Total variation cutoff for random walks on some finite groups

A Dissertation
submitted in partial fulfillment
of the requirements for the award of the
degree of
Doctor of Philosophy

by
Subhajit Ghosh

Department of Mathematics
Indian Institute of Science
Bangalore - 560012
August 2020

Declaration

I hereby declare that the entire work embodied in this thesis has been carried out by me under the supervision of **Prof. Arvind Ayyer** at the Department of Mathematics, Indian Institute of Science, Bangalore. I further declare that no part of it has been submitted for the award of any degree or diploma of any University or Institution previously.

Subhajit Ghosh
S. R. No. 10-06-01-10-31-13-1-10797

Indian Institute of Science,
Bangalore,
August, 2020.

Prof. Arvind Ayyer
(Research advisor)

Dedication

To my parents, grandparents
&
teachers.

Acknowledgments

First and foremost, I would like to express my most sincere gratitude and indebtedness to my academic advisor **Prof. Arvind Ayyer** for his continuous support, encouragement and patience throughout my graduate student life at IISc. I am thankful to him for allowing me to work under his guidance, enduring many long discussions and parsing my scribbles on his chalkboard. I convey my profound thanks to him for introducing me to the beautiful world of probability theory, combinatorics and graph theory. His insightful academic suggestions not only helped me in my research work but also motivated me in improving myself. I extend my sincere thanks to him for providing me the opportunities to visit some of the best conferences/workshops. I am highly obliged to him for going through all the drafts of my research articles including this thesis meticulously and helping me to improve them.

I would like to extend my sincere thanks and gratefulness to the faculty members of the Department of Mathematics at IISc for offering a wide variety of courses. I really benefited a lot from many of these courses. I am thankful to be a teaching assistant to **Prof. Mrinal K. Ghosh**, **Prof. Manjunath Krishnapur**, **Prof. R. Venkatesh** and my advisor. It was not only a teaching experience but also a great learning experience for me. I would like to extend my humble gratitude and sincere thanks to **Prof. Amritanshu Prasad**, **Prof. Pooja Singla** and my advisor for proposing the problem of Chapter 3. I express my special thanks to **Dr. Kartick Adhikari** for a valuable comment which helped me improve a result of Chapter 5.

It is my pleasure to have a wonderful class of colleagues/friends at the Department of Mathematics, IISc. I would like to express my thanks to all of them. My special thanks to Abu Sufian, Amar Deep Sarkar, Anwoy Maitra, Arpan Kabiraj, Arindam Mandal, Dipankar Roy, G.V.K. Teja, Haripada Sau, Himanshu Gupta, Kartick Adhikari, Kartick Ghosh, Lakshmi Priya M.E., Mayuresh Londhe, Md Ramiz Reza, Monojit Bhattacharjee, Neeladri Maitra, Nimisha Pahuja, Nishu Kumari, Pabitra Barman, Paramita Pramanick, Poornendu Kumar, Pramath A.V, Prateek Kumar Vishwakarma, Purba Banerjee, Rahul Biswas, Ramesh Chandra Sau, Ritwik Pal, Sahil Gehlawat, Samarpita Ray, Sanjoy Kumar Jhawar, Somnath Hazra, Soumitra Ghara, Shubham Rastogi and all my Integrated Ph.D. batch-mates for many helpful academic and non-academic discussions. I am also thankful to all the Mathematics Department staff at IISc for their help in carrying out all the complex administrative formalities smoothly during my stay at IISc. Needless to say about the beautiful academic environment and facilities of the Mathematics Department for carrying out my research. I would like to acknowledge the IISc integrated Ph.D. scholarship and the support in part by a UGC Centre for Advanced Study Grant.

The first two chapters of the thesis were written during my visit to **Prof. Murali K. Srinivasan** at IIT Bombay. I would like to express my sincere gratitude to him for many insightful academic discussions during the research visit. I extend my sincere thanks and gratefulness to him and my advisor for arranging this research visit. I am thankful to my colleagues/friends Anup Kumar Mandal, Arghya Pramanik, Buddhadev Hajra, Dibyendu Mondal, Hiranya Kishore Dey, Indrajit Saha and Sibaprasad Barik for making my visit pleasant, which helped me a lot in writing the first phase of this thesis.

I would like to acknowledge the facilities and beautiful atmosphere at IISc campus which makes it my second home. I would like to express my special thanks to Anshuk Howlader, Apurba Kumar Pal, Debabrata Adhikari, Debarghya Mallick, Krishna Mohan Kumar, Milan Kumar Hazra, Pradip Kumar Bera, Prodip Howlader, Priyanka De and Ratan Sarkar for making my IISc life beautiful.

I convey my profound thanks to my friends Gobinda Sau, Pranab Sarkar and Pritam Ganguly for their enormous help in various aspects. Their support in many tough situations is unforgettable. It was worth discussing the academic and non-academic issues with them. Needless to mention about the good time we have spent together at IISc. I am very fortunate to have this trio during my life at IISc.

I take this opportunity to express my gratitude to all my teachers. I extend my sincere thanks to all the teachers of Serampore college Mathematics Department during $2010-2013$, for their help in my undergraduate student life. I would like to express my sincere gratitude to Barun Babu and late Gobinda Babu for teaching me high school/pre-college mathematics during my school life.

Many others whom I have failed to mention in this abbreviated list, my sincere thanks to all of them...

Finally, I wish to express my humbleness to my parents for everything they have provided me to come to this stage of my career and pursue my future. I owe so much to them for their unconditional love and support without which I would not have gotten to where I am today. I am thankful to my grandparents, aunts specially Bandana Pal (nee Ghosh) and uncles for their love and support.

Abstract

This thesis studies mixing times for three random walk models. Specifically these are random walks on the alternating group, the group of signed permutations and the complete monomial group. The details for the models are given below:

The random walk on the alternating group: We investigate the properties of a random walk on the alternating group A_n generated by 3-cycles of the form $(i, n-1, n)$ and $(i, n, n-1)$. We call this the *transpose top-2 with random shuffle*. We find the spectrum of the transition matrix of this shuffle. We obtain the sharp mixing time by proving the total variation cutoff phenomenon at $\left(n - \frac{3}{2}\right) \log n$ for this shuffle.

The random walk on the group of signed permutations: We consider a random walk on the hyperoctahedral group B_n generated by the signed permutations of the form (i, n) and $(-i, n)$ for $1 \leq i \leq n$. We call this the *flip-transpose top with random shuffle* on B_n. We find the spectrum of the transition probability matrix for this shuffle. We prove that this shuffle exhibits the total variation cutoff phenomenon with cutoff time $n \log n$. Furthermore, we show that a similar random walk on the demihyperoctahedral group D_n generated by the identity signed permutation and the signed permutations of the form (i, n) and $(-i, n)$ for $1 \leq i < n$ also has a cutoff at $\left(n - \frac{1}{2}\right) \log n$.

The random walk on the complete monomial group: Let $G_1 \subseteq \cdots \subseteq G_n \subseteq \cdots$ be a sequence of finite groups with $|G_1| > 2$. We study the properties of a random walk on the complete monomial group $G_n \wr S_n$ generated by the elements of the form $(\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})$ and $(\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n))$ for $g \in G_n$, $1 \leq i < n$. We call this the *warp-transpose top with random shuffle* on $G_n \wr S_n$. We find the spectrum of the transition probability matrix for this shuffle. We prove that the mixing time for this shuffle is of order $n \log n + \frac{1}{2} n \log(|G_n| - 1)$. We also show that this shuffle satisfies cutoff phenomenon with cutoff time $n \log n$ if $|G_n| = o(n^\delta)$ for all $\delta > 0$.

Contents

Chapter 1

Introduction

In this thesis we study the sharp mixing time (also known as the *cutoff phenomenon*; the precise definitions will be given in Chapter 2) of certain random walks on some finite groups, using non-commutative Fourier analysis techniques. Roughly, the mixing time determines the number of transitions required for a convergent Markov chain to get close to the stationary distribution. We have used the total variation distance as the distance function between probability measures (or distributions) throughout this thesis. Similar convergence rate related questions on finite Markov chains have caught considerable attention of many probabilists, viz. random walks on combinatorial objects, random walks on algebraic structures [3, 8, 9, 10, 12, 11, 13, 14, 21, 25, 37, 49, 68]. The random walk we consider in this thesis can be described by the shuffle of some combinatorial objects (viz. cards, oriented cards).

We start with a survey of related results in Section 1.1. In Section 1.2 we describe some use of random walks on finite groups. Finally, we end this chapter by discussing the organization of this thesis in Section 1.3.

1.1 Brief literature review

In this section we briefly recall some landmark results related to the random walk problems considered in this thesis. We split this section into two subsections in order to present the card shuffling problems separately. First in Subsection 1.1.1, we introduce the card shuffling problems and describe the *riffle shuffle*. Then in Subsection 1.1.2, we discuss mixing time related known results for the random walks on some finite groups.

1.1.1 The card shuffling problems

Shuffling a deck of cards is used to randomize the arrangement of the cards in the deck. This randomization process is important for card players to avoid bias. In general casinos use shuffling machines as it gives them a few advantages (viz. increase difficulties for players to make predictions even if they are collaborating with the croupiers). The shuffling machines are carefully designed by using various complex shuffling schemes. Efficiency of a shuffling machine depends on the shuffling time of a deck. The shuffling time depends on the number of times the deck is required to shuffle for mixing up the cards. Mathematically, an arrangement of a deck of cards can be thought of as a permutation of the cards and shuffling is just permuting the cards. Thus the card shuffling problems can be modeled by considering them as random walks on the symmetric group. Indeed, card shuffling appears as one of the few examples given by Markov [69]. It then appears in the works of Émile Borel [23], and Kosambi and Rao [63]. In particular one can see the excellent historical discussion in [57]. In practice, card players mostly use the riffle shuffle, we give a brief description of this shuffle in the next paragraph. Also, there are many other shuffling schemes studied in the literature (ref. [23, 38, 39, 79]).

A model for the *riffle shuffle or dovetail shuffle* was introduced by Gilbert and Shannon and later independently by Reeds. In this model a deck of n cards is cut into two piles viz. pile A and pile B according to a binomial distribution with parameters $(n, \frac{1}{2})$. Then the two piles are riffled together by the following rule: If pile A has a cards and pile B has b cards, drop the next card from pile A with probability $\frac{a}{a+b}$ and from pile B with probability $\frac{b}{a+b}$. The dropping continues until both piles have been run through, using a new a and b at each stage. In 1992, Bayer and Diaconis [17] showed that the riffle shuffle of n cards satisfies cutoff phenomenon with cutoff time $\frac{3}{2}\log_2 n$.

1.1.2 Known results for random walks on some finite groups

We have already seen that the card shuffling problems can be described by considering them as random walks on symmetric group. The theory of random walks on finite groups took a new direction in 1981, when Diaconis and Shahshahani [40] introduced the use of non-commutative Fourier analysis techniques. Afterwards, some other techniques emerged to deal with random walks on finite groups, e.g., the coupling argument (Aldous [2] used this argument to give good bounds for random walks on the symmetric group generated by adjacent transpositions), the strong stationary time approach [70] (introduced by Aldous and Diaconis [4, 5]). A nice survey article for random walks on finite groups is [89]. Below, we note down mixing time related known results of random walks on some finite groups.

- *Random walk on n-cycle:* Diaconis [32, 33] studied the random walk on the additive group $\mathbb{Z}_n$, driven by the uniform measure on $\{1, -1\}$ i.e., a particle at any node of the n-cycle can move forward or backward with probability $\frac{1}{2}$ each. If n is odd, he showed that $O(n^2)$ steps are necessary and sufficient for the random walk to be within a fixed distance from the stationary distribution. But there is no cutoff phenomenon. He also proved that, without any condition on n, similar result holds for two lazy variants of this random walk given as follows: (i) The particle at any node of the n-cycle can stay at its position with probability $\frac{1}{2}$ and jump forward or backward with probability $\frac{1}{4}$ each. (ii) The particle at any node of the n-cycle can stay at its position with probability $\frac{1}{2}$ and jump forward with probability $\frac{1}{2}$.

- *Simple random walk on hypercube:* The elements of the additive group $\mathbb{Z}_2^n$ form the vertices of the n-dimensional hypercube. Let $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ be the vertex with a 1 in the ith position and zero elsewhere, $1 \leq i \leq n$ and $e_0 = (0, \dots, 0)$. The random walk on the n-dimensional hypercube is the random walk on the additive group $\mathbb{Z}_2^n$ driven by the probability measure defined uniformly on $\{e_0, e_1, \dots e_n\}$ i.e., particle at any vertex of the hypercube can stay at its position with probability $\frac{1}{n+1}$ and jump towards any neighbour with probability $\frac{1}{n+1}$ each [32, 33]. Random walk on the hypercube projects to the Ehrenfest model of diffusion [26, 41]. Both the random walk on the hypercube and Ehrenfest's diffusion model are well studied. It is known that $\frac{1}{4}n \log n$ steps are necessary and sufficient to reach uniformity for the above mentioned random walk on the hypercube [36].

- *Random walks on symmetric group:* Many random walk models have been considered on the symmetric group S_n. We list some notable ones below:

  1. *Random transposition:* This was considered by Diaconis and Shahshahani. Under this random walk, a permutation in S_n can either go to itself with probability $\frac{1}{n}$ or be multiplied by a transposition on the right with probability $\frac{2}{n^2}$. This can also be described as a shuffle of a deck of n cards. The shuffling scheme is the following: Either keep the deck as it is with probability $\frac{1}{n}$ or transpose two randomly chosen (distinct) cards from the deck with probability $\frac{2}{n^2}$. Diaconis and Shahshahani [40] showed that this random walk satisfies cutoff phenomenon with cutoff time $\frac{1}{2}n \log n$.
  2. *Random-to-top:* This model is an example of the Tsetlin library problem [97] when the probability of choosing any book is equally likely. We describe the Tsetlin library problem after describing this model. Under this random walk, a permutation in S_n is multiplied on the right by a permutation of the form

$(1, \ldots, i),\ \ 1 \leq i \leq n$, with probability $\frac{1}{n}$ each. As a card shuffling problem, this can be explained as follows: Choose a random card from the deck of n cards with probability $\frac{1}{n}$ and place it on the top of the deck. Aldous and Diaconis [4] showed that $O(n \log n)$ steps are sufficient for this random walk to reach near uniformity. Later Diaconis, Fill and Pitman [35], refined the earlier result [4] and showed cutoff at $n \log n$. A similar shuffle called *top-to-random* has also been studied in the literature.

3. *Tsetlin library problem:* This has arisen from the work of Tsetlin in [97]. The model is the following: Suppose a single self of a library contains n books $B_1, \ldots, B_n$. A reader can choose only one book at a time and after reading it, the reader places the book at the end of the self. The probability of book B_i to be selected and placed at the end of the self is p_i. Given $\pi \in S_n$, if $\pi_1 \ldots \pi_n$ denotes π in its one line notation, then under this random walk π can go to $\pi \cdot (k, k+1, \ldots, n)$ with probability p_{π_k} for $1 \leq k \leq n$. The stationary distribution $(\omega(\pi))_{\pi \in S_n}$ of this model and is given by $\omega(\pi) = \prod_{i=1}^{n} \frac{p_{\pi_i}}{p_{\pi_1} + \cdots + p_{\pi_i}}$. Hendricks [53, 54] studied this in the seventies. The eigenvalues of the transition matrix for the Tsetlin library problem is well studied. Donnelly [43], Kapoor and Reingold [60] and, Phatarfod [80] studied the convergence to stationarity for this random walk.

4. *Random-to-random:* In this shuffle, a random card is chosen from the deck with probability $\frac{1}{n}$ and place it in a random position chosen with probability $\frac{1}{n}$. Therefore under this shuffle a permutation in S_n can be multiplied on the right by a permutation of the form $(j, j-1, \ldots, 1)(1, \ldots, i),\ 1 \leq i, j \leq n$, with probability $\frac{1}{n^2}$. This was first considered by Diaconis and Saloff-Coste [38]. They proved that the mixing time is $O(n \log n)$. Dieker and Saliola [42], used representation theoretic arguments to find the eigenvalues and eigenvectors of this card shuffle. Some notable references on the bounds of the mixing time for this shuffle are the work of Qin and Morris [84], Saloff-Coste and Zúñiga [88], and Subag [95]. In a recent work, Bernstein and Nestoridi [20] proved cutoff for this shuffle with cutoff time $\frac{3}{4} n \log n - \frac{1}{4} n \log \log n$.

5. *Transpose top with random:* Under this random walk, a permutation in S_n can be multiplied on the right by a transposition of the form $(i, n),\ 1 \leq i \leq n$, with probability $\frac{1}{n}$ each. Thus this shuffle choose a random card from a deck of n cards with probability $\frac{1}{n}$ and transpose it with the top card of the deck. This shuffle was first studied by Flatto, Odlyzko and Wales [46]. This shuffle exhibits the total variation cutoff phenomenon with cutoff time $n \log n$ [32, 33].

6. *Adjacent transposition:* This shuffle sends a permutation π of S_n to itself with probability $\frac{1}{2}$ or to $\pi \cdot (i, i+1),\ 1 \leq i \leq n-1$ with probability $\frac{1}{2(n-1)}$. As a card shuffling problem this suggests to transpose a pair of adjacent cards with probability $\frac{1}{2}$ or do nothing with probability $\frac{1}{2}$. Aldous [2], showed that $O(n^3 \log n)$ steps are sufficient to mix the deck of n cards whereas $\Omega(n^3)$ steps are necessary. In 2004, Wilson [99] proved that $\frac{1}{\pi^2} n^3 \log n$ steps were necessary and that $\frac{2}{\pi^2} n^3 \log n$ were sufficient, and conjectured that the first was the correct answer. Lacoin [64] has proved this conjecture by showing that the deck is mixed after $\frac{1}{\pi^2} n^3 \log n(1 + o(1))$ shuffles.

- *Random walks on hyperoctahedral group:* Let B_n be the set of all arrangements of n distinct objects such that each object has two orientations viz. up and down. This set B_n forms a group under the composition of bijections, known as the *hyperoctahedral group* (Formal definition will be given in Chapter 4). We now describe two random walk models on the hyperoctahedral group B_n as follows:

  1. *The paired flip random walk:* Suppose we have n oriented (distinct) cards arranged in a row such that each card has two orientations viz. face up and face down. Then the paired flip random walk does one of the following:

     (a) Choose two cards randomly from the arrangement and transpose them after a random decision of changing their orientations or not with probability $\frac{1}{n^2}$.

     (b) Choose a card and change its orientation with probability $\frac{1}{2n^2}$.

     (c) Do nothing with probability $\frac{1}{2n}$.

     This can be thought of as a generalization of the random transposition model. Schoolfield has shown that $n \log n$ steps are sufficient for this walk to reach near uniformity [90, 91].

  2. *The arc reversal random walk:* Suppose we have an arrangement of n (distinct) labeled markers on a circle such that each marker has two orientations viz. up and down. The markers are placed in equally spaced positions on the circle. For $1 \leq i, j \leq n$, we write $[\widehat{\mathrm{i,j}}]$ to denote the arc starting at the marker at position i and ending at the marker at position j in clockwise direction. In particular, when $i = j$, then $[\widehat{\mathrm{i,j}}]$ denotes the marker at position i (or j). At every step this walk does the following: First choose integers i and j independently and uniformly from the set $\{1, \ldots, n\}$. Then detach the arc $[\widehat{\mathrm{i,j}}]$ and reattach it after a random decision of reversing it or not. By reversing an arc we mean the following: The labeling of the markers are reversed in that arc

and their orientations are changed. Schoolfield proved that $O(n \log n)$ steps are necessary and sufficient for this walk to reach near uniformity [90, 91].

- *Random transvection walk on $SL_n(\mathbb{F}_q)$:* Let $\mathbb{F}_q$ be the finite field with q elements. $SL_n(\mathbb{F}_q)$ is the group of $n \times n$ matrices with elements in $\mathbb{F}_q$ and determinant 1. A *transvection* [96] is a non-identity element of $SL_n(\mathbb{F}_q)$, which fixes all the points in a hyperplane of $(\mathbb{F}_q)^n$. An example of a transvection is $I + aE_{ij}$, where I is the identity matrix of $SL_n(\mathbb{F}_q)$, $a \in \mathbb{F}_q^*$ (the multiplicative group of $\mathbb{F}_q$) and E_{ij} is the element of $SL_n(\mathbb{F}_q)$ with the only non-zero entry 1 at the (i, j)th position. Hildebrand [55] proved cutoff for the random walk on $SL_n(\mathbb{F}_q)$ generated by random transvections with cutoff time n.

1.2 Applications of random walks on finite groups

In general, random walks on finite groups randomize the elements of the group. The convergence rate related questions for random walks on finite groups are useful in randomization algorithms. This has applications in many subjects including mathematics, computer science, biology, statistical physics and more. Some of them are listed below:

- *Application in mathematics:* The randomization processes and their convergent rate related questions are used in many areas of mathematics viz. number theory [61] (primality testing [30, 94]), data structure [44], algebraic identities [76] (polynomial [61] and matrix identity verification [98, 48], interactive proof systems [19]), mathematical programming (faster algorithms for linear programming [76], rounding linear program solutions to integer program solutions [61]), graph algorithms [76], counting and enumeration [78, 67], parallel and distributed computing [76], probabilistic existence proofs [6, 76] (existence ensured by the arrival of a combinatoral object with non-zero probability among objects drawn from a suitable probability space) and many other places.

- *Application in computer science:* Applications of randomization process in computer science includes Google page-rank algorithm [24, 65], load balancing procedure [15, 16] (e.g. distributing uniform load on the servers of a heavy traffic website), randomized binary searching [61, 93], network designing [1, 31, 62, 85], generating OTP (one time password), generating CAPTCHA (completely automated public Turing test to tell computers and humans apart), strong password suggester algorithm and more.

- *Application in statistical physics:* Metropolis Hastings, Glauber dynamics, Markov chain Monte Carlo statistical methods are based on the convergent rate related results of finite Markov chains [7, 66, 74]. There are rigorous connections of the models of finite Markov chain (which includes random walks on finite groups) with other models in statistical physics. Some examples of such models are the following: Ising model [28, 66, 82], random-cluster model [47, 52, 82], ice and dimer model [82], random graph models [6], models related to exclusion processes [28, 66] (in particular some card shuffling problems can be related to the exclusion process via a natural projection, see [64] for details), reconstruction problems on trees [72, 75], non-equilibrium particle systems [28].
- *Application in biology:* The mixing time for random walks on finite groups has many applications in biology viz. population genetics [18, 51], DNA shuffling [29, 100], chromosomes shuffling [45].

1.3 Organization of the thesis

Our models are inspired by the transpose top with random shuffle on the symmetric group S_n [32, 33, 46]. The models we consider in this thesis, are variants of the transpose top with random shuffle for the group of *even permutations*, *signed permutations* and the *complete monomial group*. Our goal is to study the mixing times for these random walk models. Our method connects these models to the Young-Jucys-Murphy elements for the underlying group. The other motivation is the landmark work of Diaconis and Shahshahani [40], which connects representation theory of finite group to probability via non-commutative Fourier analysis techniques.

In this thesis, we use the *Young-Jucys-Murphy* elements and the representation theory of the underlying group to find the spectrum of the transition matrix for the corresponding model. Then we make use of the upper bound lemma [32, Lemma 4.2] to find the sufficient number of transitions required to reach near stationary distribution. To find the number of transitions necessary to get close to the stationarity, we consider the number of fixed points of an action of the underlying group. The organization of this thesis is the following:

In Chapter 2, we discuss the background theory necessary for this thesis. Starting with a brief description of representation theory of finite group we discuss basic theory of discrete time Markov chain on finite state space and other relevant topics in this chapter.

Chapter 3 describes the *transpose top*-2 *with random shuffle* on the alternating group A_n (group of even permutations of S_n). Suppose we have a deck of cards labelled from 1 to n, such that the arrangement of the deck is a permutation in A_n. Then the transpose top-2 with random shuffle on A_n is a lazy variant of the following: First transpose the top two cards, then choose one of them and interchange it with a card randomly chosen from the remaining $(n-2)$ cards. In this chapter, we have shown that *the cutoff phenomenon for the transpose top-2 with random shuffle on A_n, with cutoff time* $\left(n-\frac{3}{2}\right)\log n$.

In Chapter 4, we consider the *flip-transpose top with random shuffle* on the group of signed permutations. First we consider this shuffle on the *hyperoctahedral group* B_n and then on the *demihyperoctahedral group* D_n. Both B_n and D_n will be defined in Chapter 4. Now we describe the flip-transpose top with random shuffle. Suppose there are n cards labelled from 1 to n and each card has two orientations namely 'face up' and 'face down'. Each arrangement of these cards can be thought of as an element of B_n (sometimes D_n too, details will be given in Chapter 4). Given a deck of n oriented cards, the flip-transpose top with random shuffle on B_n is the following: Choose a random card from the deck and transpose it with the last card after a random decision of flipping both the cards or not. The flip operation is independent of the choice of the random card. In this chapter we prove *the cutoff phenomenon for the flip-transpose top with random shuffle on B_n with cutoff time* $n\log n$. The flip-transpose top with random shuffle on D_n is a lazy variant of the following: Choose a random card from the first $n-1$ cards of the deck and transpose it with the last card after a random decision of flipping both the cards or not. The flip operation is independent of the choice of the random card in this case also. Furthermore we have shown that *the flip-transpose top with random shuffle on D_n exhibits cutoff phenomenon with cutoff time* $\left(n-\frac{1}{2}\right)\log n$.

We consider the *warp-transpose top with random shuffle* in Chapter 5. This generalizes the flip-transpose top with random shuffle (see Remark 5.17). Let $G_1\subseteq\cdots\subseteq G_n\subseteq\cdots$ be a sequence of finite groups such that $|G_1|>2$. We consider the complete monomial groups $\mathcal{G}_{\mathrm{n}}:=G_n\wr S_n$ for each positive integer n, where $\wr$ denotes the wreath product. Let e be the identity element of G_1 and hence it is the identity element of G_i for all i. Then the warp-transpose top with random shuffle is a shuffle on the complete monomial group given as follows: An element of $\mathcal{G}_{\mathrm{n}}$ can be multiplied on the right by an element of the form $(\mathrm{e},\dots,\mathrm{e},g;\mathrm{id})$ or $(\mathrm{e},\dots,\mathrm{e},g^{-1},\mathrm{e},\dots,\mathrm{e},g;(i,n))$ for $g\in G_n,\ 1\leq i<n$, with probability $\frac{1}{n|G_n|}$ each. In this chapter, we have proved that *the mixing time for the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ is* $O\left(n\log n+\frac{1}{2}n\log(|G_n|-1)\right)$. Moreover, *the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ exhibits cutoff phenomenon with cutoff time $n\log n$ if $|G_n|=o(n^{\delta})$ for all $\delta>0$.*

Chapter 2

Preliminaries

In this chapter, we develop the necessary background to explain the main objective of this thesis. We begin with the (complex) representation theory of finite groups, with special focus on symmetric group representations in Section 2.1. In Section 2.2, the basics of discrete time Markov chains on finite state spaces is discussed. Furthermore, the mixing time and cutoff phenomenon for finite Markov chain have been introduced in this section. At the end of this chapter, we discuss the non-commutative Fourier analysis and the random walks on finite groups in Section 2.3.

2.1 Representation theory background

Let V be a finite-dimensional complex vector space and $GL(V)$ be the group of all invertible linear operators from V to itself under composition of linear mappings. Unless otherwise stated, all the vector spaces considered in this chapter are finite-dimensional. Elements of $GL(V)$ can be thought of as invertible matrices over $\mathbb{C}$. Let G be a finite group. Let I denote the identity element of $GL(V)$ (i.e. the identity operator on V) and 1_{G} denote the identity element of G. A (complex) *linear representation* (ρ, V) of G is a homomorphism $\rho : G \to GL(V)$, i.e., $\rho(g_1g_2) = \rho(g_1)\rho(g_2)$ for all $g_1,\ g_2 \in G$. Therefore in particular, $\rho(1_{\text{G}}) = I$ and $\rho(g^{-1}) = \rho(g)^{-1},\ g \in G$. The representation space V is called the *G-module* corresponding to the representation ρ. Given ρ, we simply say V is a representation of G. Two useful examples are given below:

1. Let V be one-dimensional. Then $\mathbb{1} : G \to GL(V)$ defined by $\mathbb{1}(g) \mapsto (v \mapsto v)$ for all $v \in V$ and $g \in G$ is a representation of G, known as the *trivial representation* of G. In fact V one-dimensional implies that $GL(V)$ is isomorphic to $\mathbb{C}^*$ (the group of non-zero complex numbers under usual multiplication). Therefore the trivial representation $\mathbb{1} : G \to \mathbb{C}^*$ can also be defined by $\mathbb{1}(g) = 1$ for all $g \in G$.

2. Let $\mathbb{C}[G]$ be the group algebra of all formal linear combinations of the elements of G with complex coefficients, i.e. $\mathbb{C}[G] = \left\{ \sum_{g} c_g g \mid c_g \in \mathbb{C},\ g \in G \right\}$. Then the *right regular representation* $R : G \longrightarrow GL(\mathbb{C}[G])$ of G is defined by

$$R(g)\left(\sum_{h \in G} C_h h\right) = \sum_{h \in G} C_h hg, \quad C_h \in \mathbb{C}.$$

i.e., $R(g)$ is an invertible matrix over $\mathbb{C}$ of order $|G| \times |G|$. The *left regular representation* can be defined in a similar fashion.

The dimension of the vector space V is said to be the *dimension* of the representation ρ and is denoted by d_ρ. The trace of the matrix $\rho(g)$ is said to be the *character* value of ρ at g and is denoted by $\chi^\rho(g)$. A vector subspace W of V is said to be *stable* (or *invariant*) under ρ if $\rho(g)\,(W) \subset W$ for all g in G. *If W is a stable subspace of V under ρ, then there exists a complement W^0 of W in V which is stable under ρ* ([92, Theorem 1]). The representation ρ is *irreducible* if V has no non-trivial proper stable subspace. For example the trivial representation defined above is irreducible. Two representations (ρ_1, V_1) and (ρ_2, V_2) of G are are said to be *isomorphic* if there exists an invertible linear map $T : V_1 \to V_2$ such that $T \circ \rho_1(g) = \rho_2(g) \circ T$ for all $g \in G$. i.e., the following diagram commutes for all $g \in G$:

$$\begin{array}{ccc} V_1 & \xrightarrow{\rho_1(g)} & V_1 \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ V_2 & \xrightarrow{\rho_2(g)} & V_2 \end{array}$$

Let H be a subgroup of G. The *restriction* of the representation ρ to H is denoted by $\rho \downarrow_H^G$ and is defined by $\rho \downarrow_H^G (h) := \rho(h)$ for all $h \in H$. If χ^ρ is the character of ρ, then the character of the restriction $\rho \downarrow_H^G$ is denoted by $\chi^\rho \downarrow_H^G$.

Definition 2.1. Let H be a subgroup of the finite group G and $g_1H, \ldots, g_\ell H$ be all the distinct left cosets of H in G. Let ρ be a representation of H. Then the *induced representation* $\rho \uparrow_H^G$ can be defined in matrix form, as follows:

$$\rho \uparrow_H^G (g) = \begin{pmatrix} \rho(g_1^{-1}gg_1) & \rho(g_1^{-1}gg_2) & \ldots & \rho(g_1^{-1}gg_\ell) \\ \rho(g_2^{-1}gg_1) & \rho(g_2^{-1}gg_2) & \ldots & \rho(g_2^{-1}gg_\ell) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(g_\ell^{-1}gg_1) & \rho(g_\ell^{-1}gg_2) & \ldots & \rho(g_\ell^{-1}gg_\ell) \end{pmatrix},$$

where $\rho(g)$ is the zero matrix if $g \notin H$. We denote the character of the induced representation $\rho \uparrow_H^G$ by $\chi^\rho \uparrow_H^G$, where χ^ρ is the character of ρ. Note that the dimension of the induced representation $\rho \uparrow_H^G$ is $[G : H]d_\rho$.

If $V_1 \otimes V_2$ denotes the tensor product of the (complex) vector spaces V_1 and V_2, then the *tensor product of two representations* $\rho_1 : G \to GL(V_1)$ and $\rho_2 : G \to GL(V_2)$ is a representation denoted by $(\rho_1 \otimes \rho_2, V_1 \otimes V_2)$ and defined by,

$$(\rho_1 \otimes \rho_2)(g)(v_1 \otimes v_2) = \rho_1(g)(v_1) \otimes \rho_2(g)(v_2) \text{ for } v_1 \in V_1, v_2 \in V_2 \text{ and } g \in G.$$

Also the *direct sum of the representations* (ρ_1, V_1) and (ρ_2, V_2) is the representation $(\rho_1 \oplus \rho_2) : G \to GL(V_1 \oplus V_2)$ defined by,

$$(\rho_1 \oplus \rho_2)(g)(v_1 \oplus v_2) = \rho_1(g)(v_1) \oplus \rho_2(g)(v_2) \text{ for } v_1 \in V_1, v_2 \in V_2 \text{ and } g \in G.$$

Every (complex) linear representation is a direct sum of irreducible representations ([92, Theorem 1]). Moreover this decomposition is unique up to isomorphism of representation. Now we state some properties of the character of a representation without proof [81, 87, 92].

Lemma 2.2. *Let $\rho_1 : G \to GL(V_1)$ and $\rho_2 : G \to GL(V_2)$ be two representations of the finite group G. Also let χ^{ρ_1} and χ^{ρ_2} be their characters respectively. Then we have the following:*

1. *The character $\chi^{\rho_1 \oplus \rho_2}$ of the direct sum representation is equal to $\chi^{\rho_1} + \chi^{\rho_2}$.*

2. *The character $\chi^{\rho_1 \otimes \rho_2}$ of the tensor product representation is equal to $\chi^{\rho_1}\chi^{\rho_2}$.*

Lemma 2.3. *Let G be a finite group. Given the representation (ρ, V) of G and its character χ^ρ, one can easily verify the following:*

1. *If 1_{G} denotes the identity element of G and d_ρ denotes the dimension of the representation (i.e. of V), then $\chi^\rho(1_{\mathrm{G}}) = d_\rho$.*

2. *If $\overline{\chi^\rho(g)}$ denote the complex conjugate of $\chi^\rho(g)$, then $\chi^\rho(g^{-1}) = \overline{\chi^\rho(g)}$ for all $g \in G$.*

3. *The character values are constants on conjugacy classes. Hence characters are class functions.*

Lemma 2.4 (Schur's lemma [92, Proposition 5])**.** *Let (ρ_1, V_1) and (ρ_2, V_2) be two (complex) irreducible representations of the finite group G. If a linear mapping $\Phi : V_1 \to V_2$ satisfies $\rho_2(g) \circ \Phi = \Phi \circ \rho_1(g)$ for all $g \in G$, then*

1. *$\Phi = 0$, when ρ_1 and ρ_2 are non-isomorphic.*

2. *Φ is a scalar multiple of the identity, when $V_1 = V_2$ and $\rho_1 = \rho_2$.*

Definition 2.5. Let G be a finite group and $\mathscr{C}(G)$ be the complex vector space of class functions of G. Then we define the inner product $\langle \cdot, \cdot \rangle$ on $\mathscr{C}(G)$ as follows:

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}) \quad \text{for} \quad \phi, \psi \in \mathscr{C}(G).$$

Lemma 2.6. *Let (ρ, V) (respectively (ρ', V')) be a representation of the finite group G with character χ (respectively χ'). Then we have*

1. *(ρ, V) and (ρ', V') are isomorphic if and only if $\chi = \chi'$.*
2. *V is an irreducible representation of G if and only if $\langle \chi, \chi \rangle = 1$.*
3. *If V and V' are irreducible, then they are non-isomorphic if and only if $\langle \chi, \chi' \rangle = 0$.*
4. *If $V = m_1W_1 \oplus \cdots \oplus m_\ell W_\ell$ is the decomposition of V into irreducible representations of ρ, then $\langle \chi, \chi \rangle = m_1^2 + \cdots + m_\ell^2$. Moreover if χ^{W_i} denotes the irreducible character of W_i, $1 \leq i \leq \ell$, then $m_i = \langle \chi, \chi^{W_i} \rangle$ is called the multiplicity of W_i in the decomposition of V.*

Theorem 2.7 ([92, Theorem 6])**.** *The characters corresponding to the non-isomorphic irreducible representations of G form an $\langle \cdot, \cdot \rangle$-orthonormal basis of $\mathscr{C}(G)$.*

Let G be a finite group and $\widehat{G}$ be the set of equivalence classes (two representations are equivalent if they are isomorphic) of irreducible representations of G. Then from Theorem 2.7, we can conclude that $|\widehat{G}|$ is the dimension of $\mathscr{C}(G)$ i.e., the number of irreducible representations of a finite group is equal to the number of its conjugacy classes. *The regular* (*true for both left and right*) *representation of G decomposes into irreducible representations with multiplicity equal to their respective dimensions* [92, p. 18, Corollary 1]. Thus we have,

$$\mathbb{C}[G] \cong \underset{\rho \in \widehat{G}}{\oplus} d_\rho V^\rho, \tag{2.1.1}$$

where V^ρ is the irreducible G-module corresponding to $\rho \in \widehat{G}$ with dimension d_ρ. Also from Theorem 2.7 and [92, Proposition 5] we have, $|G| = \sum_{\rho \in \widehat{G}} d_\rho^2$.

Theorem 2.8 (Frobenius reciprocity [87, Theorem 1.12.6])**.** *Let H be a subgroup of the finite group G and suppose that ψ and χ are characters of H and G respectively. Then*

$$\langle \psi \uparrow_H^G, \chi \rangle = \langle \psi, \chi \downarrow_H^G \rangle,$$

where the left inner product is calculated in G and the right one in H.

2.1.1 Symmetric group representations

Recall that S_n is the group of all bijections on $[n] := \{1, \ldots, n\}$. Elements of S_n are known as *permutations* and the group operation is the composition of permutations. We first define some combinatorial objects, useful in describing the representation theory of S_n. We also prove a lemma relevant to these definitions before discussing the representation theory of S_n.

Definition 2.9. A *partition* λ of a positive integer n is a weakly decreasing finite sequence $(\lambda_1, \cdots, \lambda_r)$ of positive integers such that $\sum_{i=1}^{r} \lambda_i = n$. We write $\lambda \vdash n$ to mean λ is a partition of n. The set of all partitions of n is denoted by $\mathrm{Par}(n)$. A partition λ can be pictorially visualised as a left-justified arrangement of r rows of boxes with λ_i boxes in the ith row (English notation). This pictorial arrangement of boxes is known as the *Young diagram* of λ. For example there are five partitions of the positive integer 4 viz. (4), (3,1), (2,2), (2,1,1) and (1,1,1,1). Young diagrams corresponding to the partitions of 4 are given in Figure 2.1. We use the same notation λ to denote partition and Young diagram both. It will be clear from the context whether a partition or a Young diagram is meant. The set of all Young diagrams (there is a unique Young diagram with zero boxes) is denoted by $\mathcal{Y}$ and the set of all Young diagrams with n boxes is denoted by $\mathcal{Y}_{\mathrm{n}}$. A *Young tableau* of shape λ (or λ-*tableau*) is obtained by filling (bijectively) the

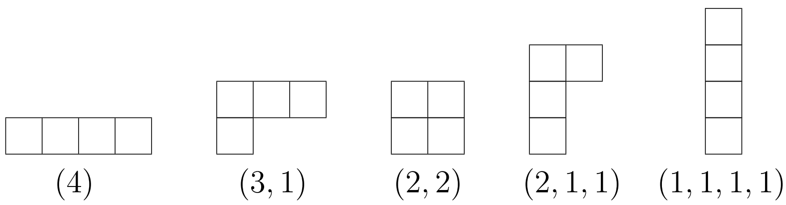


Figure 2.1: Young diagrams with 4 boxes.

numbers $1, \ldots, n$ in the boxes of the Young diagram of λ. A λ-tableau is *standard* if the entries in its boxes increase from left to right along rows and from top to bottom along columns. The set of all standard Young tableaux of a given shape λ is denoted by $\mathrm{tab}(\lambda)$ and the number of standard Young tableaux of shape λ is denoted by f^λ. For example, all standard Young tableaux of shape $(3, 1)$ are listed in Figure 2.2. The *content* of a box in row i and column j of a Young diagram is the integer $j - i$. Given $T \in \mathrm{tab}(\lambda)$, let $b_T(i)$ denote the box in T, where the number i resides. Also $c(b_T(i))$ denotes the content of the box $b_T(i)$, $1 \leq i \leq n$. Given a Young diagram λ, its *conjugate* λ' is obtained by

1	2	3
4		

1	2	4
3		

1	3	4
2		

Figure 2.2: Standard Young tableaux of shape $(3, 1)$.

reflecting λ with respect to the diagonal consisting of boxes with content 0. A Young diagram λ is *self-conjugate* if $\lambda' = \lambda$. We now prove the following lemma, which will be useful in later chapters of this thesis.

