Total variation cutoff for the flip-transpose top with random shuffle

Subhajit Ghosh

Department of Mathematics
Indian Institute of Science,
Bangalore-560 012, Karnataka, India.
E-mail address: `gsubhajit@iisc.ac.in`

Abstract. We consider a random walk on the hyperoctahedral group B_n generated by the signed permutations of the forms (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 the mixing time for this shuffle is of order $n \log n$. We also show that this shuffle exhibits the cutoff phenomenon. In the appendix, we show that a similar random walk on the demihyperoctahedral group D_n also has a cutoff at $\left(n - \frac{1}{2}\right) \log n$.

1. Introduction

Card shuffling problems are mathematically analysed by considering them as random walks on symmetric groups (Diaconis, 1988b; Flatto et al., 1985; Diaconis and Shahshahani, 1981; Saloff-Coste, 2004; Schoolfield, 2005, 2002). In this paper, our main aim is to study the properties of a random walk on Coxeter groups of type B (Björner and Brenti, 2005). This work is a generalisation of the *transpose top with random shuffle* (Flatto et al., 1985; Diaconis, 1988a) to the signed permutations. A *signed permutation* (Björner and Brenti, 2005) is a bijection π from $\{\pm 1, \ldots, \pm n\}$ 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 $[n] := \{1, \ldots, n\}$. 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 and is known as the *hyperoctahedral group* and is denoted by B_n. The subset of B_n consisting of those signed permutations having even number of negative entries in

---

their window notation form 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 arrangement 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 started 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 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 on B_n, given by

$$P(\pi) = \begin{cases} \frac{1}{2n}, & \text{if } \pi = \text{id, the identity element of } B_n, \\ \frac{1}{2n}, & \text{if } \pi = (i, n) \text{ for } 1 \le i \le n - 1, \\ \frac{1}{2n}, & \text{if } \pi = (-i, n) \text{ for } 1 \le i \le n, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

We study the flip-transpose top with random shuffle on B_n using the representation theory of B_n. However, the moves are not the same for elements of the same conjugacy class (i.e., the generating measure does not take the same value at the elements of the same conjugacy class, we abbreviate such shuffle/walk as the non-conjugacy class shuffle/walk). In general, it is not easy to study non-conjugacy class walks on finite groups using the representation theory of the underlying groups. Some examples of such non-conjugacy class walks are *the random-to-top* shuffle (Aldous and Diaconis, 1986; Diaconis et al., 1992) (this is an example of *the Tsetlin library problem* Cetlin, 1963), *the random-to-random* shuffle (Diaconis and Saloff-Coste, 1993; Ayyer et al., 2017; Dieker and Saliola, 2018; Bernstein and Nestoridi, 2019), *the one-sided transposition* shuffle on the symmetric group (Bate et al., 2019), and its generalisation to the hyperoctahedral group (Matheau-Raven, 2020). In this paper, we will show that the flip-transpose top with random shuffle on B_n satisfies the cutoff phenomenon and determine the mixing time for this random walk. Particularly if $||P^{*k} - U_{B_n}||_{\text{TV}}$ denotes the *total variation distance* between the distribution after k transitions and the *stationary distribution*, then the main results of this paper are the following:

Theorem 1.1. *For the flip-transpose top with random shuffle on B_n, we have the following:*

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

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

Theorem 1.2. *For the flip-transpose top with random shuffle on B_n, we have the following:*

(1) *For large n,* $||P^{*k} - U_{B_n}||_{\text{TV}} \ge 1 - \frac{2\left(3 + 3e^{-c} + o(1)(e^{-2c} + e^{-c} + 1)\right)}{(1 + (1 + o(1))e^{-c})^2}$, *when $k = n \log n + cn$ and $c \ll 0$.*

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

We will first recall some concepts and terminologies which we will use in this paper frequently.

1.1. *Representation theoretic 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 the composition of linear mappings. Elements of $GL(V)$ can be thought of as invertible matrices over $\mathbb{C}$. Let G be a finite group, a mapping $\rho : G \to GL(V)$ is said to be a *linear representation* of G if $\rho(g_1g_2) = \rho(g_1)\rho(g_2)$ for all g_1, g_2 in G. The dimension of the vector space V is said to be the *dimension* of the representation ρ and is denoted by d_ρ. V is called the *G-module* corresponding to the representation ρ in this case. If $\mathbb{C}[G] = \{\sum_i c_i g_i \mid c_i \in \mathbb{C},\ g_i \in G\}$, then we define the *right regular representation $R : G \longrightarrow GL(\mathbb{C}[G])$* of G by

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

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 $h \in H$. 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. The representation ρ is *irreducible* if V is non-trivial and V has no non-trivial proper stable subspace. Two representations (ρ_1, V_1) and (ρ_2, V_2) of G are said to be *isomorphic* if there exists an invertible linear map $T : V_1 \to V_2$ such that the following diagram commutes for all $g \in G$:

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

If $V_1 \otimes V_2$ denotes the tensor product of the 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,

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

We will state some results from the representation theory of finite groups without proof. For more details, see Prasad (2015); Sagan (2001); Serre (1977).

1.2. *Random walks on finite groups.* We first recall some terminology. Let p and q be two probability measures on a finite group G. The *Fourier transform $\widehat{p}$* of p at the representation ρ is defined by the matrix $\sum_{x\in G} p(x)\rho(x)$. We define the *convolution $p * q$* of p and q by

$$(p * q)(x) := \sum_{\{u,v\in G|uv=x\}} p(u)q(v).$$

It can be easily seen that $\widehat{(p * q)}(\rho) = \widehat{p}(\rho)\widehat{q}(\rho)$. For the right regular representation R, 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 discrete time Markov chain with state space G and transition probabilities $M_p(x, y) = p(x^{-1}y)$, $x, y \in G$. The transition matrix M_p is the transpose of $\widehat{p}(R)$. If p^{*k} denotes the k-fold convolution of p with itself, then the probability of reaching state y starting from state x using k transitions is $p^{*k}(x^{-1}y)$. The random walk is said to be *irreducible* if given any two states u and v there exists t (depending on u and v) such that $p^{*t}(u^{-1}v) > 0$. We now state the lemma regarding the irreducibility of the random walk on G driven by p.

Lemma 1.3 (Saloff-Coste, 2004, Proposition 2.3). *Let G be a finite group and p be a probability measure on G. The random walk on G driven by p is irreducible if and only if the support of p generates G.*

A *probability vector* (a row vector with non-negative components which sum to one) Π is said to be a *stationary distribution* of the random walk if Π is a left eigenvector of the transition matrix with eigenvalue 1. There exists a unique stationary distribution for each irreducible random walk. If the random walk on G driven by p is irreducible, then the stationary distribution for this random walk is the uniform distribution on G (Saloff-Coste, 2004, Section 2.2). From now on, we denote the uniform distribution on G by U_G. Let us consider a random walk and fix one state $x \in G$. The greatest common divisor of the set of all times when it is possible for the walk to return to the starting state x is said to be the *period* of the state x. All the states of an irreducible random walk have the same period (see Levin et al., 2009, Lemma 1.6). An irreducible random walk is said to be *aperiodic* if the common period for all its states is 1.

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

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

The total variation distance between two discrete distributions μ and ν is half the ℓ_1 distance between them (see Levin et al., 2009, Proposition 4.2). If the random walk on a finite group G driven by a probability measure on G is irreducible and aperiodic, then the distribution after the kth transition converges to the uniform measure on G as $k \to \infty$. We now define the total variation cutoff phenomenon.

