these chains are monotone and rapidly mixing. Wilson and Benjamini's papers [52, 55] led to Fill's conjecture [51]; Bhakta et al. [23] verified the conjecture in two cases.

---

[3]Fill considered the spectral gap (another measure of mixing) in his study. Here, we are interested in total variation mixing time which, in this case, is within polynomial factor of the spectral gap.

1. The simple chain. In the case where $p_{i,j} = 1/2$ for all i and j, the chain will have a simple description: Given a permutation σ, pick two adjacent elements uniformly at random, and flip a fair coin to decide whether to swap them. We call this chain, whose stationary distribution is uniform, the *simple* chain. Ironically, proving precise mixing results for this chain was not simple. Many papers targeted this problem [50, 66] and finally Wilson [52] showed the mixing time for this chain is $\Theta(n^3 \log n)$ (he proved lower and upper bounds within constant factors).

2. The constant-bias chain. After Wilson's paper, Benjamini et al. [55] studied the case where $p_{i,j} = p > 1/2$ for all i and j, and $p_{j,i} = 1-p$. Benjamini et al. [55], reduced this problem to the problem of mixing time of an asymmetric simple exclusion process (ASEP) and showed that this ASEP, and consequently the constant-bias chain, mixes in $\Theta(n^2)$ steps. We will talk more about the exclusion process chain later on in this introduction.

3. "Choose your weapon" and "league hierarchy" chains. The following two special cases were studied by Bhakta et al. [23]: the *choose your weapon chain* where $p_{i,j}$ is only dependent on i, and the *league hierarchy chain* given by a binary tree T with n leaves. Each interior node v of T is labeled with some probability $1/2 \leq q_v \leq 1$, and the leaves are labeled by numbers $1 \ldots n$. The probability of putting j ahead of i for $j > i$ is equal to $p_{i,j} = q_{j \wedge i}$ where $j \wedge i$ is the node that is the lowest common ancestor of i and j in T.

As we mentioned before, one interesting instance of the adjacent trans-

position chain which is monotone is the gladiator chain. Here we study a special case of the gladiator chain where gladiators fall into a few classes according to their strengths. (Definition 7).

The gladiator chain. In this chain each element i can be thought of as a gladiator with strength $s(i)$. Every permutation of numbers $1, 2, \ldots n$ can be thought of as a ranking of gladiators. In each step of Markov chains we choose $1 \leq k < n$ uniformly at random, i.e., we choose adjacent gladiators $\sigma(k) = i$ and $\sigma(k+1) = j$. These gladiators will fight over their position in ranking. With probability $p_{j,i} = s(i)/(s(i) + s(j))$, gladiator i will be the winner of the game and will be put i ahead of j in σ if it isn't already. With probability $1-p$, j is put ahead of i.

Corollary 9. *The gladiator chain is rapidly mixing if Fill's conjecture holds.*

Particles and the exclusion process. Let $G = \langle V, E \rangle$ be a graph and consider $m < |V|$ particles on the vertices of G. At each step of the Markov chain we pick a vertex v uniformly at random with probability $1/|V|$ and one of its adjacent vertices, w with probability $1/d(v)$. If there is a particle in one of them and not the other one, we swap the position of the particle with probability p.

If p is constant for any choice of v and w, the chain is called the *exclusion process*. The exclusion process is a well known Markov chain, and because of its applications in statistical physics it has been studied widely ([63, 65]). In fact, a special case of this chain where G is a finite line was studied by Benjamini et al. [55] as mentioned before.

Here, we consider the case where $G = \langle V, E \rangle$ is a finite line and we have $|V|$ particles of different types on the vertices of G and they swap their positions with probabilities dependent on their types (Definition 8). We call this Markov chain a *linear particle system.* To any adjacent transposition Markov chain, we can associate a linear particle system. We will explain this association in Section 1.3.2 and prove that the mixing time for an adjacent transposition chain is only polynomially larger than the mixing time of the corresponding linear particle system.

The simplest interesting case of the linear particle system whose mixing time is not well understood is the one with 3 particle types, where exchange probabilities depend only on the types. We will give a formal definition of the particle system in Section 1.3.1 and prove it is rapidly mixing under certain conditions.

Definitions and results are presented in Section 1.3.1. Section 1.3.2 contains the correspondence between the gladiator chain and the linear particle system. Section 1.3.3 contains the main proofs.

1.3.1 Definitions and Results

Definition 7. Gladiator chain. (Playing in teams) Consider the Markov chain on state space S_n that has the following properties: The set $[n]$ (i.e. gladiators) can be partitioned into subsets: $T_1, T_2, \ldots, T_k$ (k teams). We

have the following strength function: $s : [n] \to \mathbb{R}$, $s(g) = s_j$ iff $g \in T_j$. At each step of Markov chain, we choose $i \in [n-1]$ uniformly at random. Given that we are at state σ, and $\sigma(i) = g, \sigma(i+1) = g'$, we put g ahead of g' with probability $\frac{s(g)}{s(g)+s(g')}$.

This is a reversible Markov chain and the stationary distribution π is

$$\pi(\sigma) = \prod_{i=1}^{n} s(i)^{\sigma(i)}/Z. \tag{1.8}$$

(Z is a normalizing factor.)

Definition 8. The linear particle system. Assume we have k types of particles and of each type i, we have n_i indistinguishable copies. Let $n = \sum_{i=1}^{k} n_i$. Let Ω be the state space containing all the different linear arrangements of these n particles. If the current state of the Markov chain is σ, choose $i \in [1, n-1]$ uniformly at random. Let $\sigma(i)$ be of type t and $\sigma(i+1)$ be of type t'. If $t = t'$ do nothing. Otherwise, put $\sigma(i)$ ahead of $\sigma(i+1)$ w.p. $p_{t,t'}$ and put $\sigma(i+1)$ ahead of $\sigma(i)$ w.p. $1 - p_{t,t'}$.

This chain is also a reversible Markov chain.

Proposition 1. *By regarding gladiators of equal strength as indistinguishable particles, any gladiator becomes a linear particle system.*

Lemma 10. *There is an example of the particle system with 3 types of particles and non-monotone swapping probabilities that does not mix fast.*

30

Definition 9. Consider the gladiator chain. We denote the following special case by $G_q(a, b, c)$.

- The set of gladiators, can be partitioned to 3 nonintersecting teams: Arbelas (team A of a gladiators), Bestiarius (team B of b gladiators), Cestus (team C of c gladiators). Cestus are the strongest gladiators and Arbelas are weakest. [4]

- Gladiators of Arbelas have strength s_A. Gladiators of Bestiarius have strength s_B. Gladiators of Cestus have strength s_C. Furthermore, we have $s_A/s_B = s_B/s_C = q$ where $0 \leq q \leq 1$ is a constant.

The steps of $G_q(a, b, c)$ are as defined in Definition 7.

Definition 10. (The constant-ratio three particle system)

Consider the particle system chain. We denote the following special case by $\mathcal{EX}_q(a, b, c)$.

- Particles are in 3 types; type A, type B and type C. We have a indistinguishable copies of particle A, b of particle B, and c of particle C.

- $p_{A,B} = 1 - p_{B,A} = 1/(1 + 1/q)$, $p_{B,C} = 1 - p_{C,B} = 1/(1 + 1/q)$ and $p_{A,C} = 1 - p_{C,A} = 1/(1 + 1/q^2)$. For some constant $0 \leq q \leq 1$.

The steps of $\mathcal{EX}_q(a, b, c)$ are as defined in Definition 8.

---

[4]Arbelas, Bestiarius and Cestus were three of the many types of gladiators in the Roman Empire.

Theorem 11. $\mathcal{EX}_q(a, b, c)$ *is rapidly mixing for any choice of* a, b, c *and* $0 \leq q \leq 1/2$.

We will prove Theorem 11 in Section 1.3.3.

Theorem 12. $G_q(a, b, c)$ *is rapidly mixing for any choice of* a, b, c *and* $0 \leq q \leq 1/2$.

Proof. This is a consequence of Theorem 11 and of Theorem 14, which is stated and proved in the next section. $\square$

Corollary 13. *(Generalization of league hierarchies)*

Let T *be tree with* n *leaves. Each interior nodes* v *is labeled with some probability* $2/3 \leq q_v \leq 1$ *and the leaves are labeled by numbers* $1 \ldots n$. *The probability of putting* j *ahead of* i *for* $j > i$ *is equal to* $p_{i,j} = q_{j \wedge i}$ *where* $j \wedge i$ *is the node that is the lowest common ancestor of* i *and* j *in* T.

It is shown in [23] that in the case where T is a binary tree, this mixing problem can be reduced to the simple exclusion process chain. Employing Benjamini's result ([55]), they prove rapid mixing for the binary tree league hierarchies. Theorem 11 can be used to extend the results in [23] to ternary trees.

1.3.2 Gladiators and Particles

Consider the gladiator chain $\mathcal{M}$ (Definition 7). At each step of the chain, one of two things is happening:

1. Whisking: gladiators of the same team are fighting.

2. Sifting: gladiators of different teams are fighting.

If we were restricted to whisking steps the chain would be equivalent to the simple chain studied by Wilson. If we were restricted to sifting steps the chain would be the linear particle system chain. In order to study the mixing time of the gladiator chain we analyze sifting and wishing steps separately and then we employ the following decomposition theorem (Section 1.2. Theorem. 8.

Theorem 14. *Let τ and τ' be respectively the mixing time for a linear particle system and its corresponding gladiator chain. Then*

$$\tau' \leq \mathcal{O}(n^2)\tau. \tag{1.9}$$

We use conductance (Section 1.2. Theorem 5) to prove Theorem 14.

Proof. Let G be a gladiator chain and $\mathcal{M}$ its corresponding particle system with $\mathcal{K}_i$ copies of particle i and k different particles. Take $\sigma_1 \in S_{\mathcal{K}_1}, \sigma_2 \in S_{\mathcal{K}_2}, \ldots, \sigma_k \in S_{\mathcal{K}_k}$ and let $S_{\sigma_1,\sigma_2,\ldots,\sigma_k} \subseteq S_n$ be the set of all permutations in S_n in which all the gladiators corresponding to particle i preserve the ordering associated to them by σ_i. Restriction of G to $S_{\sigma_1,\sigma_2,\ldots,\sigma_k}$ is equivalent to $\mathcal{M}$. Considering all the choices of $\sigma_1 \in S_{\mathcal{K}_1}, \sigma_2 \in S_{\mathcal{K}_2}, \ldots, \sigma_k \in S_{\mathcal{K}_k}$, S_n will be partitioned into $\Pi_{i=1}^k \mathcal{K}_i!$ copies of $\mathcal{M}$.

Sifting: Let $\bar{G}$ be defined as in Theorem 8; we will show that $\bar{G}$ is rapidly mixing and applying Theorem 8. We show that the conductance of $\bar{G}$ is only polynomially smaller than conductance of the simple chain, which is known to mix in $n^3 \log n$ steps ([52]). For this purpose, it suffices to prove the following claim:

Claim. Let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be arbitrary, and let $S_{i,k} \subseteq S_{\sigma_1, \sigma_2, \ldots, \sigma_k}$ be the set of all arrangements in $S_{\sigma_1, \sigma_2, \ldots, \sigma_k}$ such that the ith and $i+1$th copy of particle k are adjacent and $\bar{S}_{i,k} = S_{\sigma_1, \sigma_2, \ldots, \sigma_k} \setminus S_{i,k}$. We have $\pi_{\sigma_1, \sigma_2, \ldots, \sigma_k}(S_{i,k}) \geq 1/(n^2 + 1)$, $\pi_{\sigma_1, \sigma_2, \ldots, \sigma_k}$ being the stationary distribution of G restricted to $S_{\sigma_1, \sigma_2, \ldots, \sigma_k}$, i.e., the stationary distribution of $\mathcal{M}$.

Proof of claim. To each $\sigma \in \bar{S}_{i,p}$, we correspond $\tau \in S_{i,p}$ such that $\pi(\sigma) \geq \pi(\tau)$. Consider τ_1 and τ_2 as follows: Let τ_1 be the arrangement that we get by taking the ith copy of particle p down until it is adjacent to the $i + 1$st copy and τ_2 the arrangement that we get by taking the $i + 1$st copy of particle p up until it is adjacent to the ith copy. By the detailed balance equation[5], we have $\pi(\tau_1)/\pi(\sigma) = \pi(\sigma)/\pi(\tau_2)$. Hence one of τ_1 or τ_2 will have a larger density of π than σ.

The mapping that we just described sends at most n^2 elements of $\bar{S}_{i,k}$ to $S_{i,k}$. Therefore, $\pi(S_{i,k}) \geq 1/(n^2 + 1)$.

$\square$

$\square$

---

[5]Detailed balance equation: $\pi(x)p_{x,y} = \pi(y)p_{y,x}$

1.3.3 The Constant-Ratio Three Particle System

In this section, we prove Theorem 11 and Lemma 10. We use the path congestion method and comparison method to prove upper bounds, and conductance to prove lower bounds. Before we proceed to our proof we introduce some terminology.

For arbitrary nonnegative natural numbers a, b, c satisfying $a+b+c = n$, let $\Omega_{a,b,c}$ be the set containing all different arrangements of a copies of particle A, b of particle B and c of particle C. Then $|\Omega_{a,b,c}| = \binom{n}{a,b,c} \leq \binom{n}{\frac{n}{3},\frac{n}{3},\frac{n}{3}}$. We denote the strength of particle A by s_A, the strength of particle B by s_B, and the strength of particle C by s_C. Consider an arbitrary arrangement $\sigma \in \Omega_{a,b,c}$. By $\sigma(i) \in A$, we mean a particle of type A is presented at index i of the arrangement σ. The probability of σ in stationarity is

$$\pi(\sigma) = w(\sigma)/\mathcal{Z}_{a,b,c} \ , \tag{1.10}$$

where $w(\sigma)$ is the weight of σ and is equal to $\Pi_{i=1}^{n} s_{\sigma(i)}^{i}$, and for arbitrary nonnegative natural numbers a, b, c $\mathcal{Z}_{a,b,c}$ is called the *partition function* and it is equal to

$$\mathcal{Z}_{a,b,c} = \sum_{\sigma \in \Omega_{a,b,c}} w(\sigma). \tag{1.11}$$

For convenience we denote $\Pi_{i=1}^{n} s_{\sigma(i)}^{j+i}$ by $w^j(\sigma)$ and similarly $\sum_{\sigma \in \Omega_{x,y,z}} w^i(\sigma)$ by $\mathcal{Z}_{a,b,c}^i$.

For any two arrangements $\sigma, \tau \in \Omega_{a,b,c}$, we denote the ratio of their

35

weights by $q_\tau(\sigma) := w(\sigma)/w(\tau) = \pi(\sigma)/\pi(\tau)$. In the case that we are interested in, $s_A/s_B = s_B/s_C = q < 1$; hence $q_\tau(\sigma)$ is always some power of q. Given an arrangement σ, we define $Q(\sigma) := \sum_{\tau \in \Omega_{a,b,c}} q_\sigma(\tau)$. If we restrict this summation to those choices of τ where all copies of particle of type X are fixed, we denote it by $Q_{\bar{X}}(\sigma)$.

We write arrangements of particles either from left to right (meaning $\sigma(1)$ stands leftmost and $\sigma(n)$ stands rightmost) or from down to top (meaning $\sigma(1)$ is the lowest and $\sigma(n)$ the highest). When we compare particles, by a lower particle we mean lower index in arrangement; by a weaker particle we mean lower in strength.

In an arrangement $\sigma \in \Omega_{a,b,c}$, we specify the ith copy of particle of type A by A_i and $\sigma(i) = A_k$ means at position i of arrangement σ we have the kth copy of type A particles. We denote the position of the ith copy of type A particles by $A_i(\sigma)$. Similar definitions for particles of type B and C holds. The position of the highest copy of particle A below position i is denoted by, $A_{\downarrow i}(\sigma)$. i.e. $A_{\downarrow i}(\sigma) = \max\{k | k < i \text{ and } \sigma(k) = A\}$. Likewise, we define $A_{\uparrow i}(\sigma)$. i.e. $A_{\uparrow i}(\sigma) = \min\{k | k > i \text{ and } \sigma(k) = A\}$. Similar definitions for particles of type B and C holds.

The sub-arrangement $\sigma[i, j]$ of an arrangement σ is the restriction of σ to the closed interval $[i, j]$; note that $\sigma[i, j] \notin \Omega_{a,b,c}$. We use $\sigma(i, n)$ to denote $\sigma[i, i + n - 1]$, so $|\sigma(i, n)| = n$. We denote by $\sigma_{-A}[i, j]$ the result of removing from $\sigma[i, j]$ all copies of particle A and adjusting indices accordingly. Hence, $\sigma[i, j] \in \Omega_{a,b,c} \Rightarrow \sigma_{-A}[i, j] \in \Omega_{0,b,c}, \sigma_{-B}[i, j] \in \Omega_{a,0,c}, \sigma_{-C}[i, j] \in \Omega_{a,b,0}$, and

$\sigma_{-A}(i, n)$ is equal to $\sigma_{-A}[i, e]$ for some e such that $|\sigma_{-A}[i, e]| = n$. Similar definitions for particles of type B and C holds.

The concatenation of two arrangements σ and τ is denoted by $\sigma|\tau$.

Example 9. Let $\sigma := AABBCABCAABACCB \in \Omega_{6,5,4}$, $\sigma(11) = B_4$, also $\sigma(11) = B$, $B_4(\sigma) = 11$, $B_{\uparrow 4}(\sigma) = 7$, $B_{\downarrow 4}(\sigma) = 3$, $\sigma[4, 8] = BCABC$, $\sigma(4, 6) = BCABCAAB$. $\sigma_{-A}[4, 8] = BCBC$, $\sigma_{-A}(4, 6) = BCBCBC$.

We want to employ the comparison technique (Section 1.2, Theorem 7), to show rapidly mixing of $\mathcal{EX}(a, b, c)$. Hence, we study an easier variation of $\mathcal{EX}(a, b, c)$ and we denote it by $\mathcal{EX}_t(a, b, c)$.

Definition 11. Let $\mathcal{EX}_t(a, b, c)$ be a Markov chain on state space $\Omega_{a,b,c}$ and $n = a + b + c$. If the current state is σ we choose natural numbers $1 \leq i < j \leq n - 1$ uniformly at random and swap them following these rules:

1. If $\sigma(i) = \sigma(j)$, do nothing.

2. If $\sigma(i) = A$ and in $\sigma(j) = C$ or vise versa and $\sigma(i+1) = \cdots = \sigma(j-1) = 1$. Then, put them in increasing order w.p. $q^{2(j-i)}/(1 + q^{2(j-i)})$. With probability $1/(1 + q^{2(j-i)})$, do nothing.

3. If $\sigma(i) = \sigma(i+1) = \cdots = \sigma(j-1) = B$ and $\sigma(j) = C$ or A or If $\sigma(j) = \sigma(j-1) = \cdots = \sigma(i+1) = B$ and $\sigma(i) = C$ or A. Then, put them in increasing order w.p. $1/(1 + q^{j-i})$. With probability $q^{j-i}/(1 + q^{j-i})$, do nothing.

37

The following picture depicts moves number 2 and 3. Later in this text we will refer to these moves by their numbers:

	C	A			B	A	B	C
	B	B			B	B	B	B
	B	B			B	B	B	B
	. ↔ .				. ↔ .		. ↔ .	
move 2:	.	.		move 3:	.	.	.	.
	.	.			.	.	.	.
	B	B			B	B	B	B
	A	C			A	B	C	B
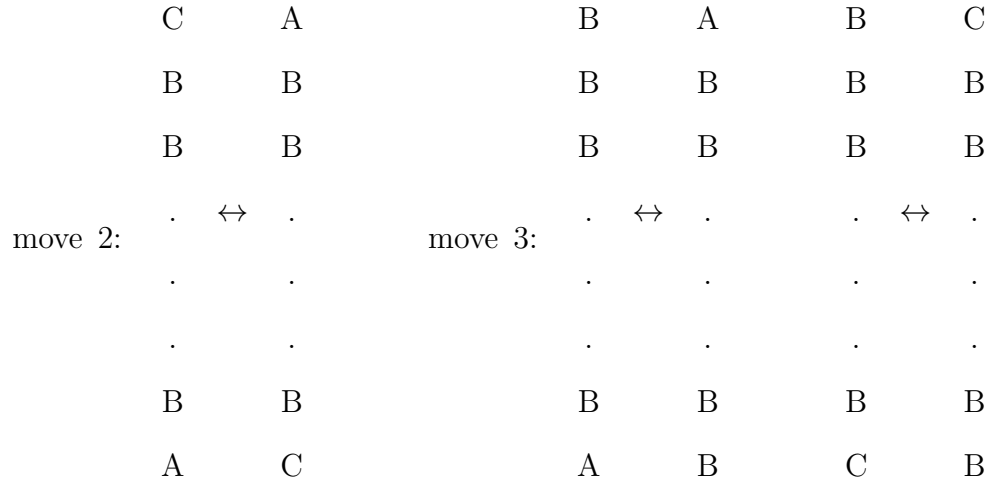
It can be easily checked that $\mathcal{EX}_t$ is reversible and its stationary distribution is the π in Equation 1.10.

Theorem 15. $\mathcal{EX}_t(a, b, c)$ *is rapidly mixing.*

We use the path congestion method (Also known as canonical paths method) in our proof. The method is one of the broad approaches usually used to solve mixing problems ([58, 57]). It was first introduced and employed by Jerrum and Sinclair to show the mixing time of a Markov chain which approximates the permanent of a matrix [56] (See Section 1.2).

Proof. For any two arbitrary states $\sigma, \tau \in \Omega$, we introduce a path $\gamma_{\sigma,\tau}$. Then, we employ the canonical paths method (Section 1.2. Theorem 6) to show that $\mathcal{EX}_t$ is rapidly mixing.

Consider arbitrary $\sigma, \tau \in \Omega_{a,b,c}$. Let $n = a+b+c$, we choose the following path from σ to τ:

Starting from σ, repeat the following steps until τ is reached.

Initially, let $i, j = 1$.

1. Let $k = B_j(\tau)$. We define the j th *block* of σ and τ to be the substring starting from i and ending in k. Note that in τ, each blocks starts right after a B and ends with a B. In the jth iteration, the goal is to change $\sigma[i, k]$ until $\sigma[1, k] = \tau[1, k]$, i.e. the first j blocks equal in σ and τ.

2. Using move 2, and starting from the lowest index i, we bring particles C or A down until the $k - i$ A and C particles have the same order in σ and τ.

3. We use move 3 and bring the jth B in σ to $B_j(\tau)$. In this process, we may need to bring several copies of particle B out of the jth block in σ. In that case, we choose a random ordering of Bs and move them with respect to that order. (Details explained in the proof.)

4. Let $i = B_j(\tau) + 1$.

5. j++.

We claim that using these paths the congestion factor for every edge in $\mathcal{EX}_t$ is polynomial.

1st iteration:

<pre>
        C   C   C   C   C   C
σ :   A   A   A   A   A   A
        B   B   B   B   B   B
        C   C   C   C   C   C
        A move 2   A move 2   A move 2   A move 3   A move 3   B
        C   ⟶   C   ⟶   A   ⟶   A   ⟶   B   ⟶   A
        B Step 1   B Step 1   B Step 1   B Step 2   A Step 2   A
        C   A   C   C   C   C
        A   C   C   C   C   C
</pre>

2nd iteration: (shown above, rightmost columns)

3rd iteration:

<pre>
                    C   A
                    A   τ :   C
                    C   C
                    B   B
        move 3   B   move 2   B
         ⟶   A   ⟶   A
        Step 2   A   Step 1   A
                    C   C
                    C   C
</pre>
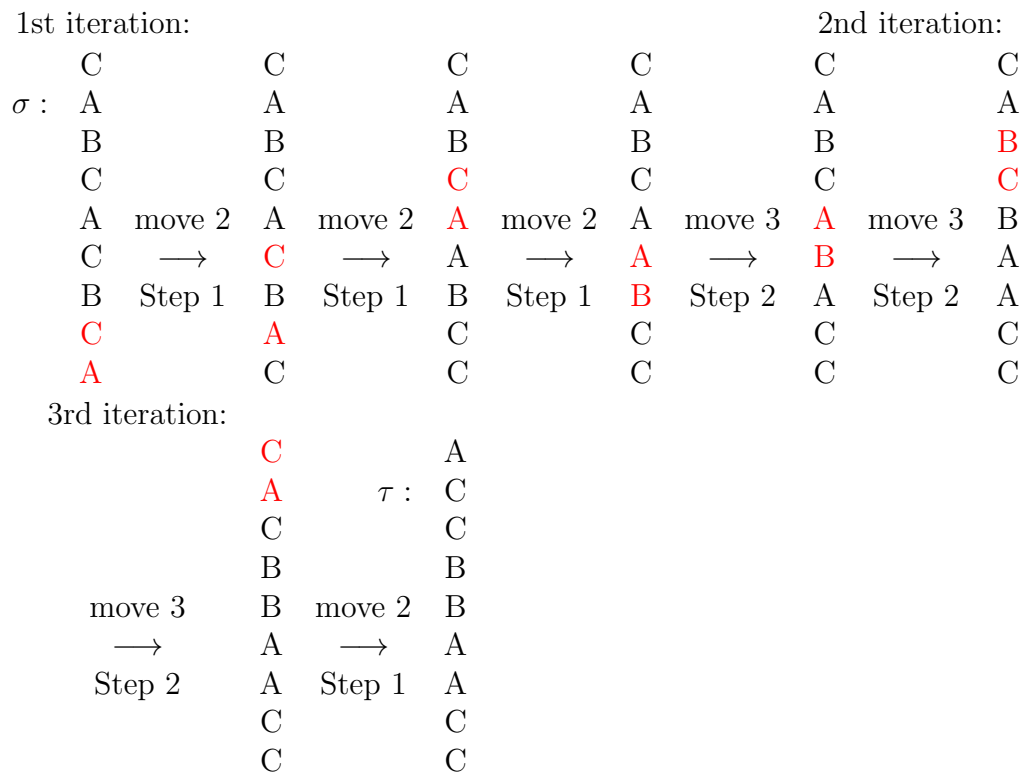
Figure 1.1: Going from σ to τ using moves 2 and 3.

There are two types of edges: those that make move 2 and those that make move 3. We show that neither of these edges are congested by a factor of more than a polynomial function of n.

Consider an edge that makes move 2. Let this edge be connecting two states α and β where α and β only differ in a transposition of type move 2. Assume the swapping C and A are the lth C and mth A and $C_l(\alpha) < A_m(\alpha)$.

Let's say $e = (\alpha, \beta)$ is on the path connecting σ to τ. i.e. $(\alpha, \beta) \in \gamma_{\sigma,\tau}$. It follows from the way we set the paths that, for some $j, A_m(\alpha) \leq j < A_{\uparrow A_m \alpha}(\alpha), \ A_m(\sigma) = j$ and for some $i, A_{\downarrow A_m(\beta)}(\beta) < i \leq A_m(\beta), \ A_m(\tau) = i$. We have, $\alpha[1, i-1] = \tau[1, i-1]$ and $\alpha[j+1, n] = \sigma[j+1, n]$. Let's say the number of Bs in $\sigma[0, i]$ is o i.e. $\sigma[0, i] \in \Omega_{m,o,l}$. There are no As in $\alpha[i, j]$, hence we assume it is in $\Omega_{0,x,y}$.

We will try to find an upper bound for the congestion factor of the edge $e = (\alpha, \beta)$. Let $g = A_m(\alpha) - C_l(\alpha)$.

$$\Phi_e = \left(1 + q^{2(g)}\right) \left(\sum_{\substack{\sigma; \\ \alpha[j+1,n]=\sigma[j+1,n]}} \frac{\pi(\sigma)}{\pi(\alpha)} \sum_{\substack{\tau; \\ \alpha[1,i-1]=\tau[1,i-1]}} \frac{\pi(\tau)}{\pi(\alpha)} \right) \pi(\alpha). \quad (1.12)$$

To make the analysis easier we divide each arrangement to 3 segments; the first segment $[1, i]$, the second $[i, j]$ and the third $[j, n]$. Let $M_t(\alpha)$ be an arrangement that you get from replacing the lowest t particles of type C with particles of type A in $\alpha[i, j]$.

41

Note that,

$$\frac{\pi(\sigma)}{\pi(\alpha)} = \frac{w(\sigma[1,i])w^i(\sigma[i,j])w^j(\sigma[j,n])}{w(\alpha[1,i])w^i(\alpha[i,j])w^j(\alpha[i,j])} = \frac{w(\sigma[1,i])w^i(\sigma[i,j])}{w(\alpha[1,i])w^i(\alpha[i,j])} \tag{1.13}$$

Let $\mathcal{S}_t$ be the set of all σs of the above form for which we have t 0s in $\sigma[i,j]$ and therefore $m+t$ 2s and $l-t$ 0s in $\sigma[1,i]$ and remember that the positions of 1s in $\sigma[i,j]$ and $\alpha[i,j]$ are the same. We have,

$$\sum_{\sigma \in S_t} \frac{w(\sigma[1,i])}{w(\alpha[1,i])} = \frac{\mathcal{Z}_{m+t,o,l-t}}{w(\alpha[1,i])}. \tag{1.14}$$

And,

$$\sum_{\sigma \in S_t} \frac{w(\sigma[i,j])}{w(\alpha[i,j])} = \frac{w^i(M_t)Q^i_{\bar{1}}(M_t(\alpha))}{w(\alpha[i,j])}. \tag{1.15}$$

In addition,

$$\sum_{\substack{\tau; \\ \alpha[1,i-1]=\tau[1,i-1]}} \frac{\pi(\tau)}{\pi(\alpha)} = \frac{\mathcal{Z}^i_{a-m,b-o,c-l}}{w^i(\alpha[i,n])} \tag{1.16}$$

Putting all of these together, Φ_e will be equal to:

$$(1+q^{2g})\left(\frac{w(\alpha)}{w(\alpha[1,j])w^i(\alpha[i,n])}\right)\left(\frac{\mathcal{Z}^i_{a-m,b-o,c-l}}{\mathcal{Z}_{a,b,c}}\right)\sum_t \mathcal{Z}_{m+t,o,l-t}w^i(M_t)Q^i_{\bar{1}}(M_t). \tag{1.17}$$

Note that $1+q^{2g} \le 2$, $\frac{w(\alpha)}{w(\alpha[1,j])w(\alpha[i,n])^i} = \frac{1}{w(\alpha[i,j])^i}$, $\alpha[i,j] = M_0(\alpha)$ and

42

C
B
C
C
B
B
B : #As in $\sigma[i,j]=0$
.
B
C
B
B
C
B

#As in $\sigma[i,j]=?$

A
B
C
A
B
B
B : #As in $\sigma[i,j]=3$
.
B
C
B
B
A $\qquad M_3(\alpha):$
B

	σ		α	β		τ
	.		.	.		.
	.		.	.		.
$A_m(\tau)=j$	A		C	C		.
	.		B	B		.
	.	$\uparrow$	C	C		.
	.	C,B	C	C		.
	.	$\downarrow$	B	B		.
	.	mth A	A	C		.
	B		B	B		.
	.		.	.		.
	B		B	B		.
	C	lth C	C	A		.
	B		B	B		.
	.	$\uparrow$	B	B		.
	.	C,B	C	C		.
	.	$\downarrow$	B	B	A	$A_m(\tau)=i$
	.		.	.		.
	.		.	.		.
	.		.	.		.