Lemma 2.10. *Let ℓ, k be any two positive integers and $a \geq b \geq 0$ be such that $b < \ell$. For $\lambda \vdash \ell$, if λ_1 denotes the largest part of λ, then*

$$\sum_{\lambda \vdash \ell} (f^\lambda)^2 \left(\frac{\lambda_1 - a}{\ell - b}\right)^{2k} < e^{-\frac{2k(a-b)}{\ell - b}} e^{\ell^2 e^{-\frac{2k}{\ell - b}}}.$$

Proof. For $\zeta \vdash (\ell - \lambda_1)$, recall that ζ_1 denotes the largest part of ζ. If $\zeta_1 \leq \lambda_1$, then we have $f^\lambda \leq \binom{\ell}{\lambda_1} f^\zeta$. Therefore $\sum\limits_{\lambda \vdash \ell} (f^\lambda)^2 \left(\dfrac{\lambda_1 - a}{\ell - b}\right)^{2k}$ is less than or equal to

$$\begin{aligned} \sum_{\lambda_1 = 1}^{\ell} \sum_{\substack{\zeta \vdash (\ell - \lambda_1) \\ \zeta_1 \leq \lambda_1}} \binom{\ell}{\lambda_1}^2 (f^\zeta)^2 \left(\frac{\lambda_1 - a}{\ell - b}\right)^{2k} &\leq \sum_{\lambda_1 = 1}^{\ell} \binom{\ell}{\lambda_1}^2 \left(\frac{\lambda_1 - a}{\ell - b}\right)^{2k} \sum_{\zeta \vdash (\ell - \lambda_1)} (f^\zeta)^2 \\ &= \sum_{u=0}^{\ell - 1} \binom{\ell}{u}^2 \left(1 - \frac{u + a - b}{\ell - b}\right)^{2k} u!. \qquad (2.1.2) \end{aligned}$$

The equality in (2.1.2) is obtained by writing $u = \ell - \lambda_1$. Using $1 - x \leq e^{-x}$ for all $x \geq 0$ and $\binom{\ell}{u} \leq \frac{\ell^u}{u!}$, the expression in the right hand side of (2.1.2) is less than or equal to

$$\sum_{u=0}^{\ell - 1} \frac{\ell^{2u}}{u!} e^{-\frac{2k}{\ell - b}(u + a - b)} < e^{-\frac{2k(a-b)}{\ell - b}} \sum_{u=0}^{\infty} \frac{1}{u!} \left(\ell^2 e^{-\frac{2k}{\ell - b}}\right)^u = e^{-\frac{2k(a-b)}{\ell - b}} e^{\ell^2 e^{-\frac{2k}{\ell - b}}}. \qquad \square$$

An immediate corollary of Lemma 2.10 follows from the fact

$$(f^\lambda)^2 \left(\frac{\lambda_1 - a}{\ell - b}\right)^{2k} = \left(\frac{\ell - a}{\ell - b}\right)^{2k}, \text{ if } \lambda = (\ell) \vdash \ell.$$

Corollary 2.11. *Following the notations and hypothesis of Lemma 2.10, we have*

$$\sum_{\substack{\lambda \vdash \ell \\ \lambda \neq (\ell)}} (f^\lambda)^2 \left(\frac{\lambda_1 - a}{\ell - b}\right)^{2k} < e^{-\frac{2k(a-b)}{\ell - b}} e^{\ell^2 e^{-\frac{2k}{\ell - b}}} - \left(\frac{\ell - a}{\ell - b}\right)^{2k}.$$

We are now in a position to discuss the representation theory of the symmetric group

S_n. Unless otherwise stated the field of scalars is considered to be $\mathbb{C}$. *The irreducible representations of S_n are indexed by the partitions of n* [87, Theorem 2.4.6]. The irreducible S_n-modules are called *Specht modules*. The Specht module indexed by $\lambda \vdash n$ is denoted by S^λ. *The basis vectors of S^λ are indexed by the standard Young tableaux of shape λ* [87, Theorem 2.5.2] and hence $\dim(S^\lambda) = f^\lambda$. An important property of symmetric group representation is the simple branching of the restriction (using Theorem 2.8 (Frobenius Reciprocity) simple branching can be proved for the induced representation also). The restriction of an irreducible S_n-module to S_{n-1} has a multiplicity-free decomposition into irreducible S_{n-1}-module [87, Theorem 2.8.3]. For example taking $n = 15$, the restriction of $S^{(5,4,4,2)}$ to S_{14} is given by the following:

$$S^{(5,4,4,2)} \downarrow_{S_{14}}^{S_{15}} = S^{(4,4,4,2)} \oplus S^{(5,4,3,2)} \oplus S^{(5,4,4,1)}.$$

Let λ^- be a partition of $n-1$ obtained by removing a single *inner corner* (a corner box in the Young diagram whose removal leaves a valid Young diagram of some partition) from the Young diagram of $\lambda \vdash n$. Then the precise statement of the branching rule is the following:

$$S^\lambda \downarrow_{S_{n-1}}^{S_n} = \bigoplus_{\lambda^-} S^{\lambda^-}.$$

A different approach to the representation theory of S_n has been initiated by Okounkov and Vershik [77] using the simple branching property. We give a brief description of their approach because that is relevant for this thesis.

A direct argument shows that the branching from S_n to S_{n-1} is multiplicity-free. Now start with an irreducible S_n-module V. The restriction of V to S_{n-1} has multiplicity-free decomposition into irreducible S_{n-1}-modules. Again, restriction of each of these irreducibles to S_{n-2} has a multiplicity-free decomposition into irreducible S_{n-2}-modules. Iterating this, we get a canonical decomposition of V into irreducible S_1-modules i.e., one-dimensional subspaces [77, Theorem 2.9]. Thus there is a canonical basis of V. This basis is named the *Gelfand-Tsetlin* basis of V. The Gelfand-Tsetlin basis vectors are the simultaneous eigenvectors of the elements of a maximal commuting subalgebra of $\mathbb{C}[S_n]$ when they act on V. This subalgebra is known as the *Gelfand-Tsetlin* subalgebra of $\mathbb{C}[S_n]$. Let Z_i denote the center of $\mathbb{C}[S_i]$, $1 \leq i \leq n$. Then the *Gelfand-Tsetlin* subalgebra is defined to be the subalgebra of $\mathbb{C}[S_n]$ generated by $Z_1 \cup \cdots \cup Z_n$ and its dimension is equal to the sum of dimensions of distinct non-isomorphic irreducible S_n-modules. It follows that any vector in the Gelfand-Tsetlin basis is uniquely (up to scalar factor) determined by the eigenvalues of the elements of the Gelfand-Tsetlin algebra on this vector. Thus the eigenvalues of the elements of a generating set of the Gelfand-Tsetlin algebra determine the Gelfand-Tsetlin vectors. The *Young-Jucys-Murphy* elements (YJM-elements)

$$Y_1 = 0 \text{ and } Y_i = (1,i) + (2,i) + \cdots + (i-1,i),\ 1 < i \leq n \tag{2.1.3}$$

form a generating set for the Gelfand-Tsetlin algebra of $\mathbb{C}[S_n]$. Now we give the action of the Young-Jucys-Murphy elements on the Gelfand-Tsetlin (basis) vectors of the irreducible S_n-module S^λ for $\lambda \vdash n$. The Gelfand-Tsetlin vectors of S^λ are indexed by the standard Young tableaux of shape λ and let $\{s_T : T \in \operatorname{tab}(\lambda)\}$ denotes the Gelfand-Tsetlin basis of S^λ. Then the action of the Young-Jucys-Murphy elements are given as follows:

$$Y_1(s_T) = 0 \text{ and } Y_i(s_T) = c(b_T(i))s_T,\ 1 < i \leq n \text{ for all } T \in \operatorname{tab}(\lambda). \tag{2.1.4}$$

Special importance is given to a particular representation, known as the *defining representation* of S_n. We will prove some results for the defining representation of S_n, which will be useful in the later chapters of this thesis.

Definition 2.12. Let $\mathbb{C}[\mathbf{n}] := \{c_1\mathbf{1} + c_2\mathbf{2} + \cdots + c_n\mathbf{n} \mid c_i \in \mathbb{C} \text{ for all } i\}$. Then the *defining representation* $\rho^{\text{def}} : S_n \to GL(\mathbb{C}[\mathbf{n}])$ of S_n is defined by

$$\rho^{\text{def}}(\pi)\left(c_1\mathbf{1} + c_2\mathbf{2} + \cdots + c_n\mathbf{n}\right) = c_1\pi(\mathbf{1}) + c_2\pi(\mathbf{2}) + \cdots + c_n\pi(\mathbf{n}) \text{ for } \pi \in S_n.$$

Throughout this thesis the character of the defining representation of S_n is denoted by χ^{def}. The defining representation of S_n splits into two irreducible Specht modules $S^{(n)}$ and $S^{(n-1,1)}$ with multiplicity one each [87, Example 2.1.8 and Theorem 2.11.2], i.e.,

$$\mathbb{C}[\mathbf{n}] = S^{(n)} \oplus S^{(n-1,1)}. \tag{2.1.5}$$

We also recall that $\rho^{\text{def}} \otimes \rho^{\text{def}} : S_n \to GL(\mathbb{C}[\mathbf{n}] \otimes \mathbb{C}[\mathbf{n}])$ is defined as follows:

$$(\rho^{\text{def}} \otimes \rho^{\text{def}})(\pi)\,(v_1 \otimes v_2) = \rho^{\text{def}}(\pi)(v_1) \otimes \rho^{\text{def}}(\pi)(v_2) \text{ for } \pi \in S_n,\ v_1 \otimes v_2 \in \mathbb{C}[\mathbf{n}] \otimes \mathbb{C}[\mathbf{n}].$$

From now on, the character of $\rho^{\text{def}} \otimes \rho^{\text{def}}$ is denoted by $\chi^{\text{def}} \otimes \chi^{\text{def}}$. The decomposition of $\rho^{\text{def}} \otimes \rho^{\text{def}}$ into irreducible S_n-modules is given below (see [87, Example 2.1.8] and [59, Lemma 2.9.16]):

$$\mathbb{C}[\mathbf{n}] \otimes \mathbb{C}[\mathbf{n}] = 2S^{(n)} \oplus 3S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1,1)}. \tag{2.1.6}$$

Lemma 2.13. *The eigenvalues of* $\displaystyle\sum_{u=1}^{n-1} \rho^{\text{def}}((u,n))$ *are given below:*

Eigenvalues:	$n-1$	$n-2$	-1
Multiplicities:	1	$n-2$	1

Proof. $\sum_{u=1}^{n-1} \rho^{\text{def}}((u,n))$ is the nth Young-Jucys-Murphy element Y_n (see (2.1.3)) of $\mathbb{C}[S_n]$. Therefore the lemma follows from (2.1.5), (2.1.4) and straightforward calculations. □

Lemma 2.14. *The eigenvalues of* $\sum_{u=1}^{n-1} \left(\rho^{\text{def}}((u,n)) \otimes \rho^{\text{def}}((u,n))\right)$ *are given as follows:*

Eigenvalues:	$n-1$	$n-2$	-1	0	-2	$n-3$
Multiplicities:	2	$3(n-2)$	3	$n-2$	$n-2$	n^2-5n+5

Proof. The matrix $\sum_{u=1}^{n-1} \left(\rho^{\text{def}}((u,n)) \otimes \rho^{\text{def}}((u,n))\right)$ is the image of Y_n (see (2.1.3)) under the diagonal action of S_n on $\mathbb{C}[\mathbf{n}] \otimes \mathbb{C}[\mathbf{n}]$. Therefore the lemma follows from (2.1.6), (2.1.4) and straightforward calculations. □

Lemma 2.15. *Let* β_i *denote the matrix* $\sum_{u=1}^{n-1} \rho^{\text{def}}((u,n)) - \rho^{\text{def}}((i,n))$ *for all* $1 \leq i < n$. *Then we have the following:*

1. *The matrices* β_i *and* β_j *are similar for* $i \neq j$ *and* $i, j \in \{1, \dots, n-1\}$.
2. *For each* $i \in \{1, \dots, n-1\}$, *the eigenvalues of* β_i *are the following:*

Eigenvalues:	$n-2$	$n-3$	-1
Multiplicities:	2	$n-3$	1

Proof. We have $\left(\rho^{\text{def}}((i,j))\right)^{-1} = \rho^{\text{def}}((i,j))$ as the transposition $(i,j) \in S_n$ is self-inverse. Thus the first part of the lemma follows from the following fact:

$$
\begin{aligned}
&\rho^{\text{def}}((i,j))\beta_i \left(\rho^{\text{def}}((i,j))\right)^{-1} \\
=& \rho^{\text{def}}((i,j))\beta_i \, \rho^{\text{def}}((i,j)) \\
=& \rho^{\text{def}}((i,j)) \left(\sum_{u=1}^{n-1} \rho^{\text{def}}((u,n))\right) \rho^{\text{def}}((i,j)) - \rho^{\text{def}}((i,j)) \, \rho^{\text{def}}((i,n)) \, \rho^{\text{def}}((i,j)) \\
=& \sum_{u=1}^{n-1} \rho^{\text{def}}((i,j)(u,n)(i,j)) - \rho^{\text{def}}((i,j)(i,n)(i,j)) = \sum_{u=1}^{n-1} \rho^{\text{def}}((u,n)) - \rho^{\text{def}}((j,n)) = \beta_j.
\end{aligned}
$$

The eigenvalues of $\beta_1, \beta_2, \dots, \beta_{n-1}$ are the same by the first part of this lemma. Therefore to prove the second part it is enough to find the eigenvalues of β_1. Let us consider the linearly independent vectors $v_1, v_2, v_3, \dots, v_{n-1}, v_n$ of $\mathbb{C}[\{\mathbf{1}, \dots, \mathbf{n}\}]$. The eigenvalues of β_1 are obtained from Table 2.1. □

Eigenvector	Action of β_1 on eigenvector	Eigenvalues
$v_1 = \mathbf{1} + \mathbf{2} + \cdots + \mathbf{n}$	$\beta_1(v_1) = (n-2)v_1$	$n-2$
$v_2 = \mathbf{1}$	$\beta_1(v_2) = (n-2)v_2$	$n-2$
$v_i = \mathbf{i} - \mathbf{2}$ for $3 \le i \le n-1$	$\beta_1(v_i) = (n-3)v_i$ for $3 \le i \le n-1$	$n-3$ for $3 \le i \le n-1$
$v_n = v_1 - v_2 - (n-1)\,\mathbf{n}$	$\beta_1(v_n) = (-1)v_n$	-1

Table 2.1: Eigenvectors and eigenvalues of β_1.

2.2 Probabilistic background

In this section we discuss discrete time Markov chain with finite state space. We state the results which will be useful in the later chapters. Most of the notations of this section are borrowed from [66].

Definition 2.16. Let Ω be a finite set. A sequence of random variables $X_0, X_1, \ldots$ is said to be a *discrete time Markov chain with state space* Ω if for all $x, y \in \Omega$, all $k \geq 1$, and all events $H_{k-1} := \underset{0 \leq s < k}{\cap} \{X_s = x_s\}$ satisfying $\mathbb{P}(H_{k-1} \cap \{X_k = x\}) > 0$, we have

$$\mathbb{P}(X_{k+1} = y \mid H_{k-1} \cap \{X_k = x\}) = \mathbb{P}(X_{k+1} = y \mid X_k = x). \tag{2.2.1}$$

This Markov chain is said to be *time-homogeneous* if

$$\mathbb{P}(X_{k+1} = y \mid X_k = x) = \mathbb{P}(X_k = y \mid X_{k-1} = x) \text{ for all } k \geq 1.$$

The matrix M whose rows and columns are indexed by Ω and entries are given by

$$M(x, y) := \mathbb{P}(X_{k+1} = y \mid X_k = x) \text{ for all } x, y \in \Omega,$$

is said to be the (one-step) *transition matrix* of this Markov chain.

Equation (2.2.1) says that given the present, the future is independent of the past. We note that the Markov chains appearing throughout this thesis are time-homogeneous. Let $\mathscr{D}_k$ denote the distribution after k transitions, i.e., $\mathscr{D}_k$ is the row vector $(\mathbb{P}(X_k{=}x))_{x\in\Omega}$. Then $\mathscr{D}_k = \mathscr{D}_{k-1}M$ for all $k \geq 1$, which implies $\mathscr{D}_k = \mathscr{D}_0 M^k$. In particular if the chain starts at $x \in \Omega$, then its distribution after k transitions is $\mathscr{D}_k = \delta_x M^k$, where δ_x is defined on Ω as follows:

$$\delta_x(u) = \begin{cases} 1 & \text{if } u = x, \\ 0 & \text{if } u \neq x. \end{cases}$$

i.e., $\mathbb{P}(X_k = y \mid X_0 = x) = M^k(x, y)$. A Markov chain is said to be *irreducible* if it is possible for the chain to reach any state starting from any state using only transitions of positive probabilities. The *period* of a state $x \in \Omega$ is defined to be the greatest common divisor of the set of all times when it is possible for the chain to return to the starting state x. *The period of all the states of an irreducible Markov chain are the same* (see [66, Lemma 1.6]). An irreducible Markov chain is said to be *aperiodic* if the common period for all its states is 1. A probability distribution Π is said to be a *stationary distribution* of the Markov chain if $\Pi M = \Pi$. Any irreducible Markov chain possesses a unique stationary distribution Π with $\Pi(x) > 0$ for all $x \in \Omega$ [66, Proposition 1.14]. Moreover if the chain is aperiodic then $\mathscr{D}_k \longrightarrow \Pi$ as $k \longrightarrow \infty$ [66, Theorem 4.9]. An irreducible and aperiodic Markov chain with transition matrix M and stationary distribution Π is said to be *time reversible* if $\Pi(x)P(x,y) = \Pi(y)P(y,x)$ for all states $x, y \in \Omega$. We now define the *total variation distance* between two probability measures and prove a lemma, which is useful for the upcoming chapters.

Definition 2.17. Let μ and ν be two probability measures on Ω. The *total variation distance* between μ and ν is defined by

$$||\mu - \nu||_{\mathrm{TV}} := \sup_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

It can be easily seen that $||\mu - \nu||_{\mathrm{TV}} = \frac{1}{2} \sum\limits_{x \in \Omega} |\mu(x) - \nu(x)|$ (see [66, Proposition 4.2]).

We now recall the definitions of *expectation* and *variance* of a random variable with respect to a probability measure for the shake of clarity and completeness.

Definition 2.18. Let $(\Omega, \mathscr{F}, \mu)$ be a probability space, X be a real valued random variable defined on Ω, and $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then the *expectation* of $f(X)$ with respect to the probability measure μ is denoted by $E_\mu(f(X))$, and is defined as follows:

$$E_\mu(f(X)) := \int_\Omega f(X) d\mu, \quad \text{provided the integral exists (i.e., finite).} \tag{2.2.2}$$

Therefore the expectation $E_\mu(X)$ of X with respect to the probability measure μ is defined by setting $f(x) = x,\ x \in \mathbb{R}$ in (2.2.2). The *variance* of X with respect to the probability measure μ is denoted by $\mathrm{Var}_\mu(X)$, and is defined by

$$\mathrm{Var}_\mu(X) := E_\mu(X^2) - (E_\mu(X))^2 .$$

In particular if Ω is countable (finite or countably infinite), the σ-algebra $\mathscr{F}$ consists of all subsets of Ω, and μ is the discrete measure on Ω, then (2.2.2) turns out to be

$$E_\mu(f(X)) := \sum_{x\in\Omega} f(x)\mu(x), \quad \text{provided the series converges.} \tag{2.2.3}$$

Lemma 2.19. *Let μ and ν be two probability measures on Ω, and X be a non-negative random variable on Ω. Also let $E_\mu(X)$ (respectively $E_\nu(X)$) denote the expectation of X with respect to the probability measure μ (respectively ν) and $\mathrm{Var}_\mu(X)$ denote the variance of X with respect to the probability measure μ. If $E_\mu(X) > 0$, then*

$$||\mu - \nu||_{\mathrm{TV}} \geq 1 - \frac{4\,\mathrm{Var}_\mu(X)}{(E_\mu(X))^2} - \frac{2E_\nu(X)}{E_\mu(X)}.$$

Proof. For any positive constant a, by Chebychev's inequality, we have

$$\mu\left(\{\omega \in \Omega : |X(\omega) - E_\mu(X)| \leq a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \geq 1 - \frac{1}{a^2}. \tag{2.2.4}$$

Now we choose a positive constant a such that $E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)} > 0$. Then by using Markov's inequality, we have

$$\nu\left(\{\omega \in \Omega : X(\omega) \geq E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \leq \frac{E_\nu(X)}{E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)}} \tag{2.2.5}$$

Now from the definition of total variation distance, we have

$$\begin{aligned}
||\mu - \nu||_{\mathrm{TV}} &= \sup_{A\subset\Omega} |\mu(A) - \nu(A)| \\
&\geq \mu\left(\{\omega \in \Omega : |X(\omega) - E_\mu(X)| \leq a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \\
&\quad - \nu\left(\{\omega \in \Omega : |X(\omega) - E_\mu(X)| \leq a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \\
&\geq \mu\left(\{\omega \in \Omega : |X(\omega) - E_\mu(X)| \leq a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \\
&\quad - \nu\left(\{\omega \in \Omega : X(\omega) \geq E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)}\}\right) \\
&\geq 1 - \frac{1}{a^2} - \frac{E_\nu(X)}{E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)}}.
\end{aligned} \tag{2.2.6}$$

The inequality (2.2.6) follows by using (2.2.4) and (2.2.5). Therefore the lemma follows by choosing $a = \frac{E_\mu(X)}{2\sqrt{\mathrm{Var}_\mu(X)}} > 0$ in (2.2.6). □

Remark 2.20. If the positive constant a is such that $E_\mu(X) - a\sqrt{\mathrm{Var}_\mu(X)} > 0$, then the inequality (2.2.6) constitutes a generalize version of Lemma 2.19.

For an irreducible and aperiodic Markov chain, an interesting quantity is the minimum number of transitions required to reach the stationary distribution Π up to a certain level of tolerance $\varepsilon > 0$. We first define the maximal distance (over $x_0 \in \Omega$) between $M^k(x_0, \cdot)$ and Π as follows:

$$d(k) := \max_{x \in \Omega} ||M^k(x, \cdot) - \Pi||_{\mathrm{TV}}.$$

Also for $\varepsilon > 0$, the ε-mixing time is defined by

$$t_{\mathrm{mix}}(\varepsilon) := \min \{k : d(k) \leq \varepsilon\}.$$

Definition 2.21. Given a sequence of discrete time Markov chains with finite state spaces, if $t_{\mathrm{mix}}^{(n)}(\varepsilon)$ denote the ε-mixing time for the nth chain, then we say that the sequence satisfies the cutoff phenomenon if for all $\varepsilon \in (0, 1)$ the following holds:

1. $\lim_{n \to \infty} t_{\mathrm{mix}}^{(n)} = \infty$,
2. $\lim_{n \to \infty} \dfrac{t_{\mathrm{mix}}^{(n)}(\varepsilon)}{t_{\mathrm{mix}}^{(n)}(1-\varepsilon)} = 1$.

In this case the cutoff time τ_n is defined to be $t_{\mathrm{mix}}^{(n)}(\varepsilon)$.

This says that for sufficiently large n the mixing time does not depend on the tolerance level ε. In other words the distribution after k transitions is very close to the stationary distribution if $k = \tau_n$ but too far from the stationary distribution if $k < \tau_n$. We follow an equivalent definition for cutoff phenomenon, which will be given in Section 2.3.

2.3 Non-commutative Fourier analysis and random walks on finite groups

Let p and q be two probability measures on a finite group G. We define the *convolution* $p * q$ of p and q by

$$(p * q)(x) := \sum_{y \in G} p(xy^{-1})q(y).$$

Let (ρ, V) be a (complex) linear representation of G. Then the *Fourier transform* $\widehat{p}$ of p at ρ is defined by the following matrix

$$\widehat{p}(\rho) := \sum_{x \in G} p(x)\rho(x).$$

It can be easily seen that $\widehat{(p*q)}(\rho) = \widehat{p}(\rho)\widehat{q}(\rho)$. In particular for the right regular representation R of G, the matrix $\widehat{p}(R)$ can be thought of as the action of the group algebra element $\sum_{g\in G} p(g)g$ on $\mathbb{C}[G]$ by multiplication on the right.

A *random walk on a finite group G driven by a probability measure p* is a Markov chain with state space G and transition probabilities $M_p(x,y) = p(x^{-1}y)$, $x,y \in G$. It can be easily seen that the transition matrix M_p is the transpose of $\widehat{p}(R)$ and the distribution after kth transition will be p^{*k} (convolution of p with itself k times) i.e., the probability of getting into state y starting at state x after k transitions is $p^{*k}(x^{-1}y)$.

Proposition 2.22 ([89, Proposition 2.3])**.** *The random walk on G driven by p is irreducible if and only if the support of p generates G.*

Proof. Let Γ be the support of p. Suppose that Γ generates G. Therefore given any two arbitrary elements $x,y\in G$, $x^{-1}y$ can be written as $\gamma_1\gamma_2\dots\gamma_t$ for $\gamma_1,\gamma_2,\dots,\gamma_t\in\Gamma$. Thus $p^{*t}(x^{-1}y)\geq p(\gamma_1)p(\gamma_2)\dots p(\gamma_t)>0$ and the random walk is irreducible.

Conversely let $g\in G$ be chosen arbitrarily. Suppose that the random walk on G driven by p is irreducible. Then there exists a positive integer s such that $M_p^s(x,xg) = p^{*s}(g)>0$ for $x\in G$. This implies that, there exist $g_1,\dots,g_s$ in the support of p such that $g=g_1\dots g_s$. Hence the proposition follows. □

Proposition 2.23 (see, for instance [89])**.** *The stationary distribution for an irreducible random walk on G driven by p, is the uniform distribution U_G on G.*

Proof. The random walk on G driven by p in irreducible. Thus it possesses a unique stationary distribution. Therefore the proposition follows from the following fact.

$$\sum_{x\in G} M_p(x,y) = \sum_{x\in G} p(x^{-1}y) = 1, \text{ for all } y\in G.$$ □

The irreducible random walk on G driven by p is *time reversible* if and only if $M_p(x,y) = M_p(y,x)$ for all $x,y\in G$ i.e., if and only if $p(x)=p(x^{-1})$ for all x in G (this condition is also known as *symmetry of p*). From now on, the uniform distribution on group G will be denoted by U_G. For the random walk on G driven by p, it is enough to focus on $||p^{*k}-U_G||_{\text{TV}}$ because,

$$||M_p^k(x,\cdot)-U_G||_{\text{TV}} = ||M_p^k(y,\cdot)-U_G||_{\text{TV}}$$

for any two elements $x,y\in G$. We now state the Diaconis-Shahshahani upper bound lemma [32, Lemma 4.2]. This lemma has been used in the upcoming chapters, to find the sufficient number of transitions required for our random walk models to reach the stationary distribution.

Lemma 2.24 (Upper bound lemma, [32, Lemma 4.2]). *Let p be a probability measure on a finite group G such that $p(x) = p(x^{-1})$ for all $x \in G$. Suppose the random walk on G driven by p is irreducible and aperiodic. Then we have the following*

$$||p^{*k} - U_G||_{\mathrm{TV}}^2 \leq \frac{1}{4} \sum_{\rho \neq \mathbb{1}} d_\rho \operatorname{Tr}\left((\widehat{p}(\rho))^{2k} \right),$$

where the sum is over all non-trivial irreducible representations ρ of G and d_ρ is the dimension of ρ.

Definition 2.25. Let $\{\mathscr{G}_n\}_0^\infty$ be a sequence of finite groups and p_n be a probability measure on $\mathscr{G}_n$ for each n. Consider the sequence of irreducible and aperiodic random walks on $\mathscr{G}_n$ driven by p_n. We say that the *total variation cutoff phenomenon* holds for the family $\{(\mathscr{G}_n, p_n)\}_0^\infty$ if there exists a sequence $\{\tau_n\}_0^\infty$ of positive real numbers such that the following hold:

1. $\lim_{n\to\infty} \tau_n = \infty$,

2. For any $\epsilon \in (0,1)$ and $k_n = \lfloor (1+\epsilon)\tau_n \rfloor$, $\lim_{n\to\infty} ||p_n^{*k_n} - U_{\mathscr{G}_n}||_{\mathrm{TV}} = 0$ and

3. For any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon)\tau_n \rfloor$, $\lim_{n\to\infty} ||p_n^{*k_n} - U_{\mathscr{G}_n}||_{\mathrm{TV}} = 1$.

Here $\lfloor x \rfloor$ denotes the floor of x (the largest integer less than or equal to x).

Informally, we will say that $\{(\mathscr{G}_n, p_n)\}_0^\infty$ has a total variation cutoff at time τ_n. Roughly the cutoff phenomenon depends on the multiplicity of the second largest eigenvalue of the transition matrix [34].

Proposition 2.26. *Following the notations of Definition 2.25, let $t_{\mathrm{mix}}^{(n)}(\varepsilon)$ be the ε-mixing time for the random walk on $\mathscr{G}_n$ driven by p_n ($\varepsilon \in (0,1)$ is arbitrary). Then the conditions of Definition 2.25 are equivalent to the following conditions:*

$$\lim_{n\to\infty} t_{\mathrm{mix}}^{(n)}(\varepsilon) = \infty \text{ and } \lim_{n\to\infty} \frac{t_{\mathrm{mix}}^{(n)}(\varepsilon)}{t_{\mathrm{mix}}^{(n)}(1-\varepsilon)} = 1.$$

Proof. Let us choose $\varepsilon \in (0,1)$ arbitrarily and assume the conditions of Definition 2.25.

Then for $\epsilon = \frac{1}{n}$, there exists positive integers N_0, N_1, N_2 such that

$$\begin{aligned} n \geq N_0 &\implies \tau_n > 1 > 0, \\ n \geq N_1 &\implies -\varepsilon < ||p_n^{*\lfloor (1+\epsilon)\tau_n \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} < \varepsilon \\ &\implies t_{\mathrm{mix}}^{(n)}(\varepsilon) \leq \lfloor (1+\epsilon)\tau_n \rfloor \leq (1+\epsilon)\tau_n, \\ n \geq N_2 &\implies \varepsilon < ||p_n^{*\lfloor (1-\epsilon)\tau_n \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} < 2 - \varepsilon \\ &\implies t_{\mathrm{mix}}^{(n)}(\varepsilon) > \lfloor (1-\epsilon)\tau_n \rfloor > (1-\epsilon)\tau_n - 1. \end{aligned}$$

Therefore we have the following:

$$1 - \frac{1}{n} - \frac{1}{\tau_n} < \frac{t_{\mathrm{mix}}^{(n)}(\varepsilon)}{\tau_n} \leq 1 + \frac{1}{n} \text{ for all } n \geq \max\{N_0, N_1, N_2\} \implies \lim_{n\to\infty} \frac{t_{\mathrm{mix}}^{(n)}(\varepsilon)}{\tau_n} = 1.$$

Hence the conditions given in the statement hold because of the arbitrariness of ε.

Conversely, assume the conditions given in the statement are true. Now choose $\varepsilon,\ \epsilon \in (0,1)$ arbitrarily. Then

$$\lfloor (1+\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor \geq t_{\mathrm{mix}}^{(n)}(\varepsilon) \implies ||p_n^{*\lfloor (1+\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} < \varepsilon \text{ for all } n \geq 1.$$

Since $\varepsilon \in (0,1)$ was chosen arbitrarily, we can conclude that,

$$\lim_{n\to\infty} ||p_n^{*\lfloor (1+\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} = 0.$$

Again $\lim_{n\to\infty} t_{\mathrm{mix}}^{(n)}(\varepsilon) = \infty$ and $\lim_{n\to\infty} \dfrac{t_{\mathrm{mix}}^{(n)}(\varepsilon)}{t_{\mathrm{mix}}^{(n)}(1-\varepsilon)} = 1$ implies $\lim_{n\to\infty} \dfrac{\lfloor (1-\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor}{t_{\mathrm{mix}}^{(n)}(1-\varepsilon)} = 1 - \epsilon$. Therefore there exists positive integer N such that

$$\begin{aligned} n \geq N &\implies (1-2\epsilon) t_{\mathrm{mix}}^{(n)}(1-\varepsilon) < \lfloor (1-\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor < t_{\mathrm{mix}}^{(n)}(1-\varepsilon) \\ &\implies ||p_n^{*\lfloor (1-\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} > 1 - \varepsilon. \end{aligned}$$

Thus arbitrariness of $\epsilon \in (0,1)$ implies the following:

$$\lim_{n\to\infty} ||p_n^{*\lfloor (1-\epsilon) t_{\mathrm{mix}}^{(n)}(\varepsilon) \rfloor} - U_{\mathscr{G}_n}||_{\mathrm{TV}} = 1.$$

Thus the conditions of Definition 2.25 hold by choosing $\tau_n = t_{\mathrm{mix}}^{(n)}(\varepsilon)$. □

Remark 2.27. Proposition 2.26 shows that Definition 2.21 and Definition 2.25 are equivalent.

Chapter 3

The transpose top-2 with random shuffle

In this chapter we study the transpose top-2 with random shuffle on A_n. We introduce the transpose top-2 with random shuffle in Section 3.1 and study its basic properties. We find the spectrum of the transition matrix of this shuffle in Section 3.2. In Section 3.3, we give an upper bound for the total variation distance of the distribution after k transitions from the stationary distribution and prove that the mixing time is $O\left(\left(n-\frac{3}{2}\right)\log n\right)$. In Section 3.4, we obtain a lower bound for the total variation distance of the distribution after k transitions from the stationary distribution and prove the cutoff phenomenon.

3.1 Introduction

Recall that the *symmetric group*, denoted S_n, is the set of all bijections of the set $[n]$. The set S_n forms a group under composition. Elements of S_n are also known as permutations. A permutation in S_n which interchanges two elements of $[n]$ and fixes the rest is called a *transposition*. A permutation in S_n is said to be an *even permutation* if it can be expressed as a product of an even number of transpositions (not necessarily disjoint). The set of all even permutations in S_n forms a subgroup of the symmetric group, known as the *alternating group* and is denoted by A_n. The *transpose top-2 with random shuffle on A_n* is a random walk on A_n driven by a probability measure P_A on A_n defined as follows:

$$P_A(\pi)=\begin{cases}\frac{1}{2n-3}, & \text{if } \pi=\text{id}, \text{ the identity permutation},\\ \frac{1}{2n-3}, & \text{if } \pi=(i,n-1,n) \text{ for } i\in\{1,\ldots,n-2\},\\ \frac{1}{2n-3}, & \text{if } \pi=(i,n,n-1) \text{ for } i\in\{1,\ldots,n-2\},\\ 0, & \text{otherwise.}\end{cases} \tag{3.1.1}$$

This shuffle can also be described as a card shuffling problem. Suppose we have a deck of cards labelled from 1 to n such that the arrangement of the deck is a permutation in A_n. Then the *transpose top-2 with random shuffle on A_n* is a lazy variant of the following: First transpose the top two cards, then choose one of them and interchange it with a card randomly chosen from the remaining $(n-2)$ cards. More formally, any permutation in A_n can either go to itself or be multiplied on the right by a 3-cycle of the form $(i, n-1, n)$ or $(i, n, n-1)$ with probability $\frac{1}{2n-3}$.

Proposition 3.1. *The transpose top-2 with random shuffle on A_n is irreducible and aperiodic.*

Proof. We know that the 3-cycles generate A_n. Let a, b, c be any three distinct integers from $\{1, 2, \ldots, n\}$. If none of a, b, c is $(n-1)$ or n, we have,

$$\begin{aligned}(a,b,c) &= (c,n,n-1)(a,b,n-1)(c,n-1,n)\\ &= (c,n,n-1)(b,n,n-1)(a,n-1,n)(b,n-1,n)(c,n-1,n).\end{aligned}$$

If one of a, b, c is $(n-1)$ or n, without loss of any generality we may assume c is either $(n-1)$ or n and we have the following,

$$\begin{aligned}&(a,b,n-1) = (b,n,n-1)(a,n-1,n)(b,n-1,n), \text{ and}\\ &(a,b,n) = (b,n-1,n)(a,n,n-1)(b,n,n-1).\end{aligned}$$

If any two of a, b, c are $(n-1)$ and n, then (a, b, c) takes the form $(\,\cdot\,, n-1, n)$ or $(\,\cdot\,, n, n-1)$. Therefore the support of the measure P_A generates A_n and hence the chain is irreducible (Proposition 2.22). Given any $\pi \in A_n$, the set of all times when it is possible for the chain to return to the starting state π contains the integer 1 ($\because P_A(\text{id}) \neq 0$). Therefore the period of the state π is 1 and hence from irreducibility all the states of this chain have period 1. Thus this chain is aperiodic. □

We have seen in Chapter 2 that there exists a unique stationary distribution for an irreducible Markov chain. Moreover if the chain is aperiodic then the distribution after the kth transition converges to the stationary distribution as $k \to \infty$ [66, Theorem 4.9]. Therefore, Proposition 3.1 gives an affirmative answer to the question of existence and uniqueness of a stationary distribution. It also says that after a large number of transitions the distribution of the chain behaves like the stationary distribution. The stationary distribution for the random walk on A_n driven by P_A is the uniform distribution on A_n (Proposition 2.23). Thus the distribution after kth transition for the transpose top-2 with random shuffle converges to U_{A_n} as $k \to \infty$.

3.2 Spectrum for the transpose top-2 with random operator

Let $\mathcal{P} = \text{id} + \sum_{i=1}^{n-2} ((i, n-1, n) + (i, n, n-1)) \in \mathbb{C}[A_n]$. Recall that $\widehat{P_A}(R)$ is the Fourier transform of the probability measure P_A at the right regular representation R of A_n, and can be written as $\widehat{P_A}(R) = \frac{1}{2n-3}\mathcal{P}$. Here we consider the action of the operator $\mathcal{P}$ on $\mathbb{C}[A_n]$ by multiplication on the right. To obtain the eigenvalues of $\mathcal{P}$, we use the representation theory of A_n. Most of the notations here are borrowed from Ruff [86]. We now define the Young-Jucys-Murphy elements for A_n and establish its connection to $\mathcal{P}$.

Definition 3.2 (Ruff [86])**.** The *Young-Jucys-Murphy* elements $X_1^A, \dots, X_n^A \in \mathbb{C}[A_n]$ are defined by $X_1^A = 0,\ X_2^A = \text{id}$ and $X_i^A = (1,2)Y_i$ for $i \geq 3$, where $Y_1, Y_2, \dots Y_n$ are the usual Young-Jucys-Murphy elements (defined in (2.1.3)) for S_n.