Definition 1.4. Let $\{\mathcal{G}_n\}_0^{\infty}$ be a sequence of finite groups and p_n be probability measures on $\mathcal{G}_n$, $n \geq 0$. For each $n \geq 0$, consider the irreducible and aperiodic random walks on $\mathcal{G}_n$ driven by p_n. We say that the *total variation cutoff phenomenon* holds for the family $\{(\mathcal{G}_n, p_n)\}_0^{\infty}$ 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 $k_n = \lfloor (1 + \epsilon)\tau_n \rfloor$, $\lim_{n \to \infty} ||p_n^{*k_n} - U_{\mathcal{G}_n}||_{\mathrm{TV}} = 0$ and

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

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

Informally, we will say that $\{(\mathcal{G}_n, p_n)\}_0^{\infty}$ 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 (Diaconis, 1996).

Proposition 1.5. *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 generates B_n, and hence the chain is irreducible by Lemma 1.3. 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 support of P). 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. $\qquad\square$

Proposition 1.5 says that the flip-transpose top with random shuffle on B_n has unique stationary distribution U_{B_n} and the distribution after the kth transition will converge to its stationary distribution as $k \to \infty$.

The plan of the rest of the paper is as follows: In Section 2, we will find the spectrum of the transition matrix $\widehat{P}(R)$. We will find an upper bound of $||P^{*k} - U_{B_n}||_{\mathrm{TV}}$ for $k \geq n\log n + cn$, $c > 0$ in Section 3. Finally, in Section 4, we will find a lower bound of $||P^{*k} - U_{B_n}||_{\mathrm{TV}}$ for $k = n\log n + cn$, $c \ll 0$ (large negative number) and show that the total variation cutoff for the shuffle on B_n occurs at $n\log n$.

In Appendix A, we give an outline of the irreducible representations of the demi-hyperoctahedral group D_n. We also give an idea for the deduction of irreducible representations of D_n from that of B_n. In Appendix B, we consider a random walk on D_n analogous to the flip-transpose top with random shuffle on B_n and show that this random walk exhibits the total variation cutoff phenomenon with cutoff at $\left(n - \frac{1}{2}\right)\log n$.

2. Spectrum of The Transition Matrix $\widehat{P}(R)$

In this section, we find the eigenvalues of the transition matrix $\widehat{P}(R)$, the Fourier transform of P at the right regular representation R of B_n. To find the eigenvalues of $\widehat{P}(R)$, we will use the representation theory of the hyperoctahedral group B_n. We briefly discuss the representation theory of B_n. For more details, one can see Mishra and Srinivasan (2016); Geissinger and Kinch (1978); Pushkarev (1997).

Definition 2.1. A *partition* λ of a positive integer n is denoted by $\lambda \vdash n$ and is defined by a finite sequence of positive integers $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell > 0$ and $|\lambda| := \sum_{i=1}^{\ell} \lambda_i = n$. The *Young diagram of shape* λ is an arrangement of n boxes into ℓ rows in a left justified way such that the ith row contains λ_i boxes for $1 \leq i \leq \ell$. We use the same notation λ to express both the partition and the Young diagram. The *content* of a box in row i and column j of a Young diagram is the integer $j - i$. Given a Young diagram λ, its *conjugate* λ' is obtained by reflecting λ with respect to the diagonal consisting of boxes with content 0. A *standard Young tableau of shape* λ is a filling of the boxes of the Young diagram of shape λ with the numbers $1, \ldots, n$ such that the numbers are increasing along each row and each column. The set of all standard Young tableaux of shape λ is denoted by $\mathrm{tab}(\lambda)$, and the number of standard Young tableaux of shape λ is denoted by d_λ.
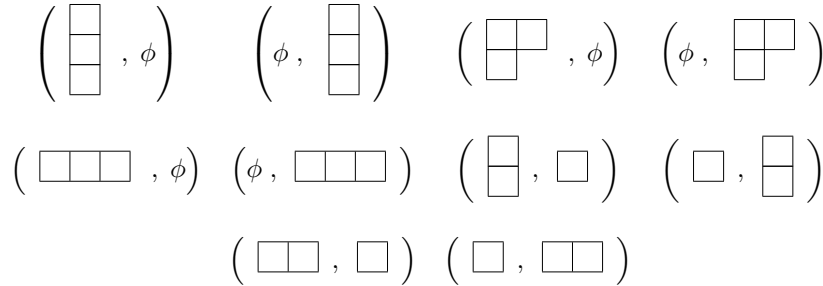
FIGURE 2.1. All elements of $\mathcal{D}_3$.

Definition 2.2. Let n be a positive integer. A (*Young*) *double-diagram* μ *with* n *boxes* is an (ordered) pair of Young diagrams such that the total number of boxes is n. We define $||\mu|| = n$. The set of all double-diagrams with n boxes is denoted by $\mathcal{D}_n$. For example, the double-diagrams with 3 boxes are listed in Figure 2.1. 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 $\mathrm{tab}_{\mathcal{D}}(n, \mu)$, where $\mu \in \mathcal{D}_n$, denote the set of all standard double-tableaux of shape μ and let $\mathrm{tab}_{\mathcal{D}}(n) = \bigcup_{\mu \in \mathcal{D}_n} \mathrm{tab}_{\mathcal{D}}(n, \mu)$.

For example an element of $\mathrm{tab}_{\mathcal{D}}(8)$ is given in Figure 2.2. Let $T \in \mathrm{tab}_{\mathcal{D}}(n, \mu)$ and $i \in [n]$. Let $b_T(i)$ be 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 2.2, we have $c(b_T(1)) = 0$, $c(b_T(2)) = 1$, $c(b_T(3)) = 0$, $c(b_T(4)) = 1$, $c(b_T(5)) = -1$, $c(b_T(6)) = -1$, $c(b_T(7)) = 0$, $c(b_T(8)) = 2$.
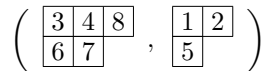
$$\left(\begin{array}{|c|c|c|} \hline 3 & 4 & 8 \\ \hline 6 & 7 \\ \cline{1-2} \end{array} \quad , \quad \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 5 \\ \cline{1-1} \end{array} \right)$$

FIGURE 2.2. An element of $\mathrm{tab}_{\mathcal{D}}(8)$.

Definition 2.3. The *Young-Jucys-Murphy* elements $X_1, X_2, \ldots, X_n$ of $\mathbb{C}[B_n]$ are defined by $X_1 = 0$ and $X_i = \sum_{k=1}^{i-1}(k, i) + \sum_{k=1}^{i-1}(-k, i)$, for all $2 \leq i \leq n$.

Definition 2.4. Let $\mu \in \widehat{B}_n$ (set of all irreducible representations of B_n) and consider the B_n-module V^μ. Since the branching is simple (Mishra and Srinivasan, 2016, 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-modules, we obtain

$$V^\mu = \bigoplus_T v_T, \qquad (2.1)$$

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 call (2.1) the *Gelfand-Tsetlin* decomposition of V^μ and each v_T in (2.1) a *Gelfand-Tsetlin* vector of V^μ. We note that if $0 \neq v_T$, then $\mathbb{C}[B_i]v_T = V^{\mu_i}$. The Gelfand-Tsetlin vectors of V^μ form a basis of V^μ.

The irreducible representations of B_n are parametrised by elements of $\mathcal{D}_\mathrm{n}$ (Mishra and Srinivasan, 2016, Lemma 6.2, Theorem 6.4). We may index the Gelfand-Tsetlin vectors of V^μ by standard double-tableaux of shape μ for $\mu \in \mathcal{D}_\mathrm{n}$ (Mishra and Srinivasan, 2016, 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 (Mishra and Srinivasan, 2016, Theorem 6.5) of the Young-Jucys-Murphy elements X_i and the signed permutation $(i, -i)$ on v_T are given by

$$X_i\, v_T = 2c(b_T(i))\, v_T \text{ for all } i \in [n],$$

$$(-i, i)\, v_T = \begin{cases} v_T & \text{if } b_T(i) \text{ is in } \mu^{(1)} \\ -v_T & \text{if } b_T(i) \text{ is in } \mu^{(2)} \end{cases} \text{ for all } i \in [n]. \tag{2.2}$$

Remark 2.5. The components of the elements of $\mathcal{D}_\mathrm{n}$ are indexed by the irreducible representations of the cyclic group of order two. In this paper, we adopt the convention that the first component of the elements of $\mathcal{D}_\mathrm{n}$ is indexed by the trivial representation.