C
B
C
A
B
B
B
.
B
A
B
B
A
A
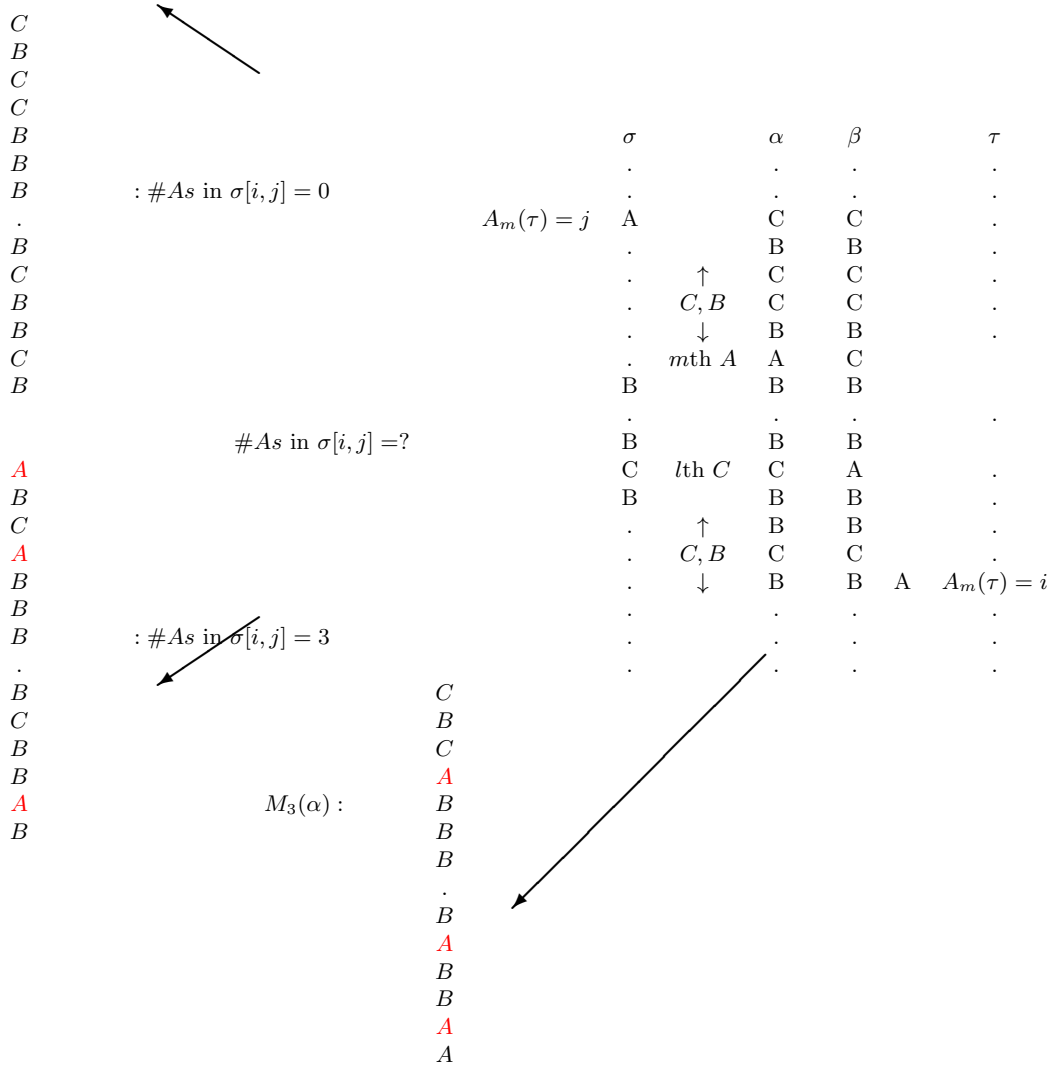
Figure 1.2: The congestion factor of move 2 in polynomial.

$\mathcal{Z}_{a,b,c} < \mathcal{Z}^i_{a-m,b-o,c-l}\mathcal{Z}_{m,o,l}$. Hence, if we only restrict the summation over all t to the case where $t=0$ we can see that Equation 2.36 is clearly less than or equal to 2. We show in a lemma that if we take the sum over all values of t, Equation 2.36 only goes up at most to a polynomial.

Lemma 16. *For any arbitrary t,*

$$w(M_t)Q_{\bar{1}}(M_t) \le q^{t(t+1)-2t}w(\alpha[i,j]). \tag{1.18}$$

and

$$\mathcal{Z}_{m+t,o,l-t} \le nq^{-ot}\mathcal{Z}_{m,o,l}. \tag{1.19}$$

Proof of Lemma 16. Using the following observations will give us the proof: We define $M_{x,y,z} := 0^x 1^y 2^z$. Clearly, $M_{x,y,z}$ has the maximum probability in $\Omega_{x,y,,z}$ and $\mathcal{Z}_{x,y,z} \simeq w(M_{x,y,x})Q(M_{x,y,z})$. For notation simplicity we use $Q(M_{x,y,z})$ and $Q_{x,y,z}$, and also $q_{M_{x,y,z}}(\sigma)$ and $q(\sigma)$ interchangeably.

Let $\binom{x+y}{y}_q$ be the q-binomial which is the generating function of number of integer partitions that fit into a rectangle of width x and length y. Note that $Q_{x,y,0} = Q_{0,y,x} = \binom{x+y}{y}_q$ (See Figure 1.3.3). We also have $Q_{x,0,y} = \binom{x+y}{y}_{q^2}$.

The following equations involving q-binomials are shown in appendix:

$$q^2 < q \Rightarrow \binom{x+y}{y}_{q^2} < \binom{x+y}{y}_q. \tag{1.20}$$

$$q < 1/2 \Rightarrow \binom{x+y}{y}_q = \binom{x+y}{x}_q < 2^x \le q^{-x}. \tag{1.21}$$

Note that $Q_{\bar{1}}(M_t) \le \binom{y}{t}_{q^2} \le q^{-2t}$ and $w(M_t)/w(\alpha[i,j]) \le q^{t(t+1)}$. Hence, we have Equation 2.6.

To prove Equation 2.5 we show that $Q_{x,y,z+t} \le nQ_{x,y,z}$ where $n = x+y+z$. Let σC^t be the concatenation of t copies of C to left of σ and let $\sigma \uparrow$ be an arrangement that you get from bringing $C_z(\sigma)$ up to the nth position.

44

Hence for $\sigma = CBBCAAB$, and $t = 3$, $\sigma C^t = CBBCAABCCC$ and $\sigma \uparrow = CBBAABC$.

$$Q_{x,y,z+t} = \sum_i \sum_{\substack{\sigma \in \Omega_{x,y,z}; \\ C_z(\sigma)=i+1}} q(\sigma[1,i])Q(\sigma[i+1,n]C^t) \leq \sum_i \sum_{\substack{\sigma \in \Omega_{x,y,z}; \\ C_z(\sigma)=i+1}} q(\sigma)\binom{i+t}{i}_q$$

$$= \sum_i \sum_{\substack{\sigma \in \Omega_{x,y,z}; \\ C_z(\sigma)=i+1}} q^i q(\sigma \uparrow)\binom{i+t}{i}_q \leq n \sum_i \sum_{\substack{\sigma \in \Omega_{x,y,z}; \\ C_z(\sigma)=n}} q(\sigma) \leq n \; Q_{x,y,z}.$$

Clearly, $Q_{x-t,y,z+t} \leq Q_{x,y,z+t}$, and $Q_{x-t,y,z+t} \leq Q_{x,y,z+t} \leq n \; Q_{x,y,z}$.

Hence,

$$\mathcal{Z}_{m-t,o,l+t} \;\; = w(M_{m-t,o,l+t})Q(M_{m-t,o,l+t})$$

$$\leq n \; w(M_{m-t,o,l+t})Q(M_{m,o,l})$$

$$\leq n \; \frac{w(M_{m-t,o,l+t})}{w(M_{m,o,l})}w(M_{m,o,l})Q(M_{m,o,l})$$

$$\leq n \; q^{-ot}\mathcal{Z}_{m,o,l}.$$

$\square$

We can now get back to Equation 2.36 and bound Φ_e.

$$\Phi_e = \left(1 + q^{2g}\right) \left(\frac{\mathcal{Z}^i_{a-m,b-o,c-l}}{\mathcal{Z}_{a,b,c}}\right) \sum_t \mathcal{Z}_{m+t,o,l-t}(\frac{w(M_t)}{w(\alpha[i,j])}Q_{\bar{1}}(M_t))^i \qquad (1.22)$$

$$\leq \left(1 + q^{2g}\right) \left(\frac{\mathcal{Z}^i_{a-m,b-o,c-l}}{\mathcal{Z}_{a,b,c}}\mathcal{Z}_{m,o,l}\right) \sum_t n(q^{2it+it(t+1)-(2t+ot)}). \qquad (1.23)$$

We know $2i + it + 1 > 2 + o$. Hence, $\Phi_e \leq 2n^2$.

Note. If $C_l(\alpha) > A_m(\alpha)$, we can similarly show that the congestion factor is less than $2n^2$.

$\tau_1 :$

C
B
C
C
C
B
B

$9 = 4 + 4 + 1 :$

$\tau_2 :$

C
C
B
C
B
C
B

$9 = 4 + 3 + 1 :$

$\tau_3 :$

C
C
C
B
B
B
C

$9 = 3 + 3 + 3 :$

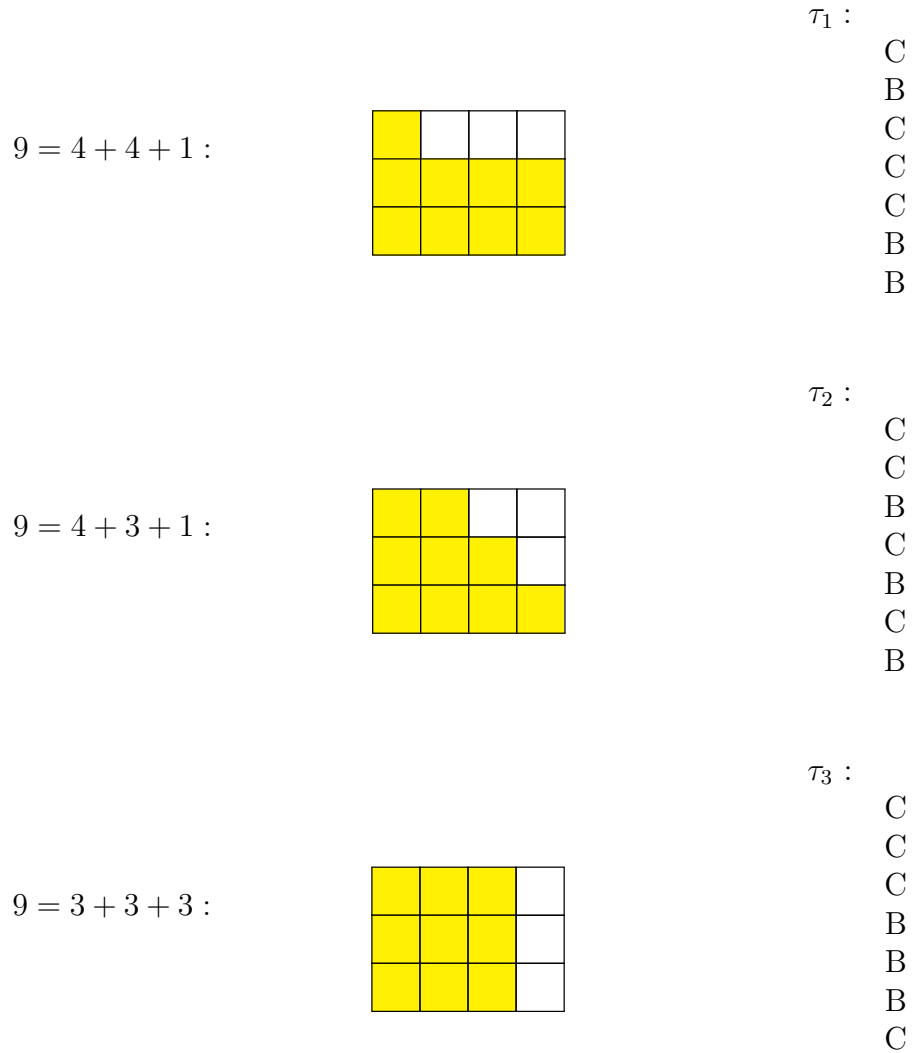Figure 1.3: Correspondence of partition functions with q-binomials: There are three integer partitions of 9 that fit into a 3×4 rectangle, and there are two arrangements of gladiators in $\Omega_{0,3,4}$ with $q(\tau_1) = q(\tau_2) = q(\tau_3) = q^9$. i.e. the coefficient for q^9 in $Q_{0,3,4}$ equals 3.

So far, we showed that any move 2 edge is only congested by a factor of a polynomial function of n. Consider an edge corresponding to move 3, namely e. We denote this edge by $e = (\alpha, \beta)$ where α and β are the same except from a B and A between which we have all Bs are swapped (Or a B and C between which we have all Bs).

Consider a state σ that used e to get to τ, and let's say we traversed e while fixing block $[i, j]$, and we had As and Cs in the block fixed and bringing the kth B to its position in τ.

Before we proceed to the proof there is a subtlety about using move 3 that needs to be explained. If A_k has to go down to reach its position in τ or if there is only one copy of it in the block there is no complication. Let's assume we have t copies of particle B in $\sigma[i, j]$. All of the t copies of B should move up and stand out of block $\sigma[i, j]$ to reach their position in τ. In order to accomplish this, we choose a subset of $S \subseteq \{1_k, \dots 1_{t+k}\}$ uniformly at random and we move the elements of S in decreasing order of their index out of the block.

Assume, when going from σ to τ we used $e = (\alpha, \beta)$ and in $\alpha[i, j]$ we have t copies of B; $B_k, \dots B_{k+t}$ and the transposition that we are making swaps the $B_{k+l}, B_{k+l+1}, \dots B_{k+d}$ with the next A. We have, $\tau[1, i] = \alpha[1, i]$, $\sigma[j+t, n] = \alpha[j+t, n]$, $\alpha_{-1}[i, j+t] = \tau(i, j)$ (See Figure 1.4) and $B_{k+i}(\alpha) < B_{l+k}(\alpha) \implies B_{k+i}(\alpha) = B_{k+i}(\sigma)$. There is some information about S that can be determined by examining α and β: $B_{k+d+1}, \dots B_{k+t} \notin S$ but S can contain any of $B_k, \dots B_{k+l}$. Hence, among the random paths connecting σ

48

σ'			σ	α	β	τ'	τ	
.			.	.	.	.	.	
.			.	.	.	.	.	
.			A	A	A	A	?	
$\sigma'[j,n] = \alpha[j,n]$			C	C	C	C	?	
.			A	A	A	B	?	
.			B	B	B	B	?	
.			C	C	C	B	?	
.			B	B	B	B	?	
.			C	B	B	B	B	j
.		∈ S	B	C	C	C	C	
$\sigma'[i,j] = \sigma[i,j]$			C	C	C	A	A	
.			C	C	C	C	C	
.		∉ S	B	B	B	C	C	
.			C	C	C	C	C	
.			A	A	A	C	C	
.			A	A	A	C	C	$\zeta[i+1,n] = \tau[i+1,n]$
.			C	C	C	C	C	
.			A	A	B	B	B	
.			C	B	B	A	A	
.			B	B	B	A	A	
.			A	B	A	C	C	B_{l+k}
.			B	C	C	A	A	
.			A	A	A	C	C	
.			C	A	A	A	A	
.			B	C	C	A	A	
.			A	A	A	C	C	
.			B	B	B	A	A	
.			C	C	C	C	C	i
.			B	B	B	B	B	
.			?	C	C	C	C	
$\sigma'[1,i] = \alpha[1,i]$			?	C	C	C	C	
.			?	B	B	B	B	
.			.	.	.	.	.	
.	$\zeta[1,i] = \sigma[1,i]$		.	.	.	.	.	
.			.	.	.	.	.	

Figure 1.4: The congestion factor for move 3 is polynomial. τ' is the arrangement that we get when all Bs are out of block $[i,j]$.

to τ, there are 2^l subsets that use e and hence the congestion they place on e is $\pi(\tau)\pi(\sigma)/2^{t-l}$.

To bound Φ_e for each e we introduce correspondence $\mathcal{F}_e : \Omega_{a,b,c} \times \Omega_{a,b,c} \to \Omega_{a,b,c}$ which satisfies,

$$\forall \zeta \in \mathcal{F}_e(\Omega_{a,b,c}); \quad \frac{\sum_{\substack{\sigma,\tau; \\ \mathcal{F}_e^{-1}(\zeta)=(\sigma,\tau)}} \pi(\sigma)\pi(\tau)}{\pi(\alpha)} \leq 2^{t-l}\pi(\zeta); \qquad (1.24)$$

where c is the number of Cs in $\alpha[i,j]$ and $\mathcal{F}_e(\sigma,\tau) \neq$ NULL $\iff$ $e = (\alpha,\beta) \in \gamma_{\sigma,\tau}$.

Let σ and τ be two ends of a path that traverses e, we define $\mathcal{F}_e$ to be the following:

$$\mathcal{F}_e(\sigma,\tau) := \zeta; \quad \zeta[i,n] := \tau[i,n], \zeta[1,i] := \sigma[1,i].$$

Let σ and τ be two ends of e and $\zeta = \mathcal{F}_e(\sigma,\tau)$.

$$\frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)} = \frac{\pi(\sigma)}{\pi(\alpha)}\frac{\pi(\tau)}{\pi(\alpha)}\pi(\alpha) = q_\alpha(\sigma)q_\alpha(\tau)\pi(\alpha). \qquad (1.25)$$

$$\frac{\pi(\zeta)}{\pi(\alpha)} = \frac{\pi(\zeta[1,i])}{\pi(\alpha[1,i])}\frac{\pi(\zeta[i,j])}{\pi(\alpha[i,j])}\frac{\pi(\zeta[j,n])}{\pi(\alpha[j,n])}. \qquad (1.26)$$

$$= \frac{\pi(\sigma[1,i])}{\pi(\alpha[1,i])}\frac{\pi(\tau[i,j])}{\pi(\alpha[i,j])}\frac{\pi(\tau[j,n])}{\pi(\alpha[j,n])} = q_{\sigma'}(\sigma)\frac{\pi(\sigma)}{\pi(\alpha)}\frac{\pi(\tau)}{\pi(\alpha)}q_{\sigma'}(\sigma)q_\alpha(\sigma)q_\alpha(\tau).$$

Where σ' is the following arrangement: $\sigma'[1,i] := \alpha[1,i]$, $\sigma'[i,j] = \sigma[i,j]$, and

$\sigma'[j, n] = \alpha[j, n]$, we have $\pi(\sigma')/\pi(\alpha) = \pi(\sigma[i, j])/\pi(\alpha[i, j])$. Hence,

$$\sum_{\sigma,\tau; \mathcal{F}(\sigma,\tau)=\zeta} \frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)} = \sum_{\substack{\sigma,\tau \\ \mathcal{F}(\sigma,\tau)=\zeta}} q_\alpha(\sigma)q_\alpha(\tau)\pi(\alpha) = \sum_{\substack{\sigma,\tau \\ \mathcal{F}(\sigma,\tau)=\zeta}} q_{\sigma'}(\sigma)\pi(\zeta). \quad (1.27)$$

Since we have $t-l$ Bs with undecided position between $j-i$ other elements we have, $\sum q_\sigma(\sigma') \leq \binom{j-i+t-l}{t-l}_q$. Using Equations 1.20 and 1.21, we have $\sum q_\sigma(\sigma') \leq 2^{t-l}$. Hence, we have Equation 1.24.

$$\Phi_{e=(\alpha,\beta)} = (1 + q^g) \sum_{\substack{\sigma,\tau \\ e \in \gamma_{\sigma,\tau}}} \frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)2^{t-l}} \leq 1. \quad (1.28)$$

We showed the for any arbitrary edge e, $\Phi_e \leq \max\{n^2, 1\}$. We also have $\pi_{min} \leq (q^{n(n+1)}/|\Omega_{a,b,c}|)$. Now we apply Theorem 6 and we will have,

$$t_\epsilon(\mathcal{EX}_t) \leq 8n^4(n^3 + \ln(\epsilon^{-1})). \quad (1.29)$$

$$t(\mathcal{EX}_t) \leq 8n^7. \quad (1.30)$$

$\square$

To show that the exclusion process chain is rapidly mixing we compare it to $\mathcal{EX}_t$.

1.3.4 Proof of Theorem 11.

We compare $\mathcal{EX}_t$ and $\mathcal{EX}$, using Theorem 7 and knowing $\mathcal{EX}_t$ is rapidly mixing will make the proof complete.

Proof. Consider an edge $e = (\sigma, \tau)$ that makes move 2 in $\mathcal{EX}_t$. Assume e is swapping $\sigma(i) = A$ and $\sigma(i + 1 + d) = C$ and $\sigma[i + 1, i + d] = B^d$. We indicate a swap of elements $\sigma(i)$ and $\sigma(i+1)$ in $\mathcal{EX}$ by t_i or $t_i(p, p')$ if we are swapping particles p and p' and $\sigma(i) = p, \sigma(i+1) = p'$.

Let $\gamma_{\sigma,\tau} := t_i(A, B), \ldots t_{i+d-1}(A, B), t_{i+d}(A, C), t_{i+d-1}(B, C), \ldots t_i(B, C)$. Consider an arbitrary $e = (\alpha, \beta)$ that is making swap $t_i(A, B)$ and $\alpha[i-t, i-1] = 1^t$. If $e \in \gamma_{\sigma,\tau}$ then, $\exists j \in [1, t]; \forall k \neq i, i - j \; \sigma(k) = \alpha(k); \sigma(i - j) = A, \sigma(i) = B$. And τ will be the state that you reach by swapping $\sigma(i-j) = A$ and $\sigma(i+1+d) = C$. Let e' be the edge in $\mathcal{EX}_t$ that connects σ to τ. We have, $\mathcal{C}(e') = \pi(\sigma)/(1 + 1/q^{2(d+j)})$ and $\pi(\sigma) = \pi(\alpha)1/q^j$.

$$A_e = \frac{\sum_{\Gamma(e)} |\gamma_{\sigma,\tau}| \mathcal{C}(e')}{\mathcal{C}(e)} = \sum_{j=1}^{t} \frac{|\gamma_{\sigma,\tau}| 1/q^j (1 + 1/q)}{1 + 1/q^{2(d+j)}} \leq 2(d+t) \sum_{j=1}^{t} \frac{1/q^j (1 + 1/q)}{1/q^{2(d+j)}}$$
(1.31)

$$A_e \leq 2(d+t)(1 + 1/q)q^{2d} \sum_{j=1}^{t} q^j = \frac{2(d+t)(1 + 1/q)q^{2d}(1 - q^{t+1})}{1 - q} \qquad (1.32)$$

Hence, $\mathcal{A}_e \leq 2(d + t) \leq n$, where n is the length of the arrangements or total number of particles.

Similarly, we can show that $\forall e \in \mathcal{EX}, \mathcal{A}_e \leq n$. $\pi_{min} \leq (q^{n(n+1)}/2^{3/2n})$. Hence using Theorem 7, Theorem 15 and Theorem 6 we have,

$$t(\mathcal{EX}) \leq n(2n^3)t(\mathcal{EX}_t) \leq 16n^{11}. \qquad\qquad (1.33)$$

$\square$

1.3.5 Proof of Lemma 10

Proof. Let $\mathcal{M}$ be a particle system of types A, B, C, each type having m copies, $p_{A,B} = p_{B,C} = \frac{1}{1+q} > 1/2$, i.e. $q < 1$ and $p_{A,C} = 1/2$ (The values for m and q will be set later.) We show that the conductance of this chain in small, employing Theorem 5 we conclude

$$t(\mathcal{M}) \geq \left(\frac{1}{4\Phi} - 1/2 \right) \geq \mathcal{O}(q^{-2m})$$

Let P be the transition matrix of $\mathcal{M}$. Consider the set $S \subseteq \Omega$ to be the set of all arrangements where all the Cs are above As. The conductance of S is denoted by $\Phi(S)$ which is an upper bound for the conductance of the chain if $\pi(S) \leq 1/2$.

Let $\rho(S) = \{x \in S | \exists y \in \Omega \setminus S; P(x,y) > 0\}$. To each $\sigma \in \rho(S)$, we assign $\sigma' \in S \setminus \rho(S)$ by taking either the topmost A (or the lowermost C in σ) and taking it down (or top) below (or above) the next $m/2$ ones; σ' will be satisfying $\pi(\sigma') = q^{m/2}\pi(\sigma)$. For a particular choice of σ' there are at most $2m$ arrangements which are assigned to it. Hence $\Phi(S) \leq q^{m/2}/2m$. It remains to show $\pi(S) \leq 1/2$.

Since swapping a A and C does not change π we have the following

equation in which k shows the position in which we can have the lowest C or topmost A:

$$\pi(S) \leq$$

$$\sum_{k=1}^{m} \frac{\binom{m+k}{k}_q \binom{2m-k}{m-k}_q}{\binom{2m}{m}_q^2}$$

$$\leq m \frac{\binom{m+m/2}{k}_q^2}{\binom{2m}{m}_q^2}$$

$$\leq m \left(\frac{(1-q^{m/2})}{(1-q^{2m})} \right)^{2m}$$

$$\leq \frac{m}{\left((1+q^{m/2})(1+q^m) \right)^{2m}}$$

Hence for having $\pi(S) \leq 1/2$ it suffices to set $q > \left((2m)^{1/2m} - 1 \right)^{1/m}$. Note that since $\lim (2m)^{1/2m} = 1$, for any $m > 2$, there is a constant value for $q < 1$ satisfying $q > \left((2m)^{1/2m} - 1 \right)^{1/m}$.

$\square$

1.4   The Interval Chain

Recall that a partially ordered set or a **poset** on $\{1, 2, \ldots n\}$ is a reflexive, transitive, and antisymmetric relation on $\{1, 2, \ldots n\}$. The spectacular variety of the applications of partially tordered sets in other fields of science together with the beauty of related combinatorial problems have been intriguing many mathematicians, computer scientists, biologists, and recently micro economists [75, 73, 84].

Notation. Like other sections of this thesis, we denote a partially ordered set by $\mathcal{P}$ and by $i \leq_\mathcal{P} j$ we mean $(i, j) \in \mathcal{P}$. If $i \not\leq_\mathcal{P} j$ nor $j \not\leq_\mathcal{P} i$, we say i and j are not comparable in $\mathcal{P}$, and we denote it by $i \parallel_\mathcal{P} j$. Recall that for a poset $\mathcal{P}$ on $\{1, 2, \ldots n\}$, the set $\mathcal{L}(\mathcal{P})$ is the subset of all linear orders l of numbers $\{1, 2, \ldots, n\}$ in which $i \leq_\mathcal{P} j \implies i \leq_l j$. The set $\mathcal{L}(\mathcal{P})$ is known as the set of linear extensions of $\mathcal{P}$. Let $l \in \mathcal{L}(\mathcal{P})$ be a linear extension of $\mathcal{P}$ with total ordering $l_1 \leq_l l_2 \leq_l \cdots \leq_l l_n$, $x, y \in \{1, 2, \ldots n\}$ are adjacent in l if and only if $\exists i l_i = x, l_{i+1} = y$. Two linear extensions $l, l' \in \mathcal{L}(P)$ are different in adjacent transposition i if $l_1 \leq_l l_2 \leq_l \ldots l_i \leq_l l_{i+1} \cdots \leq_l l_n$ and $l_1 \leq_{l'} l_2 \leq_{l'} \cdots \leq_{l'} l_{i+1} \leq_{l'} l_i \cdots \leq_{l'} l_n$. When putting no subscrit on $\leq$ we mean the ordering in real numbers.

It is known that counting the number of linear extensions is an NP-hard problem [75], and there had been a significant amount of research to find algorithms to approximate this number until a Markov chain algorithm was introduced and analyzed by Karzanov and Khachiyan in 1991 ([74]),

Wilson tightened the upper bound of its runtime in 2004 ([52]). Since the two problems of sampling uniformly from a set and approximating its size are known often to be computationally equivalent [70], many Mathematicians studied the problem of approximating $|\mathcal{L}(P)|$ by looking at the Markov chains on $\mathcal{L}(P)$ [74, 22, 52].

One of the most natural Markov chains on $\mathcal{L}(\mathcal{P})$ is the following chain introduced and analyzed by Karzanov and Khachiyan (1991) [74]:

Karzanov-Khachiyan Chain (Order Markov Chain.) Let $\mathcal{P}$ be a partially ordered set on the elements of $\{1, 2, \ldots n\}$ and $\mathcal{L}(\mathcal{P})$ the set of its linear extensions. At state $l \in \mathcal{L}(\mathcal{P})$, pick two adjacent elements l_i and l_{i+1} w.p. $1/2(n-1)$, and swap them if $l_i \parallel_\mathcal{P} l_{i+1}$. Otherwise do nothing.

It is easy to show that the above chain is ergodic and that it converges to the uniform distribution on $\mathcal{L}(P)$. Karzanov and Khachiyan demonstrated that it is rapidly mixing by analyzing its conductance (Section 1.2. Theorem 5):

Theorem 17. *[74] The conductance of the Karzanov-Khachiyan chain for a poset of n elements is at least $2^{-3/2}n^{-5/2}$.*

Using the the above analysis, Karzanov and Khachiyan obtained an upper bound of $\mathcal{O}(n^{7/2}\log^2(n))$ for the mixing time of their chain. Wilson tightened this bound to $\mathcal{O}(n^3 \log n)$[52].