Let us denote $s_i = (i, i+1)$ for $1 \leq i < n$. Then $\{s_1, \dots, s_{n-1}\}$ is a set of generators of the symmetric group S_n. A_n is generated by $t_2, \dots, t_{n-1}$, where $t_i = (1,2)s_i$ for $i \in \{2, \dots, n-1\}$ [86, Remark 2.3]. As the generators $s_1, \dots, s_{n-1}$ of the symmetric group satisfy

$$s_iY_j = Y_js_i,\ s_iY_i = Y_{i+1}s_i - \text{id} \text{ for all } 1 \leq i < n \text{ with } |i-j| > 1,$$

we have the following:

$$t_iX_i^A = X_{i+1}^A t_i - \text{id} \quad \text{for all } 3 \leq i < n.$$

Lemma 3.3. $\mathcal{P} = \text{id}$ *if* $n = 2$, $\mathcal{P} = \text{id} + X_3^A$ *if* $n = 3$ *and for* $n > 3$, *we have*

$$\mathcal{P} = t_{n-1}\left(X_n^A + X_{n-1}^A\right).$$

Proof. The cases of $n = 2, 3$ are just verification. We prove this lemma for $n > 3$.

$$\begin{aligned}
\mathcal{P} &= \text{id} + \sum_{i=1}^{n-2} ((i, n-1, n) + (i, n, n-1)) \\
&= \text{id} + \sum_{i=1}^{n-2} ((n, n-1)(n, i) + (n-1, n)(n-1, i)) \\
&= (n, n-1)\left(\sum_{i=1}^{n-1}(n, i) + \sum_{i=1}^{n-2}(n-1, i)\right) \\
&= s_{n-1}(Y_n + Y_{n-1}) = (1,2)s_{n-1}(1,2)(Y_n + Y_{n-1}) = t_{n-1}\left(X_n^A + X_{n-1}^A\right). \qquad \square
\end{aligned}$$

Remark 3.4. We note that $X_i^A X_j^A = X_j^A X_i^A$ for all $1 \leq i, j \leq n$ and the common eigenvectors for X_i^A's form a basis for the irreducible representations of A_n.

Let us recall that $\mathrm{Par}(n)$ denotes the set of all partitions of n. Let $\lambda \vdash n$. Also recall that λ' is the conjugate of the partition λ and $\mathrm{tab}(\lambda)$ denotes the set of all standard Young tableaux of shape λ. A standard Young tableau T is said to be an *upper standard Young tableau* if $c(b_T(2)) = 1$, where we recall that $c(b_T(2))$ is the content of the box in T containing 2. The collection of all upper standard tableaux of a given shape λ is denoted by $\mathrm{UStd}(\lambda)$.

Lemma 3.5. *For $n > 1$, the cardinality of $\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)$ is half the cardinality of $\underset{\lambda\vdash n}{\cup}\mathrm{tab}(\lambda)$. Moreover, for self-conjugate $\lambda \vdash n$ we have, $|\,\mathrm{UStd}(\lambda)\,| = \frac{1}{2}|\,\mathrm{tab}(\lambda)|$.*

Proof. Let us consider the set $\mathrm{LStd}(\lambda) := \{T \in \mathrm{tab}(\lambda) \mid c(b_T(2)) = -1\}$. Then by sending each element of $\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)$ to its transpose (i.e. reflecting it with respect to the diagonal containing boxes with content 0), we have a one to one correspondence between $\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)$ and $\underset{\lambda\vdash n}{\cup}\mathrm{LStd}(\lambda)$. Thus $\left|\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)\right| = \frac{1}{2}\left|\underset{\lambda\vdash n}{\cup}\mathrm{tab}(\lambda)\right|$ follows from

$$\left(\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)\right)\cap\left(\underset{\lambda\vdash n}{\cup}\mathrm{LStd}(\lambda)\right) = \phi \text{ and } \left(\underset{\lambda\vdash n}{\cup}\mathrm{UStd}(\lambda)\right)\cup\left(\underset{\lambda\vdash n}{\cup}\mathrm{LStd}(\lambda)\right) = \underset{\lambda\vdash n}{\cup}\mathrm{tab}(\lambda).$$

Also, for self-conjugate $\lambda \vdash n$, $\mathrm{UStd}(\lambda)\cup\mathrm{LStd}(\lambda) = \mathrm{tab}(\lambda)$ and the same map as above gives a bijection from $\mathrm{LStd}(\lambda)$ to $\mathrm{UStd}(\lambda)$). Therefore, $|\mathrm{UStd}(\lambda)| = \frac{1}{2}\,|\mathrm{tab}(\lambda)|$. ☐

We now describe all the irreducible representations of A_n (for more details, see [86]). Corresponding to each non-self-conjugate partition λ of n, there is an irreducible representation D_λ of A_n. Given any non-self-conjugate partition λ of n, the irreducible representations D_λ and $D_{\lambda'}$ of A_n are isomorphic. For each self-conjugate partition λ of n, there are two non-isomorphic irreducible representations D_λ^+ and D_λ^- of A_n. All the irreducible representations of A_n are given by D_λ, $\lambda \vdash n$ non-self-conjugate and $D_\lambda^\pm$, $\lambda \vdash n$ self-conjugate. The basis elements of D_λ are identified by the elements of $\mathrm{UStd}(\lambda)\cup\mathrm{UStd}(\lambda')$ for non-self-conjugate $\lambda \vdash n$ and that of $D_\lambda^\pm$ are identified by the elements of $\mathrm{UStd}(\lambda)$ for self-conjugate $\lambda \vdash n$. Therefore we have the following:

$$\dim(D_\lambda) = |\mathrm{UStd}(\lambda)| + |\mathrm{UStd}(\lambda')| = |\mathrm{tab}(\lambda)| = f^\lambda, \text{ for non-self-conjugate } \lambda \vdash n \text{ and}$$
$$\dim(D_\lambda^+) = \dim(D_\lambda^-) = |\mathrm{UStd}(\lambda)| = \frac{1}{2}\,|\mathrm{tab}(\lambda)| = \frac{1}{2}f^\lambda, \text{ for self-conjugate } \lambda \vdash n.$$

Let us consider two subsets CPar(n) and NCPar(n) of Par(n) defined as follows:

$$\mathrm{CPar}(n) = \{\lambda \in \mathrm{Par}(n) \mid \lambda = \lambda'\} \text{ and}$$
$$\mathrm{NCPar}(n) = \left\{\lambda \in \mathrm{Par}(n) \middle| \lambda \neq \lambda', \begin{array}{l}\lambda \text{ contains more boxes of positive} \\ \text{content than that of } \lambda'\end{array}\right\}.$$

Recall from (2.1.1) that, in the regular representation of a finite group G, each irreducible representation of G occurs with multiplicity equal to its dimension. Therefore from the above discussion, we have the following:

$$\mathbb{C}[A_n] \cong \left(\bigoplus_{\lambda\in\mathrm{NCPar}(n)} f^\lambda D_\lambda\right) \oplus \left(\bigoplus_{\lambda\in\mathrm{CPar}(n)} \frac{f^\lambda}{2} D_\lambda^+\right) \oplus \left(\bigoplus_{\lambda\in\mathrm{CPar}(n)} \frac{f^\lambda}{2} D_\lambda^-\right). \tag{3.2.1}$$

Now we discuss the actions of the generators $t_i,\ 3 \leq i \leq n-1$ and the Young-Jucys-Murphy elements on the irreducible representations of A_n (we don't need the action of t_2 on irreducible representations of A_n in this work). Given any partition λ of n, let us define $\alpha = (a_1, \dots, a_n) := (c(b_{T_\alpha}(1)), \dots, c(b_{T_\alpha}(n)))$, where $T_\alpha \in \mathrm{UStd}(\lambda) \cup \mathrm{UStd}(\lambda')$ $(= \mathrm{UStd}(\lambda)$, if λ is self-conjugate) and $c(b_{T_\alpha}(i))$ denotes the content of the box containing i in T_α. Ruff [86] showed that for $\lambda \in \mathrm{NCPar}(n)$, if v_α is the basis element of D_λ corresponding to $T_\alpha \in \mathrm{UStd}(\lambda) \cup \mathrm{UStd}(\lambda')$, then $X_i^A v_\alpha = a_i v_\alpha$ for all $1 \leq i \leq n$. Moreover for $3 \leq i < n$, the action of t_i on v_α is given as follows:

$$t_i v_\alpha = \frac{1}{a_{i+1} - a_i}\, v_\alpha + \sqrt{-1}\, (-1)^{\alpha,i} \sqrt{1 - \frac{1}{(a_{i+1} - a_i)^2}}\ v_{t_i\alpha}, \tag{3.2.2}$$

where

$$(-1)^{\alpha,i} = \begin{cases} 1 & \text{if } a_i < a_{i+1}, \\ -1 & \text{if } a_i > a_{i+1}, \end{cases} \qquad \text{and}$$

$t_i\alpha := (a_1, \dots, a_{i-1}, a_{i+1}, a_i, a_{i+2}, \dots a_n)$ if $a_{i+1} \neq a_i \pm 1$. We don't need $t_i\alpha$ when $a_{i+1} = a_i \pm 1$, because the coefficient of $v_{t_i\alpha}$ in the expression (3.2.2) is zero in that case.

Also for $\lambda \in \mathrm{CPar}(n)$, if v_α^+ (respectively v_α^-) is the basis element of D_λ^+ (respectively D_λ^-) corresponding to $T_\alpha \in \mathrm{UStd}(\lambda)$, then the actions of $X_i^A,\ 1 \leq i \leq n$ and $t_i,\ 3 \leq i < n$ on $v_\alpha^\pm$ are same as their respective actions on v_α in case of $\lambda \in \mathrm{NCPar}(n)$. Now we are in a position to find the eigenvalues of $\mathcal{P}$.

Theorem 3.6. *For a non-self-conjugate partition $\lambda \vdash n$, each $T_\alpha \in \mathrm{UStd}(\lambda) \cup \mathrm{UStd}(\lambda')$ provides an eigenvalue of $\mathcal{P}$. Let $\alpha = (a_1, \dots, a_n) := (c(b_{T_\alpha}(1)), \dots, c(b_{T_\alpha}(n)))$. Then*

1. *$2a_n - 1$ is an eigenvalue of $\mathcal{P}$, if $a_n = a_{n-1} + 1$.*

2. *$-(2a_n + 1)$ is an eigenvalue of $\mathcal{P}$, if $a_n = a_{n-1} - 1$.*

3. $\pm(a_n + a_{n-1})$ *are eigenvalues of* $\mathcal{P}$*, if* $a_n \neq a_{n-1} \pm 1$ *and* $a_{n-1} < a_n$.

Moreover, each eigenvalue has multiplicity f^λ.

Proof. The theorem is trivially true for $n = 2, 3$. Now we prove the theorem for $n > 3$. For any non-self-conjugate $\lambda \vdash n$, we can choose a basis element v_α of D_λ corresponding to $T_\alpha \in \mathrm{UStd}(\lambda) \cup \mathrm{UStd}(\lambda')$ such that $X_i^A v_\alpha = a_i v_\alpha$ for all $i = 1, \ldots, n$. Now a basis $\mathcal{B}$ of D_λ is the union of the following three sets:

$$\mathcal{B}_1 := \{v_\alpha \mid a_n = a_{n-1} + 1\},\ \mathcal{B}_2 := \{v_\alpha \mid a_n = a_{n-1} - 1\},\ \mathcal{B}_3 := \{v_\alpha \mid a_n \neq a_{n-1} \pm 1\}.$$

For any upper standard Young tableau $T_\alpha \in \mathcal{B}_3$, we have another upper standard Young tableau $T_{t_{n-1}\alpha} \in \mathcal{B}_3$. Therefore, the cardinality of $\mathcal{B}_3$ is even and $\mathcal{B}_3 = \{v_\alpha,\ v_{t_{n-1}\alpha} \mid a_n \neq a_{n-1} \pm 1,\ a_{n-1} < a_n\}$. Again for $a_n \neq a_{n-1} \pm 1$, from (3.2.2), we have

$$\mathbb{C}\text{-Span}\ \{v_\alpha,\ v_{t_{n-1}\alpha}\} = \mathbb{C}\text{-Span}\ \{v_\alpha,\ t_{n-1}v_\alpha\}.$$

Therefore $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \{v_\alpha,\ t_{n-1}v_\alpha \mid a_n \neq a_{n-1} \pm 1,\ a_{n-1} < a_n\}$ is a basis for D_λ. Let us consider the ordered basis $\mathcal{B}'$ of D_λ in which we first collect all the vectors from $\mathcal{B}_1$. Then all the vectors from $\mathcal{B}_2$ and finally the pair of vectors $(v_\alpha,\ t_{n-1}v_\alpha)$ from $\{v_\alpha,\ t_{n-1}v_\alpha \mid a_n \neq a_{n-1} \pm 1,\ a_{n-1} < a_n\}$. For $v_\alpha \in \mathcal{B}_1$,

$$\begin{aligned}
\mathcal{P}v_\alpha &= t_{n-1}\left(X_{n-1}^A + X_n^A\right)v_\alpha, \quad \text{by Lemma 3.3}\\
&= (a_{n-1} + a_n)t_{n-1}v_\alpha\\
&= (a_{n-1} + a_n)v_\alpha, \quad \text{by (3.2.2)}\\
&= (2a_n - 1)v_\alpha.
\end{aligned}$$

Again for $v_\alpha \in \mathcal{B}_2$,

$$\begin{aligned}
\mathcal{P}v_\alpha &= t_{n-1}\left(X_{n-1}^A + X_n^A\right)v_\alpha, \quad \text{by Lemma 3.3}\\
&= (a_{n-1} + a_n)t_{n-1}v_\alpha\\
&= -(a_{n-1} + a_n)v_\alpha, \quad \text{by (3.2.2)}\\
&= -(2a_n + 1)v_\alpha.
\end{aligned}$$

Therefore $\mathcal{P}$ acts on $\mathcal{B}_1$ and $\mathcal{B}_2$ diagonally. Now for $v_\alpha \in \mathcal{B}_3$, using $t_{n-1}X_{n-1}^A = X_n^A t_{n-1} - \mathrm{id}$ and $t_{n-1}^2 = \mathrm{id}$, the matrix for the action of $\mathcal{P}$ on $\{v_\alpha,\ t_{n-1}v_\alpha\}$ is given below,

$$\begin{pmatrix} 0 & a_{n-1} + a_n \\ a_{n-1} + a_n & 0 \end{pmatrix}. \tag{3.2.3}$$

The eigenvalues of the above 2×2 matrix (given in (3.2.3)) are $\pm(a_{n-1}+a_n)$. Therefore, the matrix of $\mathcal{P}$ with respect to the ordered basis $\mathcal{B}'$, is a block diagonal matrix, where first $|\mathcal{B}_1|$ blocks are the 1×1 matrix $(2a_n-1)$ corresponding to each α in $\{\alpha \mid a_n = a_{n-1}+1\}$, next $|\mathcal{B}_2|$ blocks are the 1×1 matrix $-(2a_n+1)$ corresponding to each α in $\{\alpha \mid a_n = a_{n-1}-1\}$ and last $|\{\alpha : a_n \neq a_{n-1}\pm 1,\ a_{n-1} < a_n\}|$ blocks are the 2×2 matrix (3.2.3) corresponding to each α in $\{\alpha : a_n \neq a_{n-1}\pm 1,\ a_{n-1} < a_n\}$. The argument for the multiplicity of the eigenvalues follows from (3.2.1). Thus the theorem follows. □

Remark 3.7. Let λ be a non-self-conjugate partition of n. Then Theorem 3.6 shows that the sets of eigenvalues obtained by considering the partitions λ and λ' are the same.

Theorem 3.8. *For a self-conjugate partition $\lambda \vdash n$, each $T_\alpha \in \mathrm{UStd}(\lambda)$ provides an eigenvalue of $\mathcal{P}$. If $\alpha = (a_1,\dots,a_n) := (c(b_{T_\alpha}(1)),\dots,c(b_{T_\alpha}(n)))$, then the eigenvalue corresponding to T_α is given as follows:*

1. $2a_n - 1$ *is an eigenvalue of* $\mathcal{P}$, *if* $a_n = a_{n-1}+1$.
2. $-(2a_n+1)$ *is an eigenvalue of* $\mathcal{P}$, *if* $a_n = a_{n-1}-1$.
3. $\pm(a_n + a_{n-1})$ *are eigenvalues of* $\mathcal{P}$, *if* $a_n \neq a_{n-1}\pm 1$ *and* $a_{n-1} < a_n$.

Moreover, each eigenvalue has multiplicity $|\,\mathrm{tab}(\lambda)| = \dim(D^+_\lambda) + \dim(D^-_\lambda)$.

Proof. The proof is similar to the proof of Theorem 3.6. Proof of this theorem follows by replacing D_λ, v_α and f^λ by $D^\pm_\lambda$, $v^\pm_\alpha$ and $\frac{f^\lambda}{2}$ respectively, in the proof of Theorem 3.6. □

Corollary 3.9. *For a partition $\lambda \vdash n$, each $T_\alpha \in \mathrm{UStd}(\lambda)\cup\mathrm{UStd}(\lambda')$ provides an eigenvalue of $\mathcal{P}$. If $\alpha = (a_1,\dots,a_n) := (c(b_{T_\alpha}(1)),\dots,c(b_{T_\alpha}(n)))$, then the eigenvalue corresponding to T_α is given as follows:*

1. $2a_n - 1$ *is an eigenvalue of* $\mathcal{P}$, *if* $a_n = a_{n-1}+1$.
2. $-(2a_n+1)$ *is an eigenvalue of* $\mathcal{P}$, *if* $a_n = a_{n-1}-1$.
3. $\pm(a_n + a_{n-1})$ *are eigenvalues of* $\mathcal{P}$, *if* $a_n \neq a_{n-1}\pm 1$ *and* $a_{n-1} < a_n$.

Moreover, each eigenvalue has multiplicity $f^\lambda := |\mathrm{tab}(\lambda)|$.

Proof. If $\lambda \vdash n$ is non-self-conjugate, then the result follows directly from Theorem 3.6. Now if $\lambda \vdash n$ is self-conjugate, then we have $\mathrm{UStd}(\lambda) = \mathrm{UStd}(\lambda)\cup\mathrm{UStd}(\lambda')$. Therefore, in the case of self-conjugate λ, this result follows from Theorem 3.8. □

Remark 3.10. Corollary 3.9, together with (3.2.1), determines the spectrum of $\mathcal{P}$ and hence that of $\widehat{P_A}(R)$.

Remark 3.11. The Cayley graph of A_n with generating set $\{(1,2,i),(1,i,2) \mid 3 \le i \le n\}$ is well studied in the graph theory literature. Let $\mathscr{A}$ denote the adjacency matrix of this graph. The second largest eigenvalue of $\mathscr{A}$ has been computed by Xueyi Huang and Qiongxiang Huang [58]. Our result provides an alternative method to obtain the eigenvalues of $\mathscr{A}$. The matrix $\mathscr{A}$ is the matrix of the right multiplication action of $(1,n)(2,n-1)\,(\mathcal{P}-\mathrm{id})\,(2,n-1)(1,n)$ on $\mathbb{C}[A_n]$. Thus $\mathscr{A}$ is similar to the matrix $\mathcal{P}-I$, where I is the identity matrix of order $\frac{n!}{2}$. Hence $\mathscr{A}$ and $\mathcal{P}-I$ have the same spectrum. Thus Corollary 3.9, together with (3.2.1), determines all the eigenvalues of $\mathscr{A}$.

Example 3.12. If $n = 4$, then eigenvalues of $\widehat{P_A}(R)$ are the following:

Eigenvalues:	1	$\frac{3}{5}$	$\frac{1}{5}$	$-\frac{1}{5}$
Multiplicities:	1	3	3	5

For the only element $T_{(0,1,2,3)} = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\ \hline\end{array}$ of $\mathrm{UStd}((4)) \cup \mathrm{UStd}((1,1,1,1))$ we have, $a_3 = 2,\ a_4 = 3$. Hence $a_4 = a_3 + 1$ and the eigenvalue of $\widehat{P_A}(R)$ corresponding to $T_{(0,1,2,3)}$ is 1 with multiplicity 1. The elements of $\mathrm{UStd}((3,1)) \cup \mathrm{UStd}((2,1,1))$ are listed below:

$$T_{(0,1,2,-1)} = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}\ , \quad T_{(0,1,-1,2)} = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1}\end{array}\ , \quad T_{(0,1,-1,-2)} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array}\ .$$

Now $a_3 = 2,\ a_4 = -1$ for $T_{(0,1,2,-1)}$ and $a_3 = -1,\ a_4 = 2$ for $T_{(0,1,-1,2)}$. Thus for both $T_{(0,1,2,-1)}$ and $T_{(0,1,-1,2)}$ we have $a_4 \ne a_3 \pm 1$. In order to satisfying $a_3 < a_4$ we choose $T_{(0,1,-1,2)}$ and the eigenvalues of $\widehat{P_A}(R)$ in this case are $\pm\frac{1}{5}$ with multiplicity 3 each. Again for $T_{(0,1,-1,-2)}$ we have $a_3 = -1,\ a_4 = -2$ which satisfies $a_4 = a_3 - 1$. Thus the eigenvalue of $\widehat{P_A}(R)$ corresponding to $T_{(0,1,-1,-2)}$ is $\frac{3}{5}$ with multiplicity 3. Finally considering the only element

$$T_{(0,1,-1,0)} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline\end{array}$$

of $\mathrm{UStd}((2,2))$ we have $a_3 = -1,\ a_4 = 0$ and thus $a_4 = a_3 + 1$. Therefore the eigenvalue of $\widehat{P_A}(R)$ corresponding to $T_{(0,1,-1,0)}$ is $-\frac{1}{5}$ with multiplicity 2.

Proposition 3.13. *Given $n \ge 4$, the eigenvalues of $\widehat{P_A}(R)$ for the $(n-1)$-dimensional irreducible representation $D_{(n-1,1)}$ (or $D_{(2,1^{n-2})}$) are given as follows:*

Eigenvalues:	$\frac{n-3}{2n-3}$	$-\frac{n-3}{2n-3}$	$\frac{2n-5}{2n-3}$
Multiplicities:	$n-1$	$n-1$	$(n-3)(n-1)$

Proof. First consider the elements

$$\begin{array}{|c|c|c|c|c|}\hline 1 & 2 & \cdots & n\text{-}1 & n \\ \hline i \\ \cline{1-1}\end{array}\ , \quad 3 \le i \le n-2$$

of $\mathrm{UStd}((n-1,1))$. Each of these elements satisfies $a_n = a_{n-1}+1$ as $a_{n-1} = n-3,\ a_n = n-2$. Therefore the eigenvalues of $\widehat{P_A}(R)$ corresponding to each of these elements are $\frac{2n-5}{2n-3}$ with multiplicity $n-1$ each. There are $n-4$ such upper standard tableaux, thus multiplicity of the eigenvalue $\frac{2n-5}{2n-3}$ is $(n-4)(n-1)$. Now considering the element

1	2	⋯	n-1
n			

of $\mathrm{UStd}((n-1,1))$, we have $a_{n-1} = n-2,\ a_n = -1$ and thus $a_n \neq a_{n-1} \pm 1$ but $a_{n-1} \not< a_n$. Therefore we do not select this upper standard tableaux. For the element

1	2	⋯	n
n-1			

of $\mathrm{UStd}((n-1,1))$, we have $a_{n-1} = -1,\ a_n = n-2$. Hence this upper standard tableaux satisfies $a_n \neq a_{n-1} \pm 1$ and $a_{n-1} < a_n$. Therefore the eigenvalues of $\widehat{P_A}(R)$ corresponding to this upper standard tableaux are $\pm\frac{n-3}{2n-3}$ with multiplicity $n-1$ each. Finally we consider the only element

1	2
3	
4	
⋮	
n-1	
n	

of $\mathrm{UStd}(2, 1^{n-2})$. This upper standard tableaux satisfies $a_n = a_{n-1} - 1$ as $a_{n-1} = -(n-3),\ a_n = -(n-2)$. Thus the eigenvalue of $\widehat{P_A}(R)$ corresponding to this upper standard tableaux is $\frac{2n-5}{2n-3}$ with multiplicity $n-1$. □

3.3 Upper bound of the mixing time

In this section we find an upper bound of $||P_A^{*k} - U_{A_n}||_{\mathrm{TV}}$ for $k \geq \left(n - \frac{3}{2}\right)(\log n + c),\ c > 0$. This gives an upper bound for the mixing time. The main theorem of this section is the following.

Theorem 3.14. *For the random walk on A_n driven by P_A, we have the following:*

1. $||P_A^{*k} - U_{A_n}||_{\mathrm{TV}} < \frac{1}{\sqrt{2}} e^{-c}$, *for* $k \geq (n - \frac{3}{2})(\log n + c)$ *and* $c > 0$.

2. $\lim\limits_{n\to\infty} ||P_A^{*k_n} - U_{A_n}||_{\mathrm{TV}} = 0$, *for any* $\epsilon \in (0,1)$ *and* $k_n = \lfloor (1+\epsilon)(n - \frac{3}{2}) \log n \rfloor$.

Proof. We know that the trace of the $(2k)$th power of a matrix is the sum of $(2k)$th powers of its eigenvalues. Now for $\lambda \vdash n$ we can say by Corollary 3.9, as T_α ranges over $\mathrm{UStd}(\lambda)\cup\mathrm{UStd}(\lambda')$, the eigenvalues of $\widehat{P_A}(R)$ are

- $\frac{2a_n-1}{2n-3} = \frac{a_n+a_{n-1}}{2n-3}$, if $a_n = a_{n-1}+1$,
- $-\frac{(2a_n+1)}{2n-3} = -\frac{a_n+a_{n-1}}{2n-3}$, if $a_n = a_{n-1}-1$,
- $\pm\frac{(a_n+a_{n-1})}{2n-3} = \pm\frac{a_n+a_{n-1}}{2n-3}$ if $a_n \neq a_{n-1}\pm 1$ and $a_{n-1} < a_n$,

where a_{n-1} (respectively a_n) is the content of $(n-1)$ (respectively n) in T_α. Lemma 2.24 implies:

$$\begin{aligned}4\left|\left|P^{*k}-U_{A_n}\right|\right|_{\mathrm{TV}}^2 \leq &\sum_{\lambda\in\mathrm{NCPar}(n)\setminus\{(n)\}}\left(f^\lambda\sum_{T_\alpha\in\mathrm{UStd}(\lambda)\,\cup\,\mathrm{UStd}(\lambda')}\left(\frac{a_n+a_{n-1}}{2n-3}\right)^{2k}\right)\\ &+\sum_{\lambda\in\mathrm{CPar}(n)}\left(\frac{f^\lambda}{2}\sum_{T_\alpha\in\mathrm{UStd}(\lambda)}\left(\frac{a_n+a_{n-1}}{2n-3}\right)^{2k}+\frac{f^\lambda}{2}\sum_{T_\alpha\in\mathrm{UStd}(\lambda)}\left(\frac{a_n+a_{n-1}}{2n-3}\right)^{2k}\right).\end{aligned} \tag{3.3.1}$$

Before coming to the main part of the proof, let us consider the leading term in (3.3.1), which corresponds to the partition $\lambda = (n-1,1)$ or equivalently its conjugate. For the partition $\lambda = (n-1,1)$, the eigenvalues are $\frac{2n-5}{2n-3}$, $-\frac{n-3}{2n-3}$ and $\frac{n-3}{2n-3}$ with algebraic multiplicities $n-3$, 1 and 1, respectively. Therefore, the term in $\sum\limits_{\lambda\in\mathrm{NCPar}(n)\setminus\{(n)\}}$ is

$$\begin{aligned}&(n-1)\left((n-3)\left(\frac{2n-5}{2n-3}\right)^{2k}+\left(-\frac{n-3}{2n-3}\right)^{2k}+\left(\frac{n-3}{2n-3}\right)^{2k}\right)\\ &=(n-1)\left((n-3)\left(1-\frac{2}{2n-3}\right)^{2k}+2\left(\frac{n-3}{2n-3}\right)^{2k}\right)=O\left(e^{-\frac{4k}{2n-3}+2\log n}\right).\end{aligned}$$

Now, if $k=(n-\frac{3}{2})(\log n+c)$, then $e^{-\frac{4k}{2n-3}+2\log n}$ is e^{-2c}, $c>0$. We show that this is the largest term, other terms being smaller.

For any upper standard Young tableau T_α of shape λ, if $\lambda=(\lambda_1,\ldots,\lambda_r)$, then $a_{n-1}+a_n\leq 2\lambda_1-3$. Then the right hand side of (3.3.1) is less than or equal to

$$\begin{aligned}&\sum_{\lambda\in\mathrm{NCPar}(n)\setminus\{(n)\}}\left(f^\lambda\sum_{T_\alpha\in\mathrm{UStd}(\lambda)\,\cup\,\mathrm{UStd}(\lambda')}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}\right)\\ &\quad+\sum_{\lambda\in\mathrm{CPar}(n)}\left(\frac{f^\lambda}{2}\sum_{T_\alpha\in\mathrm{UStd}(\lambda)}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}+\frac{f^\lambda}{2}\sum_{T_\alpha\in\mathrm{UStd}(\lambda)}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}\right).\end{aligned} \tag{3.3.2}$$

The summands in $\sum\limits_{T_\alpha\in\mathrm{UStd}(\lambda)\,\cup\,\mathrm{UStd}(\lambda')}$ and $\sum\limits_{T_\alpha\in\mathrm{UStd}(\lambda)}$ are independent of the index of the sum. Therefore, the expression in (3.3.2) becomes

$$\begin{aligned}&\sum_{\lambda\in\mathrm{NCPar}(n)\setminus\{(n)\}}\left(\left(f^{\lambda}\right)^2\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}\right)+\sum_{\lambda\in\mathrm{CPar}(n)}\left(\left(\frac{f^{\lambda}}{2}+\frac{f^{\lambda}}{2}\right)\frac{f^{\lambda}}{2}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}\right)\\ =&\sum_{\lambda\in\mathrm{NCPar}(n)\setminus\{(n)\}}\left(f^{\lambda}\right)^2\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}+\sum_{\lambda\in\mathrm{CPar}(n)}\frac{\left(f^{\lambda}\right)^2}{2}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}. \qquad (3.3.3)\end{aligned}$$

Now adding the non-negative quantities

$$\sum_{\substack{\lambda\vdash n\\ \lambda\notin\mathrm{NCPar}(n)\cup\mathrm{CPar}(n)}}\left(f^{\lambda}\right)^2\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}\quad\text{and}\quad\sum_{\lambda\in\mathrm{CPar}(n)}\frac{\left(f^{\lambda}\right)^2}{2}\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}$$

to (3.3.3), we obtain $\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}}\left(f^{\lambda}\right)^2\left(\frac{2\lambda_1-3}{2n-3}\right)^{2k}$. Therefore, the expression in (3.3.3) is less than or equal to

$$\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}}\left(f^{\lambda}\right)^2\left(\frac{\lambda_1-\frac{3}{2}}{n-\frac{3}{2}}\right)^{2k}<e^{n^2e^{-\frac{2k}{n-1.5}}}-1. \qquad (3.3.4)$$

Here (3.3.4) follows from Corollary 2.11 by taking $\ell=n$ and $a=b=\frac{3}{2}$. Therefore we have

$$4\,||P_A^{*k}-U_{A_n}||_{\mathrm{TV}}^2<\ e^{n^2e^{-\frac{2k}{n-1.5}}}-1. \qquad (3.3.5)$$

Now for $c>0$ and $k\geq(n-\frac{3}{2})(\log n+c)$, we have

$$\begin{aligned}4\,||P_A^{*k}-U_{A_n}||_{\mathrm{TV}}^2<&e^{n^2e^{-\frac{2k}{n-1.5}}}-1\\ \leq&e^{e^{-2c}}-1\leq 2e^{-2c}.\end{aligned}$$

Therefore, $||P_A^{*k}-U_{A_n}||_{\mathrm{TV}}<\frac{1}{\sqrt{2}}e^{-c}$. This proves the first part of this theorem.

Now let $\epsilon\in(0,1)$ and $k_n=\lfloor(1+\epsilon)(n-\frac{3}{2})\log n\rfloor$. Then $k_n+1\geq(1+\epsilon)(n-\frac{3}{2})\log n$. Therefore (3.3.5) implies,

$$0\leq 4\,||P_A^{*k_n}-U_{A_n}||_{\mathrm{TV}}^2<\ e^{n^2e^{-\frac{2k_n}{n-1.5}}}-1\leq e^{n^{-2\epsilon}e^{\frac{2}{n-1.5}}}-1, \qquad (3.3.6)$$

the last inequality of (3.3.6) holds because of $k_n+1\geq(1+\epsilon)(n-\frac{3}{2})\log n$. Thus the second part follows from the fact

$$\lim_{n\to\infty}e^{n^{-2\epsilon}e^{\frac{2}{n-1.5}}}-1=0. \qquad \square$$

3.4 Lower bound of the mixing time

In this section, we find a lower bound of the total variation distance $||P_A^{*k} - U_{A_n}||_{\text{TV}}$ when $k = (n - \frac{3}{2})(\log n + c)$ for $c \ll 0$. To prove the results, we consider the slow term in the upper bound lemma [32]. The slow term comes from the $(n-1)$-dimensional irreducible representation of A_n. In particular, we define a random variable on A_n giving the number of fixed points of even permutations. This random variable can be viewed as the character of the restriction of *defining* representation ρ^{def} (recall Definition 2.12 from Chapter 2) to A_n. The restriction of defining representation decomposes into two irreducible representations of A_n namely the trivial representation and the $(n-1)$-dimensional representation. Thus the character of the $(n-1)$-dimensional irreducible representation plays an important role in this section.

We have seen all the irreducible representations of A_n in Section 3.2. Let us recall from Chapter 2 that the irreducible representations of S_n are indexed by the partitions of n and the irreducible representation indexed by $\lambda \vdash n$ is S^λ. Also [81, Theorem 4.6.5] says that the restriction of the irreducible representation S^λ of S_n to A_n is an irreducible representation of A_n if $\lambda \neq \lambda'$ and a direct sum of two non-isomorphic irreducible representations of A_n if $\lambda = \lambda'$. Let ψ^λ denote the irreducible character of S_n corresponding to the irreducible representation given by the partition λ of n. If $\lambda \vdash n$ is non-self-conjugate, then we denote the irreducible character of A_n corresponding to λ by χ^λ and if $\lambda \vdash n$ is self-conjugate, then we denote the irreducible characters of A_n corresponding to λ by $\chi^\lambda_\pm$. We abbreviate the induced character $\chi \uparrow_{A_n}^{S_n}$ to $\chi \uparrow^{S_n}$ and the restricted character $\chi \downarrow_{A_n}^{S_n}$ to $\chi \downarrow_{A_n}$.

Lemma 3.15. *If λ is a non-self-conjugate partition of n, then $\psi^\lambda \downarrow_{A_n} = \chi^\lambda = \chi^{\lambda'}$. If λ is self-conjugate, then we have $\psi^\lambda \downarrow_{A_n} = \chi^\lambda_+ + \chi^\lambda_-$.*

Proof. The proof of this lemma follows directly from [81, Theorem 4.6.5]. □

Proposition 3.16. *For non-self-conjugate $\lambda \vdash n$, we have $\chi^\lambda \uparrow^{S_n} = \psi^\lambda + \psi^{\lambda'}$.*

Proof. S_n can be written as the disjoint union of A_n and $(1,2)A_n$ (two distinct left cosets of A_n in S_n). Now the proposition follows from [81, Theorem 4.4.2] and definition of induced representation. □

Proposition 3.17. *For self-conjugate $\lambda \vdash n$, we have $\chi^\lambda_\pm \uparrow^{S_n} = \psi^\lambda$.*

Proof. A_n and $(1,2)A_n$ are two distinct left cosets of A_n in S_n. Therefore the proposition follows from the definition of induced representation and [81, Theorem 4.4.2]. □

Remark 3.18. Alternative proofs of Lemma 3.15, Propositions 3.16 and 3.17 based on Clifford theory can be found in [27].