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

Theorem 2.6. *For each $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \mathcal{D}_\mathrm{n}$ satisfying*

$$m := |\mu^{(1)}| \in \left\{0, 1, \ldots, \left\lfloor \frac{n}{2} \right\rfloor\right\},$$

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 $\widehat{P}(R)$ with multiplicity $M(\mu)$ each, where

$$M(\mu) = \begin{cases} \binom{n}{m} d_{\mu^{(1)}} d_{\mu^{(2)}}, & \text{if } 0 \leq m < \frac{n}{2}, \\ \frac{1}{2}\binom{n}{m} d_{\mu^{(1)}} d_{\mu^{(2)}}, & \text{if } m = \frac{n}{2} \text{ (when } n \text{ is even).} \end{cases} \tag{2.3}$$

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 $\widehat{P}(R)$ 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 $\tilde{\mu}^{(2)}$. Without loss of generality, let us assume that $b_T(n)$ is in $\mu^{(1)}$ and $b_{\widetilde{T}}(n)$ is in $\tilde{\mu}^{(2)}$. Let us recall v_T (respectively $v_{\widetilde{T}}$) is the Gelfand-Tsetlin vector of V^μ (respectively $V^{\tilde{\mu}}$). From (2.2) we have $(-n, n)\, v_T = v_T$ and $X_n\, v_T = 2c(b_T(n))\, v_T$, which implies the following:

$$\left(\mathrm{id} + (-n, n) + X_n\right)\, v_T = (1 + 1 + 2c(b_T(n)))\, v_T = (2c(b_T(n)) + 2)\, v_T. \tag{2.4}$$

Since $\{v_T : T \in \mathrm{tab}_\mathcal{D}(n,\mu)\}$ form a basis of V^μ, the eigenvalues of the action of $\left(\mathrm{id} + (-n, n) + X_n\right)$ on V^μ can be obtained from (2.4). Now using (2.2) again we

have $(-n, n)\, v_{\widetilde{T}} = -v_{\widetilde{T}}$ and $X_n\, v_{\widetilde{T}} = 2c(b_{\widetilde{T}}(n))\, v_{\widetilde{T}}$, thus

$$(\mathrm{id} + (-n, n) + X_n)\, v_{\widetilde{T}} = (1 - 1 + 2c(b_T(n)))\, v_{\widetilde{T}} = 2c(b_T(n))\, v_{\widetilde{T}}. \qquad (2.5)$$

Therefore the eigenvalues of the action of $(\mathrm{id} + (-n, n) + X_n)$ on $V^{\tilde{\mu}}$ are obtained from (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 $\frac{1}{2n}(\mathrm{id} + (-n, n) + X_n)$ on V^{μ} and $V^{\tilde{\mu}}$ simultaneously, the eigenvalues of $\widehat{P}(R)$ 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 $\dim(V^{\mu}) = \binom{n}{m} d_{\mu^{(1)}} d_{\mu^{(2)}} = \dim(V^{\tilde{\mu}})$ if $0 \le m < \frac{n}{2}$ and the multiplicity of the eigenvalues are $\frac{1}{2}\binom{n}{m} d_{\mu^{(1)}} d_{\mu^{(2)}}$ 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 $\widehat{P}(R)$ 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 parameterised by $\mathcal{D}_{\mathrm{n}}$. $\qquad \square$

3. Upper bound of total variation distance

In this section, we will prove the theorem giving an upper bound of the total variation distance $||P^{*k} - U_{B_n}||_{\mathrm{TV}}$ for $k \ge n \log n + cn$, $c > 0$. Given a positive integer ℓ, throughout this section we write $\lambda \vdash \ell$ to denote λ is a partition of ℓ. Let us recall that $\mathrm{tab}(\lambda)$ denote the set of all standard Young tableaux of shape λ.

Lemma 3.1 (Upper bound lemma, Diaconis, 1988a, 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. Then we have the following*

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

where the sum is over all non-trivial irreducible representations D_{ρ} of G and d_{ρ} is the dimension of D_{ρ}.

Lemma 3.2. *Let m be any positive integer satisfying $1 \le m \le \frac{n}{2}$ 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*

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

with $x = 0, 1$.

Proof: The set $\mathrm{tab}_{\mathcal{D}}(n, \mu)$ is a disjoint union of the sets

$$\mathcal{T}_1 = \{(T_1, T_2) \in \mathrm{tab}_{\mathcal{D}}(n, \mu) : b_T(n) \text{ is in } T_1\} \text{ and}$$
$$\mathcal{T}_2 = \{(T_1, T_2) \in \mathrm{tab}_{\mathcal{D}}(n, \mu) : b_T(n) \text{ is in } T_2\}.$$

Therefore we have

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

(3.1)

Now the right hand side of (3.1) is equal to

$$\binom{n-1}{n-m}d_{\mu^{(2)}}\sum_{T_1\in\mathrm{tab}(\mu^{(1)})}\left(\frac{c(b_{T_1}(m))+x}{n}\right)^{2k}$$

$$+\binom{n-1}{m}d_{\mu^{(1)}}\sum_{T_2\in\mathrm{tab}(\mu^{(2)})}\left(\frac{c(b_{T_2}(n-m))+x}{n}\right)^{2k}$$

$$<\binom{n}{m}\left(d_{\mu^{(2)}}\sum_{T_1\in\mathrm{tab}(\mu^{(1)})}\left(\frac{c(b_{T_1}(m))+x}{n}\right)^{2k}+d_{\mu^{(1)}}\sum_{T_2\in\mathrm{tab}(\mu^{(2)})}\left(\frac{c(b_{T_2}(n-m))+x}{n}\right)^{2k}\right)$$

$$\leq\binom{n}{m}d_{\mu^{(2)}}d_{\mu^{(1)}}\left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k}+\left(\frac{\mu_1^{(1)'}}{n}\right)^{2k}+\left(\frac{\mu_1^{(2)}}{n}\right)^{2k}+\left(\frac{\mu_1^{(2)'}}{n}\right)^{2k}\right).$$

(3.2)

The inequality in (3.2) 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

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

$$< \left(\frac{\lambda_1}{n}\right)^{2k}+\left(\frac{\lambda_1'}{n}\right)^{2k}, \quad \text{for } x=0,1.$$

□

Lemma 3.3. *Let ℓ be a positive integer. For a partition λ of ℓ, if λ_1 denotes the largest part of λ, then*

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

Proof: For any partition ζ of $\ell-\lambda_1$ with largest part ζ_1 less than or equal to λ_1, we have $d_\lambda\leq\binom{\ell}{\lambda_1}d_\zeta$. Therefore $\displaystyle\sum_{\lambda\vdash\ell}d_\lambda^2\left(\frac{\lambda_1}{\ell}\right)^{2k}$ is less than or equal to

$$\sum_{\lambda_1=1}^{\ell}\sum_{\substack{\zeta\vdash(\ell-\lambda_1)\\\zeta_1\leq\lambda_1}}\binom{\ell}{\lambda_1}^2 d_\zeta^2\left(\frac{\lambda_1}{\ell}\right)^{2k} \leq \sum_{\lambda_1=1}^{\ell}\binom{\ell}{\lambda_1}^2\left(\frac{\lambda_1}{\ell}\right)^{2k}\sum_{\zeta\vdash(\ell-\lambda_1)}d_\zeta^2$$

$$= \sum_{\lambda_1=1}^{\ell}\binom{\ell}{\ell-\lambda_1}^2(\ell-\lambda_1)!\left(1-\frac{\ell-\lambda_1}{\ell}\right)^{2k}. \quad (3.3)$$