The Karzanov-Khachiyan chain is a random walk on a convex hull of some vertices (determined by the poset) on the permutohedron. The chain

we are interested is a random walk in the union of simplices in the unit cube of dimension n: Consider $x = (x_1, x_2, \ldots, x_n) \in [0,1]^n$ where the coordinates, x_i are ordered according to the partial order $\mathcal{P}$ defined on $\{1, 2, \ldots n\}$. i.e. $i \leq_{\mathcal{P}} j \implies x_i \leq x_j$. The set of all such xs corresponds to a union of simplices in $[0,1]^n$. Although rapidly mixing Markov chains on arbitrary convex bodies [34, 35] have been studied in various papers, Markov chains on simplices corresponding to partial orders have not been studied yet. Since simplexes especially those corresponding to linear orders as defined above, are much simpler than convex bodies in general, we became motivated to design and analyze a simpler Markov chain for the purpose of sampling points from them.

The following chain was motivated by the so called "micro-canonical hard-core lattice gas model in dimension 1" and introduced and studied by Randall and Winkler [36].

Randall-Winkler Chain . Let $\Omega = \{x = (x_1, x_2, \ldots, x_n)|0 \leq x_1 \leq x_2 \leq \ldots x_n \leq 1\}$. At state $x \in \Omega$, pick $i \in \{1, 2, \ldots, n\}$ uniformly at random. Then, pick x_i' uniformly at random from $[x_{i-1}, x_{i+1}]$. Change x_i to x_i' in x and leave the rest unchanged.

Theorem 18. *[36] The mixing time of the Randall-Winkler chain is $\mathcal{O}(n^3 \log n)$.*

In this paper we introduce a Markov chain on the set containing all the combinations of n dots[6] in the unit interval respecting a partial order P (See

---

[6]In this text, we stick to Randall and Winkler's terminology of calling the elements in $[0,1]^n$ "points". If $x = (x_1, x_2, \ldots, x_n) \in [0,1]^n$ then, we call x_is as "coordinates" or "dots in $[0,1]$".
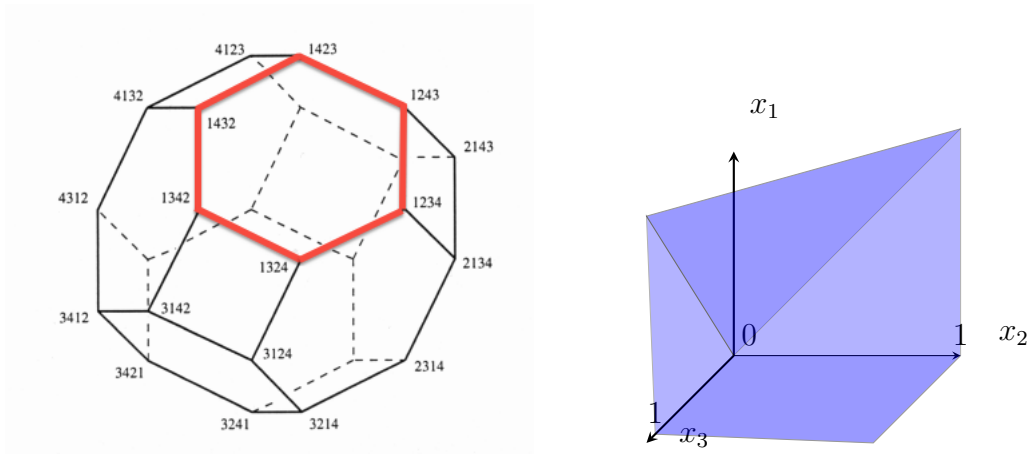
Figure 1.5:
On the left we see the convex hull corresponding to the poset $\mathcal{P} = \{1 \leq_{\mathcal{P}} 2, 1 \leq_{\mathcal{P}} 3, 1 \leq_{\mathcal{P}} 4\}$ on the permutohedron of 4 elements. (Discrete model.) On the right we see the simplex corresponding to $\mathcal{P} = \{1 \leq_{\mathcal{P}} 2, 1 \leq_{\mathcal{P}} 3\}$ in the unit cube $[0, 1]^3$ for 3 elements. (Continuous model.)

Definition 14). The two extreme cases are when $\mathcal{P}$ is a chain[7] (which is the Randall-Winkler chain) and the case where $\mathcal{P}$ is an antichain[8]. (which mixes in time $n \log n$ by the coupon collector problem.) We present positive evidences that this chain mixes rapidly for arbitrary $\mathcal{P}$.

We define the following variant of Randall-Winkler chain:

Definition 12. Let $\Omega = \{x = (x_1, x_2, \ldots, x_n) | 0 \leq x_1 \leq x_2 \leq \ldots x_n \leq 1\}$, and $\mathcal{I}$ the following Markov chain on Ω: At state $x \in \Omega$, pick $x' \in [0, 1]$ uniformly at random. Find the smallest interval $[x_i, x_{i+1}]$ with the property that $x' \in [x_i, x_{i+1}]$. (Take $x_0 = 0$ and $x_{n+1} = 1$.) Flip a fair coin and change x_i to x' or change x_{i+1} to x'. (If $x_i = 0$ or $x_{i+1} = 1$, do nothing.) Leave the rest of the dots unchanged.

At this point we have no knowledge about the missing time of $\mathcal{I}$. We present some experimental evidences that it mixes rapidly at the end of this section.

Definition 13. Let P be a partially ordered set defined on $\{1, 2, \ldots n\}$. We define Ω_P as follows:

$$\Omega_P = \{x = (x_1 \ldots x_n) | x_i \in [0, 1], i \leq_P j \implies x_i \leq x_j\} \tag{1.34}$$

Ω_P is a union of simplices in the unit cube of dimension n.

---

[7]A chain is a poset in which all the elements are comparable.
[8]An antichain is a poset in which all the elements are incomparable.

Notation. In $\mathcal{P}$, we say i covers j iff $j \leq_\mathcal{P} i, j \neq i$ and $\nexists k j \leq_\mathcal{P} k \leq_\mathcal{P} i$. We denote i covers j by $j <_\bullet i$. Consider $x = (x_1, x_2, \ldots, x_n) \in \Omega_P$, we define $\max(x_i) = x_j$ iff x_j has the minimum value among all x_ks for which we have $i <_\bullet k$ in $\mathcal{P}$. If i is maximal in $\mathcal{P}$, $\max(x_i) = 1$. likewise, $\min(x_i) = x_j$ iff x_j has the maximum value among all x_ks for which we have $k <_\bullet i$ in $\mathcal{P}$. If i is minimal in $\mathcal{P}$, $\min(x_i) = 0$.

To take approximately uniform random samples from Ω_P we propose the following Markov chain:

Definition 14. Let $\mathcal{M}$ be the following Markov chain on Ω_P:

At state $x = (x_1, x_2, \ldots x_n) \in \Omega_\mathcal{P}$, pick one of the numbers $\{1, 2, \ldots n\}$ with probability $1/2n$, and with probability $1/2$ don't do anything. If $i \in \{1, 2, \ldots, n\}$ is taken, pick a point x_i' uniformly at random from the interval $[\min(x_i), \max(x_i)]$, change the position of x_i to x_i' and keep the rest of points unchanged to move to state $x = (x_1, x_2, \ldots x_{i-1}, x_i', x_{i+1}, \ldots, x_n) \in \Omega_\mathcal{P}$.

It is easy to see than $\mathcal{M}$ is an ergodic Markov chain and since the interval $[\min(x_i), \max(x_i)]$ is unchanged we have $\forall x, x' \in \Omega_\mathcal{P}, P(x, x') = P(x', x)$ and then detailed balance affirms that the stationary distribution of $\mathcal{M}$ is the uniform distribution on $\Omega_\mathcal{P}$.

Theorem 19. *Let $t(\mathcal{I})$ be the mixing time of the Markov chain $\mathcal{I}$. Then,*

$$t(\mathcal{M}) \leq \mathcal{O}\left(n^{7/2} \log(|\mathcal{L}(\mathcal{P})|)\right) t(\mathcal{I}) \tag{1.35}$$

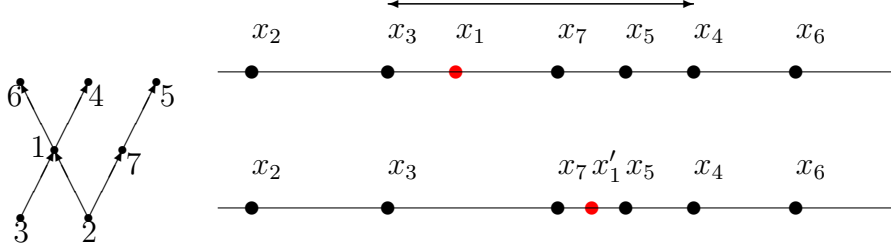We employ the decomposition technique (See Section 1.2. Theorem 8.)

Figure 1.6:
In the Markov chain $\mathcal{M}$, x_1 is picked up w.p. 1/14 then it moves uniformly at random in the interval $[x_3, x_4]$.

Proof of Theorem 19. Consider $\mathcal{L}(\mathcal{P})$ the set of all linear extensions of $\mathcal{P}$, we partition $\Omega_{\mathcal{P}}$ into $\{\Omega_l\}_{l \in \mathcal{L}(\mathcal{P})}$. For any arbitrary $l \in \mathcal{L}(\mathcal{P})$, Ω_l is the set $\{x = (x_1, x_2, \ldots, x_n) | x_i \in [0, 1], i \leq_l j \implies x_i \leq x_j\}$. By rejecting the moves outside Ω_l, we will have $\mathcal{I}$:

It remains to analyze the mixing time of $\bar{\mathcal{M}}$ of Theorem 8. This chain has the state space $\{\Omega_l\}_{l \in \mathcal{L}(\mathcal{P})}$ and transition probability of going from Ω_l to $\Omega_{l'}$ is nonzero iff l and l' only differ in a **legitimate shift**.

A legitimate shift in a linear extension of $l : l_1 \leq_l l_2 \leq_l \cdots \leq_l l_n \in \mathcal{L}_P$ is to take a sequence $l_i \leq_l l_{i+1} \leq_l \ldots l_{i+k}$ and shift it to $l_{i+1} \leq_{l'} l_{i+2} \leq_{l'} \cdots \leq_{l'} l_{i+k} \leq_{l'} l_i$ or $l_{i+k} \leq_{l'} l_i \leq_{l'} l_{i+1} \leq_{l'} l_{i+2} \leq_{l'} \cdots \leq_{l'} l_{i+k-1}$ if that makes l' another linear extension of P.

Example 10. Consider the poset P depicted in Figure 2. The following pairs of linear extensions differ in a legitimate shift: $273164\underline{5} \leftrightarrow 273\underline{5}164$, $273164\underline{5} \leftrightarrow 27\underline{5}3164$, and $2\underline{3}71564 \leftrightarrow \underline{3}275164$.

61

Although it is difficult to find the exact values for the transition probabilities of $\bar{\mathcal{M}}$, the following claim will play a key role in finding a polynomial bound for mixing time of $\bar{\mathcal{M}}$:

Claim 1. *Let l and l' be two linear extensions of $\mathcal{P}$ differing only in adjacent transposition i and $l_i \parallel_\mathcal{P} l_{i+1}$. Then the transition probability of going from Ω_l to $\Omega_{l'}$ in $\bar{\mathcal{M}}$ satisfies:*

$$P_{\bar{\mathcal{M}}}(\Omega_l, \Omega_{l'}) \geq 1/8n^2. \tag{1.36}$$

Proof of Claim 1. Consider an arbitrary $i \in \{1, 2, \ldots, n\}$ and x uniformly at random from Ω_l and assume there are k point within $[\min(x_{l_i}), \max(x_{l_i})]$. We have,

$$Pr(\frac{|x_{l_i} - x_{l_{i+1}}|}{|\max(x_{l_i}) - \min(x_{l_i}),|} \geq 1/k) = (1 - 1/k)^k \geq 1/4 \text{ (for } k \geq 2.)$$

Hence, if $x_{l_i} \parallel_\mathcal{P} x_{i+1}$, then $k \geq 2$ and in at least $1/4$th of the points $x \in \Omega_l$ we have,

$$\frac{|x_{l_i} - x_{l_{i+1}}|}{|\max(x_{l_i}) - \min(x_{l_i}),|} \geq 1/k \geq 1/n.$$

$$\sum_{y \in \Omega_{l'}} P(x, y) \geq \frac{1}{2n} \left(Pr(\frac{|x_{l_i} - x_{l_{i-1}}|}{|\max(x_{l_i} - \min(x_{l_i})|}) \geq \frac{1}{k} \right).$$

$$P_{\bar{\mathcal{M}}}(\Omega_l, \Omega_{l'}) = \sum_{x \in \Omega_l, y \in \Omega_{l'}} \pi(x)P(x,y)/\pi(\Omega_i) \geq (\frac{1}{n^2}) \sum_{x \in \Omega_l} \pi(x)/\pi(\Omega_i) \geq (\frac{1}{8n^2}).$$

□

We now compare the conductance of $\bar{\mathcal{M}}$ to the conductance of Karzanov-Khachiyan chain [74], (Theorem 17) to bound the mixing time $t(\bar{\mathcal{M}})$.

Lemma 20. $t(\bar{\mathcal{M}}) = \mathcal{O}\left(n^{5/2} \log |\mathcal{L}(P)|\right).$

Proof of Lemma 20. Let $\mathcal{K}$ be the Karzanov-Khachiyan chain and $P_{\mathcal{K}}$ its transition probability. Let l and l' be two linear extensions of $\mathcal{P}$ that differ in an adjacent transposition which swaps l_i and l_{i+1} where $l_i \parallel l_{i+1}$. We have $P_{\bar{\mathcal{M}}}(\Omega_l, \Omega_{l'}) \geq 1/8n^2$ and $P_{\bar{\mathcal{K}}}(l, l') = 1/2(n-1)$. i.e. $P_{\bar{\mathcal{M}}}(\Omega_l, \Omega_{l'}) \geq (1/4n)P_{\bar{\mathcal{K}}}(l, l').$

Hence, for any $S \subset \Omega_{\mathcal{P}}; \pi(S) \leq 1/2$ by applying Theorem 17 we have,

$$\sum_{x \in S, y \in \bar{S}} \left(\frac{\pi(x)P_{\bar{\mathcal{M}}}(x,y)}{\pi(S)} \right) \geq \sum_{x \in S, y \in \bar{S}} 1/n \left(\frac{\pi(x)P_{\mathcal{K}}(x,y)}{\pi(S)} \right) \geq (1/4n)2^{-3/2}n^{-5/2}$$
$$= 2^{-7/2}n^{-7/2}.$$
(1.37)

Using Theorem 5,

$$t(\bar{\mathcal{M}}) = \mathcal{O}\left(n^{7/2} \log |\mathcal{L}(P)|\right).$$

□

By employing Theorem 8 we will have Equation 1.35.

□

To conclude this section, we present some experimental result suggesting that the Markov chain $\mathcal{I}$ and consequently $\mathcal{M}$ mix rapidly.

1.4.1 Experimental Results

We ran a simulation of the Randall-Winkler chain and the interval chain $\mathcal{I}$ in MATLAB. For each chain, we generated point in $\mathbb{R}^n$ for different values of n, and we took a histogram of n^2 points generated between time t and $t + n^2$ (t increasing until some conditions are met), and took that as the probability distribution. We calculated the total variation distance between this histogram and the histogram for the uniform distribution.

Our experiments suggest that the mixing time of the Randall-Winkler chain and our chain are very close. In the following picture we took $n = 24$; and the number of histogram bins, b is equal to 12. For various choices of n and b we will get the same plot.
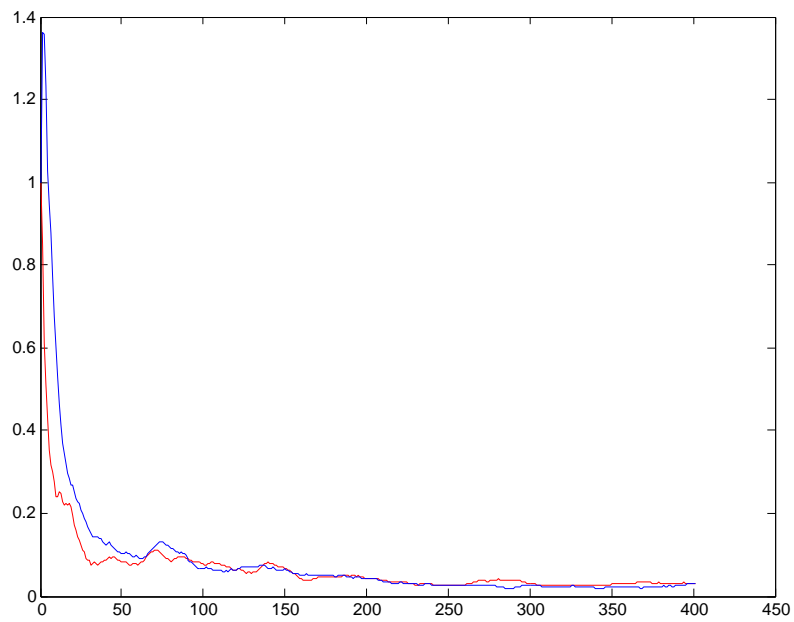
Figure 1.7: The red plot shows the total variation distance of our chain and the uniform distribution as time increases; and the blue plot shows the same quantity for Randall-Winkler chain.

1.5 Håstad's Shuffling

Consider the numbers $\{1, 2, \ldots, n\}$, and let S_n be the set containing all the permutations of these numbers. As discussed in the introduction of this chapter, shuffling these numbers is a random walk on S_n converging to the uniform distribution on S_n.

Here we study a shuffling studied by Y. Håstad (See [38]), when $n = m^2$ and the numbers $\{1, 2, \ldots n\}$ are arranged in an $m \times m$ square. We extend Håstad's result to the cases where $n = m^k$ for arbitrary k and the numbers $1, 2, \ldots, n$ are arranged in a cube of dimension k.

The shuffling in an square is as follows:

Definition 15. (Square Shuffling) Consider an $m \times m$ square. Assume we are at state σ of the chain. At each step of the Markov chain, we pick m independent permutations $(\pi_i)_{i=1}^{m}$ uniformly at random. At odd time steps and for $0 \leq i \leq m-1$, we apply π_i, to the elements in row i all at the same time. At even steps these permutations are applied to columns of the square. We call this Markov chain $\mathcal{M}_m$.

By applying these rules, we are actually taking a permutation from distribution Π and applying it is σ. We will use this notation throughout, and by $\sigma \circ \Pi^t$ we mean t repetitions of applying the aforementioned steps to σ.

The following theorem appeared in [38]:

Theorem 21. *(Håstad's theorem [38])*

Let U_{m^2} be the uniform distribution on the elements of an $m \times m$ square. We denote the standard deviation of Π^t and U_{m^2} by $\parallel (\Pi^t, U_{m^2}) \parallel_{TV}$ and we have,

$$\parallel \Pi^t, U_{m^2} \parallel_{TV} \leq O\left(m^{2-\lfloor t/3 \rfloor 1/2}(2\log m^{\lfloor t/3 \rfloor})\right).$$

Corollary 22. $\parallel \Pi^t, U_{m^2} \parallel_{TV} \leq \log^5(m)/m^4 \leq 1/4$ *for $m \geq 2$. Thus, for all m $t_{mix}(\mathcal{M}_m) \leq 15$.*

The proof of Theorem 21 which appeared in [38] relies on coupling (See subsection 1.2.1). In fact Håstad shows that for any σ_1 and σ_2 which are arrangements of numbers in an $m \times m$ square and are different only in a single transposition, if we take (τ_1, τ_2) with marginal distribution Π^3, then we can couple σ_1 and σ_2 such that $\sigma_1 \circ \tau_1 = \sigma_2 \circ \tau_2$ with probability $1 - q$, otherwise $\sigma_1 \circ \tau_1$ and $\sigma_2 \circ \tau_2$ still differ only in one transposition. In this proof $q = O((m/\log m)^{-1/2})$. We will discuss the details of this coupling in future. Using this coupling Håstad concludes:

$$\parallel \Pi^t, U_n \parallel_{TV} \leq nq^t \quad ; q = O((m/\log m)^{-1/2}). \tag{1.38}$$

Before explaining how Håstad's coupling works, we present the following definition which appeared in [38] for squares, and here we state it for rectangles of arbitrary sizes:

Definition 16. (pattern [38])

To each successive row and column permutations taken from Π^2, we associate a pattern. A pattern is an $m \times r$ rectangle containing m copies of

numbers $0, \ldots, m-1$. In each cell (i, j) of this square, we have a number α which denotes the row number where the current element in (i, j) will move to after arrow and column shuffling. Note right before another row shuffling, a pattern would contain all the important information we need.

The key idea in Håstad's proof is the association between patterns and bi-partite graph matching covers. Håstad discovers this association for squares, and we observed that it also works for rectangles. We present a very brief summary here the interested reader should refer to proof of Lemma 3.3 in [38], and observe that it also works for rectangle patterns.

The coupling. Take a uniformly random pattern p of size $m \times n$, and erase without loss of generality its contents at cell $(0, 0)$ and $(1, 1)$. To the original p we can associate a bi-partite multi graph having $2m$ vertices: $\{1^+, 2^+, \ldots m^+\} \cup \{1^-, 2^-, \ldots m^-\}$. There is an edge with label j between (i^+, α^-) if and only if cell (i, j) contains α in p. Note that each vertex has degree n, and thus this multigraph can be partitioned to n matching covers. To each matching cover will be mapped to a pattern but each pattern can be mapped will be mapped to a number of matching covers. The number of matching covers mapped to a pattern is proportional to the probability of that pattern.

Consider p and its erased cells again. If the two values we eased are the same, the patterns will be coupled after the next row permutation. Hence, we assume they have different values, i.e. there are two ways to fill this partially filled pattern. Let these two ways be p_1 and p_2. Håstad categorized the

68

matching covers associated to p_1 and p_2; and by finding a mapping between them shows:

$$\sum_p |Pr(p_2) - Pr(p_1)| \leq O\left((\frac{\log m}{m})^{1/2} \right). \tag{1.39}$$

Then, he derives Equation 1.38.

We generalize Håstad's result to higher dimension lattices:

Definition 17. (k-dimensional Cube Shuffling) Consider an m^k square. Assume we are at state σ of the chain. At each step of the Markov chain, we pick m independent permutations $(\pi_i)_{i=1}^m$ uniformly at random. Each of these permutation are permutations on an m^{k-1} cube. At time step t, we apply these permutations to the m, $k-1$-dimensional cubes, that we get by fixing the ith coordinate, i being the remainder of t divided by m.

By applying these rules, we are actually taking a permutation from distribution Π_k and applying it to σ. We will use this notation throughout, and by $\sigma \circ \Pi_k^t$ we mean t repetitions of applying the aforementioned steps to σ.

We prove that this chain, like the square chain becomes close to the uniform distribution, U_{m^k} in constant number of steps. (the constant is dependent on k but not m.)

Theorem 23. *Let U_{m^k} be the uniform distribution on the elements of a k dimensional cube of size m^k. We have,*

$$\| (\Pi_k^t, U_{m^k}) \|_{TV} \leq O(m^{k-1/2\lfloor t/3 \rfloor}(\log(m)^{1/2} + 1/m)^{\lfloor t/3 \rfloor}).$$

69

We prove Theorem 23, by induction. Hence, we need the following definitions and lemmas:

Definition 18. (Rectangle Shuffling)

Let's say we have an $m \times n$ square. At each odd time step we pick m independent permutations of $\{0, \ldots n-1\}$, $(\sigma_i)_{i=1}^m$ uniformly at random. For $1 \leq i \leq m$, we apply σ_i to the elements in row i. At even steps, we pick n independent permutations of $[m]$, $(\sigma_i)_{i=1}^n$ uniformly at random. These permutations are applied to elements of column i all at the same time.

We denote the distribution of the elements in an $m \times n$ rectangle after t steps by $\Pi_{m \times n}^t$. As we discussed Håstad's proof can be extended to work for rectangles, and we have the following lemma:

Lemma 24. *Let $U_{m \times n}$ be the uniform distribution on the elements of an $m \times n$ square. We denote the standard deviation of $\Pi_{m \times n}^t$ and $U_{m \times n}$ by $\| \Pi_{m \times n}^t - U_{m \times n} \|_{TV}$ and we have,*

$$\| \Pi_{m \times n}^t - U_{m \times n} \|_{TV} \leq O(m \times n^{1-\lfloor t/3 \rfloor 1/2} (\log n^{\lfloor t/3 \rfloor})). \qquad (1.40)$$

Definition 19. (Approximate Shuffling)

Let's say we have an $m \times n$ square. At each odd time step, we pick m independent permutations of $\{0, \ldots, n\}$, $(\sigma_i)_{i=1}^m$ uniformly at random. We apply each σ_i to the elements in row i. At even steps we pick n independent permutations of $\{0, 1, \ldots, m\}$, $(\sigma_i)_{i=1}^n$ from a random distribution $\tilde{U}_m$

where we have $\| \tilde{U}_m - U_m \|_{TV} \leq \epsilon$, and U_m is the uniform distribution on $\{0, 1 \ldots, m\}$. These permutations are applied to elements of column i all at the same time.

We denote the distribution of the elements in an $m \times n$ rectangle after t steps of approximate shuffling by $\tilde{\Pi}^t_{m \times n}$. We have the following theorem:

Lemma 25. *Let $U_{m \times n}$ be the uniform distribution on the elements of the elements of an $m \times n$ square. We have,*

$$\| \tilde{\Pi}^t - U_{m \times n} \|_{TV} \leq O(mn((\frac{\log n}{n})^{1/2} + \epsilon)^{\lfloor t/3 \rfloor}).$$

Consider all $m \times n$ patterns. Let p be any of possible $m \times n$ partial patterns in which the two cells $(0, 0)$ and $(1, 1)$ are erased. Let p_1 and p_2 be two $m \times n$ patterns that can be constructed by filling the two erased elements of p. We prove the following lemma. Then, we can show that $\tilde{\Pi}^t_{m \times n}$ allows a coupling of distance two with parameter $\epsilon + (n/\log n)^{-1/2}$.

Following Håstad's proof and employing the triangle inequality we will have:

Lemma 26. *Let p_1 and p_2 be two $m \times n$ patterns of approximate shuffling that differ in a single transposition. We have,*

$$\sum_p |Pr(p_1) - Pr(p_2)| \leq \epsilon + (n/\log n)^{-1/2}. \tag{1.41}$$

Definition 20. (Stop and Shuffling in k-Dimensional Cube)

In a k-dimensional cube , the stop and shuffling is similar to the normal shuffling of Definition 17. In the stop and shuffling of rate γ, we pick m independent permutations $(\pi_i)_{i=1}^m$ at γ consecutive time steps . Each of these permutation are permutations on an m^{k-1} cube. We apply these permutations to the m $k-1$-dimensional cubes that we obtain by fixing dimension i dimension. We go that to the next dimension after repeating this for γ times.

Lemma 27. *Given Definitions 17 and 20, we have,*

$$\| \; \underline{\Pi}_{n^k}^t - U_{n^k} \; \|_{TV} \quad \leq \quad \| \; \Pi_{n^k}^t - U_{n^k} \; \|_{TV} \; . \tag{1.42}$$

Proof. Consider a k dimensional cube of size m^k. We run an equivalent version of the Markov chain introduced in Definition 17 by taking m permutations at a time and applying them to the m, $k-1$-dimensional cubes obtained by fixing the first coordinate. i.e. rows. Instead of going to the next dimension, we rotate the cube. Since the uniform distribution is symmetric on the whole cube, the deviation distance of the distribution of the cube after t steps from the uniform distribution is equal to $\| \; \Pi_{n^k}^t - U_{n^k} \; \|_{TV}$. To imitate the stop and shuffle chain, we perform the same operation but we rotate the cube every γ times. After each $t = \gamma k$ steps we have,$\| \; \underline{\Pi}_{n^k}^t - U_{n^k} \; \|_{TV} \quad \leq \quad \| \; \Pi_{n^k}^t - U_{n^k} \; \|_{TV}$. Hence, Equation 1.42 holds. $\qquad \square$

Proof of Theorem 23.

We prove this theorem by induction. For base case, we have Håstad's

theorem.

Assume the theorem holds for $n = k - 1$ i.e. after $t = 6k + 6$ steps:

$$\| \Pi^t_{n^k} - U_{n^{k-1}} \|_{TV} \leq O(n^{-2}(\log(n)^{1/2} + 1/n)^{2k+2}) = O(n^{-2}\log(n)^{k+1}) = O(n^{-1}).$$

We perform a stop and shuffling of the k-dimensional cube at rate $6k + 6$. Note that this whole process will be equivalent with approximate shuffling a rectangle of size $n^{k-1} \times n$ where, $\| \tilde{U} - U) \|_{TV} \leq 1/n$. Therefore, using the result of theorem 19, we will have the result.

$\square$

1.6 The Top to Random Shuffle

In the top to random shuffling we take a card from the top of the deck and insert it uniformly at random into the deck. This chain was analyzed by Aldous and Diaconis ([72].)

The chain can be analyzed very similar to the coupon collector problem: Let Ω be the state space containing all permutations of numbers $\{1, 2, \ldots, n\}$. Assume without loss of generality that we start at $\sigma = (1, 2, 3, \ldots, n)$. Let T_i be the time needed to place card i after card n given that $i-1$ is already placed after n; T_i has the domestic distribution with parameter $(n-i+1)/n$ thus, $\mathbb{E}(T_i) = n/n - i + 1$.

Let T be the time it takes card $n-1$ to be places after card n. Then, $\mathbb{E}(T) = \sum_{i=1}^{n-1} \frac{n}{n-i+1} + 1$. Note that $\lim_{n\to\infty} E(T) = n \log n$. Hence, in expectation it takes $n \log n$ number of steps, and the deck will be shuffled.

We modify the top to random shuffling in the following sense:

Definition 21. (α-top to random shuffling)

At each time step we take αn of top cards, shuffle them, and place them randomly inside the deck.

Lemma 28. *The expected number of steps by which a deck of n cards mixes by the α-top to random shuffle is:*

$$\frac{H_n - 1}{H_{\alpha n}}.$$

In this formula H_n is the nth harmonic number. [9]

Proof. Let $T_{\alpha n}$ be the time that it takes to see αn distinct cards when we perform the top the random shuffle. Following the analysis in [72], $\mathbb{E}(T_i) = 1 + 1/2 + \cdots + 1/(\alpha n)$.