Proposition 3.19. *Let us recall that* χ^{def} *is the character of the defining representation* ρ^{def} *of* S_n*. Then*

$$\left(\chi^{\text{def}} \otimes \chi^{\text{def}}\right) \downarrow_{A_n} = 2\ \chi^{(n)} + 3\ \chi^{(n-1,1)} +\ \chi^{(n-2,2)} +\ \chi^{(n-2,1,1)}.$$

Proof. Recall from equation (2.1.6) of Chapter 2 that,

$$\left(\chi^{\text{def}} \otimes \chi^{\text{def}}\right) = 2\ \psi^{(n)} + 3\ \psi^{(n-1,1)} +\ \psi^{(n-2,2)} +\ \psi^{(n-2,1,1)}. \tag{3.4.1}$$

For any non-self-conjugate $\lambda \vdash n$, Theorem 2.8 (Frobenius Reciprocity) implies

$$\begin{aligned}
&\left\langle\ \chi^{\lambda},\ \left(\chi^{\text{def}} \otimes \chi^{\text{def}}\right) \downarrow_{A_n} \right\rangle \\
&= \left\langle\ \chi^{\lambda} \uparrow^{S_n},\ \chi^{\text{def}} \otimes \chi^{\text{def}} \right\rangle \\
&= \left\langle\ \chi^{\lambda} \uparrow^{S_n},\ 2\ \psi^{(n)} + 3\ \psi^{(n-1,1)} +\ \psi^{(n-2,2)} +\ \psi^{(n-2,1,1)} \right\rangle, \quad \text{by (3.4.1)} \\
&= \left\langle\ \psi^{\lambda} + \psi^{\lambda'},\ 2\ \psi^{(n)} + 3\ \psi^{(n-1,1)} +\ \psi^{(n-2,2)} +\ \psi^{(n-2,1,1)} \right\rangle.
\end{aligned} \tag{3.4.2}$$

The equality in (3.4.2) follows from Proposition 3.16. Now using orthonormality of irreducible characters of S_n, expression (3.4.2) becomes

$$\begin{cases} 2, \text{ if } \lambda = (n) \text{ or } (1^n), \\ 3, \text{ if } \lambda = (n-1,1) \text{ or } (2, 1^{n-1}), \\ 1, \text{ if } \lambda = (n-2,2) \text{ or } (2^2, 1^{n-4}), \\ 1, \text{ if } \lambda = (n-2,1,1) \text{ or } (3, 1^{n-3}). \end{cases}$$

Again, for any self-conjugate $\lambda \vdash n$, by Theorem 2.8, we have

$$\begin{aligned}
&\left\langle\ \chi^{\lambda}_{\pm},\ \left(\chi^{\text{def}} \otimes \chi^{\text{def}}\right) \downarrow_{A_n} \right\rangle \\
&= \left\langle\ \chi^{\lambda}_{\pm} \uparrow^{S_n},\ \chi^{\text{def}} \otimes \chi^{\text{def}} \right\rangle \\
&= \left\langle\ \chi^{\lambda}_{\pm} \uparrow^{S_n},\ 2\ \psi^{(n)} + 3\ \psi^{(n-1,1)} +\ \psi^{(n-2,2)} +\ \psi^{(n-2,1,1)} \right\rangle, \quad \text{by (3.4.1)} \\
&= \left\langle\ \psi^{\lambda},\ 2\ \psi^{(n)} + 3\ \psi^{(n-1,1)} +\ \psi^{(n-2,2)} +\ \psi^{(n-2,1,1)} \right\rangle, \quad \text{by Proposition 3.17} \\
&= 0, \quad \text{using orthonormality of irreducible characters of } S_n.
\end{aligned}$$

Thus the proposition follows. ☐

Lemma 3.20. *For any $i \in [n]$, the number of even permutations π in A_n which fix i (i.e. $\pi(i) = i$) is $\frac{1}{2}(n-1)!$.*

Proof. We know that the number of permutations $\pi \in S_n$ which fix $i \in \{1, \ldots, n\}$ is $(n-1)!$. Now for each i, let us consider the following sets,

$$\mathcal{S}_i = \{\pi \in S_n \mid \pi(i) = i\} \text{ and } \mathcal{A}_i = \{\pi \in A_n \mid \pi(i) = i\}.$$

Then we can define a bijection $\psi : \mathcal{A}_i \to \mathcal{S}_i \setminus \mathcal{A}_i$ by $\pi \mapsto \pi(j,k)$, for fixed $j, k \in \{1, \ldots, n\} \setminus \{i\}$ such that $j \neq k$. Therefore, the cardinality of $\mathcal{A}_i$ is same as the cardinality of $\mathcal{S}_i \setminus \mathcal{A}_i$. But we also know that the cardinality of $\mathcal{S}_i$ is $(n-1)!$. Thus $|\mathcal{S}_i| = |\mathcal{A}_i| + |\mathcal{S}_i \setminus \mathcal{A}_i|$ implies $|\mathcal{A}_i| = \frac{1}{2}(n-1)!$. □

Let us define a random variable X on A_n by $X(\pi) :=$ the number of fixed points of π, $\pi \in A_n$. Therefore, we have $X(\pi) = \chi^{\mathrm{def}} \downarrow_{A_n} (\pi)$. We now find the expectation $E_{U_{A_n}}(X)$ of X under U_{A_n}.

$$E_{U_{A_n}}(X) = \sum_{\pi \in A_n} X(\pi) \frac{2}{n!} = \frac{2}{n!} \sum_{\pi \in A_n} \chi^{\mathrm{def}} \downarrow_{A_n} (\pi) = \frac{2}{n!} \sum_{\pi \in A_n} \mathrm{Tr}\left(\rho^{\mathrm{def}} \downarrow_{A_n} (\pi)\right). \tag{3.4.3}$$

We know that for each $i \in \{1, \ldots, n\}$, the $(i,i)^{\text{th}}$ entry of the matrix $\left(\sum_{\pi \in A_n} \rho^{\mathrm{def}} \downarrow_{A_n} (\pi)\right)$ is the number of permutations in A_n which fixes i. Therefore, using Lemma 3.20 and the linearity of the trace, expression (3.4.3) becomes

$$\frac{2}{n!} \sum_{\pi \in A_n} \mathrm{Tr}\left(\rho^{\mathrm{def}} \downarrow_{A_n} (\pi)\right) = \frac{2}{n!} \sum_{i=1}^{n} \frac{(n-1)!}{2} = 1.$$

Now we find the expectation $E_k(X)$ and variance $\mathrm{Var}_k(X)$ of the same random variable X under the distribution P_A^{*k} on A_n. We know that the defining representation on S_n decomposes into the trivial representation $S^{(n)}$ and the $(n-1)$-dimensional representation $S^{(n-1,1)}$ (see equation (2.1.5) from Chapter 2). As the irreducible representations $S^{(n)}$ and $S^{(n-1,1)}$ of S_n are irreducible in A_n [81, Theorem 4.6.7], we have the following:

$$\begin{aligned} E_k(X) = \sum_{\pi \in A_n} X(\pi) P_A^{*k}(\pi) &= \sum_{\pi \in A_n} \chi^{\mathrm{def}} \downarrow_{A_n} (\pi) P_A^{*k}(\pi) \\ &= \sum_{\pi \in A_n} P_A^{*k}(\pi) \, \mathrm{Tr}\left(\rho^{\mathrm{def}} \downarrow_{A_n} (\pi)\right). \end{aligned} \tag{3.4.4}$$

Now using $\left(\sum_{\pi \in A_n} P_A(\pi) \, \rho^{\mathrm{def}} \downarrow_{A_n} (\pi)\right)^k = \sum_{\pi \in A_n} P_A^{*k}(\pi) \, \rho^{\mathrm{def}} \downarrow_{A_n} (\pi)$ and linearity of the trace, expression (3.4.4) is equal to $\operatorname{Tr}\left(\widehat{P_A}\left(\rho^{\mathrm{def}}\downarrow_{A_n}\right)\right)^k$. Therefore from Table 3.1, we have

$$E_k(X) = 1 + (n-3)\left(\frac{2n-5}{2n-3}\right)^k + \left(\frac{n-3}{2n-3}\right)^k \left(1+(-1)^k\right), \tag{3.4.5}$$

Partition of n	Eigenvalues of $\widehat{P}(R)$ corresponding to the irreducible A_n-module indexed by the partition of column 1
(n) or (1^n)	1 with algebraic multiplicity 1
$(n-1,1)$ or $(2,1^{n-1})$	$\frac{2n-5}{2n-3}$ with algebraic multiplicity $n-3$ $-\frac{n-3}{2n-3}$ with algebraic multiplicity 1 $\frac{n-3}{2n-3}$ with algebraic multiplicity 1
$(n-2,2)$ or $(2^2,1^{n-4})$	$\frac{2n-7}{2n-3}$ with algebraic multiplicity $\frac{(n-2)(n-5)}{2}$ $\frac{-1}{2n-3}$ with algebraic multiplicity 1 $\frac{n-3}{2n-3}$ with algebraic multiplicity $n-3$ $-\frac{n-3}{2n-3}$ with algebraic multiplicity $n-3$
$(n-2,1^2)$ or $(3,1^{n-3})$	$\frac{2n-7}{2n-3}$ with algebraic multiplicity $\frac{(n-3)(n-4)}{2}$ $\frac{3}{2n-3}$ with algebraic multiplicity 1 $\frac{n-5}{2n-3}$ with algebraic multiplicity $n-3$ $-\frac{n-5}{2n-3}$ with algebraic multiplicity $n-3$

Table 3.1: Eigenvalues of $\widehat{P_A}(R)$ corresponding to some irreducible representations of A_n.

$$\begin{aligned} E_k(X^2) = \sum_{\pi\in A_n} (X(\pi))^2 P_A^{*k}(\pi) &= \sum_{\pi\in A_n}\left(\chi^{\mathrm{def}}\downarrow_{A_n}(\pi)\right)^2 P_A^{*k}(\pi) \\ &= \sum_{\pi\in A_n}\left(\chi^{\mathrm{def}}\otimes\chi^{\mathrm{def}}\right)\downarrow_{A_n}(\pi) P_A^{*k}(\pi). \end{aligned} \tag{3.4.6}$$

Using Proposition 3.19, expression (3.4.6) can be written as,

$$\sum_{\pi\in A_n} P^{*k}(\pi)\left(2\ \chi^{(n)}(\pi) + 3\ \chi^{(n-1,1)}(\pi) +\ \chi^{(n-2,2)}(\pi) +\ \chi^{(n-2,1,1)}(\pi)\right). \tag{3.4.7}$$

Now if we write $\xi = \sum_{\pi \in A_n} P_A(\pi)\, \pi$, then using the definition of the character and linearity of the trace, the expression (3.4.7) equal to,

$$\begin{aligned}
&2\ \chi^{(n)}(\xi^k) + 3\ \chi^{(n-1,1)}(\xi^k) +\ \chi^{(n-2,2)}(\xi^k) +\ \chi^{(n-2,1,1)}(\xi^k) \\
=&2 + 3\left((n-3)\left(\frac{2n-5}{2n-3}\right)^k + \left(\frac{n-3}{2n-3}\right)^k \left(1+(-1)^k\right)\right) \\
&+\left(\frac{(n-2)(n-5)}{2}\left(\frac{2n-7}{2n-3}\right)^k + \left(\frac{-1}{2n-3}\right)^k + (n-3)\left(\frac{n-3}{2n-3}\right)^k\left(1+(-1)^k\right)\right) \\
&+\left(\frac{(n-3)(n-4)}{2}\left(\frac{2n-7}{2n-3}\right)^k + \left(\frac{3}{2n-3}\right)^k + (n-3)\left(\frac{n-5}{2n-3}\right)^k\left(1+(-1)^k\right)\right)
\end{aligned}$$

from Table 3.1.

Therefore from the definition of variance $\mathrm{Var}_k(X) = E_k(X^2) - (E_k(X))^2$ we have

$$\begin{aligned}
\mathrm{Var}_k(X) = 1 &+ (n-3)\left(\frac{2n-5}{2n-3}\right)^k + \left(\frac{n-3}{2n-3}\right)^k\left(1+(-1)^k\right) \\
&-\left((n-3)\left(\frac{2n-5}{2n-3}\right)^k + \left(\frac{n-3}{2n-3}\right)^k\left(1+(-1)^k\right)\right)^2 \\
&+\frac{(n-2)(n-5)}{2}\left(\frac{2n-7}{2n-3}\right)^k + \left(\frac{-1}{2n-3}\right)^k + (n-3)\left(\frac{n-3}{2n-3}\right)^k\left(1+(-1)^k\right) \\
&+\frac{(n-3)(n-4)}{2}\left(\frac{2n-7}{2n-3}\right)^k + \left(\frac{3}{2n-3}\right)^k + (n-3)\left(\frac{n-5}{2n-3}\right)^k\left(1+(-1)^k\right).
\end{aligned} \tag{3.4.8}$$

Lemma 3.21. *For* $\mathrm{Var}_k(X)$ *and* $E_k(X)$ *given in* (3.4.5) *and* (3.4.8) *respectively, the following are true*

1. *For* $c < 0,\ k = (n - \frac{3}{2})(\log n + c)$ *and large* n, *we have*

$$E_k(X) = 1 + e^{-c}\left(1 + o(1)\right) \quad \textit{and} \quad \mathrm{Var}_k(X) = 1 + e^{-c}\left(1 + o(1)\right).$$

2. *For any* $\epsilon \in (0,1),\ k_n = \lfloor (1-\epsilon)(n-\frac{3}{2})\log n \rfloor$ *and large* n, *we have*

$$E_{k_n}(X) = 1 + n^{\epsilon} + o(1) \quad \textit{and} \quad \mathrm{Var}_{k_n}(X) = 1 + n^{\epsilon} + o(1) + n^{\epsilon} o(1).$$

Proof. Throughout this proof we denote $d_k = \left(\frac{1+(-1)^k}{2^k}\right)$ for convenience. Note that $0 \le d_k < 1$ for $k \ge 1$. Let us recall that '$\approx$' denotes 'asymptotic to' i.e. $a_n \approx b_n$ means

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

First, we have

$$\begin{aligned} E_k(X) &= 1 + (n-3)\left(1 - \frac{1}{n-1.5}\right)^k + \left(1 - \frac{1.5}{n-1.5}\right)^k \left(\frac{1+(-1)^k}{2^k}\right) \\ &\approx 1 + (n-3)e^{-\frac{k}{n-1.5}} + d_k e^{-\frac{1.5k}{n-1.5}}. \end{aligned} \tag{3.4.9}$$

Now from (3.4.8), we have

$$\begin{aligned} \mathrm{Var}_k(X) &\approx 1 + (n-3)e^{-\frac{k}{n-1.5}} + d_k e^{-\frac{1.5k}{n-1.5}} - \left((n-3)e^{-\frac{k}{n-1.5}} + d_k e^{-\frac{1.5k}{n-1.5}}\right)^2 \\ &\quad + \frac{(-1)^k + 3^k}{(2n-3)^k} + (n^2 - 7n + 11)e^{-\frac{2k}{n-1.5}} + (n-3)d_k\left(e^{-\frac{1.5k}{n-1.5}} + e^{-\frac{3.5k}{n-1.5}}\right) \\ &= 1 + (n-3)e^{-\frac{k}{n-1.5}} + d_k e^{-\frac{1.5k}{n-1.5}} - (n-2)e^{-\frac{2k}{n-1.5}} - 2d_k(n-3)e^{-\frac{2.5k}{n-1.5}} \\ &\quad - d_k^2 e^{-\frac{3k}{n-1.5}} + \frac{(-1)^k + 3^k}{(2n-3)^k} + (n-3)d_k\left(e^{-\frac{1.5k}{n-1.5}} + e^{-\frac{3.5k}{n-1.5}}\right). \end{aligned} \tag{3.4.10}$$

Now for large n, if we take $k = (n - \frac{3}{2})(\log n + c)$, $c < 0$, then we have $k \geq 1$ and hence d_k is bounded above. Therefore from (3.4.9), we have $E_k(X) = 1 + e^{-c}(1 + o(1))$ and from (3.4.10), we have $\mathrm{Var}_k(X) = 1 + e^{-c}(1 + o(1))$ by straightforward calculations. This proves the first part.

Now for any $\epsilon \in (0,1)$, $k_n = \lfloor (1-\epsilon)(n - \frac{3}{2}) \log n \rfloor$ and we have $k_n \geq 1$ for large n. Hence d_{k_n} is bounded above. Therefore from (3.4.9), we have $E_{k_n}(X) = 1 + n^\epsilon + o(1)$ and from (3.4.10), we have $\mathrm{Var}_{k_n}(X) = 1 + n^\epsilon + o(1) + n^\epsilon o(1)$ by straightforward calculations. This proves the second part. □

Theorem 3.22. *For the random walk on A_n driven by A_A, we have the following:*

1. *For large n, $||P_A^{*k} - U_{A_n}||_{\mathrm{TV}} \geq 1 - \frac{6}{1+e^{-c}(1+o(1))}$, when $k = (n - \frac{3}{2})(\log n + c)$ and $c \ll 0$.*

2. *$\lim_{n\to\infty} ||P_A^{*k_n} - U_{A_n}||_{\mathrm{TV}} = 1$, for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon)(n - \frac{3}{2}) \log n \rfloor$.*

Proof. Using Lemma 2.19, $\mu = P_A^{*k}$ and $\nu = U_{A_n}$ we have,

$$||P_A^{*k} - U_{A_n}||_{\mathrm{TV}} \geq 1 - \frac{4\,\mathrm{Var}_k(X)}{(E_k(X))^2} - \frac{2}{E_k(X)}. \tag{3.4.11}$$

Now if n is large, $c \ll 0$ and $k = (n - \frac{3}{2})(\log n + c)$, then by (3.4.11) and by the first part of Lemma 3.21, we have the first part of this theorem. Again for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon)(n - \frac{3}{2}) \log n \rfloor$ from (3.4.11) and by the second part of Lemma 3.21, we have the following:

$$1 \geq ||P^{*k_n} - U_{A_n}||_{\mathrm{TV}} \geq 1 - \frac{4(1+n^{\epsilon}+o(1)+n^{\epsilon}o(1))}{(1+n^{\epsilon}+o(1))^2} - \frac{2}{1+n^{\epsilon}+o(1)}, \tag{3.4.12}$$

for large n. Therefore, the second part of this theorem follows from (3.4.12) and the following:

$$\lim_{n\to\infty} \frac{4(1+n^{\epsilon}+o(1)+n^{\epsilon}o(1))}{(1+n^{\epsilon}+o(1))^2} = 0,$$
$$\lim_{n\to\infty} \frac{2}{1+n^{\epsilon}+o(1)} = 0.$$

□

Corollary 3.23 (Total variation cutoff for transpose top-2 with random shuffle)**.** *The transpose top-2 with random shuffle on A_n exhibits the total variation cutoff phenomenon and the cutoff is at $(n-\frac{3}{2})\log n$.*

Proof. The proof follows from the second part of the Theorems 3.14 and 3.22. □

For example, if $n = 10$, the plot for $||P_A^{*k} - U_{A_{10}}||_{\mathrm{TV}}$ vs. k is given in Figure 3.1. In this case the cutoff is at $8.5 \times \log 10 = 19.572$ and Figure 3.1 confirms this.

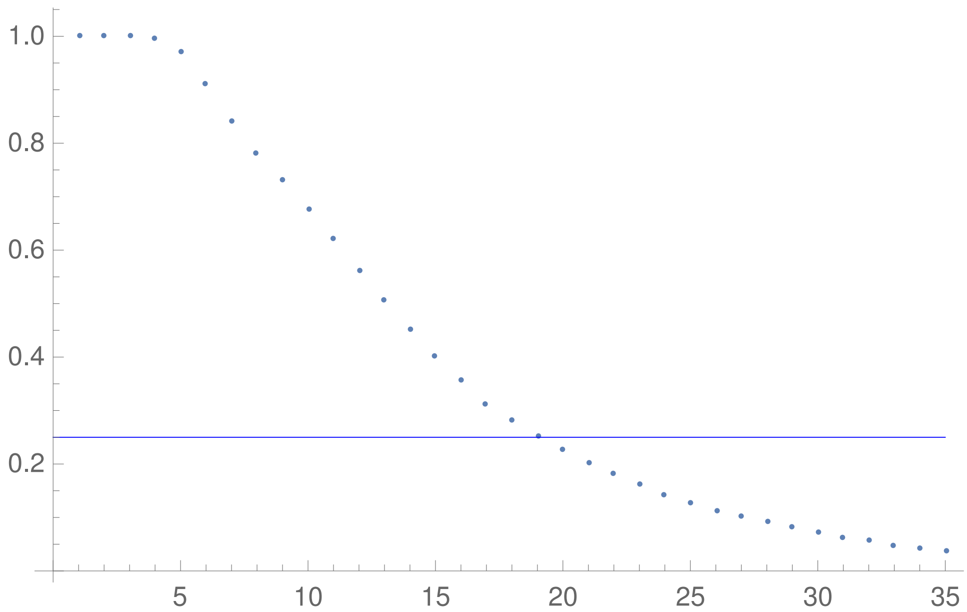


Figure 3.1: Plot for $||P_A^{*k} - U_{A_{10}}||_{\mathrm{TV}}$ vs. k for $1 \leq k \leq 35$.

Chapter 4

The flip-transpose top with random shuffle

In this chapter our main aim is to study the properties of a random walk on Coxeter groups of type B [22]. We call this the *flip-transpose top with random shuffle*. A random walk on Coxeter group of type D has also been studied analogous to the former walk. We introduce the flip-transpose top with random shuffle on the hyperoctahedral group B_n in Section 4.1. In Section 4.2, we find the spectrum of the transition matrix and give an upper bound for the total variation distance, of the distribution after k transitions from the stationary distribution. In Section 4.3, we obtain a lower bound for the total variation distance of the distribution after k transitions from the stationary distribution. We also prove the cutoff phenomenon for this shuffle with cutoff time $n \log n$ in this section. In Section 4.4, a brief description for the representation theory of the demihyperoctahedral group D_n is given. In Section 4.5, we consider a similar random walk on D_n and prove cutoff at $\left(n - \frac{1}{2}\right) \log n$.

4.1 Introduction

We begin with defining the hyperoctahedral group B_n and the demihyperoctahedral group D_n in this section. A *signed permutation* is a bijection π from $\{-n, \ldots, -1, 1, \ldots, n\}$ to itself satisfying $\pi(-i) = -\pi(i)$ for all $1 \leq i \leq n$. A signed permutation is completely determined by its image on the set $[n]$. Given a signed permutation π, we write it in window notation by $[\pi_1, \ldots, \pi_n]$, where π_i is the image of i under π. The set of all signed permutations forms a group under composition of mapping. This group is known as the *hyperoctahedral group* and is denoted by B_n [22]. The subset of B_n consisting of those signed permutations having an even number of negative entries in their window notation forms a subgroup of B_n, called the *demihyperoctahedral group* and is denoted by D_n. Suppose there are n cards labelled from 1 to n and each card has two orientations namely 'face up' and 'face down'. Given an arrangements of these n cards in a row we associate a signed permutation $[\pi_1, \pi_2, \ldots, \pi_n]$ to it in the following way: π_i is the label of the ith card (counting starts from left) with sign

$$\begin{cases} \text{positive}, & \text{if the orientation of the card is 'face up' and} \\ \text{negative}, & \text{if the orientation of the card is 'face down'.} \end{cases}$$

Thus every arrangement of the n cards in a row represents a signed permutation in its window notation. We consider the following shuffle on the set of all arrangements of these n cards in a row: Given an arrangement, either interchange the last card with a random card, or interchange the last card with a random card and flip both of them, with equal probability. We note that the random card could be the last card itself. We call this shuffle the *flip-transpose top with random shuffle.* Formally, this shuffle is the random walk on B_n driven by the probability measure P_B on B_n given by

$$P_B(\pi) = \begin{cases} \frac{1}{2n}, & \text{if } \pi = \text{id}, \text{ the identity element of } B_n, \\ \frac{1}{2n}, & \text{if } \pi = (i, n) := [1, \ldots, i-1, n, i+1, \ldots, i] \text{ for } 1 \leq i \leq n-1, \\ \frac{1}{2n}, & \text{if } \pi = (-i, n) := [1, \ldots, i-1, -n, i+1, \ldots, -i] \text{ for } 1 \leq i \leq n, \\ 0, & \text{otherwise.} \end{cases} \tag{4.1.1}$$

Proposition 4.1. *The flip-transpose top with random shuffle on B_n is irreducible and aperiodic.*

Proof. We know that the set $\{(-1, 1),\ (1, 2),\ (2, 3),\ \ldots,\ (n-1, n)\}$ generates B_n. Let i be any integer from $[n-1]$. Then

$$(i, i+1) = (i+1, n)(i, n)(i+1, n) \text{ and } (-1, 1) = (1, n)(-n, n)(1, n).$$

Therefore the support of the measure P_B generates B_n and hence the random walk is irreducible (Proposition 2.22). Given any $\pi \in B_n$, the set of all times when it is possible for the chain to return to the starting state π contains the integer 1 ($\because$ the identity element of B_n is in the support of P_B). Therefore the period of the state π is 1 and hence from irreducibility all the states of this chain have period 1. Thus this chain is aperiodic. □

Proposition 4.1 says that the flip-transpose top with random shuffles on B_n has unique stationary distribution U_{B_n} (Proposition 2.23) and the distribution after the kth transition converges to its stationary distribution as $k \to \infty$ [66, Theorem 4.9]. Throughout this chapter, id denotes the identity signed permutation.

4.2 Sufficient number of transitions to reach near stationary distribution

In this section we find the eigenvalues of the transition matrix $\widehat{P_B}(R)$, the Fourier transform of P_B at the right regular representation R of B_n. To find the eigenvalues of $\widehat{P_B}(R)$ we use the representation theory of the hyperoctahedral group B_n. Here we briefly discuss the representation theory of B_n and some necessary concepts useful for this chapter. For more details one can see [50, 73, 83].

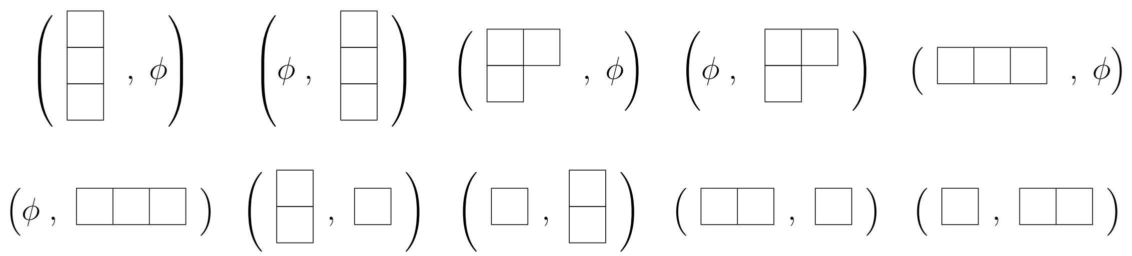


Figure 4.1: All elements of $\mathcal{D}_3$.

Definition 4.2. Let n be a positive integer. A *(Young) double-diagram with* n *boxes* μ is a pair of Young diagrams with a total number of n boxes. We define $||\mu|| = n$. The set of all double-diagram with n boxes is denoted by $\mathcal{D}_{\text{n}}$. For example, all double-diagrams with 3 boxes are listed in Figure 4.1. A *standard (Young) double-tableau of shape* μ is obtained by taking the double-diagram μ and filling its $||\mu||$ boxes (bijectively) with the numbers $1, 2, \ldots, ||\mu||$ such that the numbers in the boxes strictly increase along each row and each column of all Young diagrams occurring in μ. Let $\text{tab}_{\mathcal{D}}(n, \mu)$, where $\mu \in \mathcal{D}_{\text{n}}$, denote the set of all standard double-tableaux of shape μ and let $\text{tab}_{\mathcal{D}}(n) = \bigcup\limits_{\mu \in \mathcal{D}_{\text{n}}} \text{tab}_{\mathcal{D}}(n, \mu)$. For example an element of $\text{tab}_{\mathcal{D}}(8)$ is given in Figure 4.2. For $T \in \text{tab}_{\mathcal{D}}(n, \mu)$ and $i \in [n]$, recall that $b_T(i)$ denotes the box of the Young diagram in μ, in which the number i resides. We denote the content of the box $b_T(i)$ by $c(b_T(i))$. For the standard double-tableau given in Figure 4.2, we have $c(b_T(1)) = 0$, $c(b_T(2)) = 1$, $c(b_T(3)) = 0$, $c(b_T(4)) = 1$, $c(b_T(5)) = -1$, $c(b_T(6)) = -1$, $c(b_T(7)) = 0$ and $c(b_T(8)) = 2$.

Definition 4.3. The *Young-Jucys-Murphy* elements $X_1^B, X_2^B, \ldots, X_n^B$ of $\mathbb{C}[B_n]$ are defined by $X_1^B = 0$ and $X_i^B = \sum\limits_{k=1}^{i-1}(k, i) + \sum\limits_{k=1}^{i-1}(-k, i)$, for all $2 \leq i \leq n$.

$$\left(\begin{array}{|c|c|c|} \hline 3 & 4 & 8 \\ \hline 6 & 7 \\ \cline{1-2} \end{array} \, , \, \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 5 \\ \cline{1-1} \end{array} \right)$$

Figure 4.2: An element of $\mathrm{tab}_{\mathcal{D}}(8)$.

Let $\mu \in \widehat{B_n}$ and consider the B_n-module V^μ. Since the branching is simple [73, Section 3], the decomposition into irreducible B_{n-1}-modules is canonical and is given by

$$V^\mu = \bigoplus_\lambda V^\lambda,$$

where the sum is over all $\lambda \in \widehat{B_{n-1}}$, with $\lambda \nearrow \mu$ (i.e. there is an edge from λ to μ in the branching multi-graph). Iterating this decomposition of V^μ into irreducible B_1-submodules, we obtain

$$V^\mu = \bigoplus_T v_T, \tag{4.2.1}$$

where the sum is over all possible chains $T = \mu_1 \nearrow \mu_2 \nearrow \cdots \nearrow \mu_n$ with $\mu_i \in \widehat{B_i}$ and $\mu_n = \mu$. We note that if $0 \neq v_T$, then $\mathbb{C}[B_i]v_T = V^{\mu_i}$. We call (4.2.1) the *Gelfand-Tsetlin* decomposition of V^μ and each v_T in (4.2.1) a *Gelfand-Tsetlin* vector of V^μ. The irreducible representations of B_n are parameterized by elements of $\mathcal{D}_\mathrm{n}$ [73, Lemma 6.2, Theorem 6.4]. The Gelfand-Tsetlin vectors of V^μ form a basis of V^μ. We may index the Gelfand-Tsetlin vectors of V^μ by standard double-tableaux of shape μ for $\mu \in \mathcal{D}_\mathrm{n}$ [73, Theorem 6.5] and write the Gelfand-Tsetlin decomposition as

$$V^\mu = \bigoplus_{T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)} v_T.$$

Let $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \mathcal{D}_\mathrm{n}$ and $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$. Then the action [73, Theorem 6.5] of the Young-Jucys-Murphy elements and the signed permutation $(i,-i)$ on v_T are given by

$$\begin{aligned} X_i^B \, v_T &= 2c(b_T(i)) \, v_T \text{ for all } i \in [n], \\ (-i,i) \, v_T &= \begin{cases} v_T & \text{if } b_T(i) \text{ is in } \mu^{(1)} \\ -v_T & \text{if } b_T(i) \text{ is in } \mu^{(2)} \end{cases} \text{ for all } i \in [n]. \end{aligned} \tag{4.2.2}$$

We now come to our main problem of finding the eigenvalues of the transition matrix $\widehat{P_B}(R)$. The eigenvalues of $\widehat{P_B}(R)$ are the eigenvalues of $\frac{1}{2n}\left(\mathrm{id} + (-n,n) + X_n^B\right)$ acting on $\mathbb{C}[B_n]$ by multiplication on the right. The following theorem gives the eigenvalues of $\widehat{P_B}(R)$.

Theorem 4.4. *For each* $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \mathcal{D}_\mathrm{n}$ *satisfying* $m := |\mu^{(1)}| \in \{0, 1, \ldots, \lfloor \frac{n}{2} \rfloor\}$, *let* $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$. *Then* $\frac{c(b_T(n))+1}{n}$ *and* $\frac{c(b_T(n))}{n}$ *are eigenvalues of* $\widehat{P_B}(R)$ *with multiplicity*

$M(\mu)$ each, where

$$M(\mu)=\begin{cases} \binom{n}{m} f^{\mu^{(1)}} f^{\mu^{(2)}}, & \text{if } 0 \leq m < \frac{n}{2}, \\ \frac{1}{2}\binom{n}{m} f^{\mu^{(1)}} f^{\mu^{(2)}}, & \text{if } m = \frac{n}{2} \text{ (when } n \text{ is even)}. \end{cases} \tag{4.2.3}$$

Recall that $f^{\mu^{(i)}}$ denotes the number of standard Young tableaux of shape $\mu^{(i)}$, $i = 1, 2$.

Proof. For each $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \mathcal{D}_{\mathrm{n}}$, we have another double-diagram $\tilde{\mu}$ with n boxes such that $\tilde{\mu} = \left(\mu^{(2)}, \mu^{(1)}\right)$. We first find the eigenvalues of the matrix $\widehat{P_B}(R)$ in the irreducible B_n-modules V^{μ} and $V^{\tilde{\mu}}$. For each $T = (T_1, T_2) \in \mathrm{tab}_{\mathcal{D}}(n, \mu)$, $\widetilde{T} = (T_2, T_1) \in \mathrm{tab}_{\mathcal{D}}(n, \tilde{\mu})$. If $b_T(n)$ is in $\mu^{(1)}$, then $b_{\widetilde{T}}(n)$ is in $\mu^{(2)}$. Without loss of generality, let us assume that $b_T(n)$ is in $\mu^{(1)}$ and $b_{\widetilde{T}}(n)$ is in $\mu^{(2)}$. Let us recall that v_T (respectively $v_{\widetilde{T}}$) is the Gelfand-Tsetlin vector of V^{μ} (respectively $V^{\tilde{\mu}}$). From (4.2.2) we have

$$(-n, n)\, v_T = v_T \text{ and } X_n^B\, v_T = 2c(b_T(n))\, v_T,$$

which implies the following:

$$\left(\mathrm{id} + (-n, n) + X_n^B\right)\ v_T = (1 + 1 + 2c(b_T(n)))\ v_T = (2c(b_T(n)) + 2)\ v_T. \tag{4.2.4}$$

Since $\{v_T : T \in \mathrm{tab}_{\mathcal{D}}(n, \mu)\}$ form a basis of V^{μ}, the eigenvalues of the action of

$$\left(\mathrm{id} + (-n, n) + X_n^B\right)$$

on V^{μ} can be obtained from (4.2.4). Now using (4.2.2) we have $(-n, n)\, v_{\widetilde{T}} = -v_{\widetilde{T}}$ and $X_n^B\, v_{\widetilde{T}} = 2c(b_{\widetilde{T}}(n))\, v_{\widetilde{T}}$, thus

$$\left(\mathrm{id} + (-n, n) + X_n^B\right)\ v_{\widetilde{T}} = (1 - 1 + 2c(b_T(n)))\ v_{\widetilde{T}} = 2c(b_T(n))\ v_{\widetilde{T}}. \tag{4.2.5}$$

Therefore the eigenvalues of the action of $\left(\mathrm{id} + (-n, n) + X_n^B\right)$ on $V^{\tilde{\mu}}$ are obtained from (4.2.5), as $\{v_{\widetilde{T}} : \widetilde{T} \in \mathrm{tab}_{\mathcal{D}}(n, \tilde{\mu})\}$ form a basis of $V^{\tilde{\mu}}$. Thus considering the action of $\frac{1}{2n}\left(\mathrm{id} + (-n, n) + X_n^B\right)$ on V^{μ} and $V^{\tilde{\mu}}$ simultaneously, the eigenvalues of $\widehat{P_B}(R)$ are given by $\frac{c(b_T(n))+1}{n}$ and $\frac{c(b_T(n))}{n}$ for each $T \in \mathrm{tab}_{\mathcal{D}}(n, \mu)$.

Now we know that the multiplicity of every irreducible representation in the right regular representation is equal to its dimension. Therefore the multiplicity of the eigenvalues are $\dim(V^{\mu}) = \binom{n}{m} f^{\mu^{(1)}} f^{\mu^{(2)}} = \dim(V^{\tilde{\mu}})$ if $0 \leq m < \frac{n}{2}$ and the multiplicity of the eigenvalues are $\frac{1}{2}\binom{n}{m} f^{\mu^{(1)}} f^{\mu^{(2)}}$ if $m = \frac{n}{2}$ (when n is even). The multiplicity of the eigenvalues for the case of $m = \frac{n}{2}$ is half of the dimension of the corresponding B_n-module because of the following: In this case $m = n - m$. Thus both $\mu = (\mu^{(1)}, \mu^{(2)})$ and $\tilde{\mu} = (\mu^{(2)}, \mu^{(1)})$ are in $\mathcal{D}_{\mathrm{n}}$ such that their first component is a partition of m and the second component is a partition of $n-m$. Therefore while computing the eigenvalues of $\widehat{P_B}(R)$ by considering the irreducible B_n-modules V^{μ} and $V^{\tilde{\mu}}$, each space is counted twice. Now the proof of the theorem follows from the fact that all the irreducible representations of B_n are parameterized by $\mathcal{D}_{\mathrm{n}}$. □

We now prove the theorem giving an upper bound of the total variation distance $||P_B^{*k} - U_{B_n}||_{\mathrm{TV}}$ for $k \geq n\log n + cn,\ c > 0$. Given a positive integer n, let λ be a partition of n i.e., $\lambda \vdash n$. Let us recall that $\mathrm{tab}(\lambda)$ denote the set of all standard Young tableaux of shape λ.

Lemma 4.5. *Let m be any positive integer satisfying $1 \leq m \leq \frac{n}{2}$ and $\mu = \left(\mu^{(1)},\ \mu^{(2)}\right) \in \mathcal{D}_{\mathrm{n}}$ be such that $|\mu^{(1)}| = m,\ |\mu^{(2)}| = n-m$. If $\mu_1^{(i)}$ (respectively $\mu_1^{(i)'}$) denotes the largest part of the partition $\mu^{(i)}$ (respectively its conjugate $\mu^{(i)'}$) for $i = 1, 2$, then*

$$\sum_{T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)} \left(\frac{c(b_T(n))+1}{n}\right)^{2k} < \binom{n}{m} f^{\mu^{(2)}} f^{\mu^{(1)}} \sum_{i=1}^{2} \left(\left(\frac{\mu_1^{(i)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(i)'}}{n}\right)^{2k}\right).$$

Proof. The set $\mathrm{tab}_{\mathcal{D}}(n,\mu)$ is a disjoint union of the sets $\mathcal{T}_1 = \{(T_1,T_2) \in \mathrm{tab}_{\mathcal{D}}(n,\mu) : b_T(n) \text{ is in } T_1\}$ and $\mathcal{T}_2 = \{(T_1,T_2) \in \mathrm{tab}_{\mathcal{D}}(n,\mu) : b_T(n) \text{ is in } T_2\}$. Therefore we have

$$\sum_{T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)} \left(\frac{c(b_T(n))+1}{n}\right)^{2k} = \sum_{T \in \mathcal{T}_1} \left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \sum_{T \in \mathcal{T}_2} \left(\frac{c(b_T(n))+1}{n}\right)^{2k}. \quad (4.2.6)$$

Now the right hand side of (4.2.6) is equal to

$$\begin{aligned}
&\binom{n-1}{n-m} f^{\mu^{(2)}} \sum_{T_1 \in \mathrm{tab}(\mu^{(1)})} \left(\frac{c(b_{T_1}(m))+1}{n}\right)^{2k} + \binom{n-1}{m} f^{\mu^{(1)}} \sum_{T_2 \in \mathrm{tab}(\mu^{(2)})} \left(\frac{c(b_{T_2}(n-m))+1}{n}\right)^{2k} \\
&< \binom{n}{m} \left(f^{\mu^{(2)}} \sum_{T_1 \in \mathrm{tab}(\mu^{(1)})} \left(\frac{c(b_{T_1}(m))+1}{n}\right)^{2k} + f^{\mu^{(1)}} \sum_{T_2 \in \mathrm{tab}(\mu^{(2)})} \left(\frac{c(b_{T_2}(n-m))+1}{n}\right)^{2k} \right) \\
&< \binom{n}{m} f^{\mu^{(2)}} f^{\mu^{(1)}} \left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(1)'}}{n}\right)^{2k} + \left(\frac{\mu_1^{(2)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(2)'}}{n}\right)^{2k} \right). \qquad (4.2.7)
\end{aligned}$$

The inequality in (4.2.7) follows from the fact: If λ_1 (respectively λ_1') denotes the largest part of the partition λ (respectively its conjugate λ'), then for all $T \in \mathrm{tab}(\lambda)$ we have

$$\left(\frac{c(b_T(|\lambda|))+1}{n}\right)^{2k} \leq \max\left\{ \left(\frac{\lambda_1}{n}\right)^{2k}, \left(\frac{\lambda_1'-2}{n}\right)^{2k} \right\} < \left(\frac{\lambda_1}{n}\right)^{2k} + \left(\frac{\lambda_1'}{n}\right)^{2k}. \qquad \square$$

Theorem 4.6. *For the random walk on B_n driven by P_B, we have the following:*

1. $||P_B^{*k} - U_{B_n}||_{\text{TV}} < \sqrt{2(e+1)}\, e^{-c} + o(1)$, *for* $k \geq n\log n + cn$ *and* $c > 0$.

2. $\lim_{n\to\infty} ||P_B^{*k_n} - U_{B_n}||_{\text{TV}} = 0$, *for any* $\epsilon \in (0,1)$ *and* $k_n = \lfloor (1+\epsilon)n\log n \rfloor$.