Now writing $t = \ell - \lambda_1$ and using $1 - x \le e^{-x}$ for $x \ge 0$, the expression in (3.3) less than or equal to $\sum_{t=0}^{\ell-1} \binom{\ell}{t}^2 t! e^{-\frac{2kt}{\ell}}$. Thus we have

$$\sum_{\lambda \vdash \ell} d_\lambda^2 \left(\frac{\lambda_1}{\ell}\right)^{2k} \le \sum_{t=0}^{\ell-1} \binom{\ell}{t}^2 t! e^{-\frac{2kt}{\ell}} = \sum_{t=0}^{\ell-1} \frac{(\ell(\ell-1)\dots(\ell-t+1))^2}{t!} e^{-\frac{2kt}{\ell}}$$

$$\le \sum_{t=0}^{\ell-1} \frac{\left(\ell^2 e^{-\frac{2k}{\ell}}\right)^t}{t!} < e^{\ell^2 e^{-\frac{2k}{\ell}}}. \qquad \square$$

Proof of Theorem 1.1: 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 3.1 implies $4||P^{*k} - U_{B_n}||_{\mathrm{TV}}^2$ is bounded above by the sum of $(2k)$th powers of the non-largest eigenvalues (which are strictly less the largest eigenvalue 1) of $\widehat{P}(R)$. Thus from Theorem 2.6, we have

$$4||P^{*k} - U_{B_n}||_{\mathrm{TV}}^2$$

$$\le \left(\frac{n-1}{n}\right)^{2k} + \sum_{\substack{\lambda \vdash n \\ \lambda \ne (n)}} d_\lambda \left(\sum_{T \in \mathrm{tab}(\lambda)} \left(\left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \left(\frac{c(b_T(n))}{n}\right)^{2k}\right)\right)$$

$$+ \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{\substack{\mu^{(1)} \vdash m \\ \mu^{(2)} \vdash (n-m) \\ \mu = (\mu^{(1)}, \mu^{(2)})}} M(\mu) \left(\sum_{T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)} \left(\left(\frac{c(b_T(n))+1}{n}\right)^{2k} + \left(\frac{c(b_T(n))}{n}\right)^{2k}\right)\right).$$

$$(3.4)$$

$M(\mu)$ is defined in (2.3) and can be written as $M(\mu) = \mathbb{I}(n,m)\binom{n}{m} d_{\mu^{(1)}} d_{\mu^{(2)}}$, where

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

Using Lemma 3.2, the third term in the right hand side of (3.4) is less than the following expression

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

$$= 2\sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{\substack{\mu^{(1)} \vdash m \\ \mu^{(2)} \vdash (n-m) \\ \mu = (\mu^{(1)}, \mu^{(2)})}} 2M(\mu) \binom{n}{m} d_{\mu^{(2)}} d_{\mu^{(1)}} \left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(2)}}{n}\right)^{2k}\right) \qquad (3.5)$$

$$= 4\sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \mathbb{I}(n,m) \binom{n}{m}^2 \sum_{\substack{\mu^{(1)} \vdash m \\ \mu^{(2)} \vdash (n-m)}} d_{\mu^{(1)}}^2 d_{\mu^{(2)}}^2 \left(\left(\frac{\mu_1^{(1)}}{n}\right)^{2k} + \left(\frac{\mu_1^{(2)}}{n}\right)^{2k}\right)$$

$$= 4 \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \mathbb{I}(n,m) \binom{n}{m}^2 \left((n-m)! \sum_{\mu^{(1)} \vdash m} d^2_{\mu^{(1)}} \left(\frac{\mu_1^{(1)}}{n} \right)^{2k} + m! \sum_{\mu^{(2)} \vdash (n-m)} d^2_{\mu^{(2)}} \left(\frac{\mu_1^{(2)}}{n} \right)^{2k} \right).$$

$$(3.6)$$

The equality in (3.5) holds because

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

for $i = 1, 2$. Now the definition of $\mathbb{I}(n,m)$ and

$$\sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \mathbb{I}(n,m) \binom{n}{m}^2 m! \sum_{\mu^{(2)} \vdash (n-m)} d^2_{\mu^{(2)}} \left(\frac{\mu_1^{(2)}}{n} \right)^{2k}$$

$$= \sum_{t=\lceil \frac{n}{2} \rceil}^{n-1} \mathbb{I}(n, n-t) \binom{n}{n-t}^2 (n-t)! \sum_{\mu^{(2)} \vdash t} d^2_{\mu^{(2)}} \left(\frac{\mu_1^{(2)}}{n} \right)^{2k}$$

implies that the expression (3.6) is equal to

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

$$(3.7)$$

Replacing ℓ (respectively λ) by m (respectively $\mu^{(1)}$) in Lemma 3.3, we have

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

Therefore when $k \geq n \log n$ (which implies $k \geq m \log m$), the expression in (3.7) and hence the third term in the right hand side of (3.4) is less than

$$4e \sum_{m=1}^{n-1} \binom{n}{m}^2 (n-m)! \left(\frac{m}{n} \right)^{2k} = 4e \sum_{t=1}^{n-1} \binom{n}{t}^2 t! \left(1 - \frac{t}{n} \right)^{2k}$$

$$< 4e \sum_{t=1}^{n-1} \frac{\left(n^2 e^{-\frac{2k}{n}} \right)^t}{t!} < 4e \left(e^{n^2 e^{-\frac{2k}{n}}} - 1 \right). \quad (3.8)$$

Now we consider the second term in the right hand side of (3.4). The second term in the right hand side of (3.4) is bounded above by

$$2 \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n),(1^n)}} d^2_\lambda \left(\left(\frac{\lambda_1}{n} \right)^{2k} + \left(\frac{\lambda_1'}{n} \right)^{2k} \right) + \left(\frac{n-2}{n} \right)^{2k} + \left(\frac{n-1}{n} \right)^{2k}. \quad (3.9)$$

Now using $\displaystyle\sum_{\substack{\lambda \vdash n \\ \lambda \neq (n),(1^n)}} d_\lambda^2 \left(\frac{\lambda_1'}{n}\right)^{2k} = \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n),(1^n)}} d_\lambda^2 \left(\frac{\lambda_1}{n}\right)^{2k}$, the expression in (3.9) is equal to

$$4 \sum_{\substack{\lambda \vdash n \\ \lambda \neq (n),(1^n)}} d_\lambda^2 \left(\frac{\lambda_1}{n}\right)^{2k} + \left(1 - \frac{2}{n}\right)^{2k} + \left(1 - \frac{1}{n}\right)^{2k}$$

$$< 4\left(\sum_{\lambda \vdash n} d_\lambda^2 \left(\frac{\lambda_1}{n}\right)^{2k} - 1\right) + e^{-\frac{4k}{n}} + e^{-\frac{2k}{n}}. \tag{3.10}$$

The right hand side of the expression (3.10) and hence the second term in the right hand side of (3.4) is less than $4\left(e^{n^2 e^{-\frac{2k}{n}}} - 1\right) + e^{-\frac{2k}{n}} + e^{-\frac{4k}{n}}$ by Lemma 3.3. Thus the inequality (3.4) becomes

$$4||P^{*k} - U_{B_n}||_{\text{TV}}^2 \leq 2e^{-\frac{2k}{n}} + (4 + 4e)\left(e^{n^2 e^{-\frac{2k}{n}}} - 1\right) + e^{-\frac{4k}{n}}, \quad \text{for } k \geq n \log n. \tag{3.11}$$

Now if $k \geq n \log n + cn$ and $c > 0$, then the right hand side of (3.11) becomes

$$(4e + 4)\left(e^{e^{-2c}} - 1\right) + \frac{2e^{-2c}}{n^2} + \frac{e^{-4c}}{n^4} < (8e + 8)e^{-2c} + o(1).$$

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 (3.11) is bounded above by

$$(4e + 4)\left(e^{\frac{1}{n^{2\epsilon}}} - 1\right) + 2n^{-2(1+\epsilon)} + n^{-4(1+\epsilon)}.$$

Therefore the proof of the second part follows from

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

4. Lower bound of total variation distance

In this section, we will find a lower bound of the total variation distance $||P^{*k} - U_{B_n}||_{\text{TV}}$ for $k = \log n + cn$, $c \ll 0$. We define a group homomorphism from B_n onto the symmetric group S_n which projects the flip-transpose top with random shuffle on B_n to the transpose top with random shuffle on S_n. We begin with a variant of the transpose top with random shuffle on S_n. This will be useful in obtaining the lower bound of $||P^{*k} - U_{B_n}||_{\text{TV}}$.