Let T_i be the time by which we have seen $i\alpha n$ distinct cards when performing the top to random shuffle; and τ be the number of steps by which we will see card n in the α-top to random shuffle. Since $\mathbb{E}(\tau)$ and $\mathbb{E}(T_i)$s are all finite, and also our stopping rule does not see the future, we employ Wald's equality to derive:

$$E(T) = E(T_1 + T_2 + \cdots + T_\tau) = E(\tau)E(T') \Rightarrow E(\tau) = E(T)/E(T'). \quad (1.43)$$

Plugging in the values will give the result. $\qquad \square$

Note that this problem is equivalent to a generalized version of coupon collector problem when in each step the collector takes αn distinct coupons.

---

[9] $H_n = \sum_{i=1}^{n} 1/i$.

Chapter 2

Homomesy; Expected Value of Functions

"The orbit upon which mortals come and go,
Has no end nor beginning that we know,
And none there is to tell us in plain truth,
Whence do we come and whither do we go."
–Omar Khayyám.

2.1   Introduction and Motivation

In the previous chapter we studied various Markov chains and their mixing times. We observed that in some cases by making a bundle of steps and merging them into a single step, we can reduce the mixing time significantly.

76

e.g. Håstad's square shuffle and the generalization we presented.

Like Håstad's square lattice shuffle, the idea of "reducing randomness" and sometimes even "de-randomization" appear in various places, when studying computational models. **Systematic sweeping** and **Propp's rotor-router** model, are two examples of such models. In this introduction, we present a brief introduction of these two models, then we introduce the problem of our interest which is calculating the expected value of functions mapping a set of combinatorial objects to real numbers, given an arbitrary probability distribution.

Systematic Sweeping. Remember that in the Gibbs sampler model (Section 1.1), we have a Markov chain on a state space that has the form S^V, S being the set of values attained by elements of V. As we discussed, at state X_t of the chain the random single site update model picks a $v \in V$ uniformly at random, and it updates the value of v with respect to the conditional distribution π where any other $u \in V$ has the same value as in X_t. We update the value of v and keep the values of all the other elements of V except v the same as what they were at X_t to reach X_{t+1}.

Let $|V| = n$, note that after n random single site updates, the probability that there is at least one $v \in V$ not updated equals $1 - \frac{n!}{n^n} \simeq 1 - \frac{\sqrt{2\pi n}}{e^n}$. To avoid this situation, the systematic sweeping Gibbs sampler is suggested. In systematic sweeping, we fix an ordering of V, say $v_1, v_2, \ldots, v_n$. Instead of picking $v \in V$ uniformly at random and update it, at time t we pick v_i, i being the reminder of t divided by n, and we update it.

The Rotor-Router Model. The Rotor-Router is a de-randomized random walk coined by Jim Propp [26]. Recall that in a random walk on a graph $G = \langle V, E \rangle$ we start at some arbitrary vertex $v \in V$ and we go from v to u w.p. $p_{u,v}$ associated to the edge $(v, u) \in E$. In the Rotor-Router model, we have a chip that can be on any vertex $v \in V$. On each vertex $v \in V$ of the graph we have a number $r(v); 1 \leq r(v) \leq \deg(v)$, the numbers $1, 2, \ldots, \deg(v)$ correspond to the vertices adjacent to v, and they indicate where the chip should go if it is on edge v. At each time step t, we have a configuration consisting of the graph G, a rotor on each vertex, and a vertex v which indicates the position of the chip. At time step $t + 1$, the position of the chip will be upgraded to $r(v)$, the rotor on v will be upgraded to $r(v) + 1$ mod $\deg(v)$, and everything else will remain the same.

The rotor-router model is a deterministic automaton, and it has the property that some of the interesting statistics of it (e.g. hitting times, frequency of visits) are concentrated around the expected value of those statistics in an equivalent random walk by a discrepancy of order $1/n$. Note that by running the equivalent random walks n times, the discrepancy of an statistic from its expected value is bounded by $1/\sqrt{n}$ by the central limit theorem. (See Appendix [26] for more details.) Rotor-router models have been studied for graphs having different topologies, different initial rotor assignments, and various statistics of them have been compared to their analogous random walks. e.g. [27, 28, 29].

In this chapter our goal is to address the problem of calculating the ex-

pected value of a function $f : S \to \mathbb{R}$, S being a set of combinatorial objects, when given a probability distribution π on S. We propose a solution for this problem which is running a Markov chain on S and taking k samples from π. We identify the conditions under which this solution is efficient and present two examples in which running a Markov chain fails to serve our purpose. Thus, showing that calculating the average of $f : S \to \mathbb{R}$ cannot be always solved by taking samples from S.

We suggest that other models of computation should be taken into consideration. With the aforementioned motivation we study homomesy which is a term coined by Jim Propp and Tom Roby [7]. We study homomesy in a class of simply described lattices. Unfortunately, at this point homomesy does not provide us any known computational application, and this problem remains open.

2.2 Estimating the Expected Value of a Function Using Sampling

Consider a function $f : S \rightarrow \mathbb{R}$ and a probability distribution π on S. We are interested in finding the expected value of f, $\tilde{f} = \sum_{s \in S} f(s)\pi(s)$. In the presence of fast algorithms for sampling from S according to π, one would think of taking m samples $s_1, s_2, \ldots s_m$ and approximating $\tilde{f}$ by calculating $\sum_{i=1}^{m} s_i/m$. In this section, we investigate the efficiency of this approach. We obtain sufficient conditions under which this approach can be fruitful and bring interesting examples of #P-hard problems where either a proposed Markov chain is not rapidly mixing, or despite the existence of a rapidly mixing Markov chain, finding an approximation to the expected value of f is still an open problem.

Theorem 29. *Consider a Markov chain on set S with stationary distribution π, transition probability P, and mixing time τ Let $f : S \rightarrow \mathbb{R}$ be a real valued function with variance v and $\tilde{f} = \sum_{s \in S} f(s)\pi(s)$ the expected value of f. We can obtain an ϵ approximation for $\tilde{f}$ that satisfies Equation 2.37 by taking $m \geq 8v/\tilde{f}^2\epsilon^3$ samples $s_1, s_2, \ldots s_m$ from S and calculating $\hat{f} = 1/m \sum_{i=1}^{m} s_i$. The samples s_i are taken from S, by running the Markov chain, for at least $2\tau \log(\frac{f_{\max}}{\tilde{f}}\epsilon^{-1/2})$ number of steps. Then,*

$$Pr(|\frac{\hat{f}}{\tilde{f}} - 1| \leq \epsilon) \geq 1 - \epsilon \qquad (2.1)$$

Proof. Let $m \geq 8v/\tilde{f}^2\epsilon^3$. Fix an arbitrary T, $s_0 \in S$ and let π' be the following probability distribution on S: $\pi'(x) = P^T(s_0, x)$. We will have the following equations:

$$\mathbb{E}_{\pi'}[\hat{f}] = M\mathbb{E}_{\pi'}[f]/M = \mathbb{E}_{\pi'}[f]$$

$$\mathbb{E}_{\pi'}[f] = \sum_{s \in S} f(s)\pi'(s).$$

$$\mathbb{V}ar_{\pi'}[\hat{f}] = \sum_i \mathbb{V}ar_{\pi'}[f(s_i)/M] = \left(\sum_i \mathbb{V}ar_{\pi'}[f(s_i')]\right)/M^2 = \mathbb{V}ar_{\pi'}[f]/M$$

$$\mathbb{V}ar_{\pi'}[f] = \sum_{s \in S}(f(s) - \mathbb{E}_{\pi'}[f])^2\pi'(s).$$

By Chebyshev's inequality we have, $Pr(|\hat{f} - \mathbb{E}[\hat{f}]| \leq k\sigma) \geq 1 - \frac{1}{k^2}$ where $\sigma = \sqrt{\mathbb{V}ar[\hat{f}]}$. Hence, setting $k = \frac{\epsilon\tilde{f}}{2\sigma}$ we have,

$$
\begin{aligned}
& Pr(|\hat{f} - \mathbb{E}[\hat{f}]| \leq k\sigma) \geq 1 - \frac{1}{k^2} \\
=\ & Pr(|\hat{f} - \mathbb{E}[\hat{f}]| \leq \tilde{f}\epsilon/2) \geq 1 - \left(\frac{2\sqrt{\mathbb{V}ar[\hat{f}]}}{\tilde{f}\epsilon}\right)^2 \\
=\ & Pr(|\hat{f} - \mathbb{E}[\hat{f}]| \leq \tilde{f}\epsilon/2) \geq 1 - \left(\frac{2\sqrt{\mathbb{V}ar_{\pi'}[f]}}{\sqrt{M}\tilde{f}\epsilon}\right)^2 \\
=\ & 1 - \left(\frac{4\mathbb{V}ar_{\pi'}[f]}{M\tilde{f}^2\epsilon^2}\right).
\end{aligned}
\tag{2.2}
$$

We have, $T \geq 2\tau \log(\frac{4f_{\max}}{\tilde{f}}\epsilon^{-1/2})$. Therefore, $|\pi - \pi'|_{TV} \leq \left(\frac{\tilde{f}}{4f_{\max}}\right)^2\epsilon = \epsilon'$. Note that we have $\left(\frac{\tilde{f}}{4f_{\max}}\right)^2\epsilon = \epsilon'$ which implies $f_{\max}^2\epsilon' = \tilde{f}^2\epsilon/16$ and $f_{\max}\epsilon' \leq \tilde{f}\epsilon/16$.

Considering the following inequalities, we relate $\mathbb{V}ar_{\pi'}[f]$ to v and $\mathbb{E}_{\pi}[f]$ to $\tilde{f}$:

$$|\mathbb{E}_{\pi'}[f] - \tilde{f}| \leq \sum_{s \in S} f(s)|\pi(s) - \pi'(s)| \leq f_{\max}\epsilon' \leq \epsilon\tilde{f}/16. \qquad (2.3)$$

$$|\mathbb{E}_\pi[f^2] - \mathbb{E}_{\pi'}[f^2]| \leq \sum_{s \in S} f^2(s)|\pi(s) - \pi'(s)| \leq f_{\max}^2\epsilon'$$

$$|\mathbb{E}_{\pi'}^2[f] - \tilde{f}^2| = |(\mathbb{E}_{\pi'}[f] - \tilde{f})(\mathbb{E}_{\pi'}[f] + \tilde{f})| \leq$$
$$(f_{\max}\epsilon')(\tilde{f} + f_{\max}\epsilon' + \tilde{f}) = f_{\max}^2\epsilon'^2 + 2\tilde{f}f_{\max}\epsilon.$$

$$|\mathbb{V}ar'_\pi(f) - v| = |(\mathbb{E}_{\pi'}^2[f] - \tilde{f}^2) + (\mathbb{E}_\pi[f^2] - \mathbb{E}_{\pi'}[f^2])| \leq$$
$$|(\mathbb{E}_{\pi'}^2[f] - \tilde{f}^2)| + |(\mathbb{E}_\pi[f^2] - \mathbb{E}_{\pi'}[f^2])| \leq$$
$$f_{\max}^2\epsilon' + f_{\max}^2\epsilon'^2 + 2\tilde{f}f_{\max}\epsilon' = \qquad (2.4)$$
$$f_{\max}^2\epsilon'(1 + \epsilon' + 2\tfrac{\tilde{f}}{f_{\max}}) \leq 4f_{\max}^2\epsilon' =$$
$$\tilde{f}^2\epsilon/4.$$

Assuming $M \geq 8v/\tilde{f}^2\epsilon^3$, for $\epsilon \leq 4v/\tilde{f}^2$ we will have:

$$(4\mathbb{V}ar_{\pi'}[f])/(M\tilde{f}^2\epsilon^2) \leq$$
$$(4v + \tilde{f}^2\epsilon)/(M\tilde{f}^2\epsilon^2) \leq$$
$$4v/(M\tilde{f}^2\epsilon^2) + 1/(M\epsilon) \leq$$
$$\epsilon/2 + \epsilon/2 \leq \epsilon.$$

$$Pr(|\hat{f} - \mathbb{E}_{\pi'}[\hat{f}]| \leq \epsilon/2\tilde{f}) \geq 1 - (\tfrac{4\mathbb{V}ar_{\pi'}[f]}{M\tilde{f}^2\epsilon^2})$$
$$\geq 1 - \epsilon. \qquad (2.5)$$

Given that $M \geq 8v/\tilde{f}^2 \epsilon^3$ with a probability greater than $1 - \epsilon$ and using equations (2.3), (2.5) and triangle inequality we will have:

$$|\hat{f} - \tilde{f}| \leq |\hat{f} - \mathbb{E}_{\pi'}[\hat{f}]| + |\mathbb{E}_{\pi'}[\hat{f}] - \tilde{f}| \leq (\epsilon/2 + \epsilon/16)\tilde{f} \leq \epsilon\tilde{f}.$$
$$|\frac{\hat{f}}{\tilde{f}} - 1| \leq \epsilon. \tag{2.6}$$

□

2.2.1 Examples.

Example 11. Counting the number of Eulerian circuits in an Eulerian graph

Consider an Eulerian graph $\mathcal{G} = \langle V, E \rangle$, a connected graph in which all vertices have even degrees. We are interested in finding the number of Eulerian circuits in $\mathcal{G}$.

An Eulerian trail of a graph is a trail which traverses trough all edges once and only once. An Eulerian circuit is an Eulerian trail which starts and ends on the same vertex. We denote the set of Eulerian circuits of $\mathcal{G}$ by $\mathcal{E}u(\mathcal{G})$.

If $\mathcal{G}$ is a directed graph the exact value of $|\mathcal{E}u|$ can be calculated using BEST theorem [12], named after de Bruijn, van Aardenne-Ehrenfest, Smith and Tutte. However, not only finding the exact value of $|\mathcal{E}u|$ for an undirected graph is is known to be #P-hard by a result of Brightwell and Winkler [13],

but also approximating this value for an undirected graph in general case is still an open problem.

Finding Eulerian trails in a graph has broad applications in other fields of science such as reconstruction of DNA sequences [17] and CMOS circuit design [18]. Many Computer Scientists have been intrigued by the problem of sampling from $\mathcal{E}u(\mathcal{G})$ or counting it when $\mathcal{G}$ is undirected. However, they only succeeded to find polynomial time algorithms for this problem in restricted cases [14, 16, 15]. The problem is still open for an arbitrary, undirected graph $\mathcal{G}$.

Despite the lack of algorithms for approximating $|\mathcal{E}u(\mathcal{G})|$, Mihail and Winkler designed a rapidly mixing Markov chain to sample from the set of all Eulerian orientations of a graph. In the light of this result, one would like to define f as the following and apply Theorem 29 to approximate $\mathbb{E}(|\mathcal{E}u|)$.

Definition 22. For an arbitrary undirected graph $\mathcal{G}$, let $\mathcal{O}(\mathcal{G})$ be the set of all of its Eulerian orientations. We define $f : \mathcal{O}(\mathcal{G}) \to \mathbb{N}$ to be the function that to each orientation in $\mathcal{O}$ assigns the number of its circuits.

Unfortunately, considering the following example, we observe that applying Theorem 29, requires taking exponential number of samples, i.e. exponential runs of the chain.

Consider the multigraph G_n with vertex set $V = \{v_1, v_2, \ldots, v_n\}$ and edge set $E = \{e_{i1}, e_{i2} = \{v_i, v_{i+1}(\text{mod } n)\}, \quad 1 \leq i \leq n\}$.

There are two classes of Eulerian orientations for this graph. We call these two classes $\mathcal{O}_1$ and $\mathcal{O}_2$. in $\mathcal{O}_1$ we direct the edges of G_n in the following

way: for all i both e_{i1} and e_{i2} are directed from v_i to v_{i+1}; or, for all i both e_{i1} and e_{i2} are directed from v_{i+1} to v_i. This class has two orientations in it.

In the second class, $\mathcal{O}_2$ for each i we direct e_{i1} from v_i to v_{i+1} and e_{i2} from v_{i+1} to v_i; or, e_{i1} from v_{i+1} to v_i and e_{i2} from i to $i+1$. This class has 2^n orientations in it.

The number of Eulerian circuits associated to each orientation in $\mathcal{O}_1$ is 2^n: Choose a sequence $a_1, a_2, \dots a_n$, $a_i \in \{1, 2\}$ and for each i let b_i be the other element in $\{1, 2\}$ that is not a_i. The associated circuit will be the following:

$$e_{1,a_1}, e_{2,a_2}, \dots e_{n,a_n}, e_{1,b_1}, e_{2,b_2}, \dots e_{n,b_n}.$$

The number of Eulerian circuits associated to each orientation in $\mathcal{O}_2$ is $n+1$: One circuit is $e_{1,1}, e_{2,1}, \dots e_{n,1} e_{1,2}, e_{2,2}, \dots e_{n,2}$. For each j we have the following circuit:

$$e_{j,1}, e_{j,2}, e_{j+1,2}, e_{j+2,2} \dots e_{j-1,2}, e_{j-1,1}, e_{j-2,1}, \dots, e_{j+1,1}.$$

Let $\mathcal{O}$ be the set of orientations. We know $\mathcal{O} = \mathcal{O}_1 \cup \mathcal{O}_2$ and $|\mathcal{O}| = 2^n + n + 1$. Assuming that we have the uniform distribution on the set of orientations, we will have,

$$\tilde{f} = \frac{1}{2^n + 2}(2 \times 2^n + 2^n \times (n+1)) = \frac{2^{n-1}(n+2)}{2^{n-1}+1} = O(n).$$

$$\mathbb{V}ar[f] = \frac{2(2^n - \tilde{f})^2}{2^n + 2} + \frac{2^n(n+1-\tilde{f})^2}{2^n + 2} = O(2^n).$$

Theorem 29 requires $O(2^n/\epsilon^3)$ number of samples to estimate $\tilde{f}$ which is not efficient.

The following calculations show that it is impossible to use any algorithm sampling from $\mathcal{O}(\mathcal{G})$ to obtain a sturdy approximation for $\mathbb{E}[f]$ as defined in Definition 22.

Assume that we take p samples from $\mathcal{O}(\mathcal{G})$. With probability $(1 - \frac{1}{2^{n-1}})^p$ there is no orientation of class $\mathcal{O}_1$ among our samples. If p is polynomial, this number is very close to 1.

The approximation we obtain will be $\hat{f} = n+1$. Hence,

$$|\frac{\hat{f}}{\tilde{f}} - 1| = |\frac{(n+1)2^{n-1} + 1}{2^{n-1}(n+2)} - 1| \geq |\frac{n+1}{n+2} - 1| \geq \frac{1}{n+1}.$$

Therefore, for any $\epsilon \leq 1/(n+1)$ our method is not capable of giving us a good approximation.

Example 12. The average of real valued functions on the set of ideals in a partially ordered set.

Recall that by a partially ordered set (or a poset in abbreviation), we mean a set $\mathcal{P}$ together with a reflexive, antisymmetric, and transitive relation $\leq_{\mathcal{P}}: S \to S$. Remember that for any $x, y \in \mathcal{P}$, by $x \leq_{\mathcal{P}} y$ we mean $(x, y) \in \leq_{\mathcal{P}}$. We define an ideal in $\mathcal{P}$ as follows:

Definition 23. An order ideal I in $\mathcal{P}$ is a subset of it which satisfies the following: $\forall p \in I$ and $q \in \mathcal{P}; q \leq_{\mathcal{P}} p \implies q \in I$. The set of all order ideals in $\mathcal{P}$ is denoted by $J(\mathcal{P})$. Order ideals are also known as ideals or down sets.

Order ideals are ubiquitous in combinatorial optimization and operations research. For instance, they appear in dynamic programming algorithms [39],

86

constrained scheduling problems [40, 41], assembly line balancing problems [42], and in reliability networks [43].

A number of questions can be asked about $J(\mathcal{P})$, some natural questions being: Is it possible to enumerate $J(\mathcal{P})$, count it or take samples from it?

For arbitrary poset $\mathcal{P}$, the problem of counting (or equivalently sampling from) $J(\mathcal{P})$ is known to be a #P-hard problem [43]. The problem is tractable in some special cases, for instance, when $\mathcal{P}$ is 2-dimensional or series parallel [45, 46].

It is worth mentioning that $\langle J(\mathcal{P}), \subseteq \rangle$ constitutes a distributive lattice[1]. We study a natural single site update Gibbs sampler on $J(\mathcal{P})$. The update rule that we use is known as **toggling**. The term was coined by Striker and Williams [5] when they studied involution on sets of ideals of posets, and it appears without this name in many other places e.g. [85]. By saying we toggle an element $x \in \mathcal{P}$ in an ideal $I \in J(\mathcal{P})$ or out of it, we mean the following:

Definition 24. Consider a poset $\mathcal{P}$, an order ideal $I \in J(\mathcal{P})$, and an element $x \in \mathcal{P}$. We define the toggle in σ_x^+, toggle out σ_x^- and the toggle map σ_x as follows:

$$\sigma_x^+(I) = \begin{cases} I \cup \{x\} & \text{if } I \cup \{x\} \in J(\mathcal{P}). \\ I & \text{otherwise.} \end{cases} \tag{2.7}$$

---

[1] In fact every distributive lattice can be obtained this way i.e. the correspondence is one to one.

$$\sigma_x^-(I) = \begin{cases} I - \{x\} & \text{if } I - \{x\} \in J(\mathcal{P}). \\ I & \text{otherwise.} \end{cases} \tag{2.8}$$

$$\sigma_x(I) = \begin{cases} \sigma_x^+(I), & \text{if } x \notin I. \\ \sigma_x^-(I), & \text{if } x \in I. \end{cases} \tag{2.9}$$

Consider a poset $\mathcal{P}$. The following chain on $J(\mathcal{P})$ is a single site update Markov chain, and we denote it by $\mathcal{M}_\mathcal{P}$:

At state $I \in \mathcal{P}$,

1. Pick $p \in \mathcal{P}$ with probability $1/n$, n being the cardinality of $\mathcal{P}$.

2. Flip a biased coin. If heads, toggle p into I i.e let $I' = \sigma_x^+(I)$. If tails, toggle p out of I i.e let $I' = \sigma_x^-(I)$. Move to I'.

Despite the simplicity of the chain, it fails to mix rapidly except for a very small class of posets. It is shown by Wilson [52] that if P is a square lattice, this Markov chain is rapidly mixing. It was later proved by Greenberg et al. (2009) that the biased chain mixes rapidly when the poset is a $k-$dimensional cube lattices [85]. In Proposition 2, we present an example of a simply described poset, which we call the **butterfly** poset for which the suggested chain requires at least exponential number of steps until it becomes close to stationarity. Thus, calculating the average of any function $f : J(\mathcal{P}) \to \mathbb{R}$ using Theorem 29 will fail for the butterfly poset defined in Proposition 2.
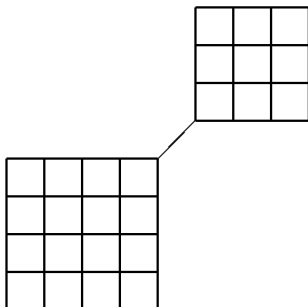
Figure 2.1: The butterfly poset when $n = 4$ and $m = 5$. The order is directed from north east to sour west meaning box j is less that box i if j is placed south west of i in the picture.

Definition 25. Let $\mathcal{Q}_{a,b} = [a] \times [b]$ ($[n] = \{1, 2, \dots, n\}$). Each element $p \in \mathcal{Q}_{a,b}$ can be presented by a pair $p = (i, j), i \in [a], j \in [b]$ and $(i_1, j_1) \leq (i_2, j_2)$ iff $i_1 \leq i_2$ and $j_1 \leq j_2$.

Proposition 2. *Consider the posets* $I = \mathcal{Q}_{n,n}$ *and* $J = \mathcal{Q}_{m,m}$, *take* x *to be the minimum element in* I *and* y *to be the maximum element in* J. *We let* $K = I \cup J$, *and* $\leq_K$ *is defined as follows: for any* i *and* j *in* K,

$$i \leq_K j \Leftrightarrow (i \leq_I j \ or \ i \leq_J j \ or \ (i = y \ and \ j = x)).$$

We call K *the butterfly poset. The conductance* Φ *of* $\mathcal{M}_K$, *satisfies:*

$$\Phi = \frac{2}{n^2 m^2} \bigg/ \binom{2n}{n} \binom{2m}{m} \sim \frac{2}{4^n 4^m ((nm)^3)(\sqrt{nm})}. \tag{2.10}$$

Hence, by Theorem 5 the mixing time of $\mathcal{M}_K$ *is at least* $\mathcal{O}(4^{m+n})$.

2.3 Some Instances of Homomesy Among Ideals of Posets

Consider a poset $\mathcal{P}$, and let $J(\mathcal{P})$ be the set containing all of the ideals in $\mathcal{P}$. The **rowmotion** operation, is an operation mapping $J(\mathcal{P})$ to itself, and it has been studied widely by combinatorists and under various names (Brouwer-Schrijver map [1], the Fon-der-Flaass map [6], the reverse map [4], and Panyushev complementation [3]). Rowmotion is defined as follows:

Definition 26. Given a set $\mathcal{P}$ with partial order $\leq_{\mathcal{P}}$, and an order ideal $I \in J(\mathcal{P})$, rowmotion is denoted by $\varrho(I)$, and it is defined to be the down set[2] of the minimal elements in $\mathcal{P}-I$.

Remark 2. Propp and Roby use Φ_J to denote rowmotion acting on order ideals and Φ_A for rowmotion acting on antichains. In this section, we discuss only actions on order ideals. Moreover, since we are using Φ for conductance, we denote rowmotion by ϱ.

Recall that another interesting operation mapping $J(\mathcal{P})$ to itself is the **toggle map** (Definition 24). We can define the rowmotion operation also as the combination of several toggles.

Proposition 3. *For all $x \in \mathcal{P}$ and $I \in J(\mathcal{P})$, $\sigma_x^2(I) = I$. If $x, y \in \mathcal{P}$ and x does not cover y nor y covers x, we have $\sigma_x \circ \sigma_y(I) = \sigma_y \circ \sigma_x(I)$.*

---

[2]In a poset $\mathcal{P}$ on elements of $\mathcal{P}$ the down set of a set $\mathcal{X} \subseteq \mathcal{P}$ is the following: $\{y \in \mathcal{P} | \exists x \in \mathcal{X}, y \leq x\}$.

We take a linear extension $(x_1, \ldots, x_n)$ of $\mathcal{P}$ to be an indexing of the elements of $\mathcal{P}$ that is $x_i < x_j$ in $\mathcal{P}$ implies $i < j$. The following proposition was demonstrated in [2].

Proposition 4. *[2] Given an arbitrary $I \in J(\mathcal{P})$ and linear extension $(x_1, \ldots, x_n)$ of $\mathcal{P}$, we have $\varrho(I) = \sigma_{x_1} \circ \sigma_{x_2} \circ \sigma_{x_3} \circ \cdots \circ \sigma_{x_n}(I)$.*

Remember the following definition from the previous section (Definition 25): Let $\mathcal{Q}_{a,b} = [a] \times [b]$ ($[n] = \{1, 2, \ldots, n\}$). Each element of the poset can be presented by a pair $(i, j), i \in [a], j \in [b]$ and $(i_1, j_1) \leq (i_2, j_2)$ iff $i_1 \leq i_2$ and $j_1 \leq j_2$.

In this paper, we are interested in the maps on $J(\mathcal{Q}_{a,b})$, as well as $J(\mathcal{U}_a)$ and $J(\mathcal{L}_a)$ where $\mathcal{U}_a$ and $\mathcal{L}_a$ are subsets of $\mathcal{Q}_{a,a}$ and defined as:

- $\mathcal{U}_a \subseteq \mathcal{Q}_{a,a}$, $\mathcal{U}_a = \{(i, j) | i, j \in [a], i \geq a+1-j\}$.

- $\mathcal{L}_a \subseteq \mathcal{Q}_{a,a}$, $\mathcal{L}_a = \{(i, j) | i, j \in [a], i \geq j\}$.

Notation. Let $\mathcal{P}$ be one of $\mathcal{Q}_{a,b}$, $\mathcal{U}_a$ or $\mathcal{L}_a$. By saying $(i, j) \in \mathcal{P}$ we are referring to the element in $[a] \times [b]$ with coordinates i and j. By saying $x = (i_1, j_1) \leq y = (i_2, j_2)$ we mean x is less than y in $\mathcal{P}$. To avoid confusion, we *never* use $(i, j) \in \mathcal{P}$ to indicate i is less than j in the partial order.

We call $\mathcal{Q}_{a,b}$ the square lattice or the product of two chains, $\mathcal{U}_a$ the upper lattice and $\mathcal{L}_a$ the left lattice. Among combinatorists $\mathcal{U}_a$ is also known as the root poset of type A_a, and $\mathcal{L}_a$ as the minuscule poset of type B_a or D_{a+1}. We employ the following terminology:
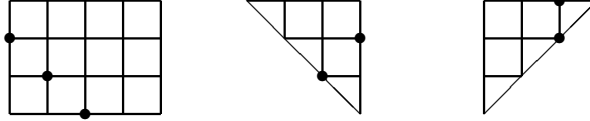
Figure 2.2: $\mathcal{Q}_{5,4}$ and $R_4(\mathcal{Q}_{5,4})$ $\mathcal{U}_4$ and $F_{-1}(\mathcal{U}_4)$ $\mathcal{L}_4$ and $C_3(\mathcal{L}_4)$

Definition 27. Let $\mathcal{P}$ be one of $\mathcal{Q}_{a,b}, \mathcal{U}_a$ or $\mathcal{L}_a$. For any arbitrary $I \in J(\mathcal{P})$,

We call the set of all points $(i,j) \in \mathcal{P}$ with constant $i+j$ a **rank**; $R_c(I) =$ $\{(i,j) \in I | i+j = c\}$.

We call the set of all points $(i,j) \in \mathcal{P}$ with constant $i-j$ a **file**; $F_c(I) =$ $\{(i,j) \in I | i-j = c\}$.

We call the sets of all points $(i,j) \in \mathcal{P}$ with constant i a **column**; $C_c(I) =$ $\{(i,j) \in I | i = c\}$.

In the case when no ideal is specified we have $R_c = R_c(\mathcal{P})$, $F_c = F_c(\mathcal{P})$ and $C_c = C_c(\mathcal{P})$; $\mathcal{P}$ should be clear from the context.