Proof. We know that the trace of the $(2k)$th power of a matrix is the sum of the $(2k)$th powers of its eigenvalues. Therefore Lemma 2.24 implies $4||P_B^{*k} - U_{B_n}||_{\text{TV}}^2$ is bounded above by the sum of $(2k)$th powers of the non-largest eigenvalues (which are strictly less the largest eigenvalue 1) of $\widehat{P_B}(R)$. Thus from Theorem 4.4 we have

$$4||P_B^{*k} - U_{B_n}||_{\text{TV}}^2 \leq \left(\frac{n-1}{n}\right)^{2k} + \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}} f^{\lambda} \sum_{T\in\text{tab}(\lambda)} \left(\left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \left(\frac{c(b_T(n))}{n}\right)^{2k}\right)$$
$$+ \sum_{m=1}^{\lfloor\frac{n}{2}\rfloor} \sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} M(\mu) \sum_{T\in\text{tab}_{\mathcal{D}}(n,\mu)} \left(\left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \left(\frac{c(b_T(n))}{n}\right)^{2k}\right). \quad (4.2.8)$$

$M(\mu)$ is defined in (4.2.3) and can be written as $M(\mu) = \mathbb{I}(n,m)\binom{n}{m} f^{\mu^{(1)}} f^{\mu^{(2)}}$, where

$$\mathbb{I}(n,m) = \begin{cases} 1 & \text{if } 0 \leq m < \frac{n}{2}, \\ \frac{1}{2} & \text{if } m = \frac{n}{2} \text{ (when } n \text{ is even).} \end{cases}$$

The third term in the right hand side of (4.2.8) is less than the following expression

$$\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor} \sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} 2M(\mu) \left(\sum_{T\in\text{tab}_{\mathcal{D}}(n,\mu)} \left(\frac{c(b_T(n))+1}{n}\right)^{2k}\right)$$
$$< \sum_{m=1}^{\lfloor\frac{n}{2}\rfloor} \sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} 2M(\mu)\binom{n}{m} f^{\mu^{(2)}} f^{\mu^{(1)}} \sum_{i=1}^{2} \left(\left(\frac{\mu_1^{(i)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(i)'}}{n}\right)^{2k}\right) \quad (4.2.9)$$
$$= \sum_{m=1}^{\lfloor\frac{n}{2}\rfloor} 2 \sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} 2M(\mu)\binom{n}{m} f^{\mu^{(2)}} f^{\mu^{(1)}} \left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(2)}}{n}\right)^{2k}\right). \quad (4.2.10)$$

Inequality in (4.2.9) follows from Lemma 4.5. The equality in (4.2.10) holds because

$$\sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} 2M(\mu)\binom{n}{m}f^{\mu^{(2)}}f^{\mu^{(1)}}\left(\frac{\mu_1^{(i)'}}{n}\right)^{2k} = \sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)\\ \mu=\left(\mu^{(1)},\,\mu^{(2)}\right)}} 2M(\mu)\binom{n}{m}f^{\mu^{(2)}}f^{\mu^{(1)}}\left(\frac{\mu_1^{(i)}}{n}\right)^{2k}$$

for $i=1,2$. Therefore the expression in (4.2.10) is equal to

$$\begin{aligned}
&4\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor}\mathbb{I}(n,m)\binom{n}{m}^2\sum_{\substack{\mu^{(1)}\vdash m\\ \mu^{(2)}\vdash(n-m)}}\left(f^{\mu^{(1)}}\right)^2\left(f^{\mu^{(2)}}\right)^2\left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k}+\left(\frac{\mu_1^{(2)}}{n}\right)^{2k}\right)\\
=&4\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor}\mathbb{I}(n,m)\binom{n}{m}^2(n-m)!\sum_{\mu^{(1)}\vdash m}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)}}{n}\right)^{2k}\\
&\qquad+4\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor}\mathbb{I}(n,m)\binom{n}{m}^2 m!\sum_{\mu^{(2)}\vdash(n-m)}\left(f^{\mu^{(2)}}\right)^2\left(\frac{\mu_1^{(2)}}{n}\right)^{2k}.
\end{aligned}\tag{4.2.11}$$

The definition of $\mathbb{I}(n,m)$ and

$$\begin{aligned}
&\sum_{m=1}^{\lfloor\frac{n}{2}\rfloor}\mathbb{I}(n,m)\binom{n}{m}^2 m!\sum_{\mu^{(2)}\vdash(n-m)}\left(f^{\mu^{(2)}}\right)^2\left(\frac{\mu_1^{(2)}}{n}\right)^{2k}\\
=&\sum_{u=\lceil\frac{n}{2}\rceil}^{n-1}\mathbb{I}(n,n-u)\binom{n}{n-u}^2(n-u)!\sum_{\mu^{(2)}\vdash u}\left(f^{\mu^{(2)}}\right)^2\left(\frac{\mu_1^{(2)}}{n}\right)^{2k},
\end{aligned}$$

implies that the expression (4.2.11) is equal to

$$4\sum_{m=1}^{n-1}\binom{n}{m}^2(n-m)!\sum_{\mu^{(1)}\vdash m}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)}}{n}\right)^{2k}.\tag{4.2.12}$$

Replacing ℓ (respectively λ) by m (respectively $\mu^{(1)}$) and choosing $a=b=0$ in Lemma 2.10, we have

$$\sum_{\mu^{(1)}\vdash m}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)}}{m}\right)^{2k} < e^{m^2e^{-\frac{2k}{m}}} \implies \sum_{\mu^{(1)}\vdash m}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)}}{m}\right)^{2k} < e,\ \text{ if } k\geq m\log m.$$

Therefore when $k\geq n\log n$ (which implies $k\geq m\log m$), the expression in (4.2.12) and hence the third term in the right hand side of (4.2.8) is less than

$$4e\sum_{m=1}^{n-1}\binom{n}{m}^2(n-m)!\left(\frac{m}{n}\right)^{2k} = 4e\sum_{u=1}^{n-1}\binom{n}{u}^2 u!\left(1-\frac{u}{n}\right)^{2k}$$
$$< 4e\sum_{u=1}^{n-1}\frac{\left(n^2e^{-\frac{2k}{n}}\right)^u}{u!} < 4e\left(e^{n^2e^{-\frac{2k}{n}}}-1\right). \quad (4.2.13)$$

Now we consider the second term in the right hand side of (4.2.8). The second term in the right hand side of (4.2.8) is bounded above by

$$2\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}} f^\lambda \sum_{T\in\text{tab}(\lambda)}\left(\frac{c(b_T(n))+1}{n}\right)^{2k}$$
$$< 2\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}}\left(f^\lambda\right)^2\left(\left(\frac{\lambda_1}{n}\right)^{2k}+\left(\frac{\lambda_1'}{n}\right)^{2k}\right)+2\left(\frac{n-2}{n}\right)^{2k}. \quad (4.2.14)$$

Now using the fact $\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}}\left(f^\lambda\right)^2\left(\frac{\lambda_1'}{n}\right)^{2k} = \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}}\left(f^\lambda\right)^2\left(\frac{\lambda_1}{n}\right)^{2k}$, the expression in the right hand side of (4.2.14) is equal to

$$4\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}}\left(f^\lambda\right)^2\left(\frac{\lambda_1}{n}\right)^{2k}+2\left(1-\frac{2}{n}\right)^{2k} < 4\left(\sum_{\lambda\vdash n}\left(f^\lambda\right)^2\left(\frac{\lambda_1}{n}\right)^{2k}-1\right)+2e^{-\frac{4k}{n}}. \quad (4.2.15)$$

The right hand side of the expression (4.2.15) and hence the second term in the right hand side of (4.2.8) is less than $4\left(e^{n^2e^{-\frac{2k}{n}}}-1\right)+2e^{-\frac{4k}{n}}$ by Lemma 2.10. Thus the inequality (4.2.8) becomes

$$4||P^{*k}-U_{B_n}||_{\text{TV}}^2 \leq e^{-\frac{2k}{n}}+(4+4e)\left(e^{n^2e^{-\frac{2k}{n}}}-1\right)+2e^{-\frac{4k}{n}}, \quad \text{for } k\geq n\log n. \quad (4.2.16)$$

Now if $k\geq n\log n+cn$ and $c>0$, then the right hand side of (4.2.16) is less than or equal to

$$(4e+4)\left(e^{e^{-2c}}-1\right)+\frac{e^{-2c}}{n^2}+\frac{2e^{-4c}}{n^4} < 2(4e+4)e^{-2c}+\frac{e^{-2c}}{n^2}+\frac{2e^{-4c}}{n^4}$$
$$= 2(4e+4)e^{-2c}+o(1).$$

This proves the first part of the theorem. Now for $\epsilon\in(0,1)$, $k_n=\lfloor(1+\epsilon)n\log n\rfloor$ implies, $k_n\geq(1+\epsilon)n\log n$. Thus the right hand side of (4.2.16) is bounded above by

$(4e+4)\left(e^{\frac{1}{n^{2\epsilon}}}-1\right)+n^{-2(1+\epsilon)}+2n^{-4(1+\epsilon)}$. Therefore the proof of the second part follows from

$$\lim_{n\to\infty}\left((4e+4)\left(e^{\frac{1}{n^{2\epsilon}}}-1\right)+\frac{1}{n^{2(1+\epsilon)}}+\frac{2}{n^{4(1+\epsilon)}}\right)=0. \qquad \square$$

4.3 Number of transitions necessary for mixing

In this section, we find a lower bound of the total variation distance $||P_B^{*k}-U_{B_n}||_{\text{TV}}$ for $k=\log n+cn,\ c\ll 0$. We also find the sharp mixing time for the flip-transpose top with random shuffle on B_n driven by P_B by proving total variation cutoff phenomenon. Throughout this section I_n denotes the identity matrix of order $n\times n$. To start with, we define a random variable X on B_n as follows:

$$X(\pi)=\text{ number of fixed points of } \pi.$$

Remark 4.7. For each $i\in[n]$, the signed permutation which fixes i will automatically fix $(-i)$. Thus X takes values on the set of non-negative even integers.

Throughout this chapter, let $E_k(X)$ be the expectation and $\text{Var}_k(X)$ be the variance of X with respect to the probability measure P_B^{*k} on B_n. Now X can also be described as follows. Let $V=\mathbb{C}[\{\textbf{-n},\dots,\textbf{-1},\mathbf{1},\dots,\mathbf{n}\}]$ be the vector space of all formal $\mathbb{C}$-linear combinations of elements of the set $\{\textbf{-n},\dots,\textbf{-1},\mathbf{1},\dots,\mathbf{n}\}$. Also let $V^+=\mathbb{C}[\{\mathbf{a_1},\mathbf{a_2},\dots,\mathbf{a_n}\}]$ and $V^-=\mathbb{C}[\{\mathbf{b_1},\mathbf{b_2},\dots,\mathbf{b_n}\}]$ be two vector subspaces of V, where $\mathbf{a_i}=\mathbf{i}+(\textbf{-i})$ and $\mathbf{b_i}=\mathbf{i}-(\textbf{-i})$ for all i $\in[n]$. We note that $\mathbf{a_{-i}}=\mathbf{a_i}$ and $\mathbf{b_{-i}}=-\mathbf{b_i}$ for all i $\in[n]$. Let us now define

$$\rho^+:B_n\to GL(V^+) \text{ by } \rho^+(\pi)(\mathbf{a_i})=\mathbf{a_{\pi(i)}} \text{ on the basis elements of } V^+, \text{ for } \pi\in B_n,$$
$$\rho^-:B_n\to GL(V^-) \text{ by } \rho^-(\pi)(\mathbf{b_i})=\mathbf{b_{\pi(i)}} \text{ on the basis elements of } V^-, \text{ for } \pi\in B_n.$$

It can be easily seen that $\rho^+(\pi)$ and $\rho^-(\pi)$ are well defined for $\pi\in B_n$. We note that ρ^+ and ρ^- are two representations of B_n. Using ρ^+ and ρ^- we can interpret X as follows:

$$X(\pi)=\text{Tr}\left(\rho^+(\pi)+\rho^-(\pi)\right),\ \text{for } \pi\in B_n. \tag{4.3.1}$$

Let $\mathbb{C}[\mathbf{n}]:=\{c_1\mathbf{1}+c_2\mathbf{2}+\cdots+c_n\mathbf{n}\mid c_i\in\mathbb{C}\text{ for all } i\}$. Also let $s_{(u,n)}$ $(1\leq u<n)$ denote the transposition in S_n interchanging u and n. Recall from Chapter 2 (Definition 2.12) that the defining representation $\rho^{\text{def}}:S_n\to GL(\mathbb{C}[\mathbf{n}])$ of S_n is defined by

$$\rho^{\text{def}}(\pi)\left(c_1\mathbf{1}+c_2\mathbf{2}+\cdots+c_n\mathbf{n}\right)=c_1\pi(\mathbf{1})+c_2\pi(\mathbf{2})+\cdots+c_n\pi(\mathbf{n}) \text{ for } \pi\in S_n.$$

From now on the matrices $\rho^+(\pi)$ (respectively $\rho^-(\pi)$) are defined with respect to the ordered bases $(\mathbf{a_1}, \ldots, \mathbf{a_n})$ (respectively $(\mathbf{b_1}, \ldots, \mathbf{b_n})$) for $\pi \in B_n$ and $\rho^{\text{def}}(\sigma)$ is defined with respect to the ordered basis $(\mathbf{1}, \ldots, \mathbf{n})$ for $\sigma \in S_n$.

Lemma 4.8. *The matrices* $\rho^+((u,n))$, $\rho^+((-u,n))$ *and* $\rho^{\text{def}}(s_{(u,n)})$ *are the same for all* $u \in \{1, \ldots, n-1\}$ *and* $\rho^+((-n,n)) = I_n$.

Proof. The lemma follows by looking at the action of each of these matrices on the basis vector. Now for each $i \in \{1, \ldots, n\}$ we have the following:

$$\rho^+((-u,n))(\mathbf{a_i}) = \begin{cases} \mathbf{a_i} & \text{if } i \neq u, n \\ \mathbf{a_n} & \text{if } i = u \\ \mathbf{a_u} & \text{if } i = n \end{cases} = \rho^+((u,n))(\mathbf{a_i}).$$

$$\rho^{\text{def}}(s_{(u,n)})(\mathbf{i}) = \begin{cases} \mathbf{i} & \text{if } i \neq u, n \\ \mathbf{n} & \text{if } i = u \\ \mathbf{u} & \text{if } i = n. \end{cases}$$

Also $\rho^+((-n,n)) = I_n$ follows trivially by looking its action on the basis elements. □

Lemma 4.9. *For each* $i \in \{1, \ldots, n\}$, *let* M_i *denote the* $n \times n$ *matrix with* (i,i)*th entry* 1 *and* 0 *elsewhere. Then for all* $u \in \{1, \ldots, n-1\}$ *we have* $\rho^-((u,n)) + \rho^-((-u,n)) = 2(I_n - M_u - M_n)$ *and* $\rho^-((-n,n)) = 2(I_n - M_n) - I_n$.

Proof. This lemma follows by looking at the action of the matrices on the basis elements. For $i \in \{1, \ldots, n\}$ we have the following:

$$\Big(\rho^-((u,n)) + \rho^-((-u,n))\Big)(\mathbf{b_i}) = \begin{cases} 2\,\mathbf{b_i} & \text{if } i \neq u, n \\ 0 & \text{if } i = u \\ 0 & \text{if } i = n \end{cases}$$

and

$$\Big(I_n + \rho^-((-n,n))\Big)(\mathbf{b_i}) = \begin{cases} 2\,\mathbf{b_i} & \text{if } i \neq n \\ 0 & \text{if } i = n. \end{cases}$$

□

Lemma 4.10. *Let* O_n *be the zero matrix of size* $n \times n$ *and recall from Lemma 2.15 that*

$$\beta_i = \sum_{u=1}^{n-1} \rho^{def}(s_{(u,n)}) - \rho^{def}(s_{(i,n)}), \quad \textit{for all } 1 \leq i < n.$$

If $\text{Blockdiag}(\beta_1, \beta_2, \ldots, \beta_{n-1}, O_n)$ *denotes the block diagonal matrix with* i*th block* β_i *for all* $i \in \{1, 2, \ldots, n-1\}$ *and* n*th block* O_n*, then*

$$\begin{aligned}&\sum_{u=1}^{n-1}\Big(\rho^-((u,n)) \otimes \rho^+((u,n)) + \rho^-((-u,n)) \otimes \rho^+((-u,n))\Big)\\&= 2\ \text{Blockdiag}(\beta_1, \beta_2, \ldots, \beta_{n-1}, O_n).\end{aligned} \tag{4.3.2}$$

Proof. Using Lemma 4.8 and Lemma 4.9 the matrix in the statement (given in (4.3.2)) can be written as

$$\begin{aligned}&\sum_{u=1}^{n-1}\Big(\rho^-((u,n)) + \rho^-((-u,n))\Big) \otimes \rho^{\text{def}}(s_{(u,n)})\\&= 2\sum_{u=1}^{n-1}(I_n - M_u - M_n) \otimes \rho^{\text{def}}(s_{(u,n)})\\&= 2\sum_{u=1}^{n-1}(I_n - M_n) \otimes \rho^{\text{def}}(s_{(u,n)}) - 2\sum_{u=1}^{n-1} M_u \otimes \rho^{\text{def}}(s_{(u,n)})\\&= 2\sum_{u=1}^{n-1}\text{Blockdiag}\Big(\rho^{\text{def}}(s_{(u,n)}), \rho^{\text{def}}(s_{(u,n)}), \ldots, \rho^{\text{def}}(s_{(u,n)}), O_n\Big)\\&\qquad - 2\ \text{Blockdiag}\Big(\rho^{\text{def}}(s_{(1,n)}), \rho^{\text{def}}(s_{(2,n)}), \ldots, \rho^{\text{def}}(s_{(n-1,n)}), O_n\Big)\\&= 2\ \text{Blockdiag}(\beta_1, \beta_2, \ldots, \beta_{n-1}, O_n).\end{aligned}$$

□

Lemma 4.11. *The eigenvalues of* $\sum_{u=1}^{n}\Big(\rho^-((u,n)) \otimes \rho^-((u,n)) + \rho^-((-u,n)) \otimes \rho^-((-u,n))\Big)$ *are given below:*

Eigenvalues:	$2n$	$2(n-1)$	0	2	-2	$2(n-2)$
Multiplicities:	1	$n-2$	1	$n-1$	$n-1$	$(n-1)(n-2)$

Proof. For simplicity let us call the matrix in the statement R. Now let us consider the following vectors of $V^- \otimes V^-$.

$$\begin{aligned}v_{1,1} &= \mathbf{b_1} \otimes \mathbf{b_1} + \cdots + \mathbf{b_n} \otimes \mathbf{b_n}\\v_{i,i} &= \mathbf{b_i} \otimes \mathbf{b_i} - \mathbf{b_1} \otimes \mathbf{b_1} \ \text{ for } i \in \{2, \ldots, n-1\}\\v_{n,n} &= v_{1,1} - n(\mathbf{b_n} \otimes \mathbf{b_n})\\v_{i,n}^+ &= \mathbf{b_i} \otimes \mathbf{b_n} + \mathbf{b_n} \otimes \mathbf{b_i} \ \text{ for } i \in \{1, \ldots, n-1\}\\v_{i,n}^- &= \mathbf{b_i} \otimes \mathbf{b_n} - \mathbf{b_n} \otimes \mathbf{b_i} \ \text{ for } i \in \{1, \ldots, n-1\}\\v_{i,j} &= \mathbf{b_i} \otimes \mathbf{b_j} \ \text{ for } i, j \in \{1, \ldots, n-1\} \text{ and } i \neq j.\end{aligned} \tag{4.3.3}$$

It can be easily seen that the vectors in (4.3.3) are linearly independent. Now the lemma follows from the following:

$$\begin{aligned}
&R(v_{1,1}) = (2n)v_{1,1} \\
&R(v_{i,i}) = (2n-2)v_{i,i} \text{ for } i \in \{2,\dots,n-1\} \\
&R(v_{n,n}) = 0 \\
&R(v^+_{i,n}) = 2v^+_{i,n} \text{ for } i \in \{1,\dots,n-1\} \\
&R(v^-_{i,n}) = (-2)v^-_{i,n} \text{ for } i \in \{1,\dots,n-1\} \\
&R(v_{i,j}) = (2n-4)v_{i,j} \text{ for } i,j \in \{1,\dots,n-1\} \text{ and } i \neq j. \qquad \square
\end{aligned}$$

Proposition 4.12. *Let X, $E_k(X)$ be defined as in the beginning of this section. Then we have, $E_k(X) = 1 + (2n-3)\left(1-\frac{1}{n}\right)^k$.*

Proof. Using (4.3.1) and the definition of expectation of a random variable we have

$$\begin{aligned}
E_k(X) &= \sum_{\pi\in B_n} X(\pi)P_B^{*k}(\pi) \\
&= \sum_{\pi\in B_n} \operatorname{Tr}\left(\rho^+(\pi)+\rho^-(\pi)\right) P_B^{*k}(\pi) \\
&= \operatorname{Tr}\left(\sum_{\pi\in B_n} \rho^+(\pi)P_B^{*k}(\pi)\right) + \operatorname{Tr}\left(\sum_{\pi\in B_n} \rho^-(\pi)P_B^{*k}(\pi)\right) \\
&= \operatorname{Tr}\left(\widehat{P_B^{*k}}(\rho^+)\right) + \operatorname{Tr}\left(\widehat{P_B^{*k}}(\rho^-)\right) \\
&= \operatorname{Tr}\left(\left(\widehat{P_B}(\rho^+)\right)^k\right) + \operatorname{Tr}\left(\left(\widehat{P_B}(\rho^-)\right)^k\right). \qquad (4.3.4)
\end{aligned}$$

Now using Lemma 4.8 we have

$$\widehat{P_B}(\rho^+) = \frac{1}{2n}\sum_{u=1}^{n}\left(\rho^+((u,n))+\rho^+((-u,n))\right) = \frac{1}{n}\left(I_n + \sum_{u=1}^{n-1}\rho^{\mathrm{def}}(s_{(u,n)})\right). \qquad (4.3.5)$$

Therefore from (4.3.5) and Lemma 2.13 the eigenvalues of $\widehat{P_B}(\rho^+)$ are the following:

Eigenvalues:	1	$\left(1-\frac{1}{n}\right)$	0
Multiplicities:	1	$n-2$	1

Again from Lemma 4.9, $\widehat{P_B}(\rho^-) = \frac{1}{2n}\sum_{u=1}^{n}\left(\rho^-((u,n))+\rho^-((-u,n))\right)$ can be written as

$$\begin{aligned}
\frac{1}{n}\left(\sum_{u=1}^{n-1}(I_n - M_u - M_n) + (I_n - M_n)\right) &= \frac{1}{n}\left(n\left(I_n - M_n\right) - \sum_{u=1}^{n-1} M_u\right) \\
&= \frac{n-1}{n}\left(I_n - M_n\right).
\end{aligned}$$

Thus the eigenvalues of $\widehat{P_B}(\rho^-)$ are given below.

Eigenvalues:	$\left(1-\frac{1}{n}\right)$	0
Multiplicities:	$n-1$	1

Hence the proposition follows from (4.3.4). □

Proposition 4.13. *Following the definitions of* X *and* $\mathrm{Var}_k(X)$*, we have*

$$\begin{aligned}\mathrm{Var}_k(X) =& 2+(4n-6)\left(1-\frac{1}{n}\right)^k+(4n^2-16n+13)\left(1-\frac{2}{n}\right)^k\\ &+(2n-3)\left(\frac{1+(-1)^k}{n^k}\right)-(4n^2-12n+9)\left(1-\frac{1}{n}\right)^{2k}.\end{aligned}$$

Proof. We first find $E_k(X^2)$. Now using (4.3.1), for each $\pi \in B_n$ we have the following:

$$\begin{aligned}(X(\pi))^2 &= \mathrm{Tr}\left(\left(\rho^+(\pi)+\rho^-(\pi)\right)\otimes\left(\rho^+(\pi)+\rho^-(\pi)\right)\right)\\ &= \mathrm{Tr}\left(\rho^+(\pi)\otimes\rho^+(\pi)\right)+2\ \mathrm{Tr}\left(\rho^-(\pi)\otimes\rho^+(\pi)\right)+\mathrm{Tr}\left(\rho^-(\pi)\otimes\rho^-(\pi)\right)\\ &= \mathrm{Tr}\left(\rho_1(\pi)\right)+2\ \mathrm{Tr}\left(\rho_2(\pi)\right)+\mathrm{Tr}\left(\rho_3(\pi)\right),\end{aligned} \tag{4.3.6}$$

where $\rho_1 : B_n \to GL(V^+\otimes V^+)$, $\rho_2 : B_n \to GL(V^-\otimes V^+)$ and $\rho_3 : B_n \to GL(V^-\otimes V^-)$ be three representations of B_n defined below.

$$\begin{aligned}\rho_1(\pi) &= \left(\rho^+\otimes\rho^+\right)(\pi)(v_i\otimes v_j)=\rho^+(\pi)(v_i)\otimes\rho^+(\pi)(v_j) \text{ for } \pi\in B_n, v_i\in V^+, v_j\in V^+,\\ \rho_2(\pi) &= \left(\rho^-\otimes\rho^+\right)(\pi)(v_i\otimes v_j)=\rho^-(\pi)(v_i)\otimes\rho^+(\pi)(v_j) \text{ for } \pi\in B_n, v_i\in V^-, v_j\in V^+,\\ \rho_3(\pi) &= \left(\rho^-\otimes\rho^-\right)(\pi)(v_i\otimes v_j)=\rho^-(\pi)(v_i)\otimes\rho^-(\pi)(v_j) \text{ for } \pi\in B_n, v_i\in V^-, v_j\in V^-.\end{aligned}$$

Now we have

$$\begin{aligned}E_k(X^2) &= \sum_{\pi\in B_n}(X(\pi))^2 P_B^{*k}(\pi)\\ &= \sum_{\pi\in B_n}\left(\mathrm{Tr}\left(\rho_1(\pi)\right)+2\ \mathrm{Tr}\left(\rho_2(\pi)\right)+\mathrm{Tr}\left(\rho_3(\pi)\right)\right)P_B^{*k}(\pi)\\ &= \mathrm{Tr}\left(\sum_{\pi\in B_n}\rho_1(\pi)P_B^{*k}(\pi)\right)+2\ \mathrm{Tr}\left(\sum_{\pi\in B_n}\rho_2(\pi)P_B^{*k}(\pi)\right)+\mathrm{Tr}\left(\sum_{\pi\in B_n}\rho_3(\pi)P_B^{*k}(\pi)\right)\\ &= \mathrm{Tr}\left(\widehat{P_B^{*k}}(\rho_1)\right)+2\ \mathrm{Tr}\left(\widehat{P_B^{*k}}(\rho_2)\right)+\mathrm{Tr}\left(\widehat{P_B^{*k}}(\rho_3)\right)\\ &= \mathrm{Tr}\left(\left(\widehat{P_B}(\rho_1)\right)^k\right)+2\ \mathrm{Tr}\left(\left(\widehat{P_B}(\rho_2)\right)^k\right)+\mathrm{Tr}\left(\left(\widehat{P_B}(\rho_3)\right)^k\right).\end{aligned} \tag{4.3.7}$$

Now from Lemma 4.8 we have

$$\begin{aligned}\widehat{P_B}(\rho_1) &= \frac{1}{2n}\left(\sum_{u=1}^{n}\Big(\rho^+((u,n))\otimes\rho^+((u,n)) + \rho^+((-u,n))\otimes\rho^+((-u,n))\Big)\right)\\ &= \frac{1}{n}\left(\sum_{u=1}^{n-1}\Big(\rho^{\text{def}}(s_{(u,n)})\otimes\rho^{\text{def}}(s_{(u,n)})\Big) + I_n\otimes I_n\right).\end{aligned} \tag{4.3.8}$$

Therefore from Lemma 2.14 the eigenvalues of $\widehat{P_B}(\rho_1)$ are:

Eigenvalues:	1	$\left(1-\frac{1}{n}\right)$	0	$\left(1-\frac{2}{n}\right)$	$\frac{1}{n}$	$-\frac{1}{n}$
Multiplicities:	2	$3(n-2)$	3	n^2-5n+5	$n-2$	$n-2$

Again from Lemma 4.8, 4.9 and 4.10 we have the following:

$$\begin{aligned}\widehat{P_B}(\rho_2) &= \frac{1}{2n}\sum_{u=1}^{n}\Big(\rho^-((u,n))\otimes\rho^+((u,n)) + \rho^-((-u,n))\otimes\rho^+((-u,n))\Big)\\ &= \frac{1}{2n}\Big(2\ \text{Blockdiag}\left(\beta_1,\beta_2,\ldots,\beta_{n-1},O_n\right) + \Big(I_n+\rho^-((-n,n))\Big)\otimes I_n\Big)\\ &= \frac{1}{n}\left(\text{Blockdiag}\left(\beta_1,\beta_2,\ldots,\beta_{n-1},O_n\right) + (I_n-M_n)\otimes I_n\right)\\ &= \frac{1}{n}\ \text{Blockdiag}\left(\beta_1+I_n,\beta_2+I_n,\ldots,\beta_{n-1}+I_n,O_n\right).\end{aligned} \tag{4.3.9}$$

Therefore from (4.3.9) and Lemma 2.15, the eigenvalues of $\widehat{P_B}(\rho_2)$ are:

Eigenvalues:	$\left(1-\frac{1}{n}\right)$	$\left(1-\frac{2}{n}\right)$	0
Multiplicities:	$2(n-1)$	$(n-3)(n-1)$	$2n-1$

Also from Lemma 4.11, the eigenvalues of $\widehat{P_B}(\rho_3)$ are:

Eigenvalues:	1	$\left(1-\frac{1}{n}\right)$	0	$\left(1-\frac{2}{n}\right)$	$\frac{1}{n}$	$-\frac{1}{n}$
Multiplicities:	1	$n-2$	1	$(n-1)(n-2)$	$n-1$	$n-1$

Hence from (4.3.7) we have

$$E_k(X^2) = 3+(8n-12)\left(1-\frac{1}{n}\right)^k + (4n^2-16n+13)\left(1-\frac{2}{n}\right)^k + (2n-3)\left(\frac{1+(-1)^k}{n^k}\right).$$

Thus the proposition follows from Proposition 4.12 and straightforward calculations. □

Proposition 4.14. *Let $E_{U_{B_n}}(X)$ denote the expectation of X with respect to the uniform distribution on B_n, $E_{U_{B_n}}(X) = 1$.*

Proof. We note that signed permutations which fix i will automatically fix $(-i)$. Let $\mathcal{B}_i$ be the set of sign permutations in B_n which fix i. Basic combinatorial arguments imply $|\mathcal{B}_i| = 2^{n-1}(n-1)!$ for all $i\in\{-n,\ldots,-1,1,\ldots,n\}$.

Let $\mathcal{R}^{\text{def}}: B_n \to GL(V)$ be the defining representation on B_n. Then we have,

$$\begin{aligned} E_{U_{B_n}}(X) = \sum_{\pi \in B_n} X(\pi) U_{B_n}(\pi) &= \frac{1}{|B_n|} \sum_{\pi \in B_n} \operatorname{Tr}\left(\mathcal{R}^{\text{def}}(\pi)\right) \\ &= \frac{1}{|B_n|} \sum_{\substack{i=-n \\ i \neq 0}}^{n} |\mathcal{B}_i| = \frac{2n|\mathcal{B}_i|}{2^n n!} = 1. \end{aligned}$$ □

Theorem 4.15. *Let X, $E_k(X)$ and $\operatorname{Var}_k(X)$ be as given above. Then we have*

1. *For large n, $||P_B^{*k} - U_{B_n}||_{\text{TV}} \geq 1 - \frac{10(1+2e^{-c}+o(1))}{(1+2e^{-c}+o(1))^2}$, when $k = n\log n + cn$ and $c \ll 0$.*

2. *$\lim_{n\to\infty} ||P_B^{*k_n} - U_{B_n}||_{\text{TV}} = 1$, for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon) n \log n \rfloor$.*

Proof. Using Lemma 2.19, $\mu = P_B^{*k}$ and $\nu = U_{B_n}$ we have,

$$||P_B^{*k} - U_{B_n}||_{\text{TV}} \geq 1 - \frac{4\operatorname{Var}_k(X)}{(E_k(X))^2} - \frac{2}{E_k(X)}. \tag{4.3.10}$$

Now if n is large, we have

$$E_k(X) \approx 1 + (2n-3)e^{-\frac{k}{n}} \tag{4.3.11}$$

from Proposition 4.12 and

$$\operatorname{Var}_k(X) \approx 2 + (4n-6)e^{-\frac{k}{n}} - 4(n-1)e^{-\frac{2k}{n}} + (2n-3)\left(\frac{1+(-1)^k}{n^k}\right) \tag{4.3.12}$$

from Proposition 4.13. Recall that ' $\approx$' means 'asymptotic to' i.e. $a_n \approx b_n$ means $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Now if n is large, $c \ll 0$ and $k = n\log n + cn$, then by (4.3.10), (4.3.11) and (4.3.12), we have the first part of this theorem. Again for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon) n \log n \rfloor$ from (4.3.10), (4.3.11) and (4.3.12), we have

$$1 \geq ||P_B^{*k_n} - U_{B_n}||_{\text{TV}} \geq 1 - \frac{10 + (20 + o(1))n^{\epsilon} + n^{2\epsilon} o(1) + o(1)}{(1 + (2 + o(1))n^{\epsilon})^2} \tag{4.3.13}$$

for large n. Therefore, the second part of this theorem follows from (4.3.13) and the fact that

$$\lim_{n\to\infty} \frac{10 + (20 + o(1))n^{\epsilon} + n^{2\epsilon} o(1) + o(1)}{(1 + (2 + o(1))n^{\epsilon})^2} = 0.$$ □

Therefore from the first part of Theorems 4.6 and 4.15 we can say that the mixing time for the flip-transpose top with random shuffle on B_n is $n \log n$. Furthermore, the second part of Theorems 4.6 and 4.15 prove the total variation cutoff for this shuffle.

4.4 Representation theory of D_n

In this section we briefly discuss the irreducible representations of D_n (detailed proofs are omitted) [71]. Our main aim is to look at the restriction of the irreducible representations of B_n to D_n.

Let us consider the one-dimensional character (or representation) $\xi : B_n \to (\{\pm 1\}, \cdot)$ of B_n. The action of ξ on the generators of B_n is defined by

$$\xi(\pi) = \begin{cases} -1, & \text{if } \pi = (-1, 1), \\ 1, & \text{if } \pi = (i, i+1) \text{ for } 1 \leq i \leq n-1. \end{cases} \tag{4.4.1}$$

It can be easily seen that $\ker(\xi) = D_n$ and the B_n-module $V \otimes \xi$ is irreducible if and only if the B_n-module V is irreducible. We have already seen in Section 4.2 that the irreducible representations of B_n are indexed by $\mathcal{D}_\mathrm{n}$. If $\mu = (\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_\mathrm{n}$, then $\tilde{\mu} = (\mu^{(2)}, \mu^{(1)}) \in \mathcal{D}_\mathrm{n}$. Now from [50, Proposition II.1.(ii)] it follows that the irreducible B_n-modules $V^\mu \otimes \xi$ and $V^{\tilde{\mu}}$ are isomorphic for $\mu \in \mathcal{D}_\mathrm{n}$.

Theorem 4.16. *For the irreducible B_n-module V^μ indexed by $\mu = (\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_\mathrm{n}$, we have the following:*

1. *If $\mu^{(1)} \neq \mu^{(2)}$, then the restriction $V^\mu \downarrow_{D_n}^{B_n}$ of V^μ to D_n is irreducible as a D_n-module. We denote this irreducible D_n-module by the same notation V^μ. Moreover if $\tilde{\mu} = (\mu^{(2)}, \mu^{(1)})$, then V^μ and $V^{\tilde{\mu}}$ are isomorphic as D_n-modules. If $\nu \in \mathcal{D}_\mathrm{n}$ be such that $\nu \neq \mu$ and $\nu \neq \tilde{\mu}$, then V^ν and V^μ are non-isomorphic as D_n-modules.*

2. *If $\mu^{(1)} = \mu^{(2)}$, then the restriction $V^\mu \downarrow_{D_n}^{B_n}$ of V^μ to D_n is a direct sum of two irreducible D_n-modules with same dimension. We denote these irreducible D_n-modules by the V_+^μ and V_-^μ.*

Proof. Recall that S_n denotes the symmetric group and A_n denotes the alternating group. The proof of this theorem follows by mimicking the steps of deducing the irreducible representations of A_n from that of S_n [81, Theorem 4.4.2, Theorem 4.6.5]. For this proof B_n (respectively D_n) will play the role of S_n (respectively A_n) and ξ will play the role of the one-dimensional sign character of S_n. □

Let $\mathcal{S}$ be the collection of subsets Γ of $\mathcal{D}_\mathrm{n}$ satisfying the following properties:

1. $\mu^{(1)} \neq \mu^{(2)}$ for each $(\mu^{(1)}, \mu^{(2)}) \in \Gamma$,

2. $(\mu^{(2)}, \mu^{(1)}) \notin \Gamma$ if and only if $(\mu^{(1)}, \mu^{(2)}) \in \Gamma$.

Let Γ_1 be the maximal element of the poset $(\mathcal{S}, \subseteq)$ and $\Gamma_2 = \{(\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_{\mathrm{n}} : \mu^{(1)} = \mu^{(2)}\}$. Then from Theorem 4.16 and the observation

$$\begin{aligned}
&\sum_{\mu\in\Gamma_1} (\dim(V^{\mu}))^2 + \sum_{\mu\in\Gamma_2} \left((\dim(V_+^{\mu}))^2 + (\dim(V_-^{\mu}))^2\right),\\
&= \sum_{\mu\in\Gamma_1} (\dim(V^{\mu}))^2 + \sum_{\mu\in\Gamma_2} \left(\left(\frac{\dim(V^{\mu})}{2}\right)^2 + \left(\frac{\dim(V^{\mu})}{2}\right)^2\right),\\
&= \frac{1}{2}\left(2\sum_{\mu\in\Gamma_1} (\dim(V^{\mu}))^2 + \sum_{\mu\in\Gamma_2} (\dim(V^{\mu}))^2\right) = \frac{|B_n|}{2} = |D_n|,
\end{aligned}$$

all the irreducible D_n-modules are given by $\{V^{\mu} : \mu \in \Gamma_1\} \cup \{V_+^{\mu}, V_-^{\mu} : \mu \in \Gamma_2\}$.

4.5 Flip-transpose top with random shuffle on D_n

Let us consider the random walk on the demihyperoctahedral group D_n driven by the probability measure P_D on D_n defined as follows:

$$P_D(\pi) = \begin{cases} \frac{1}{2n-1}, & \text{if } \pi = \text{id},\ \text{the identity element of } D_n,\\ \frac{1}{2n-1}, & \text{if } \pi = (i,n) \text{ or } (-i,n) \text{ for } 1 \leq i \leq n-1,\\ 0, & \text{otherwise.}\end{cases} \tag{4.5.1}$$

We call it the *flip-transpose top with random shuffle on* D_n. It can be easily seen that the support of P_D generates D_n and hence this random walk is irreducible (Proposition 2.22). Moreover this random walk is aperiodic too. Thus the distribution after kth transition for this random walk will converge to U_{D_n} (Proposition 2.23) as $k \to \infty$. Let us recall that $\widehat{P_D}(R)$ is the Fourier transform of P_D at the right regular representation R of D_n. The transition matrix for the random walk on D_n driven by P_D is the transpose of $\widehat{P_D}(R)$. To find the eigenvalues of $\widehat{P_D}(R)$ we will use the representation theory of D_n.

Theorem 4.17. *The eigenvalues of* $\widehat{P_D}(R)$ *are given by*

1. *If* $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \Gamma_1$*, then for each* $T \in \operatorname{tab}_{\mathcal{D}}(n,\mu)$, $\frac{2c(b_T(n))+1}{2n-1}$ *is an eigenvalue of* $\widehat{P_D}(R)$ *with multiplicity* $\dim(V^{\mu})$.

2. *If* $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \Gamma_2$*, then for each* $T \in \operatorname{tab}_{\mathcal{D}}(n,\mu)$, $\frac{2c(b_T(n))+1}{2n-1}$ *is an eigenvalue of* $\widehat{P_D}(R)$ *with multiplicity* $\frac{1}{2}\dim(V^{\mu})$.

Recall $c(b_T(n))$ *is the content of the box containing* n *in* T.

Proof. We have $\widehat{P_D}(R) = \frac{1}{2n-1}\left(X_n^B + \mathrm{id}\right)$, where X_n^B is the nth Young-Jucys-Murphy element of B_n and id is the identity element of D_n. Here we identify the elements of $D_n(\subseteq B_n)$ by the elements of B_n.

For $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \Gamma_1$, we have $\mu^{(1)} \neq \mu^{(2)}$. Therefore the restriction of irreducible B_n-module V^μ to D_n is irreducible (Theorem 4.16). Now for each $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$, let v_T be the Gelfand-Tsetlin vector of V^μ satisfying $X_n^B v_T = 2c(b_T(n))v_T$. Also we know that $\{v_T : T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)\}$ forms a basis of V^μ. Therefore the eigenvalues of $\widehat{P_D}(R)$ on the irreducible D_n-module V^μ are given by $\frac{2c(b_T(n))+1}{2n-1}$ for each $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$. Since the multiplicity of every irreducible representation in the right regular representation is equal to its dimension, therefore the multiplicity of these eigenvalues are $\dim(V^\mu)$.