Given $0 < a < 1$, we define a probability measure $\mathcal{P}_a$ on the symmetric group S_n as follows:

$$\mathcal{P}_a(\pi) = \begin{cases} a & \text{if } \pi = s_{\text{id}}, \text{ the identity element of } S_n, \\ \frac{1-a}{n-1} & \text{if } \pi = s_{(i,n)}, \ 1 \leq i < n, \end{cases} \quad \text{for } \pi \in S_n, \tag{4.1}$$

where $s_{(i,n)}$ denotes the transposition in S_n interchanging i and n. We prove a theorem which provides a lower bound of $||\mathcal{P}_a^{*k} - U_{S_n}||_{\text{TV}}$. Although the proof is straightforward and uses techniques (Diaconis, 1988a, Chapter 5(C), p. 27) from the transpose top with random shuffle on S_n, we prove it to make this paper self contained. Recall that $\widehat{\mathcal{P}}_a(R)$ denotes the Fourier transform of $\mathcal{P}_a$ at the right regular representation R of S_n. Thus $\widehat{\mathcal{P}}_a(R)$ is the transition matrix for the random

walk on S_n driven by $\mathcal{P}_a$ (since $\widehat{\mathcal{P}}_a(R)$ is symmetric). If we denote the nth Young-Jucys-Murphy element of $\mathbb{C}[S_n]$ using notation X'_n, then $\widehat{\mathcal{P}}_a(R)$ is the action of $as_{\mathrm{id}} + \frac{1-a}{n-1}X'_n$ on $\mathbb{C}[S_n]$ by multiplication on the right. Moreover given an irreducible S_n-module (Specht module) S^λ indexed by $\lambda \vdash n$, the actions of X'_n on the Gelfand-Tsetlin basis vectors of S^λ are given as follows: $X'_n(u_T) = c(b_T(n))u_T$, where u_T denotes the Gelfand-Tsetlin basis vector of S^λ indexed by $T \in \mathrm{tab}(\lambda)$ and recall that $c(b_T(n))$ is the content of the box containing n in T (Okounkov and Vershik, 1996; Ceccherini-Silberstein et al., 2010). Let us define a random variable $\mathfrak{f}$ on S_n given as follows:

$$\mathfrak{f}(\pi) := \text{numbers of fixed points of } \pi, \text{ for } \pi \in S_n.$$

The expected value of $\mathfrak{f}$ with respect to the uniform distribution U_{S_n} is given by $E_U(\mathfrak{f}) = 1$ (Diaconis, 1988a, p.=27, eq. (5.12)). Moreover if $\widehat{\mathcal{P}}_a(R)\big|_\lambda$ denotes the restriction of $\widehat{\mathcal{P}}_a(R)$ to the irreducible S_n-module S^λ, then the expected value of $\mathfrak{f}$ with respect to the distribution $\mathcal{P}_a^{*k}$ is given by

$$E_{a,k}(\mathfrak{f}) = \mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n)}\right)^k\right) + \mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n-1,1)}\right)^k\right),$$

$$\text{Diaconis, 1988a, p.=28, eq.(5.13)}$$

$$= 1 + (n-2)\left(1 - \frac{1-a}{n-1}\right)^k + \left(\frac{an-1}{n-1}\right)^k, \quad \text{see Table 4.1}$$

$$\approx 1 + (n-2)e^{-\left(\frac{1-a}{n-1}\right)k} + \left(\frac{an-1}{n-1}\right)^k. \tag{4.2}$$

Here '$\approx$' means 'asymptotic to' i.e. $a_n \approx b_n$ means $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 1$. The expectation of $\mathfrak{f}^2$ with respect to the distribution $\mathcal{P}_a^{*k}$ is given by

$$E_{a,k}(\mathfrak{f}^2) = 2\,\mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n)}\right)^k\right) + 3\,\mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n-1,1)}\right)^k\right)$$

$$+ \mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n-2,2)}\right)^k\right) + \mathrm{Tr}\left(\left(\widehat{\mathcal{P}}_a(R)\big|_{(n-2,1,1)}\right)^k\right),$$

$$\text{Diaconis, 1988a, p.=28, eq.(5.14)}$$

$$= 2 + 3\left((n-2)\left(1 - \frac{1-a}{n-1}\right)^k + \left(\frac{an-1}{n-1}\right)^k\right)$$

$$+ \left(\frac{(n-1)(n-4)}{2}\left(1 - \frac{2(1-a)}{n-1}\right)^k + (n-2)a^k\right)$$

$$+ \left(\frac{(n-2)(n-3)}{2}\left(1 - \frac{2(1-a)}{n-1}\right)^k + (n-2)\left(\frac{an+a-2}{n-1}\right)^k\right),$$

$$\text{see Table 4.1}$$

$$\approx 2 + 3(n-2)e^{-\left(\frac{1-a}{n-1}\right)k} + (n^2 - 5n + 5)e^{-2\left(\frac{1-a}{n-1}\right)k}$$

$$+ 3\left(\frac{an-1}{n-1}\right)^k + (n-2)\left(a^k + \left(\frac{an+a-2}{n-1}\right)^k\right). \tag{4.3}$$

Partition of n	Eigenvalues of $\widehat{\mathcal{P}}_a(R)$ corresponding to the partition of column 1
(n)	1 with algebraic multiplicity 1
$(n-1,1)$	$\frac{n+a-2}{n-1}$ with algebraic multiplicity $n-2$ $\frac{an-1}{n-1}$ with algebraic multiplicity 1
$(n-2,2)$	$\frac{n+2a-3}{n-1}$ with algebraic multiplicity $\frac{(n-1)(n-4)}{2}$ a with algebraic multiplicity $n-2$
$(n-2,1,1)$	$\frac{n+2a-3}{n-1}$ with algebraic multiplicity $\frac{(n-2)(n-3)}{2}$ $\frac{an+a-2}{n-1}$ with algebraic multiplicity $n-2$

TABLE 4.1. Eigenvalues of $\widehat{\mathcal{P}}_a(R)\big|_\lambda$ for $\lambda = (n), (n-1,1), (n-2,2)$, and $(n-2,1,1)$.