Example 13. The following figure shows $\mathcal{Q}_{5,4}$ and R_4 in it, $\mathcal{U}_4$ and F_{-1} in it and $\mathcal{L}_4$ and C_3 in it.

Remark 3. In Figure 13 and all the other figures in this section and Section 2.4, we draw the posets from south west to east coast, meaning if box i is located north east of box j, then i is greater than j.

We can now define toggling for the above sets.

Definition 28. Consider the poset $\mathcal{Q}_{a,b}$ and $I \in J(\mathcal{Q}_{a,b})$. Let S be one of R_c or F_c for some arbitrary c. Letting $x_1 \ldots x_m$ be some arbitrary indexing

of the elements of S, we define $\sigma_S(I) = \sigma_{x_1} \circ \sigma_{x_2} \circ \cdots \circ \sigma_{x_m}(I)$. Note that no two elements x_i, x_j of S constitute a covering pair, thus σ_S is well defined. For $S = C_c$, let $S = \{x_1, x_2, \ldots x_m\}$ where $x_1 < x_2 < \cdots < x_m$. We define $\sigma_S(I) = \sigma_{x_1} \circ \sigma_{x_2} \ldots \sigma_{x_m}(I)$.

Striker and Williams studied the class of so-called rc-posets, whose elements are partitioned into ranks and files[3]. Here, we will discuss the special rc-posets of the form $\mathcal{Q}_{a,b}$, $\mathcal{U}_a$ or $\mathcal{L}_a$. The following definitions are from [5], restricted to the product posets of interest to us.

Definition 29. [5] Consider $\mathcal{Q}_{a,b}$. Let ν be a permutation of $\{2, \ldots, a+b\}$. We define ϱ_ν to be $\sigma_{R_{\nu(a+b-1)}} \circ \sigma_{R_{\nu(a+b-2)}} \circ \cdots \sigma_{R_{\nu(1)}}$.

Having Proposition, 4 it can be concluded that the following holds:

for $\nu = (a+b, a+b-1 \ldots, 2)$, we have $\varrho_\nu = \varrho$.

Consider $\mathcal{Q}_{a,b}$, and ν a permutation of $\{2, \ldots, a+b\}$. Then, ϱ_ν is a permutation on $J(\mathcal{Q}_{a,b})$ that partitions $J(\mathcal{Q}_{a,b})$ into orbits. Striker and Williams showed that the orbit structure[4] of ϱ_ν does not depend on the choice of ν.

Definition 30. Consider $\mathcal{Q}_{a,b}$, **promotion** is a permutation $\partial : J(\mathcal{Q}_{a,b}) \to J(\mathcal{Q}_{a,b})$, defined by: $\forall I \in J(\mathcal{Q}_{a,b}), \partial(I) = \sigma_{F_{a-1}} \circ \sigma_{F_{a-2}} \circ \cdots \circ \sigma_{F_0} \circ \cdots \circ \sigma_{F_{1-b}}(I)$.

As with rowmotion, Striker and Williams [5] define a generalized version of promotion.

---

[3]Striker and Williams use the terminology "row" for what we call "rank" and "column" for what we call "file".

[4]The orbit structure of a bijection f on a set S is the multiset of the sizes of the orbits that bijection f constructs on the set S.

Definition 31. [5] Consider the lattice $\mathcal{Q}_{a,b}$, and let ν be a permutation of $\{-b+1,\ldots,a-1\}$. We define ∂_ν to be $\sigma_{F_{\nu(a+b-1)}} \circ \sigma_{F_{\nu(a+b-2)}} \circ \cdots \circ \sigma_{F_{\nu(1)}}$. By Definition 30, for $\nu = (-b+1,\ldots,a-1)$ we have $\partial_\nu = \partial$.

As with rowmotion, for any permutation ν on files of any poset $\mathcal{P}$, ∂_ν will partition $J(\mathcal{P})$ to orbits. Again, Striker and Williams [5] showed that regardless of which ν we choose, $J(\mathcal{Q}_{a,b})$ will be partitioned into the same orbit structure by ∂_ν. Moreover, the orbit structures for ∂_ν and for ϱ_ω are the same for any two permutations ν and ω:

Theorem 30. *[5] Consider the lattice $\mathcal{Q}_{a,b}$, for any arbitrary permutation ν on $\{2,\ldots,a+b\}$ and ω on $\{-b+1\ldots a-1\}$, there is an equivariant bijection between $J(\mathcal{Q}_{a,b})$ under ϱ_ν and $J(\mathcal{Q}_{a,b})$ under ∂_ω.*

Permutations defined on combinatorial structures and the associated orbit structures became more interesting after Propp and Roby introduced a phenomenon called **homomesy** [7]. Propp and Roby also discussed some instances of homomesy by studying the actions of promotion and rowmotion on the set $J(\mathcal{Q}_{a,b})$. Homomesy has attracted many combinatorics' attentions after it was defined and studied by Propp and Roby [8, 9, 10, 47, 19], and it is defined as follows:

Definition 32. [7] Consider a set S of combinatorial objects. Let $\tau : S \to S$ be a permutation that partitions S into orbits, and $f : S \to \mathbb{R}$ a statistic of the elements of S. We call the triple $\langle S, \tau, f \rangle$ **homomesic** (or we say it

exhibits homomesy) if and only if there is a constant c such that for any τ-orbit $\mathcal{O} \subset \mathcal{S}$ we have

$$\frac{1}{|\mathcal{O}|} \sum_{x \in \mathcal{O}} f(x) = c.$$

Equivalently, we can say f is homomesic or it exhibits homomesy in τ-orbits of S. If $c = 0$, the triple is called $0-$mesic.

Proposition 5. *Consider a set S and permutation $\tau : S \to S$. If $f_1, \ldots, f_n$ are homomesic functions in τ-orbits of S, then any linear combination of the f_is is also homomesic in τ-orbits of S.*

Theorem 31. *[7] Consider $f : J(\mathcal{Q}_{a,b}) \to \mathbb{R}$ defined as follows: for all $I \in \mathcal{Q}_{a,b}, f(I) = |I|$. Let $\partial, \varrho : J(\mathcal{Q}_{a,b}) \to J(\mathcal{Q}_{a,b})$ be the rowmotion and promotion operation. The triples $\langle J(\mathcal{Q}_{a,b}), \partial, f \rangle$ and $\langle J(\mathcal{Q}_{a,b}), \varrho, f \rangle$ exhibit homomesy.*

In this paper, we generalize Theorems 31 and 30 in the following sense:

Definition 33. Consider the poset $\mathcal{P}$ to be one of $\mathcal{Q}_{a,b}$, $\mathcal{U}_a$ or $\mathcal{L}_a$. For any permutation ν of $[a]$, we define the action **comotion**, $\mathcal{T}_\nu : J(\mathcal{P}) \to J(\mathcal{P})$ by:

$\forall I \in J(\mathcal{P}), \mathcal{T}_\nu(I) = \sigma_{C_{\nu(a)}} \circ \sigma_{C_{\nu(a-1)}} \circ \cdots \circ \sigma_{C_{\nu(1)}}(I).$

The following proposition can be proved by applying Proposition 3 inductively.

Proposition 6. *Let $\mathcal{P}$ be one of $\mathcal{Q}_{a,b}$, $\mathcal{U}_a$ and $\mathcal{L}_a$, the action of promotion coincides with $\mathcal{T}_{(a,a-1,\ldots 1)}$ and rowmotion coincides with $\mathcal{T}_{(1,2,\ldots a)}$.*

In what follows, Theorems 32, 33, 34 which are the main results of this paper will be stated. We will provide a roadmap to their proofs later in this introduction, and will complete the proof in Sections 2.3.1 and 2.3.2.

Theorem 32. *(Homomesy in $J(\mathcal{Q}_{a,b})$)*

1. For any arbitrary natural number a and ν a permutation on $[a]$, $\mathcal{T}_\nu$ partitions $J(\mathcal{Q}_{a,b})$ to orbits. The orbit structures of $\mathcal{T}_\nu$ on $J(\mathcal{Q}_{a,b})$ is independent of choice of ν.

2. Consider $I \in J(\mathcal{Q}_{a,b})$. We have the following homomesy results:

   - Let $g_{i,j}$, $1 \leq i \leq a$ and $1 \leq j \leq b$ be defined as follows:

$$g_{i,j} = \begin{cases} 1, & \text{if } |C_i(I)| = j \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

   For an arbitrary permutation ν of $[a]$, $1 \leq i \leq a$ and $0 \leq j \leq b$, the function $d_{i,j} = g_{i,j} - g_{a+1-i,b-j}$ is 0-mesic in $\mathcal{T}_\nu$-orbits of $J(\mathcal{Q}_{a,b})$.

   - For all $1 \leq i \leq a$, let

$$s_{i,j} = \begin{cases} 1 & \text{if } |C_i(I)| + i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

   For any arbitrary permutation ν of $[a]$ and $1 \leq j \leq b$, $s_j = \sum_{i=1}^{a} s_{i,j}$ is homomesic in $\mathcal{T}_\nu$-orbits of $J(\mathcal{Q}_{a,b})$. Moreover, the av-

96

erage of all s_j along an orbit is constant and equal to $\frac{a}{a+b}$. In other words, for all j, l, $s_l - s_j$ is 0-mesic.

Any function $f : J(\mathcal{Q}) \to \mathbb{R}$ which is a linear combination of various s_i and d_i is homomesic in $\mathcal{T}_\nu$-orbits of $J(\mathcal{Q}_{a,b})$.

Theorem 32 introduces a different family of permutations that produce the same orbit structure as ϱ and ∂; hence, it generalizes Theorem 30. It also generalizes Theorem 31 because it introduces a class of permutations and statistics whose triple with $J(\mathcal{Q}_{a,b})$ exhibit homomesy. Moreover, it will provide another proof for Theorem 31. The main idea of our proof is the correspondence between comotion and **winching** (See Definition 34). We will define winching and also its correspondence with comotion in Section 2.3.1. Then, we extend the definition of winching to winching with lower bounds and winching with zeros. Studying these two variations helps us obtain homomesy results in $J(\mathcal{U}_a)$ and $J(\mathcal{L}_a)$.

Theorem 33. *(Homomesy in $J(\mathcal{U}_a)$)*

Let a be an arbitrary natural number and ν an arbitrary permutation of $[a]$. Consider $\mathcal{T}_\nu : J(\mathcal{U}_a) \to J(\mathcal{U}_a)$ as defined in Definition 33. For each $i \in [2a]$ let $[i, 2a] = i, i+1, \ldots, 2a$ and $f : [2a] \to \mathbb{R}$ a function that has the same average in all $[i, 2a]$ where i is odd. Let $g : J(\mathcal{U}_a) \to \mathbb{R}$ be defined as: $\forall I \in J(\mathcal{U}_a), g(I) = \sum_{i=1}^{a} f(|C_i(I)| + 2i + 1)$. Then, the triple $\langle J(\mathcal{U}_a), \mathcal{T}_\nu, g \rangle$ exhibits homomesy.

Theorem 34. *(Homomesy in $J(\mathcal{L}_a)$)*

97

Let a be an arbitrary natural number and ν an arbitrary permutation of $[a]$ and $\mathcal{T}_\nu : J(\mathcal{L}_a) \to J(\mathcal{L}_a)$ be defined as in Definition 33. We have,

1. The orbit structures of $\mathcal{T}_\nu$ on $J(\mathcal{L}_{a,b})$ is independent from choice of ν.

2. For any $1 \le i \le a$ and $0 \le j \le a$ we define $s_{i,j} : J(\mathcal{L}_a) \to \mathbb{R}$ as follows:

$$s_{i,j} = \begin{cases} 1 & \text{if } |C_i| = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

For any $1 \le j \le a$ $s_j = \sum_{i=1}^{a} s_{i,j}$ is homomesic. Moreover, the average of all s_j along any $\mathcal{T}_\nu$-orbit of $J(\mathcal{L}_{a,b})$ is the same. In other words, for all j, l $s_l - s_j$ is 0-mesic.

Moreover, any function $f : J(\mathcal{L}_a) \to \mathbb{R}$ which is a linear combination of various s_i is homomesic in $\mathcal{T}_\nu$-orbits of $J(\mathcal{L}_a)$.

In Section 2.3.1 of this paper we introduce the permutation winching on the set of increasing sequences of length k. We show that there is a natural equivariant bijection between the set of ideals under comotion and the set of increasing sequences under winching.

Then, we introduce two different variations of winching and their correspondence with comotion in $J(\mathcal{U}_a)$ and $J(\mathcal{L}_a)$.

In Section 2.3.2 we will use the Theorems 32, 33 and 34 to show homomesy of some functions in the orbit structure produced by comotion in $J(\mathcal{Q}_{a,b})$, $J(\mathcal{U}_a)$ and $J(\mathcal{L}_a)$.

In Section 2.3.3 we will prove homomesy of a class of statistics when the permutation is winching and two different variations of it. These results have intrinsic interest because they are instances of homomesy. Moreover, by the correspondence between winching and comotion proof of Theorems 32, 33, and 34 will be obtained.

2.3.1 Comotion, winching and their correspondence

In the previous section, we defined the action of comotion on the set of order ideals of a poset. In this section, we define winching and show a correspondence between winching on increasing sequences and comotion on $J(\mathcal{Q}_{a,b})$. Then, we define winching with lower bounds and winching with zeros. The former corresponds to comotion on $J(\mathcal{U}_a)$ and the later corresponds to comotion on $J(\mathcal{L}_a)$.

Definition 34. Let $S_{k,m}$ be the set of all k-tuples $x = (x_1, \ldots, x_k)$ satisfying $0 < x_1 < x_2 < \cdots < x_k < m+1$. We define the map $W_i : S_{k,m} \to S_{k,m}$, called winching on index i, by $W_i(x) = y = (y_1, y_2, \ldots, y_k)$, where $y_j = x_j$ for $i \neq j$, and

$$
y_i = \begin{cases} x_i + 1, & \text{if } x_i + 1 < x_{i+1}. \\ x_{i-1} + 1, & \text{otherwise.} \end{cases}
\tag{2.14}
$$

We assume that always $x_0 = 0$ and $x_{k+1} = m+1$.

Example 14. Let $\nu = (2, 3, 1, 4)$ and $x \in S_{4,7}$ be $x = (2, 3, 5, 7)$. Then,

99

$W_\nu(x) = (1, 4, 6, 7)$.

Lemma 35. *There is a bijection $\alpha : J(\mathcal{Q}_{a,b}) \to S_{a,a+b}$ such that for any $I \in J(\mathcal{Q}_{a,b})$, $\alpha(\sigma_{C_j}(I)) = W_j(\alpha(I))$.*

Proof. Consider $I \in J(\mathcal{Q}_{a,b})$, we define $\alpha(I) = (\alpha_1, \ldots \alpha_a)$ as follows: for any $1 \le j \le a$, we have $\alpha_j(I) = |C_{a+1-j}(I)| + j$. Since $I \in J(\mathcal{Q}_{a,b})$, for any $j_1 < j_2$, take $|C_{j_1}(I)| \ge |C_{j_2}(I)|$. Therefore, $\alpha(I)$ is an increasing sequence.

Let C_j be $\{v_1, v_2, \ldots v_b\}$; $v_i = (j, i)$, and assume $|C_j(I)| = l$. We have, $n > l+1$, $\sigma_{v_n}(I) = I$, and for $n = l+1, \sigma_{v_n}(I) = I \cup \{v_n\}$ if and only if $|C_{j-1}| \ge l+1$. Also, $n < l, \sigma_{v_n}(I) = I - \{v_n\}$ if and only if $|C_{j+1}(I)| \le n-1$. For boundary cases, we assume $|C_0| = b$ and $|C_b| = 0$. Letting $K = \sigma_{C_j}(I)$ we will have,

$$
C_j(K) = \begin{cases} C_j(I) \cup \{v_{l+1}\}, & \text{if } |C_{j-1}(I)| \ge l+1. \\ C_j(I) - \{v_l, v_{l-1}, \ldots, v_{p+1}\}(p = |C_{j+1}(I)|), & \text{otherwise.} \end{cases}
$$

$$(2.15)$$

$$
\Leftrightarrow |C_j(K)| + a+1-j = \begin{cases} l+1 + a+1-j, & \\ \quad \text{if } |C_{j-1}(I)| + a-j+2 \ge l+1 + a-j+2. \\ |C_{j+1}(I)| + a+1-j, & \text{otherwise.} \end{cases}
$$

$$(2.16)$$

$$\Leftrightarrow \alpha_{a+1-j}(\sigma_{C_j}(I)) = \begin{cases} \alpha_{a-j+1}(I)+1, & \text{if } \alpha_{a-j+2}(I) > \alpha_{a-j+1}(I) + 1. \\ \\ \alpha_{a-j}(I) + 1, & \text{otherwise.} \end{cases}$$

$$(2.17)$$

$$\Leftrightarrow \alpha_{a+1-j}(\sigma_{C_j}(I)) = W_{a+1-j}(\alpha(I)). \qquad (2.18)$$

$\square$

Corollary 36. *Consider an arbitrary natural number a, ν a permutation of $[a]$, and for any $x \in S_{a,a+b}$, let $W_\nu(x) = W_{\nu(a)} \circ W_{\nu(a-1)} \circ \cdots \circ W_{\nu(1)}(x)$. The bijection α introduced in Definition 35 satisfies the following property: $\alpha(\mathcal{T}_\nu(I)) = W_\nu(\alpha(I))$.*

Theorem 37. *Consider a natural number k and an arbitrary permutation ν of $[k]$. With $W_\nu : S_{k,m} \to S_{k,m}$ defined as above we will have,*

1. *$W_\nu^m(x) = x$ for all $x \in S_{k,m}$.*

2. *The orbit structure that winching produces on the set $S_{k,m}$ is the same as the orbit structure for rotation acting on the set of 2-colored necklaces with k white beads and $n - k$ black beads, and hence independent of choice of ν.[5] (The orbit structure of necklaces if a classical problem in Combinatorics and the solution is a result of applying Pólya's Theorem [11].)*

---

[5]The definition of rotation acting on the set of 2-colored necklaces is presented in Section 2.3.3 (Definition 43).

3. *The following functions (and any linear combination of them) are homomesic in W_ν-orbits of $S_{k,m}$.*

- *Let $g_{i,j} : S_{k,m} \to \mathbb{R}, 1 \le i \le k$ and $1 \le j \le m$ be defined as follows:*

$$g_{i,j}(x) = \begin{cases} 1, & \text{if } x_i = j \\ 0, & \text{otherwise.} \end{cases} \tag{2.19}$$

For any arbitrary $1 \le i \le k$ and $1 \le j \le m$, the function $d_{i,j} = g_{i,j} - g_{k+1-i,m+1-j}$ is 0-mesic in W_ν-orbits of $S_{k,m}$.

- *For an arbitrary $1 \le j \le m$, let $f_j : S_{k,m} \to \mathbb{R}$ be defined by:*

$$f_j(x) = \begin{cases} 1, & \text{if } j \in x \\ 0, & \text{otherwise.} \end{cases} \tag{2.20}$$

For any $1 \le j \le m$, the triple $\langle S_{k,m}, W_\nu, f_j \rangle$ is homomesic and the average of f_j along W_ν orbits is k/m.

We will prove the above theorem in the next section. Given the bijection in Corrollary 36, Theorem 32 is a straightforward conclusion of Theorem 37. In addition, Theorem 31 can be concluded from the above theorem. In fact, a more general statement is shown in the next section (Corollary 44).

The following variation of winching is called **winching with lower bounds** and it corresponds to comotion on $J(\mathcal{U}_a)$.

Definition 35. Consider the sequence of lower bounds $l = (l_1, \ldots, l_k)$, $0 <$ $l_1 < \cdots < l_k < m+1$ and $S'_{k,m} = \{(x_1, x_2, \ldots x_k) \in S_{k,m} | x_i \geq l_i\}$, where $S_{k,m}$ is the set defined in Definition 34. For any index $i \in [k]$, we define the map $\underline{W}_i : S'_{k,m} \to S'_{k,m}$ called winching with lower bounds l on index i by

$$\forall w \in S'_{k,m} \quad \underline{W}_i(w) = \max\{W_i(w), l_i\},$$

where W_i is the action of winching on index i (Definition 34). Having $\mathcal{U}_a$ be the poset which is defined in Definition 25, we will have:

Lemma 38. *There is a bijection $\beta : J(\mathcal{U}_a) \to S'_{a,2a}$ such that for the lower bounds $l = (1, 3, 5, \ldots, 2a-1)$, we have: for any $I \in J(\mathcal{U}_a)$, $\beta(\sigma_{C_j}(I)) = \underline{W}_j(\beta(I))$.*

Proof. Fix arbitrary a and consider $I \in J(\mathcal{U}_a)$, we define $\beta(I) = (\beta_1, \ldots, \beta_a)$ as follows: for any $1 \leq j \leq a$, $\beta_j(I) = |C_{a+1-j}(I)| + 2j-1$. Considering the ideal $I' \in J(\mathcal{Q}_a)$, $I' = I \cup (\mathcal{Q}_{a,a} - \mathcal{U}_a)$, we will have, $\beta(I) = \alpha(I')$. Hence, β is an increasing sequence. Since $\sigma_{C_j}(I) = \sigma_{C_j}(I') - (\mathcal{Q}_{a,a} - \mathcal{U}_a)$ we have,

$$\beta_j(\sigma_{C_{a+j-1}}(I)) = |\sigma_{C_{a+j-1}}(I)| + 2j - 1 = |\sigma_{C_{a+j-1}}(I') - (\mathcal{Q}_{a,a} - \mathcal{U}_a)| + 2j - 1$$

$$(2.21)$$

$$\Rightarrow \beta_j(\sigma_{C_{a+j-1}}(I)) \quad = \max\{|\sigma_{C_{a+j-1}}(I')| - j + 1, 0\} + 2j - 1$$

$$= \max\{|\sigma_{C_{a+j-1}}(I')| + j, 2j - 1\}$$

$$(2.22)$$

$$\Rightarrow \beta_j(\sigma_{C_{a+j-1}}(I)) = \max\{(W_j(\alpha(I')))_j, 2j-1\} = \max\{(W_j(\beta(I)))_j, 2j-1\}.$$

$$(2.23)$$

$\square$

Corollary 39. *Consider any arbitrary permutation $\nu : [a] \to [a]$, the action $\mathcal{T}_\nu : J(\mathcal{U}_a) \to J(\mathcal{U}_a)$ and $I \in J(\mathcal{U}_a)$. For any $x \in S_{a,2a}$, let the lower bounds be $l = (1, 3, \ldots, 2a-1)$. Then: $\underline{W}_\nu(x) = \underline{W}_{\nu(a)} \circ \underline{W}_{\nu(a-1)} \circ \cdots \circ \underline{W}_{\nu(1)}(x)$. Bijection β introduced in 38 satisfies the following property: $\beta(\mathcal{T}_\nu(I)) = \underline{W}_\nu(\beta(I))$.*

Theorem 40. *Let ν be an arbitrary permutation of $[a]$. Consider $\underline{W}_\nu : S'_{a,b} \to S'_{a,b}$ with lower bounds $(l_1, l_2, \ldots, l_a)$. For each $i \in [a+b]$ let $[i, a+b] = i, i+1 \ldots a+b$ and $f : [a+b] \to \mathbb{R}$ a function that has the same average in all $[l_i, a+b]$, $1 \le i \le a$. Let $g : S'_{a,b} \to \mathbb{R}$ be defined as, $g(x) = \sum_{i=1}^{a} f(x_i)$. Then, the triple $\langle S'_{a,b}, \underline{W}_\nu, g \rangle$ exhibits homomesy.*

We now define the action of **winching with zeros** to study homomesy in $J(\mathcal{L}_a)$.

Definition 36. Let S_n be the set of all n-tuples $x = (x_1, \ldots, x_n)$ such that for some $0 \leq k \leq n$ $x_1 = x_2 \cdots = x_k = 0$ and $1 \leq x_{k+1} < x_{k+2} \cdots < x_n \leq n$. We define the map $\mathrm{WZ}_i : S_n \to S_n$, called winching with zeros on index i to be

$$
\mathrm{WZ}_i(x) = \begin{cases} x_i+1 & \text{if } x_i+1 < \min\{x_{i+1}, n+1\}; \\ x_{i-1}+1 & \text{if } 1 < i \text{ and } 0 < x_{i-1}; \\ 0 & \text{otherwise.} \end{cases}
$$

Lemma 41. *There is a bijection* $\gamma : J(\mathcal{L}_a) \to S_a$ *such that: for any* $I \in J(\mathcal{L}_a)$, $\gamma(\sigma_{C_j}(I)) = \mathrm{WZ}_j(\gamma(I))$.

Proof. Fix an arbitrary natural number a and consider $I \in J(\mathcal{L}_a)$. We define, $\gamma(I) = (\gamma_1, \gamma_2, \ldots, \gamma_a)$ as follows: for $1 \leq j \leq a$, $\gamma_j(I) = |C_{a+1-j}(I)|$. For any $j_1 < j_2$, we have $|C_{j_1}(I)| > |C_{j_2}(I)|$. Hence, γ will be an increasing sequence.

Let $C_j = \{v_j, v_{j+1}, \ldots v_a\}$ where for $j \leq i \leq a$, $v_i = (j, i)$. Assume $|C_j(I)| = l$, which means $C_j(I) = \{v_j, v_{j+1}, \ldots v_{j+l-1}\}$. For $n > j + l$, $\sigma_{v_n}(I) = I$. We have three cases: if $n = j + l$, we will have $\sigma_{v_n}(I) = I \cup \{v_n\}$ if and only $(j-1, j+l) \in I$ i.e. $|C_{j-1}(I)| > l + 1$. If $C_{j+1}(I) = 0$, $\sigma_{C_j}(I) = I - C_j(I)$. And if $\sigma_{C_j}(I) > 0$, then $\sigma_{C_j}(I) = I - \{v_{k+1}, \ldots, v_{j+l-1}\}$, where $k = |C_{j+1}(I)|$. Letting $\sigma_{C_j}(I) = K$, we will have:

$$C_j(K) = \begin{cases} C_j(I) \cup \{v_{j+l}\}, & \text{if } |C_{j-1}(I)| > l+1. \\ \\ \emptyset & \text{if } |C_{j-1}(I)| \le l+1 \\ & \quad \text{,and } |C_{j+1}(I)| = 0. \quad (2.24) \\ \\ C_j(I) - \{v_{k+1}, v_{k+2}, \ldots, v_{j+l-1}\}, & \text{otherwise.} \\ \quad k = |C_{j+1}(I)| > 0 \end{cases}$$

$$|C_j(K)| = \begin{cases} l+1, & \text{if } |C_{j-1}(I)| > l+1. \\ \\ 0 & \text{if } |C_{j-1}(I)| \le l+1, \text{ and } |C_{j+1}(I)| = 0. \\ \\ k+1, & \text{otherwise.} \\ \quad k = |C_{j+1}(I)| > 0 \end{cases}$$

$$(2.25)$$

$$\gamma_j(K) = \begin{cases} \gamma_j(I) + 1, & \text{if } \gamma_{j+1}(I) > l+1. \\ \\ 0 & \text{if } \gamma_{j+1}(I) \le l+1 \text{ and } \gamma_{j-1}(I) = 0. \quad (2.26) \\ \\ \gamma_{j-1} + 1, & \text{otherwise.} \end{cases}$$

$$\Leftrightarrow \gamma_{a+1-j}(\sigma_{C_j}(I)) = WZ_{a+1-j}(\gamma(I)). \quad (2.27)$$

106

$\square$

Corollary 42. *Consider any arbitrary natural number $[n]$ and permutation ν on n, the action $\mathcal{T}_\nu : J(\mathcal{L}_a) \to J(\mathcal{L}_a)$, and $I \in J(\mathcal{L}_a)$. For any $x \in S_a$, we will have: $WZ_\nu(x) = WZ_{\nu(a)} \circ WZ_{\nu(a-1)} \circ \cdots \circ WZ_{\nu(1)}(x)$. The bijection γ introduced in 41 satisfies the following property: $\gamma(\mathcal{T}_\nu(I)) = WZ_\nu(\gamma(I))$.*

Theorem 43. *Consider an arbitrary natural number n and an arbitrary permutation ν of $[n]$. With $WZ_\nu : S_n \to S_n$ defined as above we will have,*

1. *$WZ_\nu^{2n}(x) = x$ for all $x \in S_n$.*

2. *For an arbitrary $1 \le j \le n$, let $f_j : S_n \to \mathbb{R}$ be defined by:*

$$f_j(x) = \begin{cases} 1, & \text{if } j \in x \\ 0, & \text{otherwise.} \end{cases} \tag{2.28}$$

The triple $\langle S_n, WZ_\nu, f_j \rangle$ is homomesic and the average of f_j along WZ_ν-orbits is $1/2$. Moreover, any linear combination of f_js is homomesic in WZ_ν-orbits of S_n.

We will prove the above theorem in Section 2.3.3. Given the bijection in Corollary 36, Theorem 34 is a straightforward consequence of Theorem 43.

2.3.2 Some homomesy results in the comotion-orbits of $J(\mathcal{Q}_{a,b})$, $J(\mathcal{L}_a)$, and $J(\mathcal{U}_a)$.

The following homomesy results can be easily verified using Theorem 32.