Now for $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \Gamma_2$ we have $\mu^{(1)} = \mu^{(2)}$. Then the restriction of the irreducible B_n-module V^μ to D_n splits into two irreducible D_n-modules V_+^μ and V_-^μ (Theorem 4.16). In this case also v_T is the Gelfand-Tsetlin vector of V^μ and $\{v_T : T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)\}$ forms a basis of $V_+^\mu \oplus V_-^\mu$. Therefore, by similar arguments in case of $\mu^{(1)} \neq \mu^{(2)}$, the eigenvalues of $\widehat{P_D}(R)$ on the irreducible D_n-modules V_+^μ and V_-^μ are given by $\frac{2c(b_T(n))+1}{2n-1}$ for each $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$. The multiplicity of these eigenvalues are $\frac{1}{2}\dim(V^\mu)$ ($\because$ $\dim(V_+^\mu) = \dim(V_-^\mu) = \frac{1}{2}\dim(V^\mu)$). □

Theorem 4.18. *For the random walk on D_n driven by P_D, we have the following:*

1. $||P_D^{*k} - U_{D_n}||_{\mathrm{TV}} < \sqrt{e+1}\, e^{-c} + o(1)$, *for* $k \geq \left(n - \frac{1}{2}\right)(\log n + c)$ *and* $c > 0$.

2. $\lim_{n\to\infty} ||P_D^{*k_n} - U_{D_n}||_{\mathrm{TV}} = 0$, *for any* $\epsilon \in (0,1)$ *and* $k_n = \lfloor (1+\epsilon)\left(n - \frac{1}{2}\right)\log n \rfloor$.

Proof. Using Lemma 2.24 and following similar steps of Theorem 4.6, we have

$$4||P_D^{*k} - U_{D_n}||_{\mathrm{TV}}^2 \leq (2+2e)\left(e^{n^2 e^{-\frac{4k}{2n-1}}} - 1\right) + e^{-\frac{4k}{2n-1}}, \quad \text{for } k \geq \left(n - \frac{1}{2}\right)\log n. \quad (4.5.2)$$

Now if $k \geq \left(n - \frac{1}{2}\right)(\log n + c)$ and $c > 0$, then the right hand side of (4.5.2) becomes

$$(2e+2)\left(e^{e^{-2c}} - 1\right) + \frac{e^{-2c}}{n^2} < (4e+4)e^{-2c} + \frac{e^{-2c}}{n^2} = (4e+4)e^{-2c} + o(1).$$

This proves the first part of the theorem. Now for $\epsilon \in (0,1)$, $k_n = \lfloor (1+\epsilon)\left(n - \frac{1}{2}\right)\log n \rfloor$ implies, $k_n \geq (1+\epsilon)\left(n - \frac{1}{2}\right)\log n$. Thus the right hand side of (4.5.2) is bounded above by $(2e+2)\left(e^{\frac{1}{n^{2\epsilon}}} - 1\right) + \frac{1}{n^{2(1+\epsilon)}}$. Therefore the proof of the second part follows from

$$\lim_{n\to\infty}(2e+2)\left(e^{\frac{1}{n^{2\epsilon}}} - 1\right) + \frac{1}{n^{2(1+\epsilon)}} = 0.$$ □

To obtain the lower bound for the total variation distance $||P_D^{*k} - U_{D_n}||_{\text{TV}}$ we define the random variable Y as in the case of the walk on B_n driven by P_B as follows:

$$Y(\pi) = \text{ number of fixed points of } \pi, \; \pi \in D_n.$$

i.e., Y is the restriction of X to D_n. Now using the definitions of ρ^+, ρ^- and ρ^{def} and the conventions for the ordering of the bases to obtain the matrices as given in Section 4.3, we have the following:

$$\begin{aligned} Y(\pi) &= \operatorname{Tr}\left(\rho^+ \downarrow_{D_n}^{B_n}(\pi) + \rho^- \downarrow_{D_n}^{B_n}(\pi)\right) \\ &= \operatorname{Tr}\left(\rho^+(\pi) + \rho^-(\pi)\right), \text{ for } \pi \in D_n. \end{aligned}$$

Now, if $E_{U_{D_n}}(Y)$ denotes the expectation of Y with respect to the uniform distribution on D_n, then from standard combinatorial arguments one can show that $E_{U_{D_n}}(Y) = 1$.

Lemma 4.19. *The eigenvalues of*

$$\sum_{u=1}^{n-1} \left(\rho^-((u,n)) \otimes \rho^-((u,n)) + \rho^-((-u,n)) \otimes \rho^-((-u,n))\right)$$

are given as below:

Eigenvalues:	$2n-2$	$2(n-2)$	-2	2	$2(n-3)$
Multiplicities:	1	$n-2$	n	$n-1$	$(n-1)(n-2)$

Proof. The n^2 independent vectors defined in (4.3.3) are the eigenvectors in this case also. □

Proposition 4.20. *Let $E_k(Y)$ be the expectation of Y with respect to the probability measure P_D^{*k}. Then we have $E_k(Y) = 1 + (2n-3)\left(1 - \frac{2}{2n-1}\right)^k + \frac{1+(-1)^k}{(2n-1)^k}$.*

Proof. Following similar steps we used in Proposition 4.12 to get (4.3.4) and using (4.3.1) we have the following:

$$\begin{aligned} E_k(Y) &= \sum_{\pi \in D_n} Y(\pi) P_D^{*k}(\pi) \\ &= \operatorname{Tr}\left(\left(\widehat{P_D}\left(\rho^+ \downarrow_{D_n}^{B_n}\right)\right)^k\right) + \operatorname{Tr}\left(\left(\widehat{P_D}\left(\rho^- \downarrow_{D_n}^{B_n}\right)\right)^k\right). \end{aligned} \tag{4.5.3}$$

Now from Lemmas 4.8, 2.13, 4.9 and by the similar arguments given in the proof of Proposition 4.12 we have the following: The eigenvalues of $\widehat{P_D}\left(\rho^+ \downarrow_{D_n}^{B_n}\right)$ are:

Eigenvalues:	1	$\left(1 - \frac{2}{2n-1}\right)$	$-\frac{1}{2n-1}$
Multiplicities:	1	$n-2$	1

The eigenvalues of $\widehat{P_D}\left(\rho^- \downarrow_{D_n}^{B_n}\right)$ are:

Eigenvalues:	$\left(1-\frac{2}{2n-1}\right)$	1
Multiplicities:	$n-1$	1

The proposition follows from (4.5.3) and the fact that the trace of kth power of matrix is the sum of kth powers of its eigenvalues. □

Proposition 4.21. *Let* $\mathrm{Var}_k(Y)$ *be the variance of* Y *with respect to the probability measure* P_D^{*k}*. Then we have*

$$\begin{aligned}\mathrm{Var}_k(Y) =& 2+(4n-6)\left(1-\frac{2}{2n-1}\right)^k+(4n^2-16n+13)\left(1-\frac{4}{2n-1}\right)^k\\&+\frac{\left(1+(-1)^k\right)}{(2n-1)^k}\left(3n-1-\frac{2}{(2n-1)^k}-2\left(1-\frac{2}{2n-1}\right)^k\right)\\&+(n-2)\frac{\left(3^k+(-3)^k\right)}{(2n-1)^k}-\frac{3}{(2n-1)^k}+\left(\frac{3}{2n-1}\right)^k\\&-(4n^2-12n+9)\left(1-\frac{2}{2n-1}\right)^{2k}.\end{aligned}$$

Proof. We first find $E_k(Y^2)$. From (4.3.1) and by similar arguments given in the proof of Proposition 4.13 to obtain (4.3.6), we have

$$(Y(\pi))^2=\mathrm{Tr}\left(\rho_1\downarrow_{D_n}^{B_n}(\pi)\right)+2\ \mathrm{Tr}\left(\rho_2\downarrow_{D_n}^{B_n}(\pi)\right)+\mathrm{Tr}\left(\rho_3\downarrow_{D_n}^{B_n}(\pi)\right)\quad\text{for each }\pi\in D_n. \tag{4.5.4}$$

Again following similar steps we used in the proof of Proposition 4.13 to get (4.3.7) we have

$$E_k(Y^2)=\mathrm{Tr}\left(\left(\widehat{P_D}\left(\rho_1\downarrow_{D_n}^{B_n}\right)\right)^k\right)+2\ \mathrm{Tr}\left(\left(\widehat{P_D}\left(\rho_2\downarrow_{D_n}^{B_n}\right)\right)^k\right)+\mathrm{Tr}\left(\left(\widehat{P_D}\left(\rho_3\downarrow_{D_n}^{B_n}\right)\right)^k\right). \tag{4.5.5}$$

Now following Lemmas 2.14, 4.8 and similar steps in the proof of Proposition 4.13, we have the Eigenvalues of $\widehat{P_D}\left(\rho_1\downarrow_{D_n}^{B_n}\right)$ are:

Eigenvalues:	1	$\left(1-\frac{2}{2n-1}\right)$	$-\frac{1}{2n-1}$	$\left(1-\frac{4}{2n-1}\right)$	$\frac{1}{2n-1}$	$-\frac{3}{2n-1}$
Multiplicities:	2	$3(n-2)$	3	n^2-5n+5	$n-2$	$n-2$

From Lemmas 2.15 and 4.10 the eigenvalues of $\widehat{P_D}\left(\rho_2\downarrow_{D_n}^{B_n}\right)$ are:

Eigenvalues:	$\left(1-\frac{2}{2n-1}\right)$	$\left(1-\frac{4}{2n-1}\right)$	$-\frac{1}{2n-1}$	$\frac{1}{2n-1}$
Multiplicities:	$2(n-1)$	$(n-3)(n-1)$	$n-1$	n

Finally from Lemma 4.19 the eigenvalues of $\widehat{P_D}\left(\rho_3\downarrow_{D_n}^{B_n}\right)$ are:

Eigenvalues:	1	$-\frac{1}{2n-1}$	$\frac{3}{2n-1}$	$\left(1-\frac{2}{2n-1}\right)$	$\left(1-\frac{4}{2n-1}\right)$
Multiplicities:	1	n	$n-1$	$n-2$	$(n-1)(n-2)$

Therefore this proposition follows from (4.5.5) and Proposition 4.20. □

Theorem 4.22. *Following the expression of* Y*,* $E_k(Y)$ *and* $\mathrm{Var}_k(Y)$*, we have*

1. *For large* n*,* $||P_D^{*k} - U_{D_n}||_{\mathrm{TV}} \geq 1 - \frac{10(1+2e^{-c}+o(1))}{(1+2e^{-c}+o(1))^2}$*, when* $k = (n-\frac{1}{2})(\log n + c)$ *and* $c \ll 0$*.*

2. $\lim_{n\to\infty} ||P_D^{*k_n} - U_{D_n}||_{\mathrm{TV}} = 1$*, for any* $\epsilon \in (0,1)$ *and* $k_n = \lfloor (1-\epsilon)(n-\frac{1}{2})\log n \rfloor$*.*

Proof. Using Lemma 2.19, $\mu = P_D^{*k}$ and $\nu = U_{D_n}$ we have,

$$||P_D^{*k} - U_{D_n}||_{\mathrm{TV}} \geq 1 - \frac{4\,\mathrm{Var}_k(Y)}{(E_k(Y))^2} - \frac{2}{E_k(Y)}. \tag{4.5.6}$$

For large enough n, we have

$$E_k(Y) \approx 1 + (2n-3)e^{-\frac{2k}{2n-1}} \tag{4.5.7}$$

from Proposition 4.20 and

$$\mathrm{Var}_k(Y) \approx 2 + 2\left(2n - 3 - \frac{\left(1+(-1)^k\right)}{(2n-1)^k}\right)e^{-\frac{2k}{2n-1}} - 4(n-1)e^{-\frac{4k}{2n-1}} \tag{4.5.8}$$

from Proposition 4.21. Now if n is large, $c \ll 0$ and $k = (n-\frac{1}{2})(\log n + c)$, then by (4.5.6), (4.5.7) and (4.5.8), we have the first part of this theorem. Again for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon)(n-\frac{1}{2})\log n \rfloor$ from (4.5.6), (4.5.7) and (4.5.8), we have

$$1 \geq ||P_D^{*k_n} - U_{D_n}||_{\mathrm{TV}} \geq 1 - \frac{10 + (20+o(1))n^{\epsilon} + n^{2\epsilon}o(1) + o(1)}{(1+(2+o(1))n^{\epsilon})^2} \tag{4.5.9}$$

for large n. Therefore, the second part of this theorem follows from (4.5.9) and the fact that

$$\lim_{n\to\infty} \frac{10 + (20+o(1))n^{\epsilon} + n^{2\epsilon}o(1) + o(1)}{(1+(2+o(1))n^{\epsilon})^2} = 0.$$

□

Therefore the second part of Theorems 4.18 and 4.22 implies that the random walk on D_n driven by P_D satisfies the cutoff phenomenon and the total variation cutoff for this shuffle occurs at $\left(n-\frac{1}{2}\right)\log n$.

Chapter 5

The warp-transpose top with random shuffle

In this chapter, we consider a random walk on the complete monomial group $G_n \wr S_n$. This is a generalization of the flip-transpose top with random shuffle of Chapter 4 (see Remark 5.17). We call this the *warp-transpose top with random shuffle*. In Section 5.1, we introduce this shuffle. We obtain the spectrum of the transition matrix for this shuffle in Section 5.2. In Section 5.3, we give an upper bound for the total variation distance of the distribution after k transitions from the stationary distribution. We also show that the mixing time for this shuffle is $O\left(n \log n + \frac{1}{2} n \log(|G_n| - 1)\right)$, in this section. In Section 5.4, we obtain a lower bound for the total variation distance of the distribution after k transitions from the stationary distribution. In this section, we also show that this shuffle satisfies the cutoff phenomenon if $|G_n| = o(n^\delta)$ for all $\delta > 0$.

5.1 Introduction

Definition 5.1. Let G be a finite group. Recall that S_n is the symmetric group of permutations of elements of the set $[n]$. The *complete monomial group* is the wreath product of G with S_n, is a group denoted by $G \wr S_n$ and can be described as follows: The elements of $G \wr S_n$ are $(n+1)$-tuples $(g_1, g_2, \ldots, g_n; \pi)$ where $g_i \in G$ and $\pi \in S_n$. The multiplication in $G \wr S_n$ is given by

$$(g_1, \ldots, g_n; \pi)(h_1, \ldots, h_n; \eta) = (g_1 h_{\pi^{-1}(1)}, \ldots, g_n h_{\pi^{-1}(n)}; \pi\eta).$$

The identity element is $(1_G, \ldots, 1_G; \text{id})$, where 1_G is the identity element of G and id is the identity permutation. Therefore $(g_1, \ldots, g_n; \pi)^{-1} = (g^{-1}_{\pi(1)}, \ldots, g^{-1}_{\pi(n)}; \pi^{-1})$.

Let $G_1 \subseteq \cdots \subseteq G_n \subseteq \cdots$ be a sequence of finite groups such that $|G_1| > 2$. We consider the complete monomial groups $\mathcal{G}_\mathrm{n} := G_n \wr S_n$ for each positive integer n. Let e be the identity of G_n and id be the identity of S_n. For an element $\pi \in S_n$, let $\pi := (\mathrm{e}, \ldots, \mathrm{e}; \pi) \in \mathcal{G}_\mathrm{n}$ and for $g \in G_n$, let

$$g^{(i)} := (\mathrm{e}, \ldots, \mathrm{e}, \underset{\substack{\uparrow \\ i\text{th position.}}}{g}, \mathrm{e}, \ldots, \mathrm{e}; \mathrm{id}) \in \mathcal{G}_\mathrm{n}\,.$$

Unless otherwise stated from now on, $(\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n))$ denotes the element of $\mathcal{G}_\mathrm{n}$ with g^{-1} in ith position and g in nth position, for $g \in G_n,\ 1 \leq i < n$. One can check that $(g^{-1})^{(i)} g^{(n)}(i, n)$ is equal to $(\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n))$ for $g \in G_n,\ 1 \leq i < n$.

In this work we consider a random walk on the complete monomial group $\mathcal{G}_\mathrm{n}$ driven by a probability measure $P_\mathcal{G}$, defined as follows:

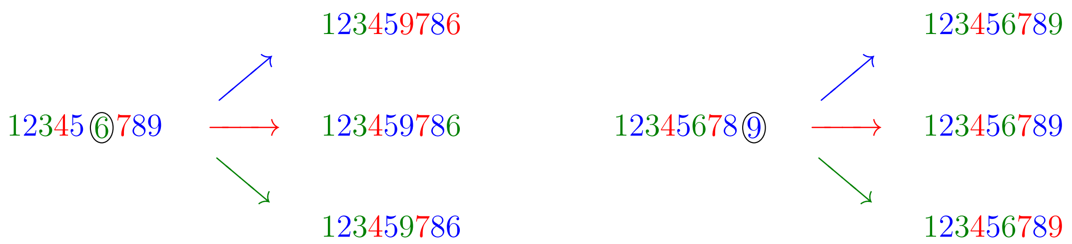
$$P_\mathcal{G}(x) = \begin{cases} \frac{1}{n|G_n|}, & \text{if } x = (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id}) \text{ for } g \in G_n, \\ \frac{1}{n|G_n|}, & \text{if } x = (\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n)) \text{ for } g \in G_n,\ 1 \leq i < n, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1.1}$$

We call this the *warp-transpose top with random shuffle* because at most times the nth component is multiplied by g and the ith component is multiplied by g^{-1} simultaneously, $g \in G_n$, $1 \leq i < n$. We now give a combinatorial description of this model as follows:

Let $\mathcal{A}_n(G)$ denote the set of all arrangements of n coloured cards in a row such that the colours of the cards are indexed by the set G. For example, if $\mathbb{Z}_2$ denotes the additive group of integers modulo 2, then elements of $\mathcal{A}_n(\mathbb{Z}_2)$ can be identified with the elements of B_n (the hyperoctahedral group). For $g, h \in G$, by saying update the colour g using colour h we mean the colour g is updated to colour $g \cdot h$. Elements of $\mathcal{G}_\mathrm{n}$ can be identified with the elements of $\mathcal{A}_n(G_n)$ as follows: The element $(g_1, \ldots, g_n; \pi) \in \mathcal{G}_\mathrm{n}$ is identified with the arrangement in $\mathcal{A}_n(G_n)$ such that the label of the ith card is $\pi(i)$, and its colour is $g_{\pi(i)}$, for each $i \in [n]$. Given an arrangement of coloured cards in $\mathcal{A}_n(G_n)$, the warp-transpose top with random shuffle on $\mathcal{G}_\mathrm{n}$ is the following: Choose a positive integer i uniformly from $[n]$. Also choose a colour g uniformly from G_n, independent of the choice of the integer i.

1. If $i = n$: update the colour of the nth card using colour g.

2. If $i < n$: first transpose the ith and n^{th} cards. Then simultaneously update the colour of the nth card using colour g and update the colour of the ith card using colour g^{-1}.



(a) (b)

Figure 5.1: Transitions for the warp-transpose top with random shuffle on $\mathbb{Z}_3 \wr S_9$. $\mathbb{Z}_3$ is the additive group of integers modulo 3, consists of the colours red, green and blue such that red represents the identity element. (a) shows transitions when the sixth card is chosen and (b) shows transitions when the last card is chosen.

Proposition 5.2. *The warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ is irreducible and aperiodic.*

Proof. The support of $P_{\mathcal{G}}$ is $\Gamma = \{(g^{-1})^{(i)}g^{(n)}(i,n), g^{(n)} \mid g \in G_n, 1 \leq i < n\}$ and it can be easily seen that $\{g^{(k)}, (i,n) \mid g \in G_n,\ 1 \leq k \leq n,\ 1 \leq i < n\}$ is a generating set of $\mathcal{G}_{\mathrm{n}}$.

$$\begin{aligned} &(g^{-1})^{(n)}\left((g^{-1})^{(i)}g^{(n)}(i,n)\right)g^{(n)} = (i,n) \text{ for each } 1 \leq i < n \text{ and } g \in G_n,\\ &(k,n)g^{(n)}(k,n) = g^{(k)} \text{ for each } 1 \leq k \leq n \text{ and for all } g \in G_n. \end{aligned} \tag{5.1.2}$$

Thus (5.1.2) implies Γ generates $\mathcal{G}_{\mathrm{n}}$ and hence the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ is irreducible (Proposition 2.22). Moreover given any $\pi \in \mathcal{G}_{\mathrm{n}}$, the set of all times when it is possible for the chain to return to the starting state π contains the integer 1 (as support of $P_{\mathcal{G}}$ contains the identity element of $\mathcal{G}_{\mathrm{n}}$). Therefore the period of the state π is 1 and hence from irreducibility all the states of this chain have period 1. Thus this chain is aperiodic. □

We know from Chapter 2 that the irreducible and aperiodic random walk on a finite group driven by some probability measure (defined on the group) converges to its stationary distribution [66, Theorem 4.9]. Moreover the stationary distribution is the uniform distribution in this case (Proposition 2.23). Thus Proposition 5.2 says that the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ converges to the uniform distribution $U_{\mathcal{G}_{\mathrm{n}}}$ as the number of transitions goes to infinity.

5.2 Spectrum of the transition matrix

In this section we find the eigenvalues of the transition matrix $\widehat{P_{\mathcal{G}}}(R)$, the Fourier transform of $P_{\mathcal{G}}$ at the right regular representation R of $\mathcal{G}_{\mathrm{n}}$. To find the eigenvalues of $\widehat{P_{\mathcal{G}}}(R)$

we will use the representation theory of the wreath product $\mathcal{G}_{\mathrm{n}}$ of a finite group G_n with the symmetric group S_n. First we briefly discuss the representation theory of $G \wr S_n$, following the notations from [73].

Recall that $\mathcal{Y}$ denotes the set of all Young diagrams (there is a unique Young diagram with zero boxes) and $\mathcal{Y}_{\mathrm{n}}$ denotes the set of all Young diagrams with n boxes. Let X be a finite set, we define $\mathcal{Y}(\mathrm{X}) = \{\mu : \mu \text{ is a map from } X \text{ to } \mathcal{Y}\}$. For $\mu \in \mathcal{Y}(\mathrm{X})$, define $||\mu|| = \sum_{x\in X} |\mu(x)|$, where $|\mu(x)|$ is the number of boxes of the Young diagram $\mu(x)$ and define $\mathcal{Y}_{\mathrm{n}}(\mathrm{X}) = \{\mu \in \mathcal{Y}(\mathrm{X}) : ||\mu|| = n\}$. Let n be a fixed positive integer. Recall that $\widehat{G}$ is the set of equivalence classes of irreducible representations of G. Given $\sigma \in \widehat{G}$, we denote by W^{σ} the corresponding irreducible G-module (the space for the corresponding irreducible representation of G). Elements of $\mathcal{Y}(\widehat{G})$ are called *Young G-diagrams* and elements of $\mathcal{Y}_{\mathrm{n}}(\widehat{G})$ are called *Young G-diagrams with n boxes.* Given $\mu \in \mathcal{Y}(\widehat{G})$ and $\sigma \in \widehat{G}$, we denote by $\mu \downarrow_{\sigma}$ the set of all Young G-diagrams obtained from μ by removing one of the inner corners in the Young diagram $\mu(\sigma)$. Let $\mu \in \mathcal{Y}(\widehat{G})$. A *Young G-tableau* of shape μ is obtained by taking the Young G-diagram μ and filling its $||\mu||$ boxes (bijectively) with the numbers $1, 2, \ldots, ||\mu||$. A Young G-tableau is said to be *standard* if the numbers in the boxes strictly increase along each row and each column of all Young diagrams occurring in μ. Let $\mathrm{tab}_G(n, \mu)$, where $\mu \in \mathcal{Y}_{\mathrm{n}}(\widehat{G})$, denote the set of all standard Young G-tableaux of shape μ and let $\mathrm{tab}_G(n) = \bigcup_{\mu\in\mathcal{Y}_{\mathrm{n}}(\widehat{G})} \mathrm{tab}_G(n, \mu)$. Let $T \in \mathrm{tab}_G(n)$ and $i \in [n]$. If i appear in the Young diagram $\mu(\sigma)$, where μ is the shape of T and $\sigma \in \widehat{G}$, we write $r_T(i) = \sigma$. Recall that $b_T(i)$ denotes the box in $\mu(\sigma)$, with the number i resides and $c(b_T(i))$ denotes the content of the box $b_T(i)$.

For example let us take $n = 10$ and G to be $\mathbb{Z}_{10}$, the additive group of integers modulo 10. Also let $\widehat{\mathbb{Z}}_{10} := \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7, \sigma_8, \sigma_9, \sigma_{10}\}$ and $\mu \in \mathcal{Y}_{10}(\widehat{\mathbb{Z}}_{10})$ be such that

$$\mu(\sigma_1) = \begin{array}{|c|c|c|}\hline & & \\ \hline & \\ \cline{1-2}\end{array}\,, \quad \mu(\sigma_2) = \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array}\,, \quad \mu(\sigma_8) = \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array}\,, \quad \mu(\sigma_{10}) = \begin{array}{|c|}\hline \\ \hline\end{array}$$

and $\mu(\sigma_i) = \phi$ for all $i \in \{3, 4, 5, 6, 7, 9\}$, where ϕ denotes the empty Young diagram (i.e. Young diagram with no boxes). Then for the element T of $\mathrm{tab}_{\mathbb{Z}_{10}}(10, \mu)$ given by

$$\mu(\sigma_1) \rightsquigarrow \begin{array}{|c|c|c|}\hline 4 & 6 & 9 \\ \hline 7 & 10 \\ \cline{1-2}\end{array}\,, \quad \mu(\sigma_2) \rightsquigarrow \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array}\,, \quad \mu(\sigma_8) \rightsquigarrow \begin{array}{|c|}\hline 3 \\ \hline 8 \\ \hline\end{array}\,, \quad \mu(\sigma_{10}) \rightsquigarrow \begin{array}{|c|}\hline 5 \\ \hline\end{array}$$

and $\mu(\sigma_i) \rightsquigarrow \phi$ for $i \in \{3, 4, 5, 6, 7, 9\}$, we have the following: $r_T(1) = \sigma_2$, $r_T(2) = \sigma_2$, $r_T(3) = \sigma_8$, $r_T(4) = \sigma_1$, $r_T(5) = \sigma_{10}$, $r_T(6) = \sigma_1$, $r_T(7) = \sigma_1$, $r_T(8) = \sigma_8$, $r_T(9) = \sigma_1$, $r_T(10) = \sigma_1$ and $c(b_T(1)) = 0$, $c(b_T(2)) = -1$, $c(b_T(3)) = 0$, $c_T(4) = 0$, $c_T(5) = 0$,

$c(b_T(6)) = 1$, $c(b_T(7)) = -1$, $c(b_T(8)) = -1$, $c(b_T(9)) = 2$, $c(b_T(10)) = 0$.

Definition 5.3. Let $\mathcal{H}_{i,n}(G)$ be the subgroup

$$\{(g_1, \ldots, g_n; \pi) \in G \wr S_n : \pi(j) = j \text{ for } i+1 \leq j \leq n\}$$

of $G \wr S_n$ for $0 \leq i \leq n$. In particular $\mathcal{H}_{0,n}(G) = \mathcal{H}_{1,n}(G) = G^n$ and $\mathcal{H}_{n,n}(G) = G \wr S_n$.

Definition 5.4. The (generalized) *Young-Jucys-Murphy* elements $X_1(G), \ldots, X_n(G)$ of $\mathbb{C}[G \wr S_n]$ or $\mathcal{H}_{n,n}(G)$ are given by $X_1(G) = 0$ and

$$\begin{aligned} X_i(G) &= \sum_{k=1}^{i-1} \sum_{g \in G} (g^{-1})^{(k)} g^{(i)} (k, i) \\ &= \sum_{k=1}^{i-1} \sum_{g \in G} (g^{-1})^{(k)} (k, i) g^{(k)}, \text{ for all } 2 \leq i \leq n. \end{aligned}$$

The Young-Jucys-Murphy elements generate a maximal commuting subalgebra of $\mathbb{C}[G \wr S_n]$, known as the *Gelfand-Tsetlin subalgebra* of $\mathcal{H}_{n,n}(G)$. We now define Gelfand-Tsetlin subspaces and the Gelfand-Tsetlin decomposition.

Let $\lambda \in \widehat{\mathcal{H}_{n,n}}(G)$ and consider the irreducible $\mathcal{H}_{n,n}(G)$-module (the space for the representation of $\mathcal{H}_{n,n}(G)$) V^λ. Since the branching is simple [73, Section 3], the decomposition into irreducible $\mathcal{H}_{n-1,n}(G)$-modules is given by

$$V^\lambda = \bigoplus_\mu V^\mu,$$

where the sum is over all $\mu \in \widehat{\mathcal{H}_{n-1,n}}(G)$, with $\mu \nearrow \lambda$ (i.e., there is an edge from μ to λ in the branching multi-graph), is canonical. Iterating this decomposition of V^λ into irreducible $\mathcal{H}_{1,n}(G)$-submodules we obtain

$$V^\lambda = \bigoplus_T V_T, \tag{5.2.1}$$

where the sum is over all possible chains $T = \lambda_1 \nearrow \lambda_2 \nearrow \cdots \nearrow \lambda_n$ with $\lambda_i \in \widehat{\mathcal{H}_{i,n}}(G)$ and $\lambda_n = \lambda$. From the definition of V_T, for $0 \neq v_T \in V_T$, we have $\mathbb{C}[\mathcal{H}_{i,n}(G)]v_T = V^{\lambda_i}$. We call (5.2.1) the *Gelfand-Tsetlin* decomposition of V^λ and each V_T in (5.2.1) a *Gelfand-Tsetlin* subspace of V^λ. The Young-Jucys-Murphy elements act by scalars on the Gelfand-Tsetlin subspaces of all irreducible $G \wr S_n$-modules. From Lemma 6.2 and Theorem 6.4 of [73], we may parametrise the irreducible representations of $G \wr S_n$ by elements of $\mathcal{Y}_n(\widehat{G})$.

Theorem 5.5 ([73, Theorem 6.5]). *Let $\mu \in \mathcal{Y}_{\mathrm{n}}(\widehat{G})$. Then we may index the Gelfand-Tsetlin subspaces of V^μ by standard Young G-tableaux of shape μ and write the Gelfand-Tsetlin decomposition as*

$$V^\mu = \bigoplus_{T \in \mathrm{tab}_G(n,\mu)} V_T,$$

where each V_T is closed under the action of G^n and as a G^n-module, is isomorphic to the irreducible G^n-module

$$W^{r_T(1)} \otimes W^{r_T(2)} \otimes \cdots \otimes W^{r_T(n)}.$$

For $i = 1, \ldots, n$; the eigenvalues of $X_i(G)$ on V_T are given by $\frac{|G|}{\dim(W^{r_T(i)})} c(b_T(i))$.

Theorem 5.6 ([73, Theorem 6.7]). *Let $\mu \in \mathcal{Y}_{\mathrm{n}}(\widehat{G})$. Write $\widehat{G} := \{\sigma_1, \ldots, \sigma_t\}$ and set $\mu^{(i)} = \mu(\sigma_i)$, $m_i = |\mu^{(i)}|, d_i = \dim(W^{\sigma_i})$ for each $1 \le i \le t$. Then*

$$\dim(V^\mu) = \binom{n}{m_1, \ldots, m_t} f^{\mu^{(1)}} \cdots f^{\mu^{(t)}} d_1^{m_1} \cdots d_t^{m_t}.$$

Recall that $f^{\mu^{(i)}}$ denotes the number of standard Young tableau of shape $\mu^{(i)}$, for each $1 \le i \le t$.

Lemma 5.7. *Let G be a finite group and $\sigma \in \widehat{G}$. Recall that $\mathbb{1}$ is the trivial representation of G. If W^σ (respectively χ^σ) denotes the irreducible G-module (respectively character) with dimension d_ρ and I_{d_σ} denotes the identity matrix of order $d_\sigma \times d_\sigma$, then the action of the group algebra element $\sum_{g \in G} g$ on W^σ is given by the following scalar matrix*

$$\sum_{g \in G} g = \frac{|G|}{d_\sigma} \langle \chi^\sigma, \chi^{\mathbb{1}} \rangle I_{d_\sigma}.$$

Proof. It is clear that $\sum_{g \in G} g$ is in the center of $\mathbb{C}[G]$. Therefore by Lemma 2.4 (Schur's lemma), we have $\sum_{g \in G} g = cI_{d_\sigma}$ for some $c \in \mathbb{C}$. The value of c can be obtained by equating the traces of $\sum_{g \in G} g$ and cI_{d_σ}. □

Remark 5.8. Our focus will be on $\mathcal{H}_{n,n}(G_n)$ i.e. $G_n \wr S_n$ for the sequence of subgroups

$$\mathcal{H}_{1,n}(G_n) \subseteq \cdots \subseteq \mathcal{H}_{i,n}(G_n) \subseteq \cdots \subseteq \mathcal{H}_{n,n}(G_n).$$

For simplicity we write the Young-Jucys-Murphy elements $X_i(G_n)$ of $G_n \wr S_n$ (i.e. $\mathcal{G}_{\mathrm{n}}$) as $X_i^{\mathcal{G}}$ for $1 \le i \le n$. Thus Theorems 5.5 and 5.6 are applicable to $\mathcal{G}_{\mathrm{n}}$.

Let $t := |\widehat{G}_n|$ and $\widehat{G}_n := \{\sigma_1, \ldots, \sigma_t\}$, where $\sigma_1 = \mathbb{1}$ (the trivial representation of G_n). We write $\mu \left(\in \mathcal{Y}_n(\widehat{G}_n)\right)$ as the tuple $(\mu^{(1)}, \ldots, \mu^{(t)})$, where $\mu^{(i)} := \mu(\sigma_i)$ for each $1 \leq i \leq t$. We also denote $m_i := |\mu^{(i)}|$, $W^{\sigma_i} :=$ the irreducible G_n-module corresponding to σ_i and $d_i = \dim(W^{\sigma_i})$ for each $1 \leq i \leq t$. Thus t, σ_i, $\mu^{(i)}$, m_i, W^{σ_i} and d_i depend on G_n i.e., on n. To avoid notational complication, the dependence of t, σ_i, $\mu^{(i)}$, m_i, W^{σ_i} and d_i on n is suppressed. We note that for $T \in \operatorname{tab}_{G_n}(n, \mu)$ the dimension of V_T is $d_1^{m_1} \cdots d_t^{m_t}$.

Theorem 5.9. *For each $\mu \in \mathcal{Y}_n(\widehat{G}_n)$, let $\widehat{P_{\mathcal{G}}}(R)\Big|_{V^\mu}$ denote the restriction of $\widehat{P_{\mathcal{G}}}(R)$ to the irreducible $\mathcal{G}_{\mathrm{n}}$-module V^μ. Then the eigenvalues of $\widehat{P_{\mathcal{G}}}(R)\Big|_{V^\mu}$ are given by,*

$$\frac{1}{n \dim(W^{r_T(n)})} \left(c(b_T(n)) + \langle \chi^{r_T(n)}, \chi^{\mathbb{1}} \rangle \right), \text{ with multiplicity } \dim(V_T) = d_1^{m_1} \cdots d_t^{m_t}$$

for each $T \in \operatorname{tab}_{G_n}(n, \mu)$.

Proof. We first find the eigenvalues of $X_n^{\mathcal{G}} + \sum\limits_{g \in G_n} (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})$. Let $I_{\dim(V_T)}$ denote the identity matrix of order $\dim(V_T) \times \dim(V_T)$. Then from Theorem 5.5 we have

$$V^\mu = \bigoplus_{T \in \operatorname{tab}_{G_n}(n,\mu)} V_T \quad \text{and} \quad X_n^{\mathcal{G}}\Big|_{V_T} = \frac{|G_n|}{\dim(W^{r_T(n)})} c(b_T(n)) I_{\dim(V_T)}. \tag{5.2.2}$$

Again from Theorem 5.5 and Lemma 5.7 we have

$$\sum_{g \in G_n} (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})\Big|_{V_T} = \frac{|G_n|}{\dim(W^{r_T(n)})} \langle \chi^{r_T(n)}, \chi^{\mathbb{1}} \rangle I_{\dim(V_T)}. \tag{5.2.3}$$

Let us recall that,

$$\widehat{P_{\mathcal{G}}}(R) = \frac{1}{n|G_n|} \sum_{g \in G_n} \left(R\left((\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})\right) + \sum_{i=1}^{n-1} R\left((\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n))\right) \right).$$

Therefore $n|G_n|\widehat{P_{\mathcal{G}}}(R)$ is nothing but the action of $X_n^{\mathcal{G}} + \sum_{g \in G_n} (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})$ on $\mathbb{C}[\mathcal{G}_{\mathrm{n}}]$ by multiplication on the right. Since $\dim(V_T) = d_1^{m_1} \cdots d_t^{m_t}$, the theorem follows from (5.2.2) and (5.2.3). □

Remark 5.10. In the regular representation of a finite group, each irreducible representation occurs with multiplicity equal to its dimension. Therefore, Theorems 5.6 and 5.9 provide the eigenvalues of $\widehat{P_{\mathcal{G}}}(R)$.

5.3 Upper bound for $||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}$

We find the order of the mixing time for the warp-transpose top with random shuffle in this section. In this section, we also find an upper bound of $||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}$ when $k \geq n\log n + Cn\log(|G_n|-1)$ for $C > 1$. Before proving the main results of this section, first we set some notations and prove a useful lemma. Given a partition ξ of the integer ℓ (here we are allowing ℓ to take value 0), throughout this section ξ_1 denotes the largest part of ξ. In particular if $\xi \vdash 0$ then $f^{\xi} = 1$ (as there is a unique Young diagram with zero boxes) and we set $\xi_1 = 0$.

Definition 5.11. Let A be a non empty set. Then the *indicator function* of A is denoted by $\mathfrak{Ind}_A$ and is defined by

$$\mathfrak{Ind}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

Lemma 5.12. *Let $\mu = (\mu^{(1)}, \dots, \mu^{(t)}) \in \mathcal{Y}_n(\widehat{G}_n)$. Recall that $\mu_1^{(j)}$ (respectively $\mu_1^{(j)'}$) denotes the largest part of $\mu^{(j)}$ (respectively its conjugate $\mu^{(j)'}$) for $1 \leq j \leq t$. Then we have*

$$\sum_{T \in \mathrm{tab}_{G_n}(n,\mu)} \left(\frac{c(b_T(n)) + \langle \chi^{r_T(n)}, \chi^{\mathbb{1}} \rangle}{n \dim(W^{r_T(n)})} \right)^{2k} < \binom{n}{m_1, \dots, m_t} f^{\mu^{(1)}} \cdots f^{\mu^{(t)}} \sum_{j=1}^{t} \left(\mathcal{M}_j^{2k} + \mathcal{M}_j'^{2k} \right) \mathfrak{Ind}_{\{m_j > 0\}},$$

where $\mathcal{M}_j := \frac{\mu_1^{(j)} - 1 + \langle \chi^{\sigma_j}, \chi^{\mathbb{1}} \rangle}{nd_j}$ and $\mathcal{M}_j' := \frac{\mu_1^{(j)'} - 1 + \langle \chi^{\sigma_j}, \chi^{\mathbb{1}} \rangle}{nd_j}$ for each $1 \leq j \leq t$.