Proposition 4.1. *For the random walk on S_n driven by $\mathcal{P}_a$, we have the following:*

$$||\mathcal{P}_a^{*k} - U_{S_n}||_{\mathrm{TV}} \geq 1 - \frac{4\left(E_{a,k}\left(\mathfrak{f}^2\right) - \left(E_{a,k}\left(\mathfrak{f}\right)\right)^2\right)}{\left(E_{a,k}(\mathfrak{f})\right)^2} - \frac{2}{E_{a,k}(\mathfrak{f})},$$

where $E_{a,k}(\mathfrak{f})$ and $E_{a,k}(\mathfrak{f}^2)$ are given in (4.2) *and* (4.3) *respectively.*

Proof: If $\mathrm{Var}_{a,k}(\mathfrak{f})$ denotes the variance of $\mathfrak{f}$ with respect to the probability measure $\mathcal{P}_a^{*k}$, then Chebychev's inequality implies that

$$\mathcal{P}_a^{*k}\left(\left\{\pi \in S_n : |\mathfrak{f}(\pi) - E_{a,k}(\mathfrak{f})| \leq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right) \geq 1 - \frac{4\,\mathrm{Var}_{a,k}(\mathfrak{f})}{\left(E_{a,k}(\mathfrak{f})\right)^2}. \qquad (4.4)$$

Again using $\mathfrak{f} \geq 0$, $E_U(\mathfrak{f}) = 1$ and the Markov's inequality, we have

$$U_{S_n}\left(\left\{\pi \in S_n : \mathfrak{f}(\pi) \geq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right) \leq \frac{2E_U(\mathfrak{f})}{E_{a,k}(\mathfrak{f})} = \frac{2}{E_{a,k}(\mathfrak{f})}. \qquad (4.5)$$

Now from the definition of total variation distance, we have

$$||\mathcal{P}_a^{*k} - U_{S_n}||_{\mathrm{TV}} = \sup_{A \subset S_n} |\mathcal{P}_a^{*k}(A) - U_{S_n}(A)|$$

$$\geq \mathcal{P}_a^{*k}\left(\left\{\pi \in S_n : |\mathfrak{f}(\pi) - E_{a,k}(\mathfrak{f})| \leq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right)$$

$$- U_{S_n}\left(\left\{\pi \in S_n : |\mathfrak{f}(\pi) - E_{a,k}(\mathfrak{f})| \leq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right)$$

$$\geq \mathcal{P}_a^{*k}\left(\left\{\pi \in S_n : |\mathfrak{f}(\pi) - E_{a,k}(\mathfrak{f})| \leq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right)$$

$$- U_{S_n}\left(\left\{\pi \in S_n : \mathfrak{f}(\pi) \geq \frac{E_{a,k}(\mathfrak{f})}{2}\right\}\right). \qquad (4.6)$$

Therefore the proposition follows from (4.6), (4.5), (4.4), and the definition of variance.   □

We now come back to our main objective of this section i.e., computation of a lower bound of $||P^{*k} - U_{B_n}||_{\mathrm{TV}}$. Let us define a homomorphism f from B_n onto S_n as follows: For $\pi \in B_n$,

$$f : \pi \mapsto (f(\pi) : i \mapsto |\pi(i)|, \text{ for } 1 \leq i \leq n). \tag{4.7}$$

i.e., $f(\pi) \in S_n$ sends i to $|\pi(i)|$ for $1 \leq i \leq n$. Here $|\pi(i)|$ denotes the absolute value of $\pi(i)$. It can be checked that the mapping f defined in (4.7) is a homomorphism (this follows directly by considering B_n as the *wreath product* $S_2 \wr S_n$). The surjectivity of f follows from the definition. The homomorphism f projects the flip-transpose top with random shuffle on B_n to the transpose top with random shuffle on S_n i.e., $Pf^{-1} = \mathcal{P}_{\frac{1}{n}}$. We now prove a lemma which will be useful in proving the main result of this section.

Remark 4.2. Although we prove the upcoming lemma for the probability distribution P on B_n, it is true if P is replaced by any other probability distribution on B_n.

Lemma 4.3. *For any positive integer k we have $\left(Pf^{-1}\right)^{*k} = P^{*k}f^{-1}$.*

Proof: We use the first principle of mathematical induction on k. The base case for $k = 1$ is true by definition. Now assume the induction hypothesis i.e., $\left(Pf^{-1}\right)^{*m} = P^{*m}f^{-1}$ for some positive integer $m > 1$. Let $\pi \in S_n$ be chosen arbitrarily. Then for the inductive step $k = m + 1$ we have the following:

$$\left(Pf^{-1}\right)^{*(m+1)}(\pi) = \left(\left(Pf^{-1}\right) * \left(Pf^{-1}\right)^{*m}\right)(\pi)$$

$$= \sum_{\{\xi,\zeta \in S_n \,:\, \xi\zeta=\pi\}} \left(Pf^{-1}\right)(\xi)\left(Pf^{-1}\right)^{*m}(\zeta)$$

$$= \sum_{\{\xi,\zeta \in S_n \,:\, \xi\zeta=\pi\}} \left(Pf^{-1}\right)(\xi)\left(P^{*m}f^{-1}\right)(\zeta),$$

$$\text{by the induction hypothesis,}$$

$$= \sum_{\substack{\xi,\zeta \in S_n \\ \xi\zeta=\pi}} P\left(f^{-1}(\xi)\right) P^{*m}\left(f^{-1}(\zeta)\right)$$

$$= \sum_{\substack{\xi,\zeta \in S_n \\ \xi\zeta=\pi}} \sum_{\substack{\xi' \in f^{-1}(\xi) \\ \zeta' \in f^{-1}(\zeta)}} P(\xi')P^{*m}(\zeta'). \tag{4.8}$$

Now using the fact that f is a homomorphism, we have the following:

$$\{(\xi',\zeta') \in f^{-1}(\xi) \times f^{-1}(\zeta) : \xi,\zeta \in S_n \text{ and } \xi\zeta = \pi\}$$
$$= \{(\xi',\zeta') \in B_n \times B_n : \xi'\zeta' \in f^{-1}(\pi)\}.$$

Therefore the expression in (4.8) becomes

$$\sum_{\{\xi',\zeta' \in B_n \,:\, \xi'\zeta' \in f^{-1}(\pi)\}} P(\xi')P^{*m}(\zeta') = \sum_{\pi' \in f^{-1}(\pi)} \sum_{\substack{\xi',\zeta' \in B_n \\ \xi'\zeta'=\pi'}} P(\xi')P^{*m}(\zeta')$$

$$= \sum_{\pi' \in f^{-1}(\pi)} P^{*(m+1)}(\pi') = \left(P^{*(m+1)} f^{-1} \right)(\pi).$$

Thus the lemma follows from the first principle of mathematical induction. $\qquad \square$

Proof of Theorem 1.2: We know that, given two probability distributions μ and ν on Ω and a mapping $\psi : \Omega \to \Lambda$, we have $||\mu - \nu||_{TV} \geq ||\mu\psi^{-1} - \nu\psi^{-1}||_{TV}$, where Λ is finite (Levin et al., 2009, Lemma 7.9). Therefore we have the following:

$$||P^{*k} - U_{B_n}||_{\mathrm{TV}} \geq ||P^{*k} f^{-1} - U_{B_n} f^{-1}||_{\mathrm{TV}}$$

$$= || \left(Pf^{-1} \right)^{*k} - U_{S_n}||_{\mathrm{TV}}, \text{ by Lemma (4.3) and } U_{B_n} f^{-1} = U_{S_n},$$

$$= ||\mathcal{P}^{*k}_{\frac{1}{n}} - U_{S_n}||_{\mathrm{TV}}, \text{ using } Pf^{-1} = \mathcal{P}_{\frac{1}{n}}. \tag{4.9}$$

Now setting $a = \frac{1}{n}$ in (4.2) and (4.3) we have

$$E_{\frac{1}{n},k}(\mathfrak{f}) \approx 1 + (n-2)e^{-\frac{k}{n}}.$$

$$E_{\frac{1}{n},k}(\mathfrak{f}^2) \approx 2 + 3(n-2)e^{-\frac{k}{n}} + (n^2 - 5n + 5)e^{-\frac{2k}{n}} + (n-2)\left(\frac{1 + (-1)^k}{n^k} \right).$$

Therefore Proposition 4.1 and (4.9) implies that

$$||P^{*k} - U_{B_n}||_{\mathrm{TV}} \geq 1 - \frac{2 \left(3 + 3(n-2)e^{-\frac{k}{n}} - 2(n-1)e^{-\frac{2k}{n}} + o(1) \right)}{\left(1 + (n-2)e^{-\frac{k}{n}} \right)^2}, \quad \text{for } k > 1. \tag{4.10}$$

Now if n is large, $c \ll 0$ and $k = n \log n + cn$, then by (4.10), 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 (4.10), we have

$$1 \geq ||P^{*k_n} - U_{B_n}||_{\mathrm{TV}} \geq 1 - \frac{2 \left(3 + 3n^\epsilon + o(1)(n^{2\epsilon} + n^\epsilon + 1) \right)}{(1 + (1 + o(1))n^\epsilon)^2}, \tag{4.11}$$

for large n. Therefore, the second part of this theorem follows from (4.11) and the fact that

$$\lim_{n \to \infty} \frac{2 \left(3 + 3n^\epsilon + o(1)(n^{2\epsilon} + n^\epsilon + 1) \right)}{(1 + (1 + o(1))n^\epsilon)^2} = 0.$$

$\qquad \square$

Therefore from the first part of Theorems 1.1 and 1.2, we can say that the mixing time for the flip-transpose top with random shuffle on B_n is $O(n \log n)$ (i.e., order of $n \log n$). Furthermore, the second part of Theorems 1.1 and 1.2 implies that this shuffle satisfies the cutoff phenomenon and the total variation cutoff for this shuffle occurs at $n \log n$.