107

Corollary 44. *Let $\mathcal{P}$ be $\mathcal{Q}_{a,b}$ or $\mathcal{L}_a$. Consider an arbitrary natural number a, an arbitrary permutation ν, and $\mathcal{T}_\nu : J(\mathcal{P}) \to J(\mathcal{P})$ as defined in 33. We define the size function, $f : J(\mathcal{P}) \to \mathbb{R}$ as, $\forall I, f(I) = |I|$. The triple $\langle J(\mathcal{P}), \mathcal{T}_\nu, f \rangle$ is homomesic for any choice of ν .*

Proof. For $\mathcal{P} = \mathcal{Q}_{a,b}$, $f = \sum_{i=1}^{a} i\, s_i - a(a+1)/2$. For $\mathcal{P} = \mathcal{L}_a$, $f = \sum_{i=1}^{a} i\, s_i$. In both cases f is a linear combination of f_i using Theorems 32 and 34 we will have the result. $\qquad\square$

Corollary 45. *Consider the lattice $\mathcal{Q}_{a,b}$ and an arbitrary permutation ν of $[a]$. Let $x \in [a] \times [b]$. We define the antipodal function $A : [a] \times [b] \to [a] \times [b]$ by $A(x) = y$ where $x = (i, j) \Leftrightarrow y = (a-i+1, b-j+1)$. For $I \in J(\mathcal{Q}_{a,b})$ and $x \in [a] \times [b]$, we define the characteristic function $\mathcal{I}_I(x) : [a] \times [b] \to \{0, 1\}$s follows:*

$$\mathcal{I}_I(x) = \begin{cases} 1 & \text{if } x \in \mathcal{I} \\ 0 & \text{otherwise} \end{cases} \tag{2.29}$$

*For any arbitrary $x \in [a] \times [b]$ let $h : J(\mathcal{Q}_{a,b}) \to \{0, 1, -1\}$ be given by $h(I) = \mathcal{I}_I(x) - (1 - \mathcal{I}_I(A(x)))$. Then h is $0-$mesic in $\mathcal{T}_\nu-$orbits of $J(\mathcal{Q}_{a,b})$. In other words, we have **central antisymmetry**, i.e. the average of number of ideals that contain x is equal to the number of ideals that do not contain $A(x)$.*

Proof. Consider arbitrary $I \in \mathcal{Q}_{a,b}$ and $x = (x_1, x_2) \in [a] \times [b]$. Then

$$I_{\mathcal{I}}(x) = 1 \Leftrightarrow (x_1, x_2) \in \mathcal{I} \Leftrightarrow |C_{x_1}(I)| \geq x_2$$
$$\Rightarrow I_{\mathcal{I}}(x) = \sum_{j=x_2}^{b} g_{x_1,j}.$$

(2.30)

Similarly,

$$1 - \mathcal{I}_I(A(x)) = 1 \Leftrightarrow (a - x_1 + 1, b - x_2 + 1) \notin I \Leftrightarrow |C_{a-x_1+1}(I)| < b - x_2 + 1$$

$$\Leftrightarrow$$

$$|C_{a-x_1+1}(I)| \leq b - x_2 \Rightarrow 1 - \mathcal{I}_I(A(x)) = \sum_{j=0}^{b-x_2} g_{a-x_1+1,j} = \sum_{j=x_2}^{b} g_{a-x_1+1,b-j}.$$

(2.31)

By Equations 2.30 and 2.31, we have $h_x(\mathcal{I}) = \sum_{j=x_2}^{b} g_{x_1,j} - g_{a-x_1+1,b-j}$. Employing Theorem 32 we deduce that h_x is $0-$mesic for any arbitrary $x \in [a] \times [b]$. $\square$

Corollary 46. *Let $\mathcal{P}$ be one of $\mathcal{Q}_a$ or $\mathcal{U}_a$. Consider arbitrary $I \in J(\mathcal{P})$. We denote the **rank-alternating** cardinality of I by $\mathcal{R}(I)$ and we define it as $\mathcal{R}(I) = \sum_{(i,j) \in I} (-1)^{i+j}$. The triple $\langle J(\mathcal{P}), \mathcal{T}_\nu, \mathcal{R} \rangle$ is homomesic for any arbitrary permutation ν of $[a]$.*

Proof. We will first consider the case when $I \in J(\mathcal{Q}_{a,b})$. In this case we have:

$$
\begin{aligned}
2\,\mathcal{R}(I) \ &= \textstyle\sum_{x=(i,j)\in\mathcal{P}}(-1)^{i+j}\mathcal{I}_I(x)\\
&= \textstyle\sum_{x=(i,j)}(-1)^{i+j}\mathcal{I}_I(x) + \sum_{x=(i,j)}(-1)^{i+j}\mathcal{I}_I(x)\\
&\Rightarrow 2\,\mathcal{R}(I) = \textstyle\sum_{x=(i,j)\in\mathcal{X}}(-1)^{i+j}\mathcal{I}_I(x) + (-1)^{2a-(i+j)+2}\mathcal{I}_I(A(x))\\
&= \textstyle\sum_{x=(i,j)\in\mathcal{X}}(-1)^{i+j}h(x) + 1.
\end{aligned}
\tag{2.32}
$$

In the case where $I \in J(\mathcal{U}_a)$ we have:

$$
\mathcal{R}(I) = (-1)^{a+1}\sum_{i;|C_i|\text{odd}} 1.
$$

We define the function $f : \mathbb{N} \to \{0,1\}$ as follows: $f(x) = 1$ iff x odd, $f(x) = 0$ otherwise. Note that the average of f in any $[i, 2a]$ that i is odd is equal to $1/2$. Therefore, by Theorem 33 we will have the result.

$\square$

2.3.3   Homomesy in winching

In this section we will prove Theorems 37, 40, and 43. The concepts of **tuple board** and **snake** are the key definitions of this section, and they help us understand the orbit structure and homomesy in winching.

Fix k, for arbitrary ν a permutation of $[k]$, let $\mathcal{F}_\nu$ be one of W_ν, $\underline{W}_\nu$ or WZ_ν. Let $S = S_k$ if $\mathcal{F} = WZ_\nu$ and $S = S_{k,m}$ otherwise. We define a tuple board as follows:

row 1 (x^1)	x^1_1	$\ldots$	x^1_k
row 2 (x^2)	x^2_1	$\ldots$	x^2_k
row 3 (x^3)	x^3_1	$\ldots$	x^3_k
$\ldots$	$\ldots$	$\ldots$	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$

Figure 2.3: A tuple board.

Definition 37. Consider $x \in S$ and ν a permutation of $[k]$. We write $x, \mathcal{F}_\nu(x), \mathcal{F}_\nu^2(x), \ldots$ in separate, consecutive rows as depicted below. Let $TB(x) = [x^1, x^2, \ldots]$ be such a table, where $TB(i, \cdot) = x^i = (x^i_1, \ldots, x^i_k)$ and $x^i = \mathcal{F}_\nu^{i-1}(x)$. We will have a board looking as follows:

$TB(x)$ is called the tuple board of x. Since $\mathcal{F}_\nu$ is a permutation, there is some n such that $\mathcal{F}_\nu^{n+1}(x) = x$. Therefore, we can also define a cylinder corresponding to the orbit containing x:

Consider $\mathcal{O}$, an $\mathcal{F}_\nu$-orbit of S which is produced by applying $\mathcal{F}_\nu$ consecutively to x. We define the **tuple cylinder** $TS(\mathcal{O})$ to be the cylinder that is produced by attaching the first and the n+1st row of $TB(x)$. Since $\mathcal{O}$ is an orbit it is more natural to think of a tuple board as a cylinder. We will use the terms interchangeably in this text.

Notice that any cell in a tuple board contains a number from the set $\{0, 1, 2, \ldots, m\}$. In what comes in the following we will introduce the notion of snakes. Given a tuple board T, any snake in it, is a sequence of adjacent cells in T that contain the numbers $1, 2, \ldots m$. The mathematical definition of a snake comes in the following:

111

Definition 38. For arbitrary $\nu = (\nu_1, \nu_2, \ldots, \nu_k)$ a permutation of $[k]$ and $x \in S$, let $TB = TB(x)$ be the tuple board of x as defined in Definition 37. Considering $T = \{TB(i,j) | 1 \le i \le n, 1 \le j \le k\}$, we define a **snake** $s = (s_f, s_{f+1}, \ldots, s_t)$ as follows: s is a maximal sequence of s_is such that each s_i is a cell in the tuple board containing i, and for $i > f$, $s_i = \mathcal{M}(s_{i-1})$, where $\mathcal{M}$ is defined as follows:

$$
\mathcal{M}(T(i,j)) = \begin{cases} T(i{+}1, j) & \text{if } T(i{+}1, j) = T(i,j) + 1. \\[1em] T(i, j{+}1) & \text{if } T(i, j{+}1) = T(i,j) + 1, \\ & \quad T(i{+}1, j) \neq T(i,j) + 1, \\ & \quad \text{and } \nu(j) < \nu(j+1). \\[1em] T(i{+}1, j{+}1) & \text{if } T(i{+}1, j{+}1) = T(i,j) + 1, \\ & \quad T(i{+}1, j) \neq T(i,j) + 1, \\ & \quad \text{and } \nu(j) > \nu(j+1). \end{cases}
\tag{2.33}
$$

Definition 39. Consider $T = TB = [x^1, x^2, \ldots, x^n]$ as defined previously for $x \in S$. In what follows row numbers in a tuple board are understood modulo n.

Consider s a snake in T. We define the **snake map** $\mathcal{S}$, a function that associates any snake with an element in $\mathbb{N}^k$ as follows: for an arbitrary snake s, $\mathcal{S}(s) = (c_1, c_2, \ldots, c_k)$, where $c_j = |\{i | T(i,j) \in s\}|$.

...	...	...	...	...
1	2	?	?	?
?	3	?	?	?
?	4	5	?	?
?	?	6	?	?
?	?	7	?	?
?	?	?	8	9
?	?	?	?	10
...	...	...	...	...

A tuple board corresponding to W_ν
$x \in S_{5,10}$ and
$\nu = (1, 2, 4, 3, 5)$.
The snake map is $(1, 3, 3, 1, 2)$.

...	...	...	...	...
?	4	?	?	?
?	?	5	?	?
?	?	6	?	?
?	?	7	?	?
?	?	8	9	10
?	?	?	?	?
...	...	...	...	...
...	...	...	...	...

A tuple board corresponding to W_ν
with lower bounds $(2, 4, 6, 7, 8)$.
$x \in S_{5,10}$ and $\nu = (1, 3, 2, 4, 5)$.
The snake map is $(0, 1, 4, 1, 1)$.

...	...	...	...	...
...	...	...	...	...
0	1	?	?	?
?	?	2	?	?
?	?	3	4	5
?	?	?	?	?
?	?	?	?	?
...	...	...	...	...
...	...	...	...	...

A tuple board corresponding
to WZ_ν. $x \in S_5$ and
$\nu = (1, 3, 2, 4, 5)$
The snake map is $(0, 1, 2, 1, 1)$.

Figure 2.4: Tuple boards and snakes corresponding to different variations of winching.

2.3.4 Proof of Theorem 37

In this subsection we prove Theorem 37.

Definition 40. Let $\bar{W}_i : S_{k,m} \to S_{k,m}$ be the following map: $\forall x = (x_1 \ldots x_k) \in S_{k,m}$, $\bar{W}_i(x) = y = (y_1, y_2, \ldots, y_k)$ where $\forall j \neq i, y_j = x_j$, and

$$y_i = \begin{cases} x_{i+1}-1, & \text{if } x_i = x_{i-1} + 1. \\ x_i-1, & \text{otherwise.} \end{cases} \tag{2.34}$$

Note that $\forall x \in S_{k,m}, \bar{W}_i \circ W_i(x) = x$. We call $\bar{W}_i$ **inverse winching** at index i.

Definition 41. For ν an arbitrary permutation of $[k]$, $\bar{W}_\nu : S^{k,m} \to S^{k,m}$ is defined by $\bar{W}_\nu = \bar{W}_{\nu(1)} \circ \bar{W}_{\nu(2)} \circ \cdots \circ W^r_{\nu(k)}$ and we have $\forall x \in S_{k,m}, \bar{W}_\nu(W_\nu(x)) = x$.

Lemma 47. *Any snake in a tuple cylinder $TS(\mathcal{O})$ (where $\mathcal{O}$ is a W_ν-orbit of $S_{k,m}$) is of length m, starts in the first column of the cylinder with s_1, and ends in the last column of the tuple cylinder with s_m.*

Proof. Consider some $x \in \mathcal{O}$ and a snake s in the tuple board $T = TB(x) = [x^1, \ldots, x^n]$. We assume that $s = (s_f, \ldots, s_t)$. Having, $x^{i+1} = W_\nu(x^i)$, it is easy to verify that unless $t = m$, we can find a cell in T to expand s. Similarly, since $x^{i-1} = \bar{W}(x^i)$. If $\nu(j) < \nu(j+1)$, we can see: unless $f = 1$, the snake s can be expanded.

□

Definition 42. Let $H : [m]^{[k]} \to [m]^{[k]}$ be defined as follows: $\forall x = (x_1, \ldots, x_k)$, $H(x) = y = (y_1, \ldots, y_k)$ where $\forall 1 \leq i < k, y_i = x_{i+1}$ and $y_k = x_1$. We call H the **left shift operator**.

Lemma 48. *Let $(p, 1)$ and $(q, 1)$ $(p < q)$ be two cells of tuple board T with value 1, such that there is no $p < i < q$ with $T(i, 1) = 1$. Consider the snake $s^p = (s_1^p \ldots s_m^p)$ starting with $s_1^p = T(p, 1)$ and $\mathcal{S}(s^p) = c^p = (c_1^p, \ldots, c_k^p)$ its snake map; and similarly consider the snake s^q and its snake map $\mathcal{S}(s^q) = c^q$ starting at $T(q, 1)$. Then,*

- *If $T(i, j) \in s^p$, we have the following:*

  - *$T(i + 1, j) \notin s^p \Rightarrow T(i + 1, j) \in s^q$.*

  - *If $j > 1$ then, $T(i, j-1) \notin s^p \Rightarrow T(i, j-1) \in s^q$.*

  In other words there is no gap between two consecutive snakes in the tuple board.

- *We have $c^q = H(c^p)$.*

Proof. In order to prove this lemma we fix $\nu = (1, 2, \ldots, k)$. The proof will be similar for any arbitrary permutation ν. To make notation simpler we drop the subscript from W meaning $\nu = (1, 2, \ldots, k)$.

Suppose that we have the action of winching $W_{(1,2,\ldots,k)}$ on $x \in S_{k,m}$ making the orbit $\mathcal{O}$ in $S_{k,m}$. Moreover, suppose the tuple board corresponding to x (equivalently, the tuple cylinder corresponding to $\mathcal{O}$) is $T = TB(x) = [x^1, x^2, \ldots, x^n]$ where $x^i = W^{i-1}(x)$ is as defined in Definition 37.

Claim 1. $c_1^p = q - p$.

Since $T(q, 1) \notin s^p$, $c_1^p \leq q - p$. Moreover $c_1^p = c_1 \leq q - p$ implies $s_{c_1+1}^p = T(p + c_1 - 1, 2) = c_1 + 1$ meaning $x_2^{c_1} = c_1 + 1$ and $x_1^{c_1} = c_1$. We have $x^{c_1+1} = W(x_1^c)$, and hence $x_1^{c_1+1} = 1$, and $T(c_1 + p, 1) \in s^q \Rightarrow c_1 + p = q \Rightarrow c_1 = q - p$.

Note that Claim 1 implies that there is no gap between the two snakes in column 1. (See Figure 2.3.4).

Claim 2. $c_1^q = c_2^p$. For simplicity, we denote c_1^p by c_1 and c_2^p by c_2.

We have $s_1^p = T(p, 1) = 1$, $s_2^p = T(p + 1, 1), \ldots, s_{c_1}^p = T(p + c_1 - 1, 1)$. (See 2.3.4)

Then, for all $1 \leq i \leq c_2$:

$$
\begin{aligned}
s_{c_1+i}^p &= T(p + c_1 - 1 + (i-1), 2) \\
\Rightarrow s_{c_1+i}^p &= T(q + (i-2), 2) && \text{(Since } q = p + c_1\text{)} && (2.35) \\
\Rightarrow x_2^{c_1+i-1} &= c_1 + i
\end{aligned}
$$

We also have

$$
s_{c_1+c_2+1}^p = T(q + c_2 - 2, 3) = T(p + c_1 + c_2 - 2, 3) \Rightarrow x_3^{c_1+c_2-1} = c_1 + c_2 + 1.
$$

$$(2.36)$$

Now consider s^q. For all $i, 1 \leq i \leq c_2 - 1$ we have that if $x_1^{c_1+i} = i$,

116

$$\left.\begin{aligned} x_1^{c_1+i} &= i \\ x_2^{c_1+i} &= c_1 + i + 1 > i \\ W(x^{c_1+i}) &= x^{c_1+i+1} \end{aligned}\right\} \Rightarrow x_2^{c_1+i+1} = i + 1 \qquad (2.37)$$

Therefore,

$$\begin{aligned} x_1^{c_1+1} = 1 \Rightarrow \forall i, 1 \le i \le c_2 - 1, \ \mathcal{M}(s_i^q) &= T(q+i, 1) \\ \Rightarrow \forall 1 \le i \le c_2, \ s_i^q &= T(q+i-1, 1). \end{aligned} \qquad (2.38)$$

From Equation 2.37 we can conclude $x_1^{c_1+c_2} = c_2$. By Equations 2.36 and 2.35, and the fact that $W(x^{c_1+c_2-1}) = x^{c_1+c_2}$, we have $x_2^{c_1+c_2} = c_2 + 1$. Hence, $\mathcal{M}(s_{c_2}^q) = T(q + c_2 - 1, 2)$.

It follows that $c_1^q = c_2^p$. Moreover, $T(i, 2) \in s^p \Rightarrow T(i, 1) \in s^q$ and $T(i-1, 2) \in s^q$ for any i (if they are not already in s^p).

Very similar to the proof of Claim 2, the following can be proved using the definitions:

Claim 3. Let $r < k$ with $\forall l, 1 \le l < r - 1, \ c_l^q = c_{l+1}^p$, then $c_r^q = c_{r+1}^p$.

Having Claims 2 and 3, by employing induction we can show: for all $i, 1 \le i < k - 1, c_i^q = c_{i+1}^p$, and there is no gap between the snakes in any of the columns. Furthermore, since all snakes have the same length, $c_k^q = c_1^p$. □

Proof of Theorem 37, Part 1. Consider an $n \times k$ tuple board T such that $T = TB(x)$ and $x \in S_{k,m}$. Let's assume that $n \ge m$ (if $n < m$, append enough copies of T to it until $n \ge m$). Let s^1 be the snake

117

that covers $T(1,1)$, s^2 the next snake immediately below s^1, and s^i the last snake right below s^{i-1}. Letting $\mathcal{S}(s^1) = c = (c_1, c_2, \ldots, c_k)$, we have $\mathcal{S}(s^i) = H^{i-1}(c)$. The numbers in the first column of T will be: $x_1, x_1 + 1, \ldots, x_1 + c_1 - 1, 1, 2, \ldots c_2, 1, 2, \ldots, c_3, \ldots$. Since $\sum_{i=1}^{k} c_i = m$, the $m + 1$st number in the first column will be x_1. Similarly, for each column i, the $m+1$st element will be x_i. Thus, $W^{m+1}(x) = x$.

$\square$

Corollary 49. *The above reasoning also shows there are exactly k snakes covering an $m \times k$ tuple cylinder.*

Corollary 50. *Fix k and n and ν a permutation of $[k]$. To each tuple cylinder T of size $k \times n$ corresponding to a $W_\nu-$orbit, we can assign a sequence $c = (c_1, c_2, \ldots, c_k)$, satisfying $\sum_{i=1}^{k} c_i = n$ where T is covered by snakes $s_1, s_2, \ldots, s_k$ and for all $1 \leq j \leq k$, there is an i such that $\mathcal{S}(s_j) = H^i(c)$. Since filling the first column of the cylinder will impose the other numbers, this correspondence is a one to one mapping.*

In order to prove Part 2 of Theorem 37, we present the definition of rotation on 2-colored necklaces with k white beads and $n - k$ black beads, then we proceed to the proof:

Definition 43. Let $N_{k,m}$ the set of all k-tuples $(x_1, x_2, \ldots, x_k)$ satisfying $1 \leq x_1 < x_2 < \cdots < x_k \leq m$ and $\sum_{i=1}^{k} x_i = m$. The action of rotation on this set is defined as $R : N_{k,m} \to N_{k,m}$, $\forall x \in N_{k,m}$, $R(x) = y$, where

Figure 2.5: Snakes in a tuple board of $W_{(1,2,\ldots,a)}$.

$y = (x_1+1, x_2+1, \ldots, x_k+1)$ if for all i, $x_i < m$. Or $y = (1, x_1+1, \ldots x_{k-1}+1)$ if $x_k = m$.

Lemma 51. *There is a map $\mathcal{K} : S_{k,m} \rightarrow N_{k,m}$ satisfying $\forall x \in S_{k,m}$, $R(\mathcal{K}(x)) = \mathcal{K}(\bar{W}_\nu(x))$.*

Proof. Consider arbitrary $x \in S_{k,m}$ and $T = TB(x)$ as in Definition 37. Let s be the snake covering $T(1,1)$. For $\mathcal{S}(s) = (c_1, c_2, \ldots, c_k)$, we define $\mathcal{K}(x) = (y_1, y_2 \ldots y_k) \in N_{k,m}$, where $y_1 = c_1 - x_1 + 1$, and for $2 \leq i \leq k$, $y_i = y_{i-1} + c_i$.

Note that $R(\mathcal{K}(x)) = \mathcal{K}(\bar{W}(x))$ if and only if $R(\mathcal{K}(W(x))) = \mathcal{K}(x)$. Let $T = TB(x)$, $T' = TB(W(x))$. Let s be the snake in T covering $T(1,1)$, $c = \mathcal{S}(s)$, and similarly let s' be the snake in T' covering $T'(1,1)$, $c' = \mathcal{S}(s')$, and $W(x) = z$. Either $c = c'$ and $x_1 + 1 = z_1$ or $c' = H(c)$, $x_1 = c_1$, and $z_1 = 1$.

$$R(\mathcal{K}(z)) = R(y_1, y_2, \ldots y_k); y_1 = c'_1 - z_1 + 1, y_{i+1} = y_i + c'_{i+1} \qquad (2.39)$$

$$= \begin{cases} R(y_1, y_2, \ldots y_k); y_1 = c_1 - (x_1 + 1) + 1, y_{i+1} = y_i + c_{i+1} \\ R(y_1, y_2, \ldots y_k); y_i | y_1 = c_2 - 1 + 1, y_{i+1} = y_i + c_{i+2} \end{cases} \qquad (2.40)$$

$$= \begin{cases} R(y_1, y_2, \ldots y_k); y_1 = c_1 - x_1, y_{i+1} = y_i + c_{i+1} \\ R(y_1, y_2, \ldots y_k); y_1 = c_2, y_{i+1} = y_i + c_{i+2(mod\ k)} \end{cases} \qquad (2.41)$$

120

$$T = TB(2,3,4,6):$$

2	3	4	6
1	2	5	7
1	3	6	7
2	4	5	6
3	4	5	7
1	2	6	7
1	3	4	5

$$T' = TB(1,2,5,7):$$

1	2	5	7
1	3	6	7
2	4	5	6
3	4	5	7
1	2	6	7
1	3	4	5
2	3	4	6

Figure 2.6: The snake covering $T(1,1)$ has snake map (2,1,3,1). Hence, $\mathcal{K}(x) = (1,2,5,6)$. The snake covering $T'(1,1)$ has snake map (1,3,1,2). Hence, $\mathcal{K}(W(x)) = (1,4,5,7)$.

Hence $R(\mathcal{K}(z))$ equals:

$$\begin{cases} \{(y'_1, y'_2, \ldots y'_k); y_1 = c_1 - x_1 + 1, y_{i+1} = y_i + c_{i+1}\}) & \text{because } \forall i, y_i < m \\ R(\{y_i | y_i = c_{i+1}(1 \leq i \leq k-1), y_k = m\}) = (1, 1 + c_2, \ldots, 1 + \sum_{i=1}^{k} c_k) \end{cases}$$

$$\tag{2.42}$$

$$= \mathcal{K}(x) \tag{2.43}$$

$\square$

Example 15. Consider $x = (2,3,4,6) \in S_{4,7}$, $W(x) = (1,2,5,6)$. In Figure 5, $TB(x)$ and $TB(W(x))$ are depicted. We see that $\mathcal{K}(W(x)) = (1,4,5,7)$, and $\mathcal{K}(x) = (1,2,5,6)$. Note that $R(1,4,5,7) = (1,2,5,6)$.

Proof of Theorem 37, Part 2. Having Lemma 51, we conclude that the orbit structures of $\langle N_{k,m}, R \rangle$ and $\langle S_{k,m}, W \rangle$ are the same.

Lemma 52. *Let T be an $m \times k$ tuple board. Consider the column r: $T_r = \{T(i,r)\}$. For any $1 \leq r \leq k$, there exists a one-to-one function $\mathcal{F} : T_r \to$*

T_{k+1-r}, *satisfying* $\mathcal{F}(x) = m+1-x$.

Proof. For any r, we construct a mapping from $\{\cup_{t=1}^{r} T_t\}$ to $\{\cup_{t=k-r+1}^{k} T_t\}$. Consider a number x in T_r. Let it be the lth element in T_r, covered by a snake having snake map $p = (c_1, c_2, \ldots, c_r, \ldots, c_k)$. Consider the snake with snake map $p' = (c_{r+1}, \ldots, c_1, c_2, \ldots, c_r)$. Let y be the $\sum_{i=1}^{r-1} c_i + l$th element from the end in this snake. Then $y = m+1 - \sum_{i=1}^{r-1} c_i + l = m+1-x$. Since $\sum_{i=1}^{r-1} c_i + l \leq \sum_{i=1}^{r} c_i$, y will be lying in one of the columns $k, \ldots k-r+1$.

Having the above mapping, we know there is also a one-to-one mapping in $\{\cup_{t=1}^{r} T_t\} \rightarrow \{\cup_{t=k-r+1}^{k} T_t\}$ and also in $\{\cup_{t=1}^{r-1} T_t\} \rightarrow \{\cup_{t=k-r}^{k} T_t\}$. Hence, there exists $\mathcal{F} : T_r \rightarrow T_{k+1-r}$ satisfying the lemma's conditions.

$\square$

Proof of Theorem 37, part 3. Considering any $m \times k$ tuple cylinder $TS(\mathcal{O})$, Lemma 48 shows that $TS(\mathcal{O})$ is totally covered by k snakes. Therefore, each element $1 \leq i \leq m$ appears k times in the cylinder and therefore the average of f_i as defined in Theorem 37 part 3 is independent of $\mathcal{O}$ and equal to k/m.

Lemma 52 shows that the number of js in any column i is equal to the number of $m-j+1$'s in column $k - i + 1$ of $TS(\mathcal{O})$. Thus, $\sum_{x \in \mathcal{O}} g_{i,j}(x) = \sum_{x \in \mathcal{O}} g_{k-i+1,m-j+1}(x)$. In other words, $\forall 1 \leq i \leq k, 1 \leq j \leq m, g_{i,j} - g_{k-i+1,m-j+1}$ is 0-mesic in W-orbits of $S_{k,m}$. $\square$

2.3.5 Proof of Theorem 40

In this subsection we will prove Theorem 40. Remember the definitions of tuple board, snake, snake map and the correspondence to the action of winching with lower bounds.

Definition 44. For the set of lower bounds $l = (l_1, \ldots l_k)$ and $i \in k$, let $\underline{\bar{W}}_i$ be defined as: $\underline{\bar{W}}_i : S_{k,m} \to S_{k,m}; \; \forall x \in S_{k,m} \; \underline{\bar{W}}_i(x) = \max\{l_i, \bar{W}_i(x)\}$, where $\bar{W}_i$ is defined as in Definition 40.

Note that $\forall x \in S_{k,m}, \bar{W}_i \circ W_i(x) = x$. We call $\bar{W}_i$ **inverse winching** at index i.

Definition 45. For arbitrary k and ν a permutation of $[k]$, $\bar{W}_\nu : S_{k,m} \to S_{k,m}$ is defined by $\bar{W}_\nu = \bar{W}_{\nu(1)} \circ \bar{W}_{\nu(2)} \circ \cdots \circ W^r_{\nu(k)}$ and we have $\forall x \in S_{k,m}, \bar{W}_\nu(W_\nu(x)) = x$.

In contrast to what we showed in Lemma 47 for W_ν-snakes , $\underline{W}_\nu$-snakes do not necessarily cover all the numbers $1, 2, \ldots, m$. As depicted in Figure 3, they only contain $l_i, \ldots m$ where l_i is one of the lower bounds. The following lemma states this formally:

Lemma 53. *Consider a tuple cylinder $TS(\mathcal{O})$ which is constructed by applying the action of $\underline{W}_\nu$ to an arbitrary $x \in S_{k,m}$. Let the lower bounds for this action be $l = (l_1, l_2, \ldots, l_k)$. Any snake in this tuple board starts in some column q and with s_{l_q}, and ends in the last column of the tuple cylinder with s_m.*

Proof. Consider $x \in \mathcal{O}$ and snake s in the tuple board $T = TB(x) = [x^1, \ldots, x^n]$. We assume that $s = (s_f \ldots s_t)$. Following the definitions of winching and snakes, we can see that unless $t = m$ we can append more cells to the tail of the snake, and if $l_i < f$ we can append more cells to the head of the snake.

$\square$

Proof of Theorem 40.

Any tuple cylinder corresponding to action of $\underline{W}_\nu$ can be partitioned to snakes that start with some s_{l_q} and end in s_m. Therefore, if f is a function that have the same average on all the numbers contained in any snake, it will have the same average over all the elements in the tuple cylinders. Therefore, we will have the result.

$\square$

2.3.6 Proof of Theorem 43

In this section, we will prove Theorem 43. Remember the definitions of tuple board, tuple cylinder, snake, snake map. Consider $x \in S_n$ and the action of WZ_ν for some arbitrary permutation ν of $[n]$. Note that in the case of winching with zeros, since we might have a bunch of zeros in our tuple board the snake does not necessarily start in column 1. However, it is a consecutive collection of numbers $1, 2, \ldots n$ as was the case in Lemma 47.

Definition 46. Let M_n be the set of all sequences $(c_1, \ldots, c_n)$ that have k

preceding 0s for some $0 \leq k \leq n-1$, $c_{k+1} \ldots c_n > 0$, and $\sum_{i=k+1}^{n} c_i = n$.
Let M_n be the set of all possible snake maps. We define the action **crawl**
$C : M_n \to M_n$ such that for any $c \in M_n$, $C(c) = c'$ where,

For $1 \leq i \leq n-1$,

$$
c_i' = \begin{cases} \max\{0, c_{i+1}-1\} & \text{If } c_1, \ldots, c_i \leq 1; \\ c_{i+1} & \text{otherwise.} \end{cases}
$$

And, $c_n' = n - \sum_{i=1}^{n-1} c_i'$.

Lemma 54. *Consider some arbitrary ν a permutation of $[n]$. Let $T_{m \times n}$ be a tuple board corresponding to WZ_ν, i a row in T containing a 1, and snake s starting at row i. Let j be the smallest number greater than i containing another 1, and s' the snake starting at row j.*

1. *We have $j = i + 2$.*

2. *Let $c = (c_1, \ldots, c_n)$ be snake map of s and $c' = (c_1', \ldots, c_n')$ be snake map of s'. We have $c' = C(c)$*

3. *Any element of the tuple board is either a 0 or it belongs to a snake.*

Proof. For simplicity we assume $\nu = (1, 2, \ldots, n)$, and drop subscript from W. The proof any arbitrary permutation ν of $[n]$ will be similar.