Proof. Let $\mathcal{T}_i = \{(T_1, \dots, T_t) \in \mathrm{tab}_{G_n}(n,\mu) \mid b_T(n) \text{ is in } T_i\}$ for each $1 \leq i \leq t$. Then $\mathrm{tab}_{G_n}(n,\mu)$ is the disjoint union of the sets $\mathcal{T}_1, \dots, \mathcal{T}_t$. Therefore we have

$$\sum_{T \in \mathrm{tab}_{G_n}(n,\mu)} \left(\frac{c(b_T(n)) + \langle \chi^{r_T(n)}, \chi^{\mathbb{1}} \rangle}{n \dim(W^{r_T(n)})} \right)^{2k} = \sum_{i=1}^{t} \sum_{T \in \mathcal{T}_i} \left(\frac{c(b_T(n)) + \langle \chi^{\sigma_i}, \chi^{\mathbb{1}} \rangle}{nd_i} \right)^{2k} \mathfrak{Ind}_{\{m_i > 0\}}$$

and this is equal to,

$$\begin{aligned} &\sum_{i=1}^{t} \binom{n-1}{m_1, .., m_i - 1, .., m_t} \frac{f^{\mu^{(1)}} \cdots f^{\mu^{(t)}}}{f^{\mu^{(i)}}} \sum_{T_i \in \mathrm{tab}(\mu^{(i)})} \left(\frac{c(b_T(m_i)) + \langle \chi^{\sigma_i}, \chi^{\mathbb{1}} \rangle}{nd_i} \right)^{2k} \mathfrak{Ind}_{\{m_i > 0\}} \\ &< \sum_{i=1}^{t} \binom{n}{m_1, \dots, m_t} \frac{f^{\mu^{(1)}} \cdots f^{\mu^{(t)}}}{f^{\mu^{(i)}}} \sum_{T_i \in \mathrm{tab}(\mu^{(i)})} \left(\mathcal{M}_i^{2k} + \mathcal{M}_i'^{2k} \right) \mathfrak{Ind}_{\{m_i > 0\}}. \end{aligned} \tag{5.3.1}$$

The inequality in (5.3.1) holds because $T_i \in \mathrm{tab}(\mu^{(i)})$ implies the following:

$$\begin{aligned}
&\left(\frac{c(b_T(m_i)) + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}\\
&\leq \max\left\{\left(\frac{\mu_1^{(i)} - 1 + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}, \left(\frac{\mu_1^{(i)\prime} - 1 - \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}\right\}\\
&\leq \max\left\{\left(\frac{\mu_1^{(i)} - 1 + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}, \left(\frac{\mu_1^{(i)\prime} - 1 + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}\right\}, \qquad \text{as } \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle = 0 \text{ or } 1\\
&< \left(\frac{\mu_1^{(i)} - 1 + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k} + \left(\frac{\mu_1^{(i)\prime} - 1 + \langle\chi^{\sigma_i}, \chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k} = \mathcal{M}_i^{2k} + \mathcal{M}_i^{\prime 2k}.
\end{aligned}$$

Therefore the result follows from (5.3.1) and

$$\sum_{T_i \in \mathrm{tab}(\mu^{(i)})} \left(\mathcal{M}_i^{2k} + \mathcal{M}_i^{\prime 2k}\right) = f^{\mu^{(i)}}\left(\mathcal{M}_i^{2k} + \mathcal{M}_i^{\prime 2k}\right). \qquad \square$$

Proposition 5.13. *For the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$, we have*

$$\begin{aligned}
4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}^2 \;<\; & 2\left(e^{n^2 e^{-\frac{2k}{n}}} - 1\right) + e^{-\frac{4k}{n}} + 2e\left(e^{n^2(|G_n|-1)e^{-\frac{2k}{n}}} - 1\right)\\
& + 2(|\widehat{G}_n| - 1)\left(e^{-\frac{2k}{n}} e^{n^2 e^{-\frac{2k}{n}}} + \frac{e}{n^2}\left(e^{n^2(|G_n|-1)e^{-\frac{2k}{n}}} - 1\right)\right),
\end{aligned}$$

for all $k \geq \max\{n,\ n\log n\}$.

Proof. Let us recall that $\widehat{G}_n = \{\sigma_1, \ldots, \sigma_t\}$ and $\sigma_1 = \mathbb{1}$, the trivial representation of G_n. Given $\mu \in \mathcal{Y}_n(\widehat{G}_n)$, throughout this proof we write $\mu = (\mu^{(1)}, \ldots, \mu^{(t)})$, where $\mu^{(i)} = \mu(\sigma_i),\ \mu^{(i)} \vdash m_i,\ \sum_{i=1}^{t} m_i = n$. Now using Lemma 2.24, we have

$$4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}^2 \leq \sum_{\mu \in \mathcal{Y}_n(\widehat{G}_n):\ \mu(\mathbb{1}) \neq (n)} \dim(V^{\mu})\,\mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(R)\Big|_{V^{\mu}}\right)^{2k}\right). \tag{5.3.2}$$

First we partition the set $\mathcal{Y}_n(\widehat{G}_n)$ into two disjoint subsets $\mathcal{A}_1$ and $\mathcal{A}_2$ as follows:

$$\mathcal{A}_1 = \bigcup_{1 \leq i \leq t} \mathcal{B}_i, \text{ where } \mathcal{B}_i = \{\mu \in \mathcal{Y}_n(\widehat{G}_n) \mid m_i = n,\ m_k = 0 \text{ for all } k \in \{1, \ldots, t\} \setminus \{i\}\}$$

$$\mathcal{A}_2 = \{\mu \in \mathcal{Y}_n(\widehat{G}_n) \mid \sum_{k=1}^{t} m_k = n,\ 0 \leq m_k \leq n-1\}.$$

It can be easily seen that $\mathcal{B}_i$'s are disjoint. Therefore by using Theorem 5.9 and Remark

5.10, the inequality (5.3.2) become

$$\begin{aligned}
4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_n}||_{\mathrm{TV}}^2 \leq & \sum_{\substack{\mu\in\mathcal{B}_1\\ \mu(\mathbb{1})\neq(n)}} \dim(V^{\mu}) \sum_{T\in\mathrm{tab}_{G_n}(n,\mu)} \left(\frac{c(b_T(n))+1}{nd_1}\right)^{2k} d_1^n \\
& + \sum_{i=2}^{t}\sum_{\mu\in\mathcal{B}_i} \dim(V^{\mu}) \sum_{T\in\mathrm{tab}_{G_n}(n,\mu)} \left(\frac{c(b_T(n))}{nd_i}\right)^{2k} d_i^n \\
& + \sum_{\mu\in\mathcal{A}_2} \dim(V^{\mu}) \sum_{T\in\mathrm{tab}_{G_n}(n,\mu)} \left(\frac{c(b_T(n)) + \langle \chi^{r_T(n)}, \chi^{\mathbb{1}}\rangle}{n\dim(W^{r_T(n)})}\right)^{2k} d_1^{m_1}\cdots d_t^{m_t}.
\end{aligned} \tag{5.3.3}$$

The first two terms in the right hand side of (5.3.3) are equal to

$$\begin{aligned}
& \sum_{\substack{\lambda\vdash n\\ \lambda_1\neq n}} f^{\lambda} d_1^n \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))+1}{nd_1}\right)^{2k} d_1^n + \sum_{i=2}^{t}\sum_{\lambda\vdash n} f^{\lambda} d_i^n \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))}{nd_i}\right)^{2k} d_i^n \\
= & \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}} f^{\lambda} \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \left(\frac{n-2}{n}\right)^{2k} + \sum_{i=2}^{t} \frac{d_i^{2n}}{d_i^{2k}} \sum_{\lambda\vdash n} f^{\lambda} \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))}{n}\right)^{2k}.
\end{aligned} \tag{5.3.4}$$

Now recalling λ_1 (respectively λ_1') is the largest part of λ (respectively its conjugate), we have the following:

$$\begin{aligned}
\left(\frac{c(b_T(n))+x}{n}\right)^{2k} & \leq \max\left\{\left(\frac{\lambda_1-1+x}{n}\right)^{2k}, \left(\frac{\lambda_1'-1+x}{n}\right)^{2k}\right\}, \\
& < \left(\frac{\lambda_1-1+x}{n}\right)^{2k} + \left(\frac{\lambda_1'-1+x}{n}\right)^{2k}, \quad \text{for } T\in\mathrm{tab}(\lambda) \text{ and } x\geq 0.
\end{aligned}$$

This implies

$$\begin{aligned}
\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}} f^{\lambda} \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))+1}{n}\right)^{2k} & < \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}} \left(f^{\lambda}\right)^2 \left(\left(\frac{\lambda_1}{n}\right)^{2k} + \left(\frac{\lambda_1'}{n}\right)^{2k}\right) \\
& < 2 \sum_{\substack{\lambda\vdash n\\ \lambda\neq(n),(1^n)}} \left(f^{\lambda}\right)^2 \left(\frac{\lambda_1}{n}\right)^{2k} \quad \text{and} \\
\sum_{\lambda\vdash n} f^{\lambda} \sum_{T\in\mathrm{tab}(\lambda)} \left(\frac{c(b_T(n))}{n}\right)^{2k} & < \sum_{\lambda\vdash n} \left(f^{\lambda}\right)^2 \left(\left(\frac{\lambda_1-1}{n}\right)^{2k} + \left(\frac{\lambda_1'-1}{n}\right)^{2k}\right) \\
& < 2\sum_{\lambda\vdash n} \left(f^{\lambda}\right)^2 \left(\frac{\lambda_1-1}{n}\right)^{2k}.
\end{aligned}$$

Thus using $1-x\leq e^x$ for $x\geq 0$, $k\geq n$, and $d_i\geq 1$ for all $1\leq i\leq t$, the expression in (5.3.4) is bounded above by

$$\begin{aligned}
&2\sum_{\substack{\lambda\vdash n\\ \lambda\neq(n)}}(f^{\lambda})^2\left(\frac{\lambda_1}{n}\right)^{2k}+\left(1-\frac{2}{n}\right)^{2k}+2\sum_{i=2}^{t}\sum_{\lambda\vdash n}(f^{\lambda})^2\left(\frac{\lambda_1-1}{n}\right)^{2k}\\
&<2\left(e^{n^2e^{-\frac{2k}{n}}}-1\right)+e^{-\frac{4k}{n}}+2(t-1)e^{-\frac{2k}{n}}e^{n^2e^{-\frac{2k}{n}}}. \qquad (5.3.5)
\end{aligned}$$

The inequality in (5.3.5) follows from Corollary 2.11 (by taking $\ell=n, a=b=0$) and Lemma 2.10 (by taking $\ell=n, a=1, b=0$). Now recalling $\mathcal{M}_j:=\frac{\mu_1^{(j)}-1+\langle\chi^{\sigma_j},\chi^{\mathbb{1}}\rangle}{nd_j}$, $\mathcal{M}'_j:=\frac{\mu_1^{(j)'}-1+\langle\chi^{\sigma_j},\chi^{\mathbb{1}}\rangle}{nd_j}$, and using Lemma 5.12, the third term in the right hand side of (5.3.3) is less than

$$\sum_{\mu\in\mathcal{A}_2}\binom{n}{m_1,\dots,m_t}^2(f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2d_1^{2m_1}\dots d_t^{2m_t}\sum_{j=1}^{t}\left(\mathcal{M}_j^{2k}+\mathcal{M}_j'^{2k}\right)\mathfrak{Ind}_{\{m_j>0\}}. \quad (5.3.6)$$

We deal with (5.3.6) by considering two separate cases namely $j=1$ and $1<j\leq t$. Now using

$$\sum_{\mu^{(1)}\vdash m_1}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)'}}{nd_1}\right)^{2k}=\sum_{\mu^{(1)}\vdash m_1}\left(f^{\mu^{(1)}}\right)^2\left(\frac{\mu_1^{(1)}}{nd_1}\right)^{2k},$$

the partial sum corresponding to $j=1$ in (5.3.6) is equal to,

$$\begin{aligned}
&\sum_{m_1=1}^{n-1}\sum_{\substack{(m_2,\dots,m_t)\\ \sum m_k=n-m_1\\ 0\leq m_k\leq n-1}}2\sum_{\substack{\mu^{(i)}\vdash m_i\\ 1\leq i\leq t}}\binom{n}{m_1}^2\binom{n-m_1}{m_2,\dots,m_t}^2(f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2d_1^{2m_1}\dots d_t^{2m_t}\left(\frac{\mu_1^{(1)}}{nd_1}\right)^{2k}\\
=&2\sum_{m_1=1}^{n-1}(d_2^2+\cdots+d_t^2)^{n-m_1}\binom{n}{m_1}^2(n-m_1)!\left(\frac{1}{d_1}\right)^{2k-2m_1}\left(\frac{m_1}{n}\right)^{2k}\sum_{\mu^{(1)}\vdash m_1}(f^{\mu^{(1)}})^2\left(\frac{\mu_1^{(1)}}{m_1}\right)^{2k}\\
<&2\sum_{m_1=1}^{n-1}(d_2^2+\cdots+d_t^2)^{n-m_1}\binom{n}{m_1}^2(n-m_1)!\left(\frac{1}{d_1}\right)^{2k-2m_1}\left(\frac{m_1}{n}\right)^{2k}e^{m_1^2e^{-\frac{2k}{m_1}}}. \qquad (5.3.7)
\end{aligned}$$

The inequality in (5.3.7) follows from Lemma 2.10 (by taking $\ell=m_1, a=b=0$). As $k\geq n\log n$, we have $k\geq m_1\log m_1$. Thus writing $n-m_1$ by u, the expression in (5.3.7) is less than or equal to

$$2e\sum_{u=1}^{n-1}\left(\frac{d_2^2+\cdots+d_t^2}{d_1^2}\right)^u\left(\frac{1}{d_1}\right)^{2k-2n}\binom{n}{u}^2u!\left(1-\frac{u}{n}\right)^{2k} \qquad (5.3.8)$$

Now using $1-x\leq e^{-x}$ for all $x\geq 0$ and $d_1=1$ the expression in (5.3.8) is less than or equal to

$$2e\sum_{u=1}^{n-1}\frac{1}{u!}\left(n^2\left(\frac{|G_n|}{d_1^2}-1\right)e^{-\frac{2k}{n}}\right)^u < 2e\left(e^{\left(n^2\left(\frac{|G_n|}{d_1^2}-1\right)e^{-\frac{2k}{n}}\right)}-1\right). \tag{5.3.9}$$

Now using the notation $m_1,\ldots,\widehat{m_j},\ldots,m_t$ to denote $m_1,\ldots,m_{j-1},m_{j+1},\ldots,m_t$, and

$$\sum_{\mu^{(j)}\vdash m_j}\left(f^{\mu^{(j)}}\right)^2\left(\frac{\mu_1^{(j)'}}{nd_j}\right)^{2k} = \sum_{\mu^{(j)}\vdash m_j}\left(f^{\mu^{(j)}}\right)^2\left(\frac{\mu_1^{(j)}}{nd_j}\right)^{2k},$$

the partial sum corresponding to $1<j\leq t$ in (5.3.6) turns out to be

$$\sum_{m_j=1}^{n-1}\sum_{\substack{(m_1,\ldots,\widehat{m_j},\ldots,m_t)\\ \sum m_k=n-m_j\\ 0\leq m_k\leq n-1}} 2\sum_{\substack{\mu^{(i)}\vdash m_i\\ 1\leq i\leq t}}\binom{n}{m_j}^2\binom{n-m_j}{m_1,\ldots,\widehat{m_j},\ldots,m_t}^2(f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2d_1^{2m_1}\ldots d_t^{2m_t}\zeta^{2k}, \tag{5.3.10}$$

where $\zeta=\frac{\mu_1^{(j)}-1}{nd_j}$. The expression given in (5.3.10) is equal to the following

$$\begin{aligned}&2\sum_{m_j=1}^{n-1}d_j^{2m_j}(d_1^2+\cdots+d_t^2-d_j^2)^{n-m_j}\binom{n}{m_j}^2(n-m_j)!\left(\frac{m_j}{nd_j}\right)^{2k}\sum_{\mu^{(j)}\vdash m_j}(f^{\mu^{(j)}})^2\left(\frac{\mu_1^{(j)}-1}{m_j}\right)^{2k}\\ &< 2\sum_{m_j=1}^{n-1}(d_1^2+\cdots+d_t^2-d_j^2)^{n-m_j}\binom{n}{m_j}^2(n-m_j)!\left(\frac{1}{d_j}\right)^{2k-2m_j}\left(\frac{m_j}{n}\right)^{2k}e^{-\frac{2k}{m_j}}e^{m_j^2e^{-\frac{2k}{m_j}}}.\end{aligned} \tag{5.3.11}$$

The inequality in (5.3.11) follows from Lemma 2.10 (by taking $\ell=m_j, a=1, b=0$). As $k\geq n\log n$, we have $k\geq m_j\log m_j,\ m_j\log n$. Thus writing $n-m_j$ by v, the expression in (5.3.11) is less than or equal to

$$\frac{2e}{n^2}\sum_{v=1}^{n-1}\left(\frac{d_1^2+\cdots+d_t^2-d_j^2}{d_j^2}\right)^v\left(\frac{1}{d_j}\right)^{2k-2n}\binom{n}{v}^2v!\left(1-\frac{v}{n}\right)^{2k} \tag{5.3.12}$$

Thus using $1-x\leq e^{-x}$ for all $x\geq 0$ and $d_j^{2k-2n}\geq 1$ for all $j\in\{1,\ldots,t\}$, the expression in (5.3.12) is less than or equal to

$$\frac{2e}{n^2}\sum_{v=1}^{n-1}\frac{1}{v!}\left(n^2\left(\frac{|G_n|}{d_j^2}-1\right)e^{-\frac{2k}{n}}\right)^v < \frac{2e}{n^2}\left(e^{\left(n^2\left(\frac{|G_n|}{d_j^2}-1\right)e^{-\frac{2k}{n}}\right)}-1\right). \tag{5.3.13}$$

Therefore the proposition follows from (5.3.3), (5.3.5), (5.3.9), (5.3.13) and $\frac{1}{d_j}\leq 1$ for all $j \in \{1, \dots, t\}$. □

Theorem 5.14. *For the random walk on $\mathcal{G}_{\rm n}$ driven by $P_{\mathcal{G}}$, we have the following:*

1. *Let $C > 1$. Then for $k \geq n \log n + Cn \log(|G_n| - 1)$, we have*

$$||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\rm n}}||_{\rm TV} < \sqrt{1+2e}\, 2^{-C} + o(1).$$

2. *Let $\epsilon \in (0,1)$. If $|G_n| = o(n^\delta)$ for all $\delta > 0$, then $k_n = \lfloor (1+\epsilon) n \log n \rfloor$ implies*

$$\lim_{n\to\infty} ||P_{\mathcal{G}}^{*k_n} - U_{\mathcal{G}_{\rm n}}||_{TV} = 0.$$

Proof. Using $k \geq n \log n + Cn \log(|G_n| - 1)$ and Proposition 5.13 we have the following:

$$\begin{aligned} 4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\rm n}}||_{\rm TV}^2 &< 2\left(e^{\left(\frac{1}{|G_n|-1}\right)^{2C}} - 1\right) + \frac{1}{n^4(|G_n|-1)^{4C}} + 2e\left(e^{\left(\frac{1}{|G_n|-1}\right)^{2C-1}} - 1\right) \\ &\quad + 2(|\widehat{G}_n| - 1)\left(\frac{1}{n^2(|G_n|-1)^{2C}} e^{\left(\frac{1}{|G_n|-1}\right)^{2C}} + \frac{e}{n^2}\left(e^{\left(\frac{1}{|G_n|-1}\right)^{2C-1}} - 1\right)\right) \\ &< \frac{4}{(|G_n|-1)^{2C}} + \frac{1}{n^4(|G_n|-1)^{4C}} + \frac{4e}{(|G_n|-1)^{2C-1}} \\ &\quad + \frac{2e^{\left(\frac{1}{|G_n|-1}\right)^{2C}}}{n^2(|G_n|-1)^{2C-1}} + \frac{4e}{n^2(|G_n|-1)^{2C-2}}. \end{aligned} \tag{5.3.14}$$

The inequality in (5.3.14) follows from the fact $e^x - 1 < 2x$ for all $x \in (0, \frac{1}{2}] \subset (0, \log 2)$ and $|\widehat{G}_n| \leq |G_n|$. Also since $|G_n| - 1 \geq |G_1| - 1 \geq 2$ and $C > 1$ implies

$$\frac{1}{|G_n|-1}, \frac{1}{(|G_n|-1)^{2C-1}}, \frac{1}{(|G_n|-1)^{2C}} \leq \frac{1}{2}.$$

Thus (5.3.14) become

$$4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\rm n}}||_{\rm TV}^2 < \frac{4}{2^{2C}} + \frac{4e}{2^{2C-1}} + \frac{1}{n^4 2^{4C}} + \frac{2e^{\frac{1}{2}}}{n^2 2^{2C-1}} + \frac{4e}{n^2 2^{2C-2}}$$

$$\implies ||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\rm n}}||_{\rm TV} < \frac{\sqrt{1+2e}}{2^C} + \frac{1}{n\, 2^C}\sqrt{\frac{1}{n^2 2^{2C+2}} + e^{\frac{1}{2}} + 4e}\,.$$

Thus the first part of the theorem follows.

Again $k_n = \lfloor (1+\epsilon) n \log n \rfloor$ implies $k_n + 1 > (1+\epsilon) n \log n$ and thus $e^{-\frac{2k_n}{n}} < e^{\frac{2}{n}} n^{-2(1+\epsilon)}$.

Therefore using Proposition 5.13 we have

$$
\begin{aligned}
0 &\leq 4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}^2 \\
&< 2\left(e^{e^{\frac{2}{n}}n^{-2\epsilon}} - 1\right) + e^{\frac{4}{n}}n^{-4(1+\epsilon)} + 2e\left(e^{e^{\frac{2}{n}}(|G_n|-1)n^{-2\epsilon}} - 1\right) \\
&\qquad + 2(|\widehat{G}_n| - 1)\left(\frac{e^{\frac{2}{n}}}{n^{2(1+\epsilon)}}\, e^{e^{\frac{2}{n}}n^{-2\epsilon}} + \frac{e}{n^2}\left(e^{e^{\frac{2}{n}}(|G_n|-1)n^{-2\epsilon}} - 1\right)\right)
\end{aligned}
\tag{5.3.15}
$$

Therefore the second part follows form the fact that the right hand side of the inequality (5.3.15) converges to zero as $n \to \infty$ and $|\widehat{G}_n| \leq |G_n| = o(n^\delta)$ for all $\delta > 0$. $\square$

Remark 5.15. The mixing window should be negligible for the existence of cutoff. Thus the first part of Theorem 5.14 suggests the following condition for the existence of cutoff.

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{n\log(|G_n|-1)}{n\log n} = 0 \\
\iff &|G_n| = o(n^\delta) \text{ for all } \delta > 0.
\end{aligned}
$$

Theorem 5.16. *The mixing time for the warp-transpose top with random shuffle on* $\mathcal{G}_{\mathrm{n}}$ *is* $O\left(n\log n + \frac{1}{2}n\log(|G_n|-1)\right)$.

Proof. Let $a > \frac{1}{2}$ and $k = n\log n + \frac{1}{2}n\log(|G_n|-1) + an\log(|\widehat{G}_n|-1)$. Then from Proposition 5.13 we have the following:

$$
\begin{aligned}
4\,||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}^2 <& 2\left(e^{\frac{1}{(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}}} - 1\right) + \frac{1}{n^4\,(|G_n|-1)^2\left(|\widehat{G}_n|-1\right)^{4a}} \\
&+ 2e\left(e^{\frac{1}{\left(|\widehat{G}_n|-1\right)^{2a}}} - 1\right) \\
&+ 2(|\widehat{G}_n|-1)\left(\frac{1}{n^2(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}} e^{\frac{1}{(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}}} + \frac{e}{n^2}\left(e^{\frac{1}{\left(|\widehat{G}_n|-1\right)^{2a}}} - 1\right)\right)
\end{aligned}
\tag{5.3.16}
$$

Now using $|G_n| \geq |G_1| > 2$ and $|\widehat{G}_n| > 2$ we have

$$
0 < \frac{1}{(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}} < \frac{1}{\left(|\widehat{G}_n|-1\right)^{2a}} \leq \frac{1}{2^{2a}} < \frac{1}{2} < \log 2, \text{ for all } n \geq 1. \tag{5.3.17}
$$

Therefore using $e^x - 1 < 2x$ for $0 < x < \log 2$ and (5.3.17), the right hand side of (5.3.16) is bounded above by

$$\frac{4}{(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}}+\frac{1}{n^4(|G_n|-1)^2\left(|\widehat{G}_n|-1\right)^{4a}}+\frac{4e}{\left(|\widehat{G}_n|-1\right)^{2a}}$$
$$+\frac{2(|\widehat{G}_n|-1)}{n^2(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}}e^{\frac{1}{(|G_n|-1)\left(|\widehat{G}_n|-1\right)^{2a}}}+\frac{4e(|\widehat{G}_n|-1)}{n^2\left(|\widehat{G}_n|-1\right)^{2a}}$$

$$<\frac{4}{2^{2a}}+\frac{1}{n^4 2^{4a}}+\frac{4e}{2^{2a}}+\frac{4}{n^2 2^{2a}}+\frac{4e}{n^2\left(|\widehat{G}_n|-1\right)^{2a-1}}. \tag{5.3.18}$$

Finally using $a > \frac{1}{2}$, $|\widehat{G}_n| > 2$, (5.3.16), and (5.3.18) we have

$$||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\text{n}}}||_{\text{TV}} < \sqrt{1+e}\; 2^{-a} + o(1), \quad \text{for } n \geq 1. \tag{5.3.19}$$

Thus (5.3.19) implies that the mixing time for the warp-transpose top with random shuffle on $\mathcal{G}_{\text{n}}$ is $O\left(n\log n + \frac{1}{2}n\log(|G_n|-1)\right)$. □

Remark 5.17. By taking $G_n = S_2$ for all $n \geq 1$, the warp-transpose top with random shuffle on $\mathcal{G}_{\text{n}}$ boils down to the flip-transpose top with random shuffle on B_n. As $|S_2| \not> 2$, Theorem 5.16 and the first part of Theorem 5.14 are not applicable to the case of flip-transpose top with random shuffle. Although the second part of Theorem 5.14 implies the mixing time for the flip-transpose top with random shuffle is $O(n\log n)$, it fails to give the correct mixing window.

5.4 Lower bound for $||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\text{n}}}||_{\text{TV}}$

In this section, we will find a lower bound of the total variation distance $||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\text{n}}}||_{\text{TV}}$ when $k = n\log n + cn$ for $c \ll 0$. To establish the theorem giving this lower bound we use tools from representation theory of finite groups. We also use the action of nth Young-Jucys-Murphy element of the symmetric group S_n on irreducible S_n-modules. Let us recall that $[n] := \{1,\ldots,n\}$ and $[n-1] := \{1,\ldots,n-1\}$. Now we define an auxiliary representation $\mathcal{R}$ of $\mathcal{G}_{\text{n}}$ and a random variable X on $\mathcal{G}_{\text{n}}$.

Let $V = \mathbb{C}[G_n \times [n]]$ be the complex vector space of all formal linear combinations of elements of $G_n \times [n]$ and $GL(V)$ be the set of all invertible linear maps from V to itself. We now define the representation $\mathcal{R} : \mathcal{G}_{\text{n}} \longrightarrow GL(V)$ on the basis elements of V by

$$\mathcal{R}(g_1,\ldots,g_n;\pi)\left((h,i)\right) = \left(g_{\pi(i)}h, \pi(i)\right).$$

The random variable X counts the number of fixed points of the action of $\mathcal{R}$ i.e. X is the character $\chi^{\mathcal{R}}$ of $\mathcal{R}$. Let $E_k(X)$ be the expectation and $\mathrm{Var}_k(X)$ be the variance of X with respect to the probability measure $P_{\mathcal{G}}^{*k}$ on $\mathcal{G}_{\mathrm{n}}$. Also $E_U(X)$ denotes the expectation of X with respect to the uniform distribution on $\mathcal{G}_{\mathrm{n}}$. Our goal is to compute $E_k(X)$, $\mathrm{Var}_k(X)$, $E_U(X)$ and use them.

Proposition 5.18. *The expectation $E_U(X)$ of X with respect to $U_{\mathcal{G}_{\mathrm{n}}}$ is 1.*

Proof. For notational simplicity let us denote $A = G_n \times [n]$. Also for $a \in A$, $\mathrm{Fix}_{\mathcal{G}_{\mathrm{n}}}(a)$ denotes the set consisting of elements $\mathcal{G}_{\mathrm{n}}$, which fixes a. Then $|\mathrm{Fix}_{\mathcal{G}_{\mathrm{n}}}(a)| = |G_n|^{n-1}(n-1)!$ for all $a \in A$. Because if $a = (x, i)$ then a is fixed by $|G_n|^{n-1}(n-1)!$ elements of $\mathcal{G}_{\mathrm{n}}$ of the form:

$$(*, \dots, *, \underset{\substack{\uparrow \\ i\text{th position.}}}{\mathrm{e}}, *, \dots, *; \tilde{\pi}) \in \mathcal{G}_{\mathrm{n}} \text{ such that } \tilde{\pi}(i) = i.$$

Here $*$ can be chosen independently from G_n and for each of these choice $\tilde{\pi}$ can be chosen from S_n such that $\tilde{\pi}(i) = i$. Therefore using the definition of expectation we have

$$E_U(X) = \sum_{g \in \mathcal{G}_{\mathrm{n}}} U_{\mathcal{G}_{\mathrm{n}}}(g) X(g) = \frac{1}{|\,\mathcal{G}_{\mathrm{n}}\,|} \sum_{g \in \mathcal{G}_{\mathrm{n}}} \chi^{\mathcal{R}}(g) = \frac{1}{|G_n|^n n!} \sum_{g \in \mathcal{G}_{\mathrm{n}}} \mathrm{Tr}\left(\mathcal{R}(g)\right).$$

Now the lemma follows from the following fact

$$\begin{aligned} \sum_{g \in \mathcal{G}_{\mathrm{n}}} \mathrm{Tr}\left(\mathcal{R}(g)\right) = \mathrm{Tr}\left(\sum_{g \in \mathcal{G}_{\mathrm{n}}} \mathcal{R}(g)\right) &= \sum_{a \in A} |\mathrm{Fix}_{\mathcal{G}_{\mathrm{n}}}(a)| \\ &= \sum_{a \in A} |G_n|^{n-1}(n-1)! = |G_n|^n n!. \end{aligned}$$ □

Let V^+ be the subspace of V spanned by $\{\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_\mathrm{n}\}$, where v_is are defined as follows:

$$\mathrm{v}_\mathrm{i} = \sum_{g \in G_n} (g, i), \quad \text{for } 1 \leq i \leq n.$$

Also let V^- be the subspace of V spanned by $\{\mathrm{v}_\mathrm{i}{}^g \mid 1 \leq i \leq n, g \in G_n \setminus \{\mathrm{e}\}\}$, where $\mathrm{v}_\mathrm{i}{}^g$s are defined by

$$\mathrm{v}_\mathrm{i}{}^g = \mathrm{v}_\mathrm{i} - |G_n|(g, i) = \sum_{h \in G_n} (h, i) - |G_n|(g, i), \quad \text{for } 1 \leq i \leq n \text{ and } g \in G_n.$$

It can be seen that $V = V^+ \oplus V^-$ and both $V^+$and V^- are invariant under $\mathcal{R}$. Now we define the following sub-representations of $\mathcal{R}$

$$\mathcal{R}^+ : \mathcal{G}_{\mathrm{n}} \to GL(V^+) \text{ defined by } \mathcal{R}^+(g) = \mathcal{R}(g)\Big|_{V^+} \text{ for all } g \in \mathcal{G}_{\mathrm{n}},$$
$$\mathcal{R}^- : \mathcal{G}_{\mathrm{n}} \to GL(V^-) \text{ defined by } \mathcal{R}^-(g) = \mathcal{R}(g)\Big|_{V^-} \text{ for all } g \in \mathcal{G}_{\mathrm{n}}\,.$$

Using $\mathcal{R}^+$ and $\mathcal{R}^-$, X can be written as follows:

$$X(g) = \mathrm{Tr}\left(\mathcal{R}(g)\right) = \mathrm{Tr}\left(\mathcal{R}^+(g)\right) + \mathrm{Tr}\left(\mathcal{R}^-(g)\right) \text{ for all } g \in \mathcal{G}_{\mathrm{n}}\,.$$

Unless otherwise stated from now on we have the following notational assumptions for this chapter:

- I_n (respectively O_n) denotes the identity (respectively zero) matrix of order $n \times n$.
- M_i denotes the matrix of order $n \times n$ with 1 at (i,i)th position and 0 elsewhere.
- $\rho^{\mathrm{def}}(\pi)$ (recall Definition 2.12 from Chapter 2) denotes the matrix of the action of π on $\mathbb{C}[\mathbf{n}]$ with respect to the ordered basis $(\mathbf{1}, \dots, \mathbf{n})$ for $\pi \in S_n$.
- $\mathcal{R}^+(g)$ denotes the matrix of its action on V^+ with respect to the ordered basis $(\mathrm{v}_1, \mathrm{v}_2 \dots, \mathrm{v}_{\mathrm{n}})$ for $g \in \mathcal{G}_{\mathrm{n}}$.
- $\mathcal{R}^-(g)$ denotes the matrix of its action on V^- with respect to the ordered basis

  $$\bigcup_{h \in G_n \setminus \{\mathrm{e}\}} \left(\mathrm{v}_1{}^h, \mathrm{v}_2{}^h \dots, \mathrm{v}_{\mathrm{n}}{}^h\right) \text{ for } g \in \mathcal{G}_{\mathrm{n}}\,.$$

  In this case the ordered basis is the union of $(|G_n| - 1)$ ordered bases indexed by elements of $G_n \setminus \{\mathrm{e}\}$.

Lemma 5.19. *For any $(g_1, g_2, \dots, g_n; \pi) \in \mathcal{G}_{\mathrm{n}}$, the matrices $\mathcal{R}^+((g_1, g_2, \dots, g_n; \pi))$ and $\rho^{\mathrm{def}}(\pi)$ are the same. Moreover the eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}^+)$ are given by*

Eigenvalues:	1	$1 - \frac{1}{n}$	0
Multiplicities:	1	$n-2$	1

Proof. Following the definition of $\mathcal{R}^+$ and ρ^{def} we have,

$$\mathcal{R}^+((g_1, g_2, \dots, g_n; \pi))(\mathrm{v}_{\mathrm{i}}) = \mathcal{R}((g_1, g_2, \dots, g_n; \pi))(\mathrm{v}_{\mathrm{i}}) = \mathrm{v}_{\pi(\mathrm{i})} \text{ and } \rho^{\mathrm{def}}(\pi)(\mathbf{i}) = \pi(\mathbf{i}),$$

for $1 \leq i \leq n$. Thus the first part of the lemma follows. Also the second part of this theorem follows from Lemma 2.13 and the fact

$$\widehat{P_{\mathcal{G}}}(\mathcal{R}^+) = \frac{1}{n}\left(\sum_{u=1}^{n-1} \rho^{\mathrm{def}}((u,n)) + \rho^{\mathrm{def}}(\mathrm{id})\right). \qquad \square$$

Lemma 5.20. *The eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}^-)$ are given by*

Eigenvalues:	$1-\frac{1}{n}$	0
Multiplicities:	$(n-1)(\lvert G_n\rvert-1)$	$(\lvert G_n\rvert-1)$

Proof. Let I_* denote the identity matrix of order $(|G_n|-1)\times(|G_n|-1)$. Also recall that

$$\begin{aligned} h^{(n)} &:= (\mathrm{e},\dots,\mathrm{e},h;\mathrm{id}) \text{ for } h\in G_n \quad\text{and}\\ (h^{-1})^{(u)}h^{(n)}(u,n) &= (\mathrm{e},\dots,\mathrm{e},\underset{\substack{\uparrow\\ u\text{th position.}}}{h^{-1}},\mathrm{e},\dots,\mathrm{e},h;(u,n)) \text{ for } h\in G_n. \end{aligned}$$

For $g\in G_n\setminus\{\mathrm{e}\}$ and $1\le u<n$, we have the following

$$\begin{gathered} \sum_{h\in G_n}\mathcal{R}^-\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)(\mathrm{v_i}^g)=\begin{cases}|G_n|\,\mathrm{v_i}^g & \text{if } i\neq u,n\\ 0 & \text{if } i=u\\ 0 & \text{if } i=n\end{cases}\\ \text{and}\\ \sum_{h\in G_n}\mathcal{R}^-(h^{(n)})(\mathrm{v_i}^g)=\begin{cases}|G_n|\,\mathrm{v_i}^g & \text{if } i\neq n\\ 0 & \text{if } i=n.\end{cases} \end{gathered} \tag{5.4.1}$$

Therefore (5.4.1) implies

$$\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)=I_*\otimes(I_n-M_u-M_n) \text{ for all } u\in[n-1]. \tag{5.4.2}$$

$$\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left(h^{(n)}\right)=I_*\otimes(I_n-M_n). \tag{5.4.3}$$

Now using (5.4.2), (5.4.3) and the definition of $\widehat{P_{\mathcal{G}}}(\mathcal{R}^-)$ we have,

$$\begin{aligned} \widehat{P_{\mathcal{G}}}(\mathcal{R}^-) &= \frac{1}{n}\left(\sum_{u=1}^{n-1}\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)+\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left(h^{(n)}\right)\right)\\ &= \frac{1}{n}\left(\sum_{u=1}^{n-1}I_*\otimes(I_n-M_u-M_n)+I_*\otimes(I_n-M_n)\right)\\ &= \frac{1}{n}\,I_*\otimes\left(n(I_n-M_n)-\sum_{u=1}^{n-1}M_u\right)=\left(1-\frac{1}{n}\right)\,I_*\otimes(I_n-M_n)\,. \end{aligned} \tag{5.4.4}$$

Therefore the lemma follows from (5.4.4). ∎

Proposition 5.21. *Recall that $E_k(X)$ is the expectation of X with respect to the probability measure $P_{\mathcal{G}}^{*k}$ on $\mathcal{G}_{\mathrm{n}}$. Then*

$$E_k(X) = 1 + ((n-1)|G_n| - 1)\left(1 - \frac{1}{n}\right)^k.$$

Proof. In the proof we use the fact that the trace of kth power of a matrix is the sum of the kth powers of its eigenvalues. We also know that $X(g) = \mathrm{Tr}\,(\mathcal{R}^+(g)) + \mathrm{Tr}\,(\mathcal{R}^-(g))$ for all $g \in \mathcal{G}_{\mathrm{n}}$. Thus from the definition of expectation we have

$$\begin{aligned} E_k(X) &= \sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\left(\mathrm{Tr}\left(\mathcal{R}^+(g)\right) + \mathrm{Tr}\left(\mathcal{R}^-(g)\right)\right) \\ &= \mathrm{Tr}\left(\sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\mathcal{R}^+(g)\right) + \mathrm{Tr}\left(\sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\mathcal{R}^-(g)\right) \qquad (5.4.5) \\ &= \mathrm{Tr}\left(\widehat{P_{\mathcal{G}}^{*k}}(\mathcal{R}^+)\right) + \mathrm{Tr}\left(\widehat{P_{\mathcal{G}}^{*k}}(\mathcal{R}^-)\right) \\ &= \mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}^+)\right)^k\right) + \mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}^-)\right)^k\right) \\ &= 1 + (n-2)\left(1-\frac{1}{n}\right)^k + (n-1)(|G_n|-1)\left(1-\frac{1}{n}\right)^k \\ &= 1 + ((n-1)|G_n|-1)\left(1-\frac{1}{n}\right)^k \qquad (5.4.6) \end{aligned}$$

The equality in (5.4.5) holds because Tr is linear. The equality in (5.4.6) follows from Lemmas 5.19 and 5.20. □