Remark 4.4. Let $0 \leq \alpha \leq 1$. A generalisation of the flip-transpose top with random shuffle on B_n can be considered, which we call *the biased flip-transpose top with random shuffle on B_n*. Given an arrangement of n distinct oriented cards in a row, choose a card uniformly at random and choose the last card. Then perform one of the following moves:

(1) Transpose the chosen cards with probability $\frac{\alpha}{2}$.
(2) Transpose the chosen cards after flipping both the cards with probability $\frac{\alpha}{2}$.

(3) Transpose the chosen cards after flipping one of the cards with probability $\frac{1-\alpha}{2}$.

This is the random walk on B_n driven by the probability measure P_α on B_n, defined below.

$$P_\alpha(\pi) = \begin{cases} \frac{1}{n} \cdot \frac{\alpha}{2}, & \text{if } \pi = (i,n) \text{ or } (-i,n) \text{ for } 1 \le i \le n, \text{ here } (n,n) := \text{id}, \\ \frac{1}{n} \cdot \frac{1-\alpha}{2}, & \text{if } \pi = (-n,n)(i,n) \text{ or } (-i,i)(i,n) \text{ for } 1 \le i \le n, \\ 0, & \text{otherwise.} \end{cases}$$

(4.12)

The Fourier transform $\widehat{P}_\alpha(R)$ of P_α at the right regular representation R is the transition matrix for this biased variant. Recall that the nth Young-Jucys-Murphy element of B_n is X_n. Then $\widehat{P}_\alpha(R)$ is the action of

$$\frac{1}{2n} \left(\alpha \operatorname{id} + (1-\alpha)(-n,n) \right) \left(\operatorname{id} + (-n,n) + X_n \right)$$

on $\mathbb{C}[B_n]$ by multiplication on the right. It can be easily seen that $\widehat{P}_\alpha(R)$ and $\widehat{P}(R)$ have the same set of eigenvectors when they act on the irreducible B_n-modules. Therefore using the arguments given in the proof of Theorem 2.6, we can obtain the eigenvalues of $\widehat{P}_\alpha(R)$ as follows: *For each $\mu = \left(\mu^{(1)}, \mu^{(2)} \right) \in \mathcal{D}_n$ satisfying $m := |\mu^{(1)}| \in \{0, 1, \ldots, \lfloor \frac{n}{2} \rfloor\}$, let $T \in \operatorname{tab}_\mathcal{D}(n, \mu)$. Then $\frac{c(b_T(n))+1}{n}$ and $\frac{c(b_T(n))}{n}(2\alpha - 1)$ are eigenvalues of $\widehat{P}_\alpha(R)$ with multiplicity $M(\mu)$ each.* Now using the fact $-1 \le 2\alpha - 1 \le 1$, we can conclude that $||P^{*k} - U_{B_n}||_{\text{TV}}$ and $||P_\alpha^{*k} - U_{B_n}||_{\text{TV}}$ have the same upper bound. Thus Theorem 1.1 is true if P is replaced by P_α. Moreover, the same mapping f defined in (4.7) projects the biased flip-transpose top with random shuffle on B_n to the transpose top with random shuffle on S_n. Therefore $||P^{*k} - U_{B_n}||_{\text{TV}}$ and $||P_\alpha^{*k} - U_{B_n}||_{\text{TV}}$ have the same lower bound. Thus Theorem 1.2 is true if P is replaced by P_α. *Hence the biased flip-transpose top with random shuffle on B_n satisfies total variation cutoff phenomenon with cutoff time $n \log n$.*

Appendix A. Representation theory of demihyperoctahedral group D_n

In this section, we briefly discuss the irreducible representations of D_n (detailed proofs are omitted). 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

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

(A.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 2 that the irreducible representations of B_n are indexed by $\mathcal{D}_n$. If $\mu = (\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_n$, then $\tilde{\mu} = (\mu^{(2)}, \mu^{(1)}) \in \mathcal{D}_n$. Now from Geissinger and Kinch (1978, Proposition II.1.(ii)), it follows that the irreducible B_n-modules $V^\mu \otimes \xi$ and $V^{\tilde{\mu}}$ are isomorphic for $\mu \in \mathcal{D}_n$.

Theorem A.1. *For the irreducible B_n-module V^μ indexed by $\mu = (\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_n$, 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 $\nu \in \mathcal{D}_n$ 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 the same dimension. We denote these irreducible D_n-modules by the V_+^μ and V_-^μ.*

Proof: The proof follows by mimicking the steps of deducing the irreducible representations of A_n from that of S_n (Prasad, 2015, Theorem 4.4.2, Theorem 4.6.5). Here A_n denotes the alternating group. 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. $\square$

Let $\mathcal{S}$ be the collection of subsets Γ of $\mathcal{D}_n$ 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 $\Gamma_2 = \{(\mu^{(1)}, \mu^{(2)}) \in \mathcal{D}_n : \mu^{(1)} = \mu^{(2)}\}$. Then from Theorem A.1 and the observation

$$\sum_{\mu \in \Gamma_1} (\dim(V^\mu))^2 + \sum_{\mu \in \Gamma_2} \left(\left(\dim(V_+^\mu)\right)^2 + \left(\dim(V_-^\mu)\right)^2 \right)$$

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

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

Appendix B. A random walk on D_n analogous to the walk on B_n driven by P

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

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

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

Theorem B.1. *The eigenvalues of $\widehat{Q}(R)$ are given by*

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

(2) *If* $\mu = \left(\mu^{(1)}, \mu^{(2)}\right) \in \Gamma_2$, *then for each* $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$, $\frac{2c(b_T(n))+1}{2n-1}$ *is an eigenvalue of* $\widehat{Q}(R)$ *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 $\widehat{Q}(R) = \frac{1}{2n-1}(X_n + \mathrm{id})$, where X_n 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 A.1). Now for each $T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)$, let v_T be the Gelfand-Tsetlin vector of V^μ satisfying $X_n v_T = 2c(b_T(n))v_T$. Also, we know that $\{v_T : T \in \mathrm{tab}_{\mathcal{D}}(n,\mu)\}$ forms a basis of V^μ. Therefore the eigenvalues of $\widehat{Q}(R)$ 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 A.1). 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 $\widehat{Q}(R)$ 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 $\frac{1}{2}\dim(V^\mu)$ ($\because \dim(V_+^\mu) = \dim(V_-^\mu) = \frac{1}{2}\dim(V^\mu)$). $\qquad\square$

Theorem B.2. *For the random walk on* D_n *driven by* Q, *we have the following:*

