Séminaire Lotharingien de Combinatoire **84B** (2020)
Article #69, 12 pp.

Proceedings of the 32nd Conference on Formal Power Series and Algebraic Combinatorics (Online)

Cutoff for the warp-transpose top with random shuffle

Subhajit Ghosh*[1]

[1]*Department of Mathematics, Indian Institute of Science, Bangalore, India*

Abstract. We consider a random walk on the complete monomial group $G_n \wr S_n$ generated by the elements of the forms $(\mathrm{e},\ldots,\mathrm{e},g;\mathrm{id})$ 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 $n\log n + \frac{1}{2}n\log(|G_n|-1)$ and under some condition on $|G_n|$, this shuffle exhibits the cutoff phenomenon.

Keywords: random walk, complete monomial group, mixing time, cutoff, Young–Jucys–Murphy elements

1 Introduction

Random walks on finite groups are well-studied topic in probability theory. Under certain natural conditions, a random walk converges to a unique stationary distribution. Random walks on finite groups can be used in various situations to approximate, understand and sample from their stationary distribution. For more details about random walks on finite groups, see [10, 1, 2]. The topic of interest in this case is the mixing time i.e., the number of steps required to reach near the stationary distribution upto a given tolerance. To study the convergence of random walks, it is helpful to know the eigenvalues and eigenvectors of the transition matrix. In the eighties, the theory of random walks on finite groups obtained its own independence, its own problems and techniques. In an seminal work of Diaconis and Shahshahani [4], they introduced the use of non-commutative Fourier analytic techniques. Before describing the random walk we are going to consider, let us first recall the definition of the complete monomial group.

Let $\{G_n\}_n$ be a sequence of finite groups (Assume $|G_n| > 1$ for all n) and S_n be the symmetric group. The complete monomial group is the *wreath product* of G_n with S_n, is a group denoted by $\mathcal{G}_\mathrm{n} = G_n \wr S_n$ and can be described as follows: The elements of $\mathcal{G}_\mathrm{n}$ are $(n+1)$-tuples $(g_1, g_2, \ldots, g_n; \pi)$ where $g_i \in G_n$ and $\pi \in S_n$. The multiplication in $\mathcal{G}_\mathrm{n}$ is given by $(g_1,\ldots,g_n;\pi)(h_1,\ldots,h_n;\eta) = (g_1h_{\pi^{-1}(1)},\ldots,g_nh_{\pi^{-1}(n)};\pi\eta)$. Therefore $(g_1,\ldots,g_n;\pi)^{-1} = (g^{-1}_{\pi(1)},\ldots,g^{-1}_{\pi(n)};\pi^{-1})$. Let e be the identity of G_n and id be the

*gsubhajit@iisc.ac.in. Supported by a UGC Centre for Advanced Study grant

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 $g^{(i)} := (\mathrm{e}, \ldots, \mathrm{e}, g, \mathrm{e}, \ldots, \mathrm{e}; \mathrm{id}) \in \mathcal{G}_{\mathrm{n}}$ (g is in i^{th} position). Let $(\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n))$ be the element of $\mathcal{G}_{\mathrm{n}}$ with g^{-1} in i^{th} position and g in n^{th} 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, defined as follows:

$$P(x) = \begin{cases} \frac{1}{n|G_n|} & \text{if } x = (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id}) \text{ for } g \in G_n, \\ \frac{1}{n|G_n|} & \text{if } x = (\mathrm{e}, \ldots, \mathrm{e}, g^{-1}, \mathrm{e}, \ldots, \mathrm{e}, g; (i, n)) \text{ for } g \in G_n,\ 1 \leq i < n, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

We call this the *warp-transpose top with random shuffle* because at most times the n^{th} component is multiplied by g and the i^{th} component is multiplied by g^{-1} simultaneously, $g \in G_n$, $1 \leq i < n$. We now state the main theorem of this paper.

Theorem 1.1. *The mixing time for the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ is of order $n \log n + \frac{1}{2} n \log(|G_n| - 1)$. Moreover if $|G_n| = o(n^{\delta})$ for all $\delta > 0$, then this shuffle satisfies the cutoff phenomenon.*

1.1 Preliminaries

Given a finite group G, the group algebra $\mathbb{C}[G]$ be the set of all formal linear combinations of the elements of G with complex coefficients. $\mathbb{C}[G]$ can be thought of as a vector space over $\mathbb{C}$ with basis G. If we denote $V = \mathbb{C}[G]$ then the *right regular representation* $R : G \longrightarrow GL(V)$ is defined by $g \mapsto \left(\sum_{h \in G} C_h h \mapsto \sum_{h \in G} C_h hg\right)$, where $C_h \in \mathbb{C}$ i.e., $R(g)$ is an invertible matrix over $\mathbb{C}$ of order $|G| \times |G|$. We will assume basic definitions and terminologies of finite group representations. We use some results from representation theory of finite groups without recalling the proof. For details about finite group representation see [8, 9, 11].

Let p and q be two probability measures on a finite group G. We define the *convolution* $p * q$ of p and q by $(p * q)(x) := \sum_{y \in G} p(xy^{-1}) q(y)$. The *Fourier transform* $\widehat{p}$ of p at the right regular representation R is defined by the matrix $\sum_{x \in G} p(x) R(x)$. 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 $\mathbb{C}[G]$ from the right. It can be easily seen that $\widehat{(p * q)}(R) = \widehat{p}(R) \widehat{q}(R)$.