Our goal now is to find the expectation $E_k(X^2)$ of X^2 with respect to the probability measure $P_{\mathcal{G}}^{*k}$. For any $g \in \mathcal{G}_{\mathrm{n}}$, let us first observe the following:

$$\begin{aligned} (X(g))^2 &= \left(\mathrm{Tr}\left(\mathcal{R}^+(g)\right) + \mathrm{Tr}\left(\mathcal{R}^-(g)\right)\right)^2 \\ &= \left(\mathrm{Tr}\left(\mathcal{R}^+(g)\right)\right)^2 + 2\left(\mathrm{Tr}\left(\mathcal{R}^-(g)\right)\right)\left(\mathrm{Tr}\left(\mathcal{R}^+(g)\right)\right) + \left(\mathrm{Tr}\left(\mathcal{R}^-(g)\right)\right)^2 \\ &= \mathrm{Tr}\left(\mathcal{R}^+(g)\otimes\mathcal{R}^+(g)\right) + 2\,\mathrm{Tr}\left(\mathcal{R}^-(g)\otimes\mathcal{R}^+(g)\right) + \mathrm{Tr}\left(\mathcal{R}^-(g)\otimes\mathcal{R}^-(g)\right). \qquad (5.4.7) \end{aligned}$$

Expression (5.4.7) suggests us to define three representations $\mathcal{R}_1 : \mathcal{G}_{\mathrm{n}} \to GL(V^+ \otimes V^+)$, $\mathcal{R}_2 : \mathcal{G}_{\mathrm{n}} \to GL(V^- \otimes V^+)$ and $\mathcal{R}_3 : \mathcal{G}_{\mathrm{n}} \to GL(V^- \otimes V^-)$ of $\mathcal{G}_{\mathrm{n}}$. Precisely given as follows:

$$\mathcal{R}_1(g) = \left(\mathcal{R}^+\otimes\mathcal{R}^+\right)(g)(v_i\otimes v_j) = \mathcal{R}^+(g)(v_i)\otimes\mathcal{R}^+(g)(v_j) \text{ for } g\in\mathcal{G}_{\mathrm{n}}, v_i\in V^+, v_j\in V^+,$$
$$\mathcal{R}_2(g) = \left(\mathcal{R}^-\otimes\mathcal{R}^+\right)(g)(v_i\otimes v_j) = \mathcal{R}^-(g)(v_i)\otimes\mathcal{R}^+(g)(v_j) \text{ for } g\in\mathcal{G}_{\mathrm{n}}, v_i\in V^-, v_j\in V^+,$$
$$\mathcal{R}_3(g) = \left(\mathcal{R}^-\otimes\mathcal{R}^-\right)(g)(v_i\otimes v_j) = \mathcal{R}^-(g)(v_i)\otimes\mathcal{R}^-(g)(v_j) \text{ for } g\in\mathcal{G}_{\mathrm{n}}, v_i\in V^-, v_j\in V^-.$$

Lemma 5.22. *$E_k(X^2)$ can be expressed as follows*

$$\mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_1)\right)^k\right) + 2\,\mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_2)\right)^k\right) + \mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\right)^k\right).$$

Proof. Using (5.4.7), i.e.,

$$(X(g))^2 = \mathrm{Tr}\left(\mathcal{R}_1(g)\right) + 2\,\mathrm{Tr}\left(\mathcal{R}_2(g)\right) + \mathrm{Tr}\left(\mathcal{R}_3(g)\right)$$

and linearity of Tr we have

$$\begin{aligned}
E_k(X^2) &= \sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\left(\mathrm{Tr}\left(\mathcal{R}_1(g)\right) + 2\,\mathrm{Tr}\left(\mathcal{R}_2(g)\right) + \mathrm{Tr}\left(\mathcal{R}_3(g)\right)\right)\\
&= \mathrm{Tr}\left(\sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\mathcal{R}_1(g)\right) + 2\,\mathrm{Tr}\left(\sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\mathcal{R}_2(g)\right) + \mathrm{Tr}\left(\sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}^{*k}(g)\mathcal{R}_3(g)\right)\\
&= \mathrm{Tr}\left(\widehat{P_{\mathcal{G}}^{*k}}(\mathcal{R}_1)\right) + 2\,\mathrm{Tr}\left(\widehat{P_{\mathcal{G}}^{*k}}(\mathcal{R}_2)\right) + \mathrm{Tr}\left(\widehat{P_{\mathcal{G}}^{*k}}(\mathcal{R}_3)\right)\\
&= \mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_1)\right)^k\right) + 2\,\mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_2)\right)^k\right) + \mathrm{Tr}\left(\left(\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\right)^k\right).
\end{aligned}$$

□

Lemma 5.23. *The eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_1)$ are given as follows:*

Eigenvalues:	1	$1-\frac{1}{n}$	0	$\frac{1}{n}$	$-\frac{1}{n}$	$1-\frac{2}{n}$
Multiplicities:	2	$3(n-2)$	3	$n-2$	$n-2$	n^2-5n+5

Proof. Let us recall that $\mathcal{R}^+((g_1,g_2,\ldots,g_n;\pi)) = \rho^{\mathrm{def}}(\pi)$ for all $(g_1,g_2,\ldots,g_n;\pi)\in\mathcal{G}_{\mathrm{n}}$ and $\widehat{P_{\mathcal{G}}}(\mathcal{R}_1) = \sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}(g)\mathcal{R}_1(g) = \sum_{g\in\mathcal{G}_{\mathrm{n}}} P_{\mathcal{G}}(g)\left(\mathcal{R}^+(g)\otimes\mathcal{R}^+(g)\right)$. Therefore we have

$$\begin{aligned}
\widehat{P_{\mathcal{G}}}(\mathcal{R}_1) &= \frac{1}{n|G_n|}\sum_{u=1}^{n-1}\sum_{h\in G_n}\mathcal{R}^+\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)\otimes\mathcal{R}^+\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)\\
&\qquad + \frac{1}{n|G_n|}\sum_{h\in G_n}\mathcal{R}^+\left(h^{(n)}\right)\otimes\mathcal{R}^+\left(h^{(n)}\right)\\
&= \frac{1}{n}\left(\sum_{u=1}^{n-1}\rho^{\mathrm{def}}((u,n))\otimes\rho^{\mathrm{def}}((u,n)) + \rho^{\mathrm{def}}(\mathrm{id})\otimes\rho^{\mathrm{def}}(\mathrm{id})\right).
\end{aligned}$$

Thus the lemma follows from Lemma 2.14. □

Lemma 5.24. *The eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_2)$ are given as follows:*

Eigenvalues:	$1-\frac{1}{n}$	$1-\frac{2}{n}$	0
Multiplicities:	$2(n-1)(\lvert G_n\rvert-1)$	$(n-3)(n-1)(\lvert G_n\rvert-1)$	$(2n-1)(\lvert G_n\rvert-1)$

Proof. Let us recall that $\mathcal{R}^+((g_1, g_2, \dots, g_n; \pi)) = \rho^{\rm def}(\pi)$ for all $(g_1, g_2, \dots, g_n; \pi) \in \mathcal{G}_{\rm n}$. Now using the definition of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_2)$ we have

$$\begin{aligned}
\widehat{P_{\mathcal{G}}}(\mathcal{R}_2) &= \sum_{g\in\mathcal{G}_{\rm n}} P_{\mathcal{G}}(g)\mathcal{R}_2(g) = \sum_{g\in\mathcal{G}_{\rm n}} P_{\mathcal{G}}(g) \left(\mathcal{R}^-(g) \otimes \mathcal{R}^+(g)\right)\\
&= \frac{1}{n|G_n|}\sum_{u=1}^{n-1}\sum_{h\in G_n} \mathcal{R}^-\left((h^{-1})^{(u)}h^{(n)}(u,n)\right) \otimes \mathcal{R}^+\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)\\
&\qquad\qquad + \frac{1}{n|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left(h^{(n)}\right)\otimes\mathcal{R}^+\left(h^{(n)}\right)\\
&= \frac{1}{n}\sum_{u=1}^{n-1}\left(\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left((h^{-1})^{(u)}h^{(n)}(u,n)\right)\right)\otimes\rho^{\rm def}\left((u,n)\right)\\
&\qquad\qquad + \frac{1}{n}\left(\frac{1}{|G_n|}\sum_{h\in G_n}\mathcal{R}^-\left(h^{(n)}\right)\right)\otimes\rho^{\rm def}({\rm id})\\
&= \frac{1}{n}\sum_{u=1}^{n-1} I_*\otimes(I_n - M_u - M_n)\otimes\rho^{\rm def}((u,n)) + \frac{1}{n}\, I_*\otimes(I_n - M_n)\otimes\rho^{\rm def}({\rm id}) \qquad (5.4.8)\\
&= \frac{1}{n}\, I_*\otimes\left(\sum_{u=1}^{n-1}(I_n - M_u - M_n)\otimes\rho^{\rm def}((u,n)) + (I_n - M_n)\otimes\rho^{\rm def}({\rm id})\right)\\
&= \frac{1}{n}\, I_*\otimes\left((I_n - M_n)\otimes\left(\sum_{u=1}^{n-1}\rho^{\rm def}\left((u,n)\right) + \rho^{\rm def}\left({\rm id}\right)\right) - \sum_{u=1}^{n-1} M_u\otimes\rho^{\rm def}((u,n))\right) \quad (5.4.9)
\end{aligned}$$

The equality in (5.4.8) follows from (5.4.2) and (5.4.3). If $\text{Blockdiag}\left(A_1, A_2, \dots, A_n\right)$ denotes the block diagonal matrix with ith block A_i for all $i \in [n-1]$. Then recalling

$$\beta_i = \sum_{u=1}^{n-1}\rho^{\rm def}\left((u,n)\right) - \rho^{\rm def}((i,n)) \text{ for all } 1 \leq i < n$$

from Lemma 2.15, the right hand side of (5.4.9) can be written as

$$\frac{1}{n}\, I_*\otimes\text{Blockdiag}\left(I_n + \beta_1, I_n + \beta_2, \dots, I_n + \beta_{n-1}, O_n\right).$$

Therefore the lemma follows from Lemma 2.15. □

Let us first prove a lemma which will be useful in finding the eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$. This lemma is true in general not only for our setting. We prove this lemma using the *Geršhgorin disc theorem*, stated as follows:

Theorem 5.25 (Geršhgorin disc theorem [56, Theorem 6.1.1])**.** *Let $A = (a_{ij})_{n\times n}$ be a matrix with complex entries. Also let $r'_i(A) = \sum_{j\neq i}|a_{ij}|,\ 1 \leq i \leq n$ denote the deleted absolute row sums of A, and consider the n Geršhgorin discs $\{z \in \mathbb{C} : |z - a_{ii}| <$*

$r'_i(A)\}$, $1 \le i \le n$. *Then the eigenvalues of A are in the union of Geršhgorin discs*

$$G(A) := \bigcup_{1\le i\le n} \{z \in \mathbb{C} : |z - a_{ii}| < r'_i(A)\}.$$

Furthermore, if the union of k of the n discs that comprise $G(A)$ forms a set $G_k(A)$ that is disjoint from the remaining $n-k$ discs, then $G_k(A)$ contains exactly k eigenvalues of A, counted according to their algebraic multiplicities.

Lemma 5.26. *Let G be a finite group. Recall that $V = \mathbb{C}[G]$ is the complex vector space with basis G. Also let (ρ, V) be the left regular representation of G i.e.*

$$\rho(g) \mapsto \left(\sum_{h\in G} C_h h \mapsto \sum_{h\in G} C_h gh\right), \quad g \in G,\ C_h \in \mathbb{C}.$$

Then the eigenvalues of $\frac{1}{|G|}\sum_{g\in G} \rho(g)\otimes\rho(g^{-1})$ are in the closed unit disc

$$\overline{\mathbb{D}} := \{z \in \mathbb{C} : |z| \le 1\}.$$

Proof. Before proving the lemma let us recall the definition of a *stochastic matrix*. A real square matrix is said to be *stochastic* if all its entries are from the interval $[0,1]$ and sum of the elements in each row is 1.

It can be easily seen that $\rho(g)$ is a permutation matrix and thus a stochastic matrix for each $g \in G$. Therefore $\rho(g)\otimes\rho(g^{-1})$ is a stochastic matrix for each $g \in G$. It is known that the average of stochastic matrices are stochastic. Therefore $\frac{1}{|G|}\sum_{g\in G}\rho(g)\otimes\rho(g^{-1})$ is a stochastic matrix. Hence the lemma follows from Theorem 5.25. □

Corollary 5.27. *Let G be a finite group and V' be the vector space spanned by the complex linear combinations of elements of the set $\{v^g | g \in G\}$, where $v^g = \sum_{h\in G} h - |G|g$. Also let L be the representation of G defined by left regular action on V'. Then the eigenvalues of $\frac{1}{|G|}\sum_{g\in G} L(g)\otimes L(g^{-1})$ are in the closed unit disc $\overline{\mathbb{D}} := \{z \in \mathbb{C} : |z| \le 1\}$.*

Proof. Let $V = \mathbb{C}[G]$ and ρ denotes the left regular representation of G. Also note that $V'\otimes V'$ is a subspace of $V\otimes V$ invariant under $\frac{1}{|G|}\sum_{g\in G}\rho(g)\otimes\rho(g^{-1})$. Then the corollary follows follows from the fact $\frac{1}{|G|}\sum_{g\in G} L(g)\otimes L(g^{-1}) = \frac{1}{|G|}\sum_{g\in G}\rho(g)\otimes\rho(g^{-1})\Big|_{V'\otimes V'}$. □

Lemma 5.28. *Let L be the representation of G_n defined by left regular action on the vector space spanned by the set $\{v^g | g \in G\}$, where $v^g = \sum_{h\in G_n} h - |G_n|g$. If the eigenvalues* *of* $\frac{1}{|G_n|}\sum_{g\in G_n} L(g)\otimes L(g^{-1})$ *are* λ_i *for* $1 \leq i \leq (|G_n|-1)^2$*, then the eigenvalues of* $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$ *are given in* Table 5.1*. Also we note that* $|\lambda_i| \leq 1$ *for* $1 \leq i \leq (|G_n|-1)^2$ *in this case.*

Eigenvalues	*Multiplicities*
1	$\lvert G_n\rvert - 1$
$1-\frac{1}{n}$	$(n-1)(\lvert G_n\rvert-1)^2 - (\lvert G_n\rvert-1)$
$1-\frac{2}{n}$	$(n-1)(n-2)(\lvert G_n\rvert-1)^2$
$\pm\frac{\lambda_i}{n},\ 1 \leq i \leq (\lvert G_n\rvert-1)^2$	$(n-1)$
0	$(\lvert G_n\rvert-1)^2$

Table 5.1: Eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$

Proof. We prove this lemma by splitting $V^-\otimes V^-$ into three subspaces which are invariant under the action of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$. The subspaces are given as follows.

- Let W be the subspace of $V^-\otimes V^-$ spanned by the set of vectors

$$\bigcup_{1\leq i<n}\left(\left\{ \mathrm{v_i}^x\otimes \mathrm{v_n}^y : x,y\in G_n\setminus\{\mathrm{e}\}\right\}\cup\left\{ \mathrm{v_n}^x\otimes \mathrm{v_i}^y : x,y\in G_n\setminus\{\mathrm{e}\}\right\}\right).$$

- Let W' be the subspace of $V^-\otimes V^-$ spanned by the set of vectors

$$\left\{ \mathrm{v_i}^g\otimes \mathrm{v_j}^h : g,h\in G_n\setminus\{\mathrm{e}\}, i\neq j, 1\leq i,j<n\right\}.$$

- Let W'' be the subspace of $V^-\otimes V^-$ spanned by the set of vectors

$$\left\{ \mathrm{v_i}^g\otimes \mathrm{v_i}^h : g,h\in G_n\setminus\{\mathrm{e}\}, 1\leq i\leq n\right\}.$$

It can be easily seen that $V^-\otimes V^- = W\oplus W'\oplus W''$. We also note that $\dim(W) = 2(n-1)(|G_n|-1)^2$, $\dim(W') = (n-1)(n-2)(|G_n|-1)^2$ and $\dim(W'') = n(|G_n|-1)^2$. Since W, W' and W'' are all invariant under the action of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$, we can write $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$ as block diagonal matrix with respect to a certain choice for ordering of the basis elements. Our goal is to use that block diagonal decomposition and obtain the eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$.

Eigenvalue	Eigenvectors corresponding to eigenvalue given in column 1	number of (independent) vectors in column 2
1	$\sum_{\substack{h\in G_n\\ 1\leq i\leq n}} \mathrm{v_i}^h \otimes \mathrm{v_i}^{hx}$, for $x \in G_n \setminus \{\mathrm{e}\}$.	$\lvert G_n\rvert - 1$
0	$\frac{n}{n-1} \sum_{\substack{h\in G_n\\ 1\leq i\leq n-1}} \mathrm{v_i}^h \otimes \mathrm{v_i}^{hx} - n \sum_{h\in G_n} \mathrm{v_n}^h \otimes \mathrm{v_n}^{hx}$, for $x \in G_n \setminus \{\mathrm{e}\}$.	$\lvert G_n\rvert - 1$
0	$\mathrm{v_n}^g \otimes \mathrm{v_n}^{gx} - \mathrm{v_n}^h \otimes \mathrm{v_n}^{hx}$, for $x \in G_n \setminus \{g^{-1}\}$ and $h \in G_n \setminus \{g, x^{-1}\}$.	$(\lvert G_n\rvert - 1)(\lvert G_n\rvert - 2)$
$1 - \frac{1}{n}$	$\mathrm{v_i}^g \otimes \mathrm{v_i}^{gx} - \mathrm{v_i}^h \otimes \mathrm{v_i}^{hx}$, for $x \in G_n \setminus \{g^{-1}\}$, $h \in G_n \setminus \{g, x^{-1}\}$ and $1 \leq i < n$.	$(n-1)(\lvert G_n\rvert - 1)(\lvert G_n\rvert - 2)$
$1 - \frac{1}{n}$	$\mathrm{v_i}^g \otimes \mathrm{v_i}^{gx} - \mathrm{v_1}^g \otimes \mathrm{v_1}^{gx}$ for $x \in G_n \setminus \{g^{-1}\}$, and $2 \leq i < n$.	$(n-2)(\lvert G_n\rvert - 1)$

Table 5.2: Eigenvectors and eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\Big|_{W''}$.

First let us notice that the subspaces W_i of $V^- \otimes V^-$ spanned by the set of vectors

$$\left(\left\{ \mathrm{v_i}^x \otimes \mathrm{v_n}^y : x, y \in G_n \setminus \{\mathrm{e}\}\right\} \cup \left\{ \mathrm{v_n}^x \otimes \mathrm{v_i}^y : x, y \in G_n \setminus \{\mathrm{e}\}\right\}\right)$$

is also invariant under the action of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)$ for each $i \in [n-1]$ and $W = W_1 \oplus \cdots \oplus W_{n-1}$. We now focus on $W = W_1 \oplus \cdots \oplus W_{n-1}$. For any $1 \leq i < n$, consider the ordered basis $\mathscr{B}_i$ of W_i as follows: In $\mathscr{B}_i$ first list all elements of the form $\mathrm{v_i}^x \otimes \mathrm{v_n}^y$ and then list the elements of the form $\mathrm{v_n}^x \otimes \mathrm{v_i}^y$ by maintaining the same ordering of the pair (x, y). Using this ordered basis $\mathscr{B}_i$, $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\Big|_{W_i}$ will be of the form

$$\frac{1}{n}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \left(\frac{1}{\lvert G_n\rvert} \sum_{g\in G_n} L(g) \otimes L(g^{-1})\right). \tag{5.4.10}$$

Here the matrix $\frac{1}{\lvert G_n\rvert} \sum_{g\in G_n} L(g) \otimes L(g^{-1})$ is written with respect to a ordered basis following the same ordering as that of (x, y) in $\mathrm{v_i}^x \otimes \mathrm{v_n}^y$ (or $\mathrm{v_n}^x \otimes \mathrm{v_i}^y$). Since (5.4.10) is true for all

$i \in [n-1]$, using the Corollary 5.27 we can conclude that the eigenvalues of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\Big|_W$ are given by: $\pm\frac{\lambda_i}{n}$ with multiplicity $(n-1)$ each for $1 \leq i \leq (|G_n|-1)^2$.

We now focus on the subspace W' of $V^- \otimes V^-$. It can be easily seen that

$$\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)(\mathrm{v_i}^g \otimes \mathrm{v_j}^h) = \left(1 - \frac{2}{n}\right)(\mathrm{v_i}^g \otimes \mathrm{v_j}^h), \quad \text{for } g, h \in G_n \setminus \{\mathrm{e}\}, 1 \leq i, j < n \text{ and } i \neq j.$$

Therefore $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\Big|_{W'}$ is the scalar matrix $\left(1-\frac{2}{n}\right) I_{\dim(W')}$.

The eigenvalues and eigenvectors of $\widehat{P_{\mathcal{G}}}(\mathcal{R}_3)\Big|_{W''}$ are given in Table 5.2, where g is a fixed non identity element of G_n. Hence the lemma follows. $\square$

Proposition 5.29. *Recall that* $\mathrm{Var}_k(X)$ *is the variance of* X *with respect to the probability measure* $P_{\mathcal{G}}^{*k}$ *on* $\mathcal{G}_{\mathrm{n}}$*. Then*

$$\begin{aligned}\mathrm{Var}_k(X) &= |G_n| + \Big((n-1)|G_n|^2 - |G_n|\Big)\left(1-\frac{1}{n}\right)^k \\ &+ \Big((n-1)(n-2)(|G_n|-1)^2 + 2(n-3)(n-1)(|G_n|-1) + n^2 - 5n + 5\Big)\left(1-\frac{2}{n}\right)^k \\ &+ \frac{\left(1+(-1)^k\right)}{n^k}\left(n - 2 + (n-1)\sum_{i=1}^{(|G_n|-1)^2}\lambda_i^k\right) - ((n-1)|G_n|-1)^2\left(1-\frac{1}{n}\right)^{2k}.\end{aligned}$$

Where λ_i *are defined in Lemma 5.28.*

Proof. This proposition follows from the definition of variance, Proposition 5.21, Lemmas 5.22, 5.23, 5.24, 5.28 and straightforward calculations. $\square$

Theorem 5.30. *For the random walk on* $\mathcal{G}_{\mathrm{n}}$ *driven by* $P_{\mathcal{G}}$*, we have the following:*

1. *For large* n*, we have*

$$||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} > 1 - \frac{2\left(2+\frac{1}{|G_n|}\right)\left(e^{-c}+\frac{1}{|G_n|}\right) + o(1)(1+e^{-c}+e^{-2c})}{\left(\frac{1}{|G_n|} + (1+o(1))e^{-c}\right)^2},$$

   when $k = n\log n + cn$ *and* $c \ll 0$.

2. *For any* $\epsilon \in (0,1)$ *and* $k_n = \lfloor (1-\epsilon)n\log n \rfloor$*, we have*

$$\lim_{n\to\infty} ||P_{\mathcal{G}}^{*k_n} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} = 1.$$

Proof. Using Lemma 2.19, $\mu = P_{\mathcal{G}}^{*k}$ and $\nu = U_{\mathcal{G}_{\mathrm{n}}}$ we have,

$$||P_{\mathcal{G}}^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} \geq 1 - \frac{4\,\mathrm{Var}_k(X)}{(E_k(X))^2} - \frac{2}{E_k(X)}. \tag{5.4.11}$$

Now from Propositions 5.21 and 5.29 we have

$$\begin{aligned} E_k(X) &\approx 1 + ((n-1)|G_n| - 1)\, e^{-\frac{k}{n}}, \quad \text{for } k \geq 1. && (5.4.12)\\ \mathrm{Var}_k(X) &\approx |G_n| + \Big((n-1)|G_n|^2 - |G_n|\Big)\, e^{-\frac{k}{n}} - (n-1)|G_n|^2 e^{-\frac{2k}{n}} \\ &\quad + \frac{\left(1+(-1)^k\right)}{n^{k-1}}\left(1 + (|G_n|-1)^2\right), \text{ for } k \geq 1. && (5.4.13) \end{aligned}$$

Here '$\approx$' means 'asymptotic to' i.e. $a_n \approx b_n$ means $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$. Now if $c \ll 0$ and $k = n\log n + cn$, then by (5.4.11), (5.4.12) and (5.4.13), we have the first part of this theorem.

Now for any $\epsilon \in (0,1)$ and $k_n = \lfloor (1-\epsilon) n \log n \rfloor$ from (5.4.12) and (5.4.13) we have

$$\begin{aligned} E_k(X) &= 1 + ((1+o(1))|G_n| + o(1))\, n^{\epsilon} \\ \mathrm{Var}_k(X) &= |G_n| + \Big((1+o(1))|G_n|^2 + o(1)|G_n|\Big)\, n^{\epsilon} \\ &\quad - \frac{|G_n|^2\,(1+o(1))}{n^{1-2\epsilon}} + o(1)\left(1 + (|G_n|-1)^2\right). \end{aligned}$$

Therefore the second part of this theorem follows from (5.4.11) and the following:

$$\begin{aligned} &\lim_{n\to\infty} \frac{|G_n| + ((1+o(1))|G_n|^2 + o(1)|G_n|)\, n^{\epsilon} - \frac{|G_n|^2(1+o(1))}{n^{1-2\epsilon}} + o(1)\left(1+(|G_n|-1)^2\right)}{\left(1 + ((1+o(1))|G_n| + o(1))\, n^{\epsilon}\right)^2} = 0 \\ &\lim_{n\to\infty} \frac{1}{1 + ((1+o(1))|G_n| + o(1))\, n^{\epsilon}} = 0. \end{aligned}$$

□

The second part of Theorems 5.14 and 5.30 prove the following theorem.

Theorem 5.31. *The warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ exhibits cutoff phenomenon with cutoff time $n \log n$ if $|G_n| = o(n^{\delta})$ for all $\delta > 0$.*

Remark 5.32. All the results of this chapter will be the same if we consider any sequence of groups $\{G_n\}_1^{\infty}$ with the condition $|G_n| > 2$ for all $n \geq 1$, instead of considering the ascending sequence of groups $G_1 \subseteq G_2 \subseteq \cdots$ with $|G_1| > 2$. The proofs are similar.

Bibliography

[1] Sheldon B. Akers and Balakrishnan Krishnamurthy. A group-theoretic model for symmetric interconnection networks. *IEEE Trans. Comput.*, 38(4):555–566, 1989.

[2] David Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin, 1983.

[3] David Aldous. A Metropolis-type optimization algorithm on the infinite tree. volume 22, pages 388–412. 1998. Average-case analysis of algorithms.

[4] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly*, 93(5):333–348, 1986.

[5] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987.

[6] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., New York, 1992. With an appendix by Paul Erdős, A Wiley-Interscience Publication.

[7] Sø ren Asmussen and Peter W. Glynn. *Stochastic simulation: algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.

[8] Arvind Ayyer and Svante Linusson. Reverse juggling processes. *Random Structures Algorithms*, 55(1):56–72, 2019.

[9] Arvind Ayyer, Anne Schilling, Benjamin Steinberg, and Nicolas M. Thiéry. Markov chains, $\mathscr{R}$-trivial monoids and representation theory. *Internat. J. Algebra Comput.*, 25(1-2):169–231, 2015.

[10] Arvind Ayyer, Anne Schilling, and Nicolas M. Thiéry. Spectral gap for random-to-random shuffling on linear extensions. *Exp. Math.*, 26(1):22–30, 2017.

[11] Arvind Ayyer and Pooja Singla. Random motion on finite rings, II: noncommutative rings. *arXiv preprint arXiv:1807.04082*, 2018.

[12] Arvind Ayyer and Pooja Singla. Random motion on finite rings, I: Commutative rings. *Algebr. Represent. Theory*, 23(3):583–604, 2020.

[13] Arvind Ayyer and Benjamin Steinberg. Random walks on rings and modules. *Algebr. Comb.*, 3(2):309–329, 2020.

[14] Arvind Ayyer and Volker Strehl. Stationary distribution and eigenvalues for a de Bruijn process. In *Advances in combinatorics*, pages 101–120. Springer, Heidelberg, 2013.

[15] Yossi Azar, Andrei Z. Broder, and Anna R. Karlin. On-line load balancing. *Theoret. Comput. Sci.*, 130(1):73–84, 1994.

[16] Yossi Azar, Bala Kalyanasundaram, Serge Plotkin, Kirk R. Pruhs, and Orli Waarts. On-line load balancing of temporary tasks. *J. Algorithms*, 22(1):93–110, 1997.

[17] Dave Bayer and Persi Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, 2(2):294–313, 1992.

[18] Robert G Beiko, Jonathan M Keith, Timothy J Harlow, and Mark A Ragan. Searching for convergence in phylogenetic Markov chain monte carlo. *Systematic Biology*, 55(4):553 – 565, 2006.

[19] Mihir Bellare, Oded Goldreich, and Shafi Goldwasser. Randomness in interactive proofs. *Comput. Complexity*, 3(4):319–354, 1993.

[20] Megan Bernstein and Evita Nestoridi. Cutoff for random to random card shuffle. *Ann. Probab.*, 47(5):3303–3320, 2019.

[21] Pat Bidigare, Phil Hanlon, and Dan Rockmore. A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Math. J.*, 99(1):135–174, 1999.

[22] Anders Björner and Francesco Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.

[23] Émile Borel and André Chéron. *Théorie Mathématique du Bridge à la Portée de Tous.* 134 *Tableaux de Probabilités avec Leurs Modes d'Emploi. Formules Simples. Applications. Environ* 4000 *Probabilités.* Monographies des Probabilités, Fasc. V. Gauthier-Villars, Paris, 1940.

[24] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1):107 – 117, 1998. Proceedings of the Seventh International World Wide Web Conference.

[25] Kenneth S. Brown. Semigroups, rings, and Markov chains. *J. Theoret. Probab.*, 13(3):871–938, 2000.

[26] Krzysztof Burdzy and Wilfrid S. Kendall. Efficient Markovian couplings: examples and counterexamples. *Ann. Appl. Probab.*, 10(2):362–409, 2000.

[27] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli. On the representation theory of the alternating groups. *Int. J. Group Theory*, 2(1):187–198, 2013.

[28] Mu Fa Chen. *From Markov chains to nonequilibrium particle systems.* World Scientific Publishing Co., Inc., River Edge, NJ, 1992.

[29] Jon Cohen. How dna shuffling works. *Science*, 293(5528):237–237, 2001.

[30] Ivan Damgå rd, Peter Landrock, and Carl Pomerance. Average case error estimates for the strong probable prime test. *Math. Comp.*, 61(203):177–194, 1993.

[31] Khaled Day and Anand Tripathi. A comparative study of topological properties of hypercubes and star graphs. *IEEE Trans. Parallel Distrib. Systems*, 5(1):31–38, 1994.

[32] Persi Diaconis. Applications of non-commutative fourier analysis to probability problems. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pages 51–100. Springer, 1988.

[33] Persi Diaconis. Group representations in probability and statistics. 11:vi+198, 1988.

[34] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, 93(4):1659–1664, 1996.

[35] Persi Diaconis, James Allen Fill, and Jim Pitman. Analysis of top to random shuffles. *Combin. Probab. Comput.*, 1(2):135–155, 1992.

[36] Persi Diaconis, R. L. Graham, and J. A. Morrison. Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms*, 1(1):51–72, 1990.

[37] Persi Diaconis and Susan P. Holmes. Random walks on trees and matchings. *Electron. J. Probab.*, 7:no. 6, 17, 2002.

[38] Persi Diaconis and Laurent Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.*, 21(4):2131–2156, 1993.

[39] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.*, 3(3):696–730, 1993.

[40] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981.

[41] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.*, 18(1):208–218, 1987.

[42] A. B. Dieker and F. V. Saliola. Spectral analysis of random-to-random Markov chains. *Adv. Math.*, 323:427–485, 2018.

[43] Peter Donnelly. The heaps process, libraries, and size-biased permutations. *J. Appl. Probab.*, 28(2):321–335, 1991.

[44] Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms.* Cambridge University Press, Cambridge, 2009.

[45] Rick Durrett. Shuffling chromosomes. *J. Theoret. Probab.*, 16(3):725–750, 2003.

[46] L. Flatto, A. M. Odlyzko, and D. B. Wales. Random shuffles and group representations. *Ann. Probab.*, 13(1):154–178, 1985.

[47] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.

[48] Rusiņš Freivalds. Probabilistic machines can use less running time. In *Information processing 77 (Proc. IFIP Congr., Toronto, Ont., 1977)*, pages 839–842. IFIP Congr. Ser., Vol. 7, 1977.

[49] Jason Fulman. Convergence rates of random walk on irreducible representations of finite groups. *J. Theoret. Probab.*, 21(1):193–211, 2008.

[50] L. Geissinger and D. Kinch. Representations of the hyperoctahedral groups. *J. Algebra*, 53(1):1–20, 1978.

[51] Harry Gonshor. An application of random walk to a problem in population genetics. *The American Mathematical Monthly*, 94(7):668–671, 1987.

[52] Geoffrey Grimmett. The random-cluster model. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 73–123. Springer, Berlin, 2004.

[53] W. J. Hendricks. The stationary distribution of an interesting Markov chain. *J. Appl. Probability*, 9:231–233, 1972.

[54] W. J. Hendricks. An extension of a theorem concerning an interesting Markov chain. *J. Appl. Probability*, 10:886–890, 1973.

[55] Martin Hildebrand. Generating random elements in $\mathrm{SL}_n(\mathbf{F}_q)$ by random transvections. *J. Algebraic Combin.*, 1(2):133–150, 1992.

[56] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[57] B. Hostinsky. *Méthodes générales du calcul des probabilités*. Gauthier-Villars, 1931.

[58] Xueyi Huang and Qiongxiang Huang. The second largest eigenvalues of some Cayley graphs on alternating groups. *J. Algebraic Combin.*, 50(1):99–111, 2019.

[59] Gordon James and Adalbert Kerber. *The representation theory of the symmetric group*, volume 16 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1981. With a foreword by P. M. Cohn, With an introduction by Gilbert de B. Robinson.

[60] Sanjiv Kapoor and Edward M. Reingold. Stochastic rearrangement rules for self-organizing data structures. *Algorithmica*, 6(2):278–291, 1991.

[61] Richard M. Karp. An introduction to randomized algorithms. volume 34, pages 165–201. 1991. Combinatorics and theoretical computer science (Washington, DC, 1989).

[62] A. E. Kiasari and H. Sarbazi-Azad. Analytic performance comparison of hypercubes and star graphs with implementation constraints. *J. Comput. System Sci.*, 74(6):1000–1012, 2008.

[63] D. D. Kosambi and U. V. Ramamohana Rao. The efficiency of randomization by card shuffling. *Journal of the Royal Statistical Society. Series A (General)*, 121(2):223–233, 1958.

[64] Hubert Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.*, 44(2):1426–1487, 2016.

[65] Amy N. Langville and Carl D. Meyer. *Google's PageRank and beyond: the science of search engine rankings.* Princeton University Press, Princeton, NJ, 2012. Paperback edition of the 2006 original.

[66] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[67] Dror Marcus and Yuval Shavitt. Rage–a rapid graphlet enumerator for large networks. *Computer Networks*, 56(2):810–819, 2012.

[68] Stuart Margolis, Franco Saliola, and Benjamin Steinberg. Combinatorial topology and the global dimension of algebras arising in combinatorics. *J. Eur. Math. Soc. (JEMS)*, 17(12):3037–3080, 2015.

[69] AA Markov. Extension of the law of large numbers to dependent events (russian) bull. *Soc. Phy. Math. Kazanstan*, 2:135–156, 1906.

[70] Peter Matthews. Strong stationary times and eigenvalues. *J. Appl. Probab.*, 29(1):228–233, 1992.

[71] S. J. Mayer. On the characters of the Weyl group of type D. *Math. Proc. Cambridge Philos. Soc.*, 77:259–264, 1975.

[72] Marc Mézard and Andrea Montanari. Reconstruction on trees and spin glass transition. *J. Stat. Phys.*, 124(6):1317–1350, 2006.

[73] Ashish Mishra and Murali K. Srinivasan. The Okounkov-Vershik approach to the representation theory of $G \sim S_n$. *J. Algebraic Combin.*, 44(3):519–560, 2016.

[74] Michael Molloy. The Glauber dynamics on colourings of a graph with high girth and maximum degree. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 91–98. ACM, New York, 2002.

[75] Elchanan Mossel. Reconstruction on trees: beating the second eigenvalue. *Ann. Appl. Probab.*, 11(1):285–300, 2001.

[76] Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms.* Cambridge University Press, Cambridge, 1995.

[77] Andrei Okounkov and Anatoly Vershik. A new approach to representation theory of symmetric groups. *Selecta Math. (N.S.)*, 2(4):581–605, 1996.

[78] Rasmus Pagh and Charalampos E. Tsourakakis. Colorful triangle counting and a MapReduce implementation. *Inform. Process. Lett.*, 112(7):277–281, 2012.

[79] Robin Pemantle. Randomization time for the overhand shuffle. *J. Theoret. Probab.*, 2(1):37–49, 1989.

[80] R. M. Phatarfod. On the matrix occurring in a linear search problem. *J. Appl. Probab.*, 28(2):336–346, 1991.

[81] Amritanshu Prasad. *Representation theory: a combinatorial viewpoint*, volume 147. Cambridge University Press, 2015.

[82] James Gary Propp and David Bruce Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

[83] I. A. Pushkarev. On the theory of representations of the wreath products of finite groups and symmetric groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 240(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 2):229–244, 294–295, 1997.

[84] Chuan Qin and Ben Morris. Improved bounds for the mixing time of the random-to-random shuffle. *Electron. Commun. Probab.*, 22:Paper No. 22, 7, 2017.

[85] Sanjay Ranka, Jhy-Chun Wang, and Nangkang Yeh. Embedding meshes on the star graph. *J. Parallel Distrib. Comput.*, 19(2):131–135, 1993.

[86] Oliver Ruff. Weight theory for alternating groups. *Algebra Colloq.*, 15(3):391–404, 2008.

[87] Bruce E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.

[88] L. Saloff-Coste and J. Zúñiga. Refined estimates for some basic random walks on the symmetric and alternating groups. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:359–392, 2008.

[89] Laurent Saloff-Coste. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 263–346. Springer, Berlin, 2004.

[90] Clyde H. Schoolfield, Jr. Generating a random signed permutation with random reversals. *J. Theoret. Probab.*, 18(4):911–931, 2005.

[91] Clyde Henry Schoolfield, Jr. *Random walks on wreath products of groups and Markov chains on related homogeneous spaces.* ProQuest LLC, Ann Arbor, MI, 1999. Thesis (Ph.D.)–The Johns Hopkins University.

[92] Jean-Pierre Serre. *Linear representations of finite groups.* Springer-Verlag, New York-Heidelberg, 1977. Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.

[93] Daniel Dominic Sleator and Robert Endre Tarjan. Self-adjusting binary search trees. *J. Assoc. Comput. Mach.*, 32(3):652–686, 1985.

[94] R. Solovay and V. Strassen. A fast Monte-Carlo test for primality. *SIAM J. Comput.*, 6(1):84–85, 1977.

[95] Eliran Subag. A lower bound for the mixing time of the random-to-random insertions shuffle. *Electron. J. Probab.*, 18:no. 20, 20, 2013.

[96] Michio Suzuki. *Group theory. I*, volume 247 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, 1982. Translated from the Japanese by the author.

[97] M. L. Tsetlin. Finite automata and the simulation of the simplest forms of behavior. *Uspehi Mat. Nauk*, 18(4 (112)):3–28, 1963.

[98] D. J. A. Welsh. Randomised algorithms. *Discrete Appl. Math.*, 5(1):133–145, 1983.

[99] David Bruce Wilson. Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.*, 14(1):274–325, 2004.

[100] Huimin Zhao and Frances H. Arnold. Optimization of DNA Shuffling for High Fidelity Recombination. *Nucleic Acids Research*, 25(6):1307–1308, 03 1997.