Proof of Part 1. First, we argue that $c_1 \leq 1$. We know that $\sum_{i=1}^{n} c_i =$

n. If $c_1 \neq 0$ then we have

$$c_2, \ldots, c_n > 0 \Rightarrow c_2 + \cdots + c_n \geq n-1 \Rightarrow c_1 \leq n - (n-1) = 1.$$

Since there is a 1 in row i if $T(i,1) = 0$, $T(i,2) = 0$ or $T(i,2) = 1$. In both cases, applying winching will derive, $T(i+1,1) = 0$. If $T(i,1) = 1$ since $c_1 \leq 1$ we have $T(i,2) = 2$, and $T(i+1,1) = 0$. Thus, in both cases $T(i+1,1) = 0$. Consider the first column where s turns down and let it be column j. We have, $c_1, \ldots, c_{l-1} \leq 1$, and $T(i,l) + 1 = T(i+1,l)$. Note that for any $1 \leq k \leq l$, $T(i,k) = 0$ or $T(i,k) = T(i,k-1) + 1$. Hence, applying WZ to $y = T(i,.)$ will result in $T(i+1,k) = 0$ for $1 \leq k \leq l-1$, and $T(i+1,l) > 1$. Since the rest of entries in row $i+1$ will be in increasing order, there will not be any 1 in this row. In the $i + 2$nd row, we will have $T(i+2,k) = 0$ for $1 \leq k \leq l-2$, and $T(i+2,l-1) = 1$. As a result, $j = i + 2$.

Proof of Part 2. As we argued in part 1, s' starts at $T(i+2,l-1)$. Since elements of s and s' will be nonzero from this point to the right then by the same argument presented for regular winching in the proof of Lemma 48) we can show the snakes will move in parallel to each other. Note that in row $i+1$ we have zeros until we get to $T(i+1,l)$. In row $i+2$ we have zeros in $T(i+2,1), \ldots, T(i+2,l-2)$, i.e. $c'_1 = \cdots = c'_{l-2} = 0$. At column $l-1$, s' will start and move parallel to s but there is a zero between s and s' in column $l-1$. Therefore, $c'_{l-1} = c_l - 1$, and there is no zero between the remaining part of s and s', thus $c'_k = c_{k+1}$, for $l \leq k \leq n-1$. The rest of s' will continue

in column n.

Proof of Part 3. It is clear from the above argument that the space between the initial segments of any two snakes is filled with zeros. $\square$

Lemma 55.

For any $c \in M_n$ we have, $C^n(c) = c$.

We will need the following definitions and lemmas to prove Lemma 55.

Definition 47. Consider the set M_n. We define the one-to-one map $\mathcal{F} :$ $M_n \to \{0,1\}^{2n}$ as follows:

for all $c \in M_n$, $\mathcal{F}(c) = b = (b_1, \ldots, b_{2n})$ where for $1 \leq i \leq n$,

$$
b_i = \begin{cases} 1 & \text{If } \exists k; c_1 + \ldots + c_k = i; \\ 0 & \text{Otherwise.} \end{cases}
$$

And, for $n < i \leq 2n$, $b_i = \neg b_{i-n+1}$.

Lemma 56. *$\mathcal{F}$ is one-to-one.*

Proof. Assume $\mathcal{F}(x) = \mathcal{F}(y) = w$, and let j be the smallest index where $w_j = 1$. We have $x_1 = y_1 = j$. The next nonzero index will determine that $x_2 = y_2$ and likewise, we can verify that all entries of x and y are equal. $\square$

Definition 48. Let $\mathcal{B}_n \subset \{0,1\}^{2n}$ be the set of all $b \in \{0,1\}^{2n}$ such that for all, $1 \leq i \leq n, b_i = \neg b_{n+i}$. We define the action of rotation $\mathcal{R} : \mathcal{B} \to \mathcal{B}$ on this set as follows: Partition b into maximal blocks of $1^k 0$, remove the leftmost block, and put it on the right.

127

Example 16. Let $b = (110010001101)$. The partitioning of b will be $(110.0.10.0.0.110.1)$. Therefore, $\mathcal{R}(b) = (0.10.0.0.110.1.110)$.

Lemma 57. *For $c \in M_n$, we have $\mathcal{F}(C(c)) = \mathcal{R}(\mathcal{F}(c))$*

Proof. Consider an arbitrary $c = (c_1, \ldots, c_n) \in M_n$. Let's say we have $c_1 = \cdots = c_{k-1} = 0$, and c_k is the leftmost nonzero element in c. Consider the set $\mathcal{A} = \{a_1 = c_k, a_2 = c_k + c_{k+1}, \ldots, a_{n-k} = \sum_{i=k}^{n} c_i\}$. Let $C(c) = c'$ and $b = (b_1, \ldots, b_n)$, the binary word representing $\mathcal{A}$. In other words for all $a, a \in \mathcal{A} \Leftrightarrow b_a = 1$.

Similarly, let $\mathcal{A}' = \{c_{k'}, c_{k'} + c_{k'+1}, \ldots, \sum_{i=k'}^{n} c_i'\}$ where k' is the leftmost nonzero element in c' and b' be $\mathcal{A}'$s binary representation.

According to the definition of crawl, we know that, if $c_k = \cdots = c_{l-1} = 1$, we will have $c'_{k-1}, \ldots, c'_{l-2} = 0$ and $c'_{l-1} = c_l - 1$, where l is the leftmost element greater than 1. This means that if have $a_1 = 1, a_2 = 2, a_3 = 3, \ldots a_{l-1} = l$, they should be removed from $\mathcal{A}$ to make $\mathcal{A}'$. In other words, any set of consecutive elements starting from a 1 will be removed in $\mathcal{A}'$. Moreover, c_l will be decremented which means a_1 and the rest of the elements in $\mathcal{A}$ will be decreased by l except the last one which should always be an n. Now, let's see how b will change accordingly. We remove consecutive elements starting with a 1 from $\mathcal{A}$ which means we remove the preceding 1s from b until we hit a 0. All the other elements will be decreased by l which means they should be shifted to left by l positions. This is equivalent to removing the first block from b. Now, we need to add $b'_{n-l+1}, \ldots, b'_{n-1} = 0$. And $b'_n = 1$ because c'_n should be increased by l to make the length of the snake equal to n. This

128

whole process is removing the leftmost block and adding its negation to the right, which is equivalent to a rotation of a block in $\mathcal{F}(c)$. $\square$

Lemma 58. $\forall x \in \mathcal{B}_n, \quad \mathcal{R}^n(x) = x.$

Proof. Consider any arbitrary x, any block in x has a single 0. Moreover, the number of zeros in x is n. Therefore, after n rotations x will get back to its initial state. $\square$

Proof of Lemma 55. From Lemma 58 and 57 and the fact that $\mathcal{F}$ is a one-to-one function we have, $\forall c \in \mathcal{M}_n, \quad C^n(c) = n.$

Proof of Theorem 43 Part 1. By employing Lemma 54 we can verify that the snakes appear in alternating rows. By Lemma 55 we know that each snake gets back to itself after n crawls. Thus, $T(1,.) = T(2n+1,.)$ where T is a tuple board corresponding to WZ, and $WZ^{2n}(x) = x$.

$\square$

Proof of Theorem 43 Part 2. Part 2. Using Lemma 54 part 3 we know that half of any tuple board is filled with zeros, and the rest is filled by equal repetitions of numbers 1 to n. In addition, there are n snakes in any tuple board and in any snake j appears once and only once. Therefore, each element will appear n times in the tuple board and the average of $f_j = 1/2$ for each j.

2.4 Expected Jaggedness of Ordered Ideals

Recall the definition of $\mathcal{Q}_{a,b}$ (Definition 25) from Section 2.3, and consider $J(\mathcal{Q}_{a,b})$. We define a probability distribution on this set as follows:

Definition 49. An ideal occurs with probability proportional to the number of $a \times b$ standard Young tableaux[6] with which it is compatible. (We say a path s is compatible with a tableau T if all the labels of T northwest of s are smaller than all of the labels of T southeast of s. We use English notation throughout.) We will call this distribution μ_{lin}, the *linear* distribution on lattice paths, since it comes from linear orderings of the ab boxes in the grid.

Example 17. The following picture shows the six lattice paths in a 2×2 grid occur in μ_{lin} with the probabilities shown below:
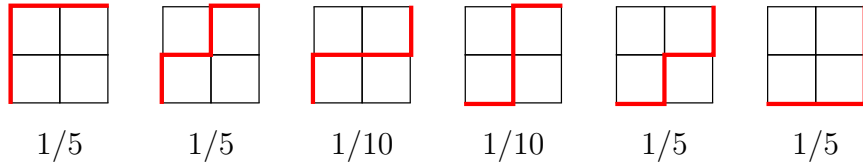
Note that each ideal we can assign a lattice path (presenting its border). We define the *jaggedness* of a lattice path to be equal to the number of turns in the lattice path:

Definition 50. The *jaggedness* of an order ideal $I \in J(\mathcal{P})$, denoted jag(I), is the total number of elements $p \in \mathcal{P}$ which can be toggled into I or out of I.

---

[6]See Appendix 2.5 for an introduction to Young tableaux.

We define, for each $p \in \mathcal{P}$, two indicator random variables $\mathcal{T}_p^+, \mathcal{T}_p^-$: $J(P) \to \mathbb{R}$ that record whether p can be toggled in (respectively toggled out) of an order ideal. Explicitly, for $I \in J(\mathcal{P})$, we define

$$\mathcal{T}_p^+(I) := \begin{cases} 1 & \text{if } p \text{ can be toggled in to } I, \\ 0 & \text{otherwise} \end{cases},$$

$$\mathcal{T}_p^-(I) := \begin{cases} 1 & \text{if } p \text{ can be toggled out of } I, \\ 0 & \text{otherwise}. \end{cases}$$

These random variables are highly related to Striker's toggleability [48]. Indeed, her toggleability statistic $\mathcal{T}_p$ simply decomposes as $\mathcal{T}_p = \mathcal{T}_p^+ - \mathcal{T}_p^-$. Note furthermore that jag $= \sum_{p \in P}(\mathcal{T}_p^+ + \mathcal{T}_p^-)$. Here, we will show how certain conditions on $\mathcal{T}_p$ imply conditions on jag, as in the following main definition of the section.

Definition 51. Let μ be a probability distribution on $J(\mathcal{P})$. Given an element $p \in \mathcal{P}$, we say that μ is *toggle-symmetric at p* if

$$Pr_\mu(p \text{ can be toggled in to } I) = Pr_\mu(p \text{ can be toggled out of } I).$$

Equivalently, μ is toggle-symmetric at p if

$$\mathbb{E}_\mu(\mathcal{T}_p) = \mathbb{E}_\mu(T_p^+) - \mathbb{E}_\mu(T_p^-) = 0.$$

131

We say that μ is *toggle-symmetric* if it is toggle-symmetric at every $p \in \mathcal{P}$.

We would like to introduce toggle-symmetric probability distributions as an interesting class of distributions on order ideals of posets. Throughout, we fix a poset $\mathcal{P}$ with $n = |P|$.

We may ask: what is the expected jaggedness of a lattice path, chosen according to μ_{lin}? The answer is surprisingly simple. The following theorem appeared in [49]:

Theorem 59. *([49]) The expected jaggedness of a lattice path in an $a \times b$ grid, chosen under the distribution μ_{lin}, is exactly $2ab/(a+b)$, the harmonic mean of a and b.*

The purpose of this chapter is to give a generalization of Theorem 59, in particular explaining the seeming coincidence above, and putting it in its proper combinatorial context: order ideals in arbitrary posets, and *toggle-symmetric* probability distributions on them. This last is a class of probability distributions that we would like to put forth as an interesting property to study, especially in relation to the developing area of dynamical algebraic combinatorics. We define toggle-symmetric distributions on order ideals of posets. The word "toggle" refers to the procedure of adding or removing an element from a set if it is permissible to do so. The term was coined by Striker-Williams [48]. Our results have direct applications to homomesy results for order ideals under special compositions of toggles, as we will discuss.

In Section 2.4.1, we prove our main result: a formula for expected jaggedness that applies to all skew Young diagrams, not just rectangles, and any toggle-symmetric distribution. Here is our main theorem:

Theorem 60. *Let σ be a connected skew shape with height a and width b. Let μ be any toggle-symmetric probability distribution on the subshapes of σ. Then the expected jaggedness of a subshape of σ with respect to μ is*

$$\frac{2ab}{a+b} \left(1 + \sum_{c \in C(\sigma)} \delta(c) \, Pr_\mu(c) \right).$$
$$(2.44)$$

Here $C(\sigma)$ is the set of *outward corners* of σ, $Pr_\mu(c)$ is the probability according to μ that the outward corner c is included in a subshape (a.k.a. lattice path), and $\delta(c)$ is a notion of signed displacement that we will define in Section 2.4.1. For now, the main point is that the expected jaggedness can be calculated as the harmonic mean of a and b, plus a sum of correction terms that can be completely understood in terms of μ and σ. (When σ is a rectangle, there are no correction terms and (2.44) gives the harmonic mean exactly, for *any* toggle-symmetric distribution.)

There are several key differences between Theorem 60 and the corresponding result [49, Theorem 2.8] of Chan et al. First, our theorem applies to *any* toggle-symmetric distribution. Moreover, it is fully symmetric with respect to interchanging rows and columns, which is not the case in [49]. Indeed, our result makes explicit that the only dependence is on the outer corners and their displacements. This will allow us to immediately derive that for any

balanced shape, the expected jaggedness is always the harmonic mean; see Corollary 64.

We also note that our results, combined with theorems of Striker [48], have direct applications to homomesy results under the operations of rowmotion on posets. In particular, they allow us to recover and generalize a theorem of Propp and Roby [7] on homomesies for antichain cardinalities. We present an example that illustrates Theorem 60, and then we will proceed to the proof.

Example 18. Consider the Young diagram shape $\sigma = (3, 1)$. The seven subshapes of σ, equivalently the lattice paths in σ, are depicted below. The numbers below each path indicate that subshape's jaggedness, along with the probability of that subshape's occurrence according to the linear distribution. Then we can calculate directly that $\mathbb{E}_{\mu_{\mathrm{lin}}}(\mathrm{jag}) = 34/15$.

(1; 1/5)　(2; 1/15)　(3; 2/15)　(3; 1/5)　(2; 1/15)　(3; 2/15)　(2; 1/5)
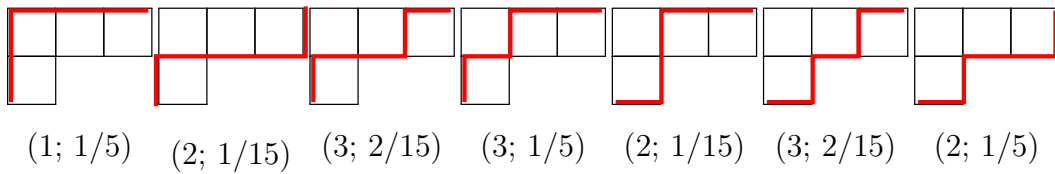
Now, let us use Theorem 60 instead to compute $\mathbb{E}_{\mu_{\mathrm{lin}}}(\mathrm{jag})$. Note that the corner c occurring at $(1, 1)$ is the only outward corner of σ. Its displacement $\delta(c)$ is $-1/6$, as in Definition 53. Finally, we have $Pr_{\mu_{\mathrm{lin}}}(c) = 1/3$. Plugging

these values into (2.44) yields

$$\mathbb{E}(\text{jag}) = (12/5)\,(1 - 1/6 \cdot 1/3) = 34/15,$$

as expected.

2.4.1 Proof of Theorem 60.

In this section, we prove a general result giving a formula for the expected jaggedness of an order ideal in a poset P for any toggle-symmetric distribution whenever $\mathcal{P}$ is the poset corresponding to a skew Young diagram. We first define the terms that are used in the statement of Theorem 60.

We assume that the reader is familiar with definitions of Young diagram, Young tableaux, and skew shapes. This definitions are provided in Appendix 2.5.

Let $\sigma = \lambda/\nu$ be a skew shape. Throughout, we let a denote the *height* of σ, i.e., the number of rows in σ, and let b denote the *width* of σ, i.e., the number of columns. In order to refer to the boxes of σ and their corners, we will fix coordinates as follows. Place σ in an $a \times b$ rectangle. Our convention will be that the northwest corner of the rectangle is $(0,0)$ and the southeast corner is (a, b). The corners of the boxes of σ are then various lattice points in this rectangle. Furthermore, we will extend this coordinate system to the boxes of σ by writing $[i, j]$ for the box whose *southeast* corner is (i, j). For example, the upper-leftmost box of a Young diagram is the box $[1, 1]$.
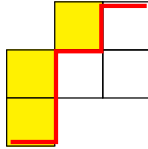
Figure 2.7: With $\sigma = (3,3,1)/(1)$, we depict $(2,1,1)/(1) \in \mathcal{J}(\sigma)$ shaded in yellow and its associated lattice path in red.

Associated to any skew shape σ is a poset $\mathcal{P}_\sigma$ whose elements are the boxes of σ and with $[i,j] \leq [k,l]$ if and only if $i \leq k$ and $j \leq l$. The order ideals in $J(\mathcal{P}_\sigma)$ are subshapes of σ.

We also often identify a subshape $\rho/\nu \in J(\mathcal{P}_\sigma)$ with its lattice path, which is the sequence of steps of the form $(-1,0)$ and $(0,1)$ connecting the point $(a,0)$ to $(0,b)$ (in the coordinate system defined above) given by the southeast border of ρ. In this way $J(\mathcal{P}_\sigma)$ is in bijection with the set of lattice paths connecting $(a,0)$ to $(0,b)$ that stay within the diagram of σ. For an example of this bijection see Figure 2.7.

Definition 52. Let σ be a skew shape. We say σ is *connected* if the poset $\mathcal{P}_\sigma$ is connected. Suppose σ is connected. Then an *outward corner* of σ is two consecutive steps along the boundary of σ that do not border a common box of σ. We say that a corner *occurs* at the lattice point (i,j) where its two steps meet. We write $C(\sigma)$ for the set of outward corners of σ.

Note that because σ is a skew shape, the outward corners of σ are either *northwest corners* or *southeast corners*, i.e., they comprise part of the northwest border of σ or the southeast border, respectively.
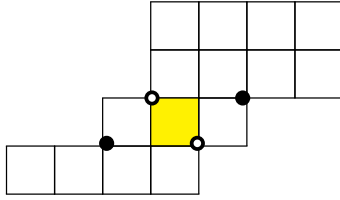
136

Figure 2.8: A diagram explaining our notation for corners. Box $[3, 4]$ of σ is shaded yellow and the points where corners $c \in C(\sigma)$ occur are marked with a circle; the points where $c \in C_{34}(\sigma)$ occur are open circles.

The following notation will be convenient for us: given a box $[i, j] \in \sigma$, we define $C_{ij}(\sigma)$ to be equal to the following set:

{corners $c \in C(\sigma)$ occurring strictly northwest or strictly southeast of $[i, j]$}.

When we say that a corner c occurs "strictly northwest" or "strictly southeast" of a box $[i, j]$, we mean that it occurs strictly northwest (respectively strictly southeast) of the *center* of that box. For example, a corner at the point (i, j) occurs strictly southeast of the box $[i, j]$. Figure 2.8 illustrates our notation for corners.

For $c \in C(\sigma)$ and μ a probability distribution on $J(\mathcal{P}_\sigma)$ we use the notation $Pr_\mu(c)$ to mean the probability with respect to μ that a subshape of σ, thought of as a lattice path, includes the two steps of the corner c.

Definition 53. Let σ be a connected skew shape with height a and width b. The *main anti-diagonal of σ* is the line joining $(a, 0)$ to $(0, b)$. For $(i, j) \in \mathbb{R}^2$ let $\vec{d}(i, j)$ denote the vector from (i, j) to the main anti-diagonal of σ (and

137

orthogonal to it). For an outward corner $c \in C(\sigma)$ that occurs at (i, j) we define the *displacement* of c to be

$$
\delta(c) := \begin{cases}
\text{the unique } x \in \mathbb{R} \text{ with } \vec{d}(0, 0) = x \cdot \vec{d}(i, j) & \text{if } c \text{ is a northwest corner,} \\[2mm]
\text{the unique } x \in \mathbb{R} \text{ with } \vec{d}(a, b) = x \cdot \vec{d}(i, j) & \text{if } c \text{ is a southeast corner.}
\end{cases}
$$

Note that $\delta(c)$ is a signed quantity. Explicitly,

$$
\delta(c) = \begin{cases}
1 - \frac{i}{a} - \frac{j}{b} & \text{if } c \text{ is a northwest corner,} \\[2mm]
-1 + \frac{i}{a} + \frac{j}{b} & \text{if } c \text{ is a southeast corner.}
\end{cases}
$$

In the rest of this section we will prove Theorem 60. In order to that, we define a set of random variables R_{ij} that we refer to as *rooks*. The proof of the main theorem involves strategically placing rooks on our skew shape σ.

Let $[i, j]$ be a box in σ. We write $\mathcal{T}_{ij}^{+}$ and $\mathcal{T}_{ij}^{-}$ for the toggle-indicator random variables $T_{[i,j]}^{+}$ and $T_{[i,j]}^{-}$ on $J(\mathcal{P}_\sigma)$. We define the rook random variable $R_{ij} : J(\mathcal{P}_\sigma) \to \mathbb{R}$ as follows:

$$
R_{ij} := \sum_{\substack{i' \leq i,\, j' \leq j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^{+} + \sum_{\substack{i' \geq i,\, j' \geq j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^{-} - \sum_{\substack{i' < i,\, j' < j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^{-} - \sum_{\substack{i' > i,\, j' > j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^{+}. \qquad (2.45)
$$

The equation defining R_{ij} is complicated and it is best understood by a drawing as in Figure 2.9. In this figure, we record the coefficients of the terms $T_{i'j'}^{+}$ and $T_{i'j'}^{-}$ in R_{ij} in the northwest and southeast corners, respectively, of

1	1						
-1							
1	1						
-1							
1	1						
	1	1	1	1	1		
	-1	-1	-1	-1			
	1	1	1	1			
	-1	-1	-1	-1			
	1	1	1	1			

Figure 2.9: An example of a "rook" at the box $[3, 2]$.

the box $[i', j']$. The reason we call R_{ij} a rook is explained by the next lemma, which says that for a toggle-symmetric distribution μ only the toggleability statistics corresponding to boxes in the same row or column as $[i, j]$ contribute to the expectation $\mathbb{E}_\mu(R_{ij})$.

Lemma 61. *Let σ be a skew shape and μ a toggle-symmetric probability distribution on $J(\mathcal{P}_\sigma)$. Then for any $[i, j] \in \sigma$ we have*

$$\mathbb{E}_\mu(R_{ij}) = \sum_{[i',j] \in \sigma} \mathbb{E}_\mu(\mathcal{T}_{i',j}^+) + \sum_{[i,j'] \in \sigma} \mathbb{E}_\mu(\mathcal{T}_{i,j'}^+).$$

Proof. Expanding formula (2.45),

$$R_{ij} = \sum_{\substack{i'<i, j'<j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^+ - \sum_{\substack{i'<i, j'<j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^- + \sum_{\substack{i'>i, j'>j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^+ - \sum_{\substack{i'>i, j'>j \\ [i',j'] \in \sigma}} \mathcal{T}_{i'j'}^-$$

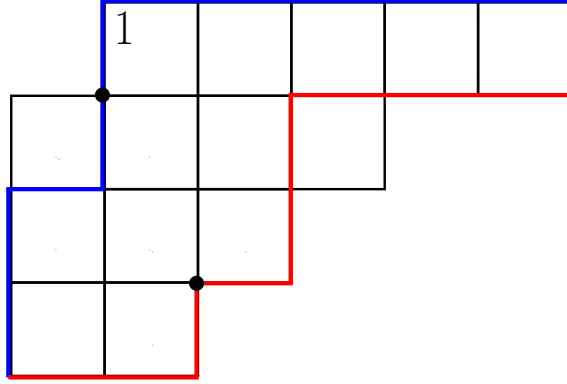$$+ \sum_{[i,j'] \in \sigma} \mathcal{T}_{i',j}^+ + \sum_{[i,j'] \in \sigma} \mathcal{T}_{i,j'}^+.$$

Figure 2.10: This figure illustrates how each subshape may contribute to $\mathbb{E}(R_{ij})$. Here $[i, j] = [3, 2]$ and the points where corners $c \in C_{ij}$ occur are marked with a circle. Two lattice paths $\rho_1, \rho_2 \in J(\mathcal{P}_\sigma)$ are drawn in blue and red; we can verify that $R_{ij}(\rho_k) = 1 + \#C_{ij}(\rho_k)$ for $k = 1, 2$.

Since μ is a toggle symmetric distribution, by linearity of expectation we get

$$
\mathbb{E}_\mu \left(\sum_{\substack{i' < i,\, j' < j \\ [i', j'] \in \sigma}} \mathcal{T}^+_{i'j'} - \sum_{\substack{i' < i,\, j' < j \\ [i', j'] \in \sigma}} \mathcal{T}^-_{i'j'} \right) = 0;
$$

$$
\mathbb{E}_\mu \left(\sum_{\substack{i' > i,\, j' > j \\ [i', j'] \in \sigma}} \mathcal{T}^+_{i'j'} - \sum_{\substack{i' > i,\, j' > j \\ [i', j'] \in \sigma}} \mathcal{T}^-_{i'j'} \right) = 0.
$$

Hence the claimed expression for $\mathbb{E}_\mu(R_{ij})$ indeed holds. $\qquad\square$

Lemma 62. *Let σ be a connected skew shape and μ a probability distribution*

on $J(\mathcal{P}_\sigma)$. Then for any $[i,j] \in \sigma$ we have

$$\mathbb{E}_\mu(R_{ij}) = 1 + \sum_{c \in C_{ij}(\sigma)} Pr_\mu(c).$$

Proof. Let $\rho \in J(\mathcal{P}_\sigma)$. Let $C_{ij}(\rho)$ be the set of all corners $c \in C_{ij}(\sigma)$ included in the lattice path ρ. We observe that $R_{ij}(\rho) = 1 + \#C_{ij}(\rho)$. This observation is again best understood by a picture, as in Figures 2.9 and 2.10. In Figure 2.9, the set $C_{ij}(\sigma)$ is empty, and the claim that $R_{ij}(\rho) = 1$ for any lattice path ρ drawn through the skew shape corresponds to the observation that the turns in ρ always have total weight 1 (with the weights as drawn). As usual, we identify lattice paths and subshapes.

The more general formula $R_{ij}(\rho) = 1 + \#C_{ij}(\rho)$ then corresponds to the fact that any outward corner $c \in C_{ij}(\sigma)$ used by ρ is no longer labeled -1, simply because there is no box at c to be toggled in or toggled out. This is illustrated in Figure 2.10.

But

$$\mathbb{E}_\mu(\#C_{ij}(\rho)) = \sum_{c \in C_{ij}(\sigma)} Pr_\mu(c)$$

and hence the claimed expression for $\mathbb{E}_\mu(R_{ij})$ indeed holds. $\qquad\square$

Lemma 63. *For any connected skew shape σ with height a and width b there exist integral coefficients $r_{ij} \in \mathbb{Z}$ for $[i,j] \in \sigma$ such that*

- *for all $1 \le i \le a$, $\sum_{[i,j'] \in \sigma} r_{i,j'} = b$;*

- *for all $1 \le j \le b$, $\sum_{[i',j] \in \sigma} r_{i',j} = a$.*

141

Proof. If we interpret the coefficient r_{ij} as the number (possibly negative) of rooks placed at box $[i, j] \in \sigma$, the equalities say that each row should be attacked by a total of b rooks and each column by a total of a rooks. There are many possible such placements. Here is one. Let $\mathcal{B} := \{[i, j] \in \sigma :$ $[i+1, j+1] \notin \sigma\}$ denote the set of boxes in the *southeast border strip of σ*. We claim there is a unique choice of r_{ij} satisfying the desired equalities with $r_{ij} = 0$ if $[i, j] \notin \mathcal{B}$. Let $b_1, b_2, \ldots, b_m$ be the elements of $\mathcal{B}$ in the unique order so that b_1 is southwesternmost, b_m is northeasternmost, and b_k is adjacent to b_{k+1} for all $1 \leq k < m$. Then for each $1 \leq k \leq m$, exactly one of the following holds:

1. b_l is not in the same row as b_k for all $l > k$;

2. b_l is not in the same column as b_k for all $l > k$.

Thus for $k = 1, \ldots, m$ with $b_k = [i_k, j_k]$, we can choose the corresponding coefficients r_{i_k, j_k} in order: when we are in case (I) we choose r_{i_k, j_k} so that $\sum_{[i_k, j] \in \sigma} r_{i_k, j} = b$; when we are in case (II) we choose r_{i_k, j_k} so that $\sum_{[i, j_k] \in \sigma} r_{i, j_k} = a$. For each row or column, there is at least one b_k in that row or column, so in the end all the equations will be satisfied. The result is an assignment of coefficients that looks like Figure 2.11. $\square$

Proof of Theorem 60. Let r_{ij} be the coefficients from Lemma 63. Note that the sum of all coefficients is $\sum_{[i,j] \in \sigma} r_{ij} = ab$. Also, for any $[i', j'] \in \sigma$ the sum of coefficients in its row and its column is $\sum_{\substack{[i,j] \in \sigma \\ i=i'}} r_{ij} + \sum_{\substack{[i,j] \in \sigma \\ j=j'}} r_{ij} = a + b$.
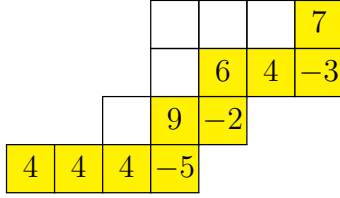
Figure 2.11: An example of a rook placement that satisfies Lemma 63. Here $a = 4$ and $b = 7$; the southeast border strip is shaded in yellow.