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

Proof: Using Lemma 3.1 and following similar steps of Theorem 1.1, we have

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

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

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

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

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

Now we obtain a lower bound for the total variation distance $||Q^{*k} - U_{D_n}||_{\mathrm{TV}}$. Recall the homomorphism f defined in (4.7) and set $f' = f\big|_{D_n}$, the restriction of f to D_n. Then f' projects the random walk on D_n driven by Q to the random walk on S_n driven by $\mathcal{P}_{\frac{1}{2n-1}}$. Thus we have $U_{D_n} f'^{-1} = U_{S_n}$ and $Qf'^{-1} = \mathcal{P}_{\frac{1}{2n-1}}$. Now

using the arguments used in the proof of Lemma 4.3, we can conclude that

$$\mathcal{P}^{*k}_{\frac{1}{2n-1}} = \left(Qf'^{-1}\right)^{*k} = Q^{*k}f'^{-1}.$$

Therefore using Levin et al. (2009, Lemma 7.9), we have the following:

$$||Q^{*k} - U_{D_n}||_{\mathrm{TV}} \geq ||Q^{*k}f'^{-1} - U_{D_n}f'^{-1}||_{\mathrm{TV}} = \left|\left|\mathcal{P}^{*k}_{\frac{1}{2n-1}} - U_{S_n}\right|\right|_{\mathrm{TV}}. \tag{B.3}$$

Theorem B.3. *For the random walk on D_n driven by Q, we have the following:*

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

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

Proof: Setting $a = \frac{1}{2n-1}$ in (4.2) and (4.3) we have

$$E_{\frac{1}{2n-1},k}(\mathfrak{f}) \approx 1 + (n-2)e^{-\frac{2k}{2n-1}} + \left(-\frac{1}{2n-1}\right)^k.$$

$$E_{\frac{1}{2n-1},k}\left(\mathfrak{f}^2\right) \approx 2 + 3(n-2)e^{-\frac{2k}{2n-1}} + (n^2 - 5n + 5)e^{-\frac{4k}{2n-1}}$$

$$+ 3\left(-\frac{1}{2n-1}\right)^k + (n-2)\left(\left(\frac{1}{2n-1}\right)^k + \left(\frac{-3}{2n-1}\right)^k\right).$$

Therefore Proposition 4.1 and (B.3) implies that

$$||Q^{*k} - U_{D_n}||_{\mathrm{TV}} \geq 1 - \frac{2\left(3 + 3(n-2)e^{-\frac{2k}{2n-1}} - 2(n-1)e^{-\frac{4k}{2n-1}} + o(1)\right)}{\left(1 + (n-2)e^{-\frac{2k}{2n-1}} + o(1)\right)^2}, \quad \text{for } k > 1.$$

$$\tag{B.4}$$

Now if n is large, $c \ll 0$ and $k = (n - \frac{1}{2})(\log n + c)$, then by (B.4), 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 (B.4), we have

$$1 \geq ||Q^{*k_n} - U_{D_n}||_{\mathrm{TV}} \geq 1 - \frac{2\left(3 + 3n^\epsilon + o(1)(n^{2\epsilon} + n^\epsilon + 1)\right)}{(1 + (1 + o(1))n^\epsilon + o(1))^2} \tag{B.5}$$

for large n. Therefore, the second part of this theorem follows from (B.5) and the fact that

$$\lim_{n\to\infty} \frac{2\left(3 + 3n^\epsilon + o(1)(n^{2\epsilon} + n^\epsilon + 1)\right)}{(1 + (1 + o(1))n^\epsilon + o(1))^2} = 0.$$

$\square$

Therefore from the first part of Theorems B.2 and B.3, we can say that the mixing time for the random walk on D_n driven by Q is $O\left(\left(n - \frac{1}{2}\right)\log n\right)$. Furthermore, the second part of Theorems B.2 and B.3 implies that this shuffle satisfies the cutoff phenomenon and the total variation cutoff for this shuffle occurs at $\left(n - \frac{1}{2}\right)\log n$.

Acknowledgements

References

Aldous, D. and Diaconis, P. Shuffling cards and stopping times. *Amer. Math. Monthly*, **93** (5), 333–348 (1986). MR841111.

Ayyer, A., Schilling, A., and Thiéry, N. M. Spectral gap for random-to-random shuffling on linear extensions. *Exp. Math.*, **26** (1), 22–30 (2017). MR3599002.

Bate, M. E., Connor, S. B., and Matheau-Raven, O. Cutoff for a one-sided transposition shuffle. *ArXiv Mathematics e-prints* (2019). arXiv: 1907.12074.

Bernstein, M. and Nestoridi, E. Cutoff for random to random card shuffle. *Ann. Probab.*, **47** (5), 3303–3320 (2019). MR4021252.

Björner, A. and Brenti, F. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York (2005). ISBN 978-3540-442387; 3-540-44238-3. MR2133266.

Ceccherini-Silberstein, T., Scarabotti, F., and Tolli, F. *Representation theory of the symmetric groups. The Okounkov-Vershik approach, character formulas, and partition algebras*, volume 121 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2010). ISBN 978-0-521-11817-0. MR2643487.

Cetlin, M. L. Finite automata and the simulation of the simplest forms of behavior. *Uspehi Mat. Nauk*, **18** (4 (112)), 3–28 (1963). MR0159734.

Diaconis, P. Applications of noncommutative Fourier analysis to probability problems. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pp. 51–100. Springer, Berlin (1988a). MR983372.

Diaconis, P. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA (1988b). ISBN 0-940600-14-5. MR964069.

Diaconis, P. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, **93** (4), 1659–1664 (1996). MR1374011.

Diaconis, P., Fill, J. A., and Pitman, J. Analysis of top to random shuffles. *Combin. Probab. Comput.*, **1** (2), 135–155 (1992). MR1179244.

Diaconis, P. and Saloff-Coste, L. Comparison techniques for random walk on finite groups. *Ann. Probab.*, **21** (4), 2131–2156 (1993). MR1245303.

Diaconis, P. and Shahshahani, M. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, **57** (2), 159–179 (1981). MR626813.

Dieker, A. B. and Saliola, F. V. Spectral analysis of random-to-random Markov chains. *Adv. Math.*, **323**, 427–485 (2018). MR3725883.

Flatto, L., Odlyzko, A. M., and Wales, D. B. Random shuffles and group representations. *Ann. Probab.*, **13** (1), 154–178 (1985). MR770635.

Geissinger, L. and Kinch, D. Representations of the hyperoctahedral groups. *J. Algebra*, **53** (1), 1–20 (1978). MR491917.

Levin, D. A., Peres, Y., and Wilmer, E. L. *Markov chains and mixing times*. American Mathematical Society, Providence, RI (2009). ISBN 978-0-8218-4739-8. MR2466937.

Matheau-Raven, O. Random walks on the symmetric group: Cutoff for one-sided transposition shuffles. *ArXiv Mathematics e-prints* (2020). arXiv: 2012.05118.

Mishra, A. and Srinivasan, M. K. The Okounkov-Vershik approach to the representation theory of $G \sim S_n$. *J. Algebraic Combin.*, **44** (3), 519–560 (2016). MR3552899.

Okounkov, A. and Vershik, A. A new approach to representation theory of symmetric groups. *Selecta Math. (N.S.)*, **2** (4), 581–605 (1996). MR1443185.

Prasad, A. *Representation theory. A combinatorial viewpoint*, volume 147 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Delhi (2015). ISBN 978-1-107-08205-2. MR3287258.

Pushkarev, I. A. 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). MR1691647.

Sagan, B. E. *The symmetric group. Representations, combinatorial algorithms, and symmetric functions*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition (2001). ISBN 0-387-95067-2. MR1824028.

Saloff-Coste, L. Random walks on finite groups. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pp. 263–346. Springer, Berlin (2004). MR2023654.

Schoolfield, C. H., Jr. Random walks on wreath products of groups. *J. Theoret. Probab.*, **15** (3), 667–693 (2002). MR1922442.

Schoolfield, C. H., Jr. Generating a random signed permutation with random reversals. *J. Theoret. Probab.*, **18** (4), 911–931 (2005). MR2289938.

Serre, J.-P. *Linear representations of finite groups*. Springer-Verlag, New York-Heidelberg (1977). ISBN 0-387-90190-6. MR0450380.