Using Lemma 61, we get

$$\mathbb{E}\left(\sum_{[i,j]\in\sigma} r_{ij}R_{i,j}\right) = \sum_{[i,j]\in\sigma} r_{ij}\left(\sum_{[i',j]\in\sigma} \mathbb{E}(\mathcal{T}^+_{i',j}) + \sum_{[i,j']\in\sigma} \mathbb{E}(\mathcal{T}^+_{i,j'})\right) \qquad (2.46)$$

$$= \sum_{[i,j]\in\sigma}\left(\sum_{[i',j]\in\sigma} r_{i',j} + \sum_{[i,j']\in\sigma} r_{i,j'}\right)\mathbb{E}(\mathcal{T}^+_{i,j})$$

$$= (a+b)\sum_{[i,j]\in\sigma} \mathbb{E}(\mathcal{T}^+_{i,j})$$

On the other hand, by Lemma 62,

$$\mathbb{E}\left(\sum_{[i,j]\in\sigma} r_{ij}R_{i,j}\right) = \sum_{[i,j]\in\sigma} r_{ij}\left(1 + \sum_{c\in C_{ij}(\sigma)} Pr_\mu(c)\right) \qquad (2.47)$$

$$= \sum_{[i,j]\in\sigma} r_{ij} + \sum_{[i,j]\in\sigma} r_{ij}\sum_{c\in C_{ij}(\sigma)} Pr_\mu(c)$$

$$= ab + \sum_{c\in C(\sigma)}\left(\sum_{\substack{[i,j]\in\sigma \text{ with} \\ C_{ij}(\sigma)\ni c}} r_{ij}\right)Pr_\mu(c)$$

As depicted in Figure 2.12, for any corner $c \in C(\sigma)$ occurring at (x,y) and

143

for any $[i, j] \in \sigma$, we have $c \in C_{ij}(\sigma)$ if and only if $(x \geq i$ and $y \geq j)$ or $(x < i$ and $y < j)$. Let $c \in C(\sigma)$ be a southeast corner occurring at (x, y). We have

$$\sum_{\substack{[i,j]\in\sigma \text{ with} \\ C_{ij}(\sigma)\ni c}} r_{ij} = \sum_{[i,j]\in\sigma} r_{ij} - \sum_{\substack{[i,j]\in\sigma \\ i\leq x; j>y}} r_{ij} - \sum_{\substack{[i,j]\in\sigma \\ i>x; j\leq y}} r_{ij}$$

$$= ab - (a - x)b - (b - y)a$$

$$= ab \left(\frac{x}{a} + \frac{y}{b} - 1 \right).$$

With similar calculations we can see for any $c \in C(\sigma)$ a northeast corner occurring at (x, y) we also have $\sum_{\substack{[i,j]\in\sigma \\ C_{ij}(\sigma)\ni c}} r_{ij} = ab(1 - \frac{x}{a} - \frac{y}{b})$. In other words, for $c \in C(\sigma)$,

$$\sum_{\substack{[i,j]\in\sigma \\ c\in C_{ij}(\sigma)}} r_{ij} = ab \cdot \delta(c). \tag{2.48}$$

Putting equations (2.46), (2.47) and (2.48) together yields

$$(a + b) \sum_{[i,j]\in\sigma} \mathbb{E}(\mathcal{T}_{i,j}^+(\sigma)) = ab \left(1 + \sum_{c\in C(\sigma)} \delta(c) Pr(c) \right).$$

But since μ is a toggle-symmetric measure, $\mathbb{E}_\mu(\text{jag}) = 2\sum_{[i,j]\in\sigma} \mathbb{E}(\mathcal{T}_{i,j}^+(\sigma))$. Hence the claimed formula for $\mathbb{E}_\mu(\text{jag})$ holds. $\square$

Let us say a skew shape σ is *balanced* if it is connected and $\delta(c) = 0$ for all $c \in C(\sigma)$. In other words, a connected skew shape is balanced if all outward corners occur at the main anti-diagonal. An immediate corollary of
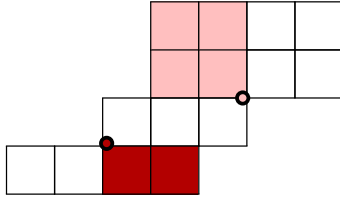
144

Figure 2.12: In the above diagram, let X_1 be the set of pink boxes and X_2 the set of dark red boxes. Let c_1 be the corner occurring at $(2,5)$ (in pink) and c_2 the corner at $(3,2)$ (in dark red). Then $[i,j] \in X_1$ if and only if $c_1 \in C_{i,j}(\sigma)$ and $[i,j] \in X_2$ if and only if $c_2 \in C_{i,j}(\sigma)$.

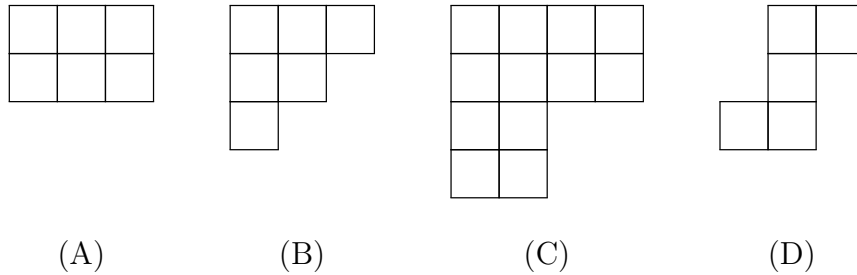(A)   (B)   (C)   (D)

Figure 2.13: Examples of balanced skew shapes.

our main theorem is the following:

Corollary 64. *Let σ be a balanced skew shape with height a and width b. Let μ be any toggle-symmetric probability distribution on $J(\mathcal{P}_\sigma)$. Then the expected jaggedness of a subshape in $J(\mathcal{P}_\sigma)$ with respect to the distribution μ is $\frac{2ab}{a+b}$.*

Some examples of balanced skew shapes are depicted in Figure 2.13. They include rectangles like (A), staircases like (B), "stretched" staircases (i.e., staircases where we have replaced each box by a $k \times l$ rectangle) like (C), as well as other more general shapes like (D). There are a total of $3^{\gcd(a,b)-1}$ balanced skew shapes with height a and width b for any $a, b \geq 1$.

145

Along the same lines, say that a skew shape σ is *abundant* if all of its northwest corners occur on or above its antidiagonal and all of its southeast corners occur on or below its antidiagonal. Let us say that σ is *deficient* if all of its northwest corners occur on or below the antidiagonal and all of its southeast corners occur on or above the antidiagonal. Then we immediately get:

Corollary 65. *Let σ be a skew shape of height a and width b and μ be any toggle-symmetric probability distribution on $J(\mathcal{P}_\sigma)$.*

- *If σ is abundant, then the expected jaggedness of a subshape in $J(\mathcal{P}_\sigma)$ with respect to the distribution μ is at least $2ab/(a+b)$.*

- *If σ is deficient, then the expected jaggedness of a subshape with respect to μ is at most $2ab/(a+b)$.*

2.4.2 Connections to antichain cardinality homomesy

In this subsection we give an application of our main result to the study of homomesies in combinatorial maps. Recall the definitions of rowmotion, and homomesy from Section 2.3.

Let $\mathcal{P}$ be a poset. To any $I \in J(\mathcal{P})$ we associate the antichain $A(I)$ of $\mathcal{P}$ consisting of the maximal elements of I. The **antichain cardinality statistic** is the map $f : J(\mathcal{P}) \to \mathbb{R}$ given by $f(I) = |A(I)|$.

Corollary 66. *If $\mathcal{P}$ is the poset associated to the skew shape σ and μ is any toggle-symmetric distribution, then*

$$\mathbb{E}_\mu(f(I)) = \frac{ab}{a+b}\left(1 + \sum_{c \in C(\sigma)} \delta(c)\, Pr_\mu(c)\right).$$

Proof. This result was already obtained in the proof of Theorem 60. Explicitly, the antichain cardinality statistic is just $\sum_{p \in P_\sigma} \mathcal{T}_p^-$, so the average of this statistic is

$$\mathbb{E}_\mu\left(\sum_{p \in P_\sigma} \mathcal{T}_p^-\right) = \frac{1}{2}\left(\mathbb{E}_\mu\left(\sum_{p \in P_\sigma} \mathcal{T}_p^-\right) + \mathbb{E}_\mu\left(\sum_{p \in P_\sigma} \mathcal{T}_p^+\right)\right)$$

$$= \frac{1}{2}\mathbb{E}_\mu(\mathrm{jag})$$

where $\mathbb{E}_\mu(\sum_{p \in P_\sigma} \mathcal{T}_p^-) = \mathbb{E}_\mu(\sum_{p \in P_\sigma} \mathcal{T}_p^+)$ thanks to the toggle-symmetry of μ. Now apply Theorem 60. $\qquad\square$

Corollary 67. *For $\mathcal{P}_\sigma$ the poset corresponding to a balanced skew shape σ of height a and width b, the antichain cardinality statistic is $\frac{ab}{a+b}$-mesic with respect to the action of rowmotion on $J(\mathcal{P}_\sigma)$.*

Proof. Let $\mathcal{O} \subseteq J(\mathcal{P}_\sigma)$ be a ϱ-orbit and let μ be the distribution on $J(\mathcal{P}_\sigma)$ that is uniform on $\mathcal{O}$. By a result of Striker [48] we know that μ is toggle-symmetric. Thus, by Corollaries 64 and 66 we conclude that $\mathbb{E}_\mu(\#A(I)) = \frac{ab}{a+b}$. $\qquad\square$

147

The above corollary is a result of Propp and Roby [7] if σ has a rectangular shape, and our Theorem provides another proof to it. For non-rectangular, balanced σ the homomesy result appears to be new.

2.5 Conclusion

Remember the butterfly poset K and the set of its ideas $J(K)$ from Example 11 in Section 2.2. We observed in Section 2.2 that the natural Markov chain defined by toggling is not rapidly mixing for the butterfly poset.

Let $f : J(K) \to \mathbb{R}$ be a function that is homomesic in rowmotion orbits of $\mathcal{Q}_{n,n}$. We are interested to find the average of f in K. (See Theorem 32, Section 2.3.)

We have no knowledge whether or not $\langle K, \varrho, f \rangle$ exhibits homomesy, and in fact for particular choices of f, for instance the size of the ideal, we can observe that it does not. However, looking at the ϱ-orbits of K, we notice that all the orbits are the same orbits as in $\mathcal{Q}_{n,n}$ or in $\mathcal{Q}_{m,m}$ except the one orbit that contains $\emptyset$. Let this orbit be $\mathcal{O}$. Applying rowmotion to $\emptyset$ repeatedly, we will obtain an orbit of size $2(n + m) - 2$. By Theorem 32, Section 2.3, we know the average of f on other orbits, and also size of the orbits containing $\emptyset$ in $\mathcal{Q}_{n,n}$ and $\mathcal{Q}_{m,m}$ it remains to find the average of f in $\mathcal{O}$ and we will be able to find the average of f in $J(\mathcal{Q})$. Considering the fact that $|\mathcal{O}|$ is polynomial in n and m, this can be done in polynomial time.

The above example suggests that in cases were calculating the average of $f : S \to \mathbb{R}$ is an NP-hard problem, we might be able to use homomesy, and with a similar approach, simplify S by removing those orbits of it in which we have some knowledge of the average and then, approximating the average

of f on remaining parts. Unfortunately, at this point we do not have any example to which we can apply this idea.

Integer Partitions, Young Tableaux, and $q-$binomials

The contents of this appendix can be found in Richard Stanley's course: "Topics in Algebraic Combinatorics", Chapter 6 (See [59].)

For an integer n, an integer partition of it is $\lambda = (\lambda_1 \geq \ldots \geq \lambda_k)$ which is a sequence of weakly decreasing positive integers satisfying $n = \sum_{i=1}^k \lambda_i$.

Definition 54. A *Young diagram* is a finite collection of left justified boxes with the row lengths weakly decreasing. The size of a Young diagram is the number of boxes in it.

Note that to any integer partition λ there is a Young diagram where λ_i is the number of boxes in the ith row.

Given two partitions λ and ν of the numbers ℓ and n respectively, we say that $\nu \subseteq \lambda$ if $\nu_i \leq \lambda_i$ for all i. We use the usual convention that $\lambda_i = 0$ if i is greater than the number of parts of λ. We use English notation when drawing partitions, so for instance the Young diagram corresponding to the

partition $\lambda = (4, 3)$ is
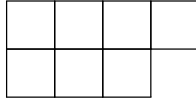
Definition 55. Let $\nu \subseteq \lambda$ be two partitions. The diagram obtained by subtracting the Young diagram of ν from the Young diagram of λ is called a *skew Young diagram* or *skew shape*. We will write $\sigma = \lambda/\nu$ for this shape.

A Young diagram or a skew diagram can also be understood as a lattice. Take the partial order to be: box x is less than box y if and only if x is located north west of y.

Definition 56. A *Young tableau* (A skew tableau) is a one to one an assignment of numbers $1, 2, \ldots, n$ to the bode in a Young diagram (skew diagram) where n is the size of it.

Note that any Young (skew) tableau can be understood as a linear ordering of the lattice associated to the corresponding diagram.

The Gaussian binomial coefficients or q-binomial coefficients are polynomials defined in analogy to the binomial coefficients. For $r \leq m$ they are defined by:

$$\binom{m}{r}_q = \frac{(1 - q^m)(1 - q^{m-1}) \ldots (1 - q^{m-r+1})}{(1 - q)(1 - q^2) \ldots (1 - q^r)}. \tag{49}$$

It is known that the generating function for the number of Young diagrams fitting in an $m \times n$ box or equivalently the number of integer partitions having at most n parts each smaller than m equals $\binom{m}{n}_q$.

Gaussian binomial coefficients have many properties similar to binomial coefficients such as the following:

$$\binom{m}{r}_q = \binom{m}{m-r}_q.$$

$$\binom{m}{r}_q = \binom{m-1}{r}_q + q^{m-r}\binom{m-1}{r-1}_q.$$

Proposition 7. *If $q < 1/2$ we will have,*

$$\binom{m}{r}_q = \frac{(1-q^m)(1-q^{m-1})...(1-q^{m-r+1})}{(1-q)(1-q^2)...(1-q^r)}$$

$$= \prod_{i=1}^{r}(1 - q^{m-i+1})/(1 - q^i)$$

$$< \prod_{i=1}^{r} 1/(1 - q)$$

$$< 2^r < \left(\tfrac{1}{q}\right)^r.$$

Proposition 8. *The q-binomial is always a polynomial of q with positive coefficients. Hence,*

$$q < q' \Rightarrow \binom{m}{r}_q < \binom{m}{r}_{q'}. \tag{50}$$

153

Bibliography

[1] A. Brouwer and A. Schrijver. *On the Period of an Operator Defined on Antichains.* (1974) Math. Centrum report ZW 24/74. Pages 1–13.

[2] P. Cameron and D. Fon-Der-Flaass. *Orbits of Antichains Revisited.* (1995) *European Journal of Combinatorics* Vol. 16. Pages 545–554.

[3] D. Armstrong, C. Stump, and H. Thomas. *A Uniform Bijection Between Nonnesting and Noncrossing Partitions.* (2013) Trans. Amer. Math. Soc. Vol. 365. Pages 4121–4151.

[4] D. Panyushev. *On Orbits of Antichains with Positive Roots.* (2009) European Journal of Combinatorics. Vol. 30. Pages 586–594.

[5] J. Striker and N. Williams. *Promotion and Rowmotion.* European Journal of Combinatorics Vol. 33. Pages 1919–1942.

[6] D. B. Rush, and X. Shi. *On Orbits of Order Ideals of Minuscule Posets.* (2013) DMTCS Proceedings, 25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC).

[7] J. Propp and T. Roby. *Homomesy in Products of Two Chains.* (2013) DMTCS Proceedings, 25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC).

[8] M. Chan, S. Haddadan, S. Hopkins, and L Moci. *The Expected Jaggedness of Order Ideals.* `arXiv:1507.00249`.

[9] D. Einstein and J. Propp. *Piecewise-linear and Birational Toggling.* (2014) DMTCS Proceedings, 26th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC).

[10] S. Hopkins and I. Zhang. *A Note on Statistical Averages for Oscillating Tableaux.* (2015) *The Electronic Journal of Combinatorics.* Vol. 22. Pages 2–48.

[11] G. Pólya *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen.* (1937) Acta Mathematica. Vol. 68 (1). Pages 145–254.

[12] W. T. Tutte. *Graph Theory.* (2001) Cambridge University Press.

[13] G. R. Brightwell, P. Winkler. *Counting Eulerian circuits is #P-complete.* (2005) Analytic Algorithmics and Combinatorics (ANALCO). Pages. 259–262.

[14] P. Tetali, S. Vempala. *Random Sampling of Euler Tours.* (2001) Algorithmica. Vol. 30. Pages 376–385.

[15] P. Chebolua, M. Cryanb, and R. Martina. *Exact Counting of Euler Tours for Generalized Series-Parallel Graphs.* (2012) Journal of Discrete Algorithms. Vol. 10. Pages. 110–122.

[16] P. J. Creed. *Counting and Sampling Problems on Eulerian Graphs.* PhD thesis of P. J. Creed, Edinburgh Research Archive `https://www.era.lib.ed.ac.uk/handle/1842/4759`

[17] P. A. Pevzner, H. Tang, and M.S. Waterman. *An Eulerian trail approach to DNA fragment assembly.* (2001) Proceedings of the National Academy of Sciences of United States of America. Vol. 98 (17). Pages. 9748–9753.

[18] R. Kuntal. *Optimum Gate Ordering of CMOS Logic Gated Using Euler Path Approach: Some Insights and Explanations.'* (2001) Journal of Computing and Information Technology. Vol. 15 (1). Pages. 85–92.

[19] D. Einstein, M. Farber, E. Gunawan, M. Joseph, M. Macauley, J. Propp, and S. Rubinstein-Salzedo *Noncrossing Partitions, Toggles, and Homomesies.* `arXiv:1510.06362`

[20] S. Haddadan and P. Winkler. Mixing of Permutations by Biased Transposition. `arXiv:1604.00870`

[21] M. Chan, S. Haddadan, S. Hopkins, and L. Moci. *The Expected Jaggedness of Order Ideals.* `arXiv:1507.00249`

[22] R. Bubley and M. Dyer. *Faster Random Generation of Linear Extensions.* (1999) Discrete Mathematics. Vol. 201. Pages. 81–88

[23] P. Bhakta, S. Miracle, D. Randall and A. Streib. *Mixing Times of Markov Chains for Self-Organizing Lists and Biased Permutations.* (2013) 24th Symposium on Discrete Algorithms (SODA).

[24] G. R. Brightwell and P. Winkler. *Counting Linear Extensions.* (1991) ORDER. Vol. 8. Pages. 225–242.

[25] S. Haddadan. *Some Instances of Homomesy among Ideals of Posets*, arXiv:1410.4819

[26] A. E. Holroyd and J. Propp. *Rotor walks and Markov chains.* (2010) Algorithmic Probability and Combinatorics. AMS. Pages.105–126.

[27] O. Angel and A. E. Holroyd. *Rotor Walks on General Trees.* (2011) SIAM Journal of Discrete Mathematics. Vol. 25(1). Pages. 423–446.

[28] T. Friedrich, T. Sauerwald. *The Cover Time of Deterministic Random Walks.* (2010) The Electronic Journal of Combinatorics. Vol. 17.

[29] B. Doerr , T. Friedrich , and T. Sauerwald. *Quasirandom Rumor Spreading.* Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms. Pages 773–781.

[30] A. Karzanov and L. Khachiyan. *On the Conductance of Order Markov Chains* (1991) ORDER Vol. 8. Pages. 7–15.

[31] M. Jerrum. *A Very Simple Algorithm for Estimating the Number of k-Colourings of a Low-Degree Graph.* (1995) Random Structures and Algorithms. Vol. 7. Pages. 157–165.

[32] E. Vigoda. *Improved Bounds for Sampling Colorings.* (1999) 54th Annual Symposium on Foundations of Computer Science. Pages. 51–59.

[33] M. Dyer and A. Frieze. *Computing the Volume of Convex Bodies: A Case Where Randomness Provably Helps.* (1991) In Probabilistic Combinatorics and its Applications, Proceedings of AMS Symposia in Applied Mathematics. Vol. 44. Pages. 123–170.

[34] L. Lovász, *Hit-and-Run Mixes Fast.* (1998)Mathematical Programming Vol. 86. Pages 443–461.

[35] R. Kannan, L. Lovász and M. Simonovits. *Random Walks and an $\mathcal{O}^*(n^5)$ Volume Algorithm.* (1997) Vol. 11. Random Structures and Algorithms. Pages 1–50.

[36] D. Randall, P. Winkler. *Mixing Points on an Interval.* (2005) Analytic Algorithmics and Combinatorics (ANALCO). Pages. 216–221.

[37] O. Håggström, *Finite Markov Chains and Algorithmic Applications Matematisk statistik.* Cambridge University Press.

[38] J. Håstad. *The Square Lattice Shuffle.* (2006) Random structures and Algorithms. Vol 29. Pages. 466-474.

[39] M. Held and R. Karp. *A Dynamic Programming Approach to Sequencing Problems.* (1962). SIAM 10. Pages. 196–210.

[40] M. Held, R. Karp, and R. Shareshian. *Assembly line balancing.* (1967). Operational Research. Vol. 11. Pages. 442–458.

[41] E. L. Lawler. *Efficient Implementation of Dynamic Programming Algorithms for Sequencing Problems.* (1979). Report BW 106/79, Stichting Mathematisch Centrum, Amsterdam.

[42] R. H. Möhring. *Scheduling Problems with Singular Solution.* (1982). Annals of Discrete Mathematics. Vol. 16. Pages. 225–239.

[43] J. S. Provan and M. O. Ball. *The Complexity of Counting Cuts and of Computing the Probability that a Graph is Connected.* (1983). SIAM Journal of Computing. Vol. 12. Pages. 777–788.

[44] G. Steiner. *On estimating the Number of Order Ideals in Partial Orders, with Some Applications.* (1993). Journal of Statistical Planning and Inference. Vol. 34. Pages. 281–290.

[45] D. G. Corneil, H. Lerchs, and L. Stewart. *Complement Reducible Graphs.*(1981) Discrete Applied Mathematics. Vol. 3. Pages. 163–174.

[46] G. Steiner. *An Algorithm to Generate the Ideals of a Partial Order.* (1986). Operational Research Letters. Vol. 5. Pages. 317–320.

[47] J. Bloom, O. Pechenikb, and D. Saracinoc. *Proofs and Generalizations of a Homomesy Conjecture of Propp and Roby.* (2016) Discrete Mathematics Vol. 339. Pages 194–206.

[48] J. Striker. *The Toggle Group, Homomesy, and the Razumov-Stroganov Correspondence.* (2015) the Electronic Journal of Combinatorics. Vol. 22(2).

[49] M. Chan, A. López Martón, N. Pflueger, and M. Teixidor i Bigas. *Genera of Brill-Noether curves and staircase paths in Young tableaux—.* (2015) `arXiv:1506.00516`.

[50] P. Diaconis and M. Shahshahani. *Generating a Random Permutation with Random Transpositions.* (1981) Probability Theory and Related Fields. Vol. 57. Pages. 159–179.

[51] J. Fill. *An Interesting Spectral Gap Problem.* (2003) Unpublished manuscript.

[52] D. Wilson, *Mixing times of Lozenge Tiling and Card Shuffling Markov Chains.* (2004) The Annals of Applied Probability. Vol. 14. Pages. 274–325.

[53] P. Diaconis and L. Saloff-Coste. *Comparison Techniques for Random Walks on Finite Groups.* (1993) The Annals of Applied Probability. Vol. 21. Pages. 2131–2156.

[54] D. Randall and P. Tetali, *Analyzing Glauber Dynamics by Comparison of Markov Chains.* (2000) Journal of Mathematical Physics.

[55] I. Benjamini, N. Berger, C. Hoffman and E. Mossel. *Mixing times of the biased card shuffling and the asymmetric exclusion process.* (2005) Transactions of the American Mathematical Society. Vol. 357. Pages 3013–3029.

[56] M. Jerrum and A. Sinclair. *Approximating the Permanent.* (1989) SIAM Journal on Computing. Vol. 18. Pages. 1149–1178.

[57] A. Sinclair. *Improved Bounds for Mixing Rates of Markov cCains and Multicommodity Flow.* (1992) Combinatorics, Probability, and Computing. Vol. 1. Pages. 351–370.

[58] M. Mihail and P. Winkler. *On the number of Eulerian Orientations of a Graph.* (1995) Algorithmica Vol. 6. Pages. 402–414.

[59] R. P. Stanley, *Topics in Algebraic Combinatorics.* (2012) Course notes for Mathematics, Citeseer.
http://www-math.mit.edu/∼rstan/algcomb/algcomb.pdf.

[60] A. Sinclair. *Algorithms for Random Generation and Counting: a Markov Chain Approach. Ph.D Thesis.* University of Edinburgh, June 1988.

[61] P. Diaconis and D. Strook. *Geometric Bounds for Eigenvalues of Markov Chains.* (1991).Annals of Applied Probability. Pages. 36–61.

[62] A. Sinclair and M. Jerrum, *Approximate Counting, Uniform Generation, and Rapidly Mixing Markov Chains.* (1989) Information and Computation. Vol. 82. Page. 93–133.

[63] R. I. Oliveira. *Mixing of the Symmetric Exclusion Processes in Terms of the Corresponding Single-Particle Random Walk.* (2013) Annals of Probability. Vol. 41. Pages. 871-913.

[64] R. Bubley and M. E. Dyer. *Path coupling: a Technique for Proving Rapid Mixing in Markov Chains.* (1997) Proceedings of the 38th IEEE Symposium on Foundations of Computer Science (FOCS), Pages. 223–231.

[65] B. Morris. *The Mixing Time for Simple Exclusion.* (2006) Annals of Applied Probability. Vol. 16.

[66] P. Diaconis and L. Saloff-Coste. *Comparison Techniques for Random Walks on Finite Groups.* (1993) The Annals of Applied Probability. Vol. 21. Pages. 2131-2156.

[67] R. Martin and D. Randall. *Disjoint Decomposition of Markov Chains and Sampling Circuits in Cayley Graphs.* (2006) Combinatorics, Probability, and Computing . Vol.15. Pages. 411-448.

[68] R. Rivest. *On Self-Organizing Sequential Search Heuristics.* (1976) Communications of the ACM. Vol. 19 (2).

[69] J. H. Hester and D. S. Hirschberg. *Self-Organizing Linear Search*(1985) ACM Computing Surveys. Vol. 19. Pages. 295–311.

[70] M .R. Jerrum and A. J. Sinclair. *Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains.* (989) Journal Information and Computation. Vol. 82. Pages 93–133.

[71] M. Mihail and P. Winkler, *On the Number of Eulerian Orientations of a Graph* (1995) Algorithmica. Vol. 16. Pages 402–414.

[72] D. Aldous, P. Diaconis *Shuffling Cards and Stopping Times.* (1986) Math'l Monthly. Vol. 93(5). Pages. 333–348.

[73] N. Caspard, B. Leclerc, and B. Monjardet. *Finite Ordered Sets, Concepts, Results and Uses.* (2012) Cambridge University Press.

[74] A. Karzanov and L. Khachiyan. *On the Conductance of Order Markov Chains.* (1991) Order 8, Pages 7–15.

[75] G. Brightwell and P. Winkler. *Counting Linear Extensions.* (1991) Order. Vol. 8. Pages 225-242.

[76] D. Randall and P. Winkler. *Mixing Points on a Circle.* (2005) 9th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM). Vol. 3624. Pages 426–435.

[77] R. Martin and D. Randall. *Disjoint Decomposition of Markov Chains and Sampling Circuits in Cayley graphs.* (2006) Combinatorics, Probability and Computing. Vol. 15, Pages 411–448.

[78] V. Guruswami. *Rapidly mixing Markov chains: A Comparison of Techniques. (A Survey)*

[79] M. Mihail. *Conductance and Convergence of Markov chains: a Combinatorial Treatment of Expanders.* (1989) Proceeding of the 30th Foundations of Computer Science, 30th Annual Symposium (FOCS). Pages 526–531.

[80] D. A. Levin, Y. Peres , and E. L. Wilmer. *Markov Chains and Mixing Times.* (2009) American Mathematical Society. http://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf

[81] D. Bayer and P. Diaconis *Trailing the Dovtail Shuffle to its Lair.*(1992) The Annals of Applied Probability Vol. 2. Pages. 294–313.

[82] N. Madras and D. Randall. *Markov chain Decomposition for Convergence Rate Analysis.* (2002). Annals of Applied Probability. Vol. 12. Pages. 58–606.

[83] P. L. Erdös, I Miklás, and Z. Toroczkai. *A Decomposition Based Proof for Fast Mixing of a Markov chain over Balanced Realizations of a Joint Degree Matrix.* (2015) SIAM Journal on Discrete Mathematics.

[84] H. Badingera, W. H. Reutera. *Measurement of Fiscal Rules: Introducing the Application of Partially Ordered Set (POSET) Theory.* (2015) Journal of Macroeconomics, Vol. 43. Pages. 108–123.

[85] S. Greenberg, A. Pascoe, and D. Randall. *Sampling Biased Lattice Configurations using Exponential Metrics.* (2009) 20th Symposium on Discrete Algorithms (SODA).

[86] A. Sly. *Computational transition at the uniqueness threshold Conference* (2010). In Foundations of Computer Science, 51st Annual Symposium (FOCS). Pages 287–296.

[87] D. Weitz. *Counting Independent Sets up to the Tree Threshold.* (2006). STOC : Proceedings of the thirty-eighth annual ACM symposium on Theory of computing.Pages. 140–149.

[88] A. Sinclair, P. Srivastava , and M. Thurley *Approximation Algorithms for Two-State Anti-Ferromagnetic Spin Systems on Bounded Degree Graphs* (2014) Journal of Statistical Physics. Vol. 155. Issue 4. Pages. 666–686.

[89] M. E. Dyer, A. M. Frieze, and M. R. Jerrum. *On Counting Independent Sets in Sparse Graphs.*(1999) In 40th Annual Symposium on Foundations of Computer Science. Pages. 210–217.

[90] E. Vigoda *A note on the Glauber dynamics for sampling Independent Sets.* (2001). Electronic Journal of Combinatorics. Vol. 8(1).

[91] M. Luby and E. Vigoda. *Fast convergence of the Glauber Dynamics for Sampling Independent Sets.* (1999) Random Structures and Algorithms. Vol.15(3-4). Pages. 229–241.

[92] D. Aldous. *Random walks on finite groups and rapidly mixing Markov chains.*(1983) Śeminnaire de Probabilit és. Springer Lecture Notes in Mathematics. Pages. 243-297.