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 k^{th} 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 $p^{*k}(x^{-1} y)$. A random walk 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. One can easily check that the random walk on G driven by p is irreducible if and only if the support of p generates G. A probability distribution Π is said to be a *stationary distribution* of a random walk with transition matrix M if $\Pi M = \Pi$. Any irreducible random walk possesses a unique stationary distribution. The stationary distribution for an irreducible random walk on G driven by p, is the uniform distribution U_G on G. From now on, the uniform distribution on G will be denoted by U_G. Given a random walk (discrete time, finite state space) the *period* of a state x 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 random walk are the same (see [6, 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 measures on Ω. The *total variation distance* between μ and ν is defined by

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

We note that $||\mu - \nu||_{\text{TV}} = \frac{1}{2}\sum_{x\in\Omega} |\mu(x) - \nu(x)|$ (see [6, Proposition 4.2]). If the irreducible random walk on a finite group G driven by a probability measure p (defined on G) is aperiodic, then the distribution after k^{th} transition converges to U_G in total variation distance as $k \to \infty$.

Definition 1.2. 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 tending to infinity such that the following hold:

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

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

Here $\lfloor x \rfloor$ denotes the floor of x (the largest integer less than or equal to x). See [3] for more details on cutoff phenomenon.

Proposition 1.3. *The warp-transpose top with random shuffle on $\mathcal{G}_\text{n}$ is irreducible and aperiodic.*

Proof. The support of P 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 $\mathcal{G}_\text{n}$. Thus (1.2) implies Γ generates $\mathcal{G}_\text{n}$ and hence the warp-transpose top with random shuffle on $\mathcal{G}_\text{n}$ is irreducible.

$$\begin{aligned} &(g^{-1})^{(n)} \left((g^{-1})^{(i)} g^{(n)}(i,n)\right) g^{(n)} = (i,n) \text{ for each } 1 \leq i < n \text{ and } g \in G_n, \\ &(k,n) g^{(n)} (k,n) = g^{(k)} \text{ for each } 1 \leq k \leq n \text{ and for all } g \in G_n. \end{aligned} \tag{1.2}$$

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 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. $\square$

Proposition 1.3 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. In Section 2 we will find the spectrum of $\widehat{P}(R)$. We will prove the theorem which gives an upper bound of $||P^{*k} - U_{\mathcal{G}_\mathrm{n}}||_{\mathrm{TV}}$ in Section 3. In Section 4, lower bound of $||P^{*k} - U_{\mathcal{G}_\mathrm{n}}||_{\mathrm{TV}}$ will be discussed and Theorem 1.1 will be proved. Throughout this article $[n]$ denotes the set $\{1, \ldots, n\}$ for any positive integer n.

2 Spectrum of the transition matrix

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 $\mathcal{G}_\mathrm{n}$. To find the eigenvalues of $\widehat{P}(R)$ 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 $\mathcal{G}_\mathrm{n}$, following the notation from [7].

Definition 2.1. Let $\mathcal{Y}$ denote the set of all Young diagrams (there is a unique Young diagram with zero boxes) and $\mathcal{Y}_\mathrm{n}$ denote the set of all Young diagrams with n-boxes. For a finite set X, 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 $\widehat{G_n}$ denote the (finite) set of equivalence classes of finite dimensional complex irreducible representations of G_n. Given $\sigma \in \widehat{G_n}$, we denote by W^σ the corresponding irreducible G_n-module. Here by irreducible G_n-module we mean the space for the corresponding irreducible representation of G_n. Elements of $\mathcal{Y}(\widehat{G_n})$ are called *Young G_n-diagrams* and elements of $\mathcal{Y}_n(\widehat{G_n})$ are called *Young G_n-diagrams with n boxes*. Given $\mu \in \mathcal{Y}(\widehat{G_n})$ and $\sigma \in \widehat{G_n}$, we denote by $\mu \downarrow_\sigma$ the set of all Young G_n-diagrams obtained from μ by removing one of the inner corners in the Young diagram $\mu(\sigma)$. Let $\mu \in \mathcal{Y}$. A *Young tableau* of shape μ is obtained by taking the Young diagram μ and filling its $|\mu|$ boxes (bijectively) with the numbers $1, 2, \ldots, |\mu|$. A Young tableau is said to be *standard* if the numbers in the boxes strictly increase along each row and each column of the Young diagram of μ. The set of all standard Young tableaux of shape μ is denoted by $\mathrm{tab}(\mu)$. Let $\mu \in \mathcal{Y}(\widehat{G_n})$. A *Young G_n-tableau* of shape μ is obtained by taking the Young G_n-diagram μ and filling its $||\mu||$ boxes (bijectively) with the numbers $1, 2, \ldots, ||\mu||$. A Young G_n-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 $\mathrm{tab}_{G_n}(n, \mu)$, where $\mu \in \mathcal{Y}_n(\widehat{G_n})$, denote the set of all standard Young G_n-tableaux of shape μ and let $\mathrm{tab}_{G_n}(n) = \cup_{\mu\in\mathcal{Y}_n(\widehat{G_n})} \mathrm{tab}_{G_n}(n,\mu)$. Let $T \in \mathrm{tab}_{G_n}(n)$ and $i \in [n]$. If i appear in the Young diagram $\mu(\sigma)$, where μ is the shape of T and $\sigma \in \widehat{G_n}$, we write $r_T(i) = \sigma$. The *content* of a box in row p and column q of a Young diagram is the integer $q - p$. Let $b_T(i)$ be the box in $\mu(\sigma)$, with the number i resides. Also $c(b_T(i))$ denotes the content of the box $b_T(i)$. The (generalized) *Young–Jucys–Murphy* elements $X_1, X_2, \ldots, X_n$ of $\mathbb{C}[\mathcal{G}_\mathrm{n}]$ are given by $X_1 = 0$ and

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

We now define Gelfand–Tsetlin subspaces and the Gelfand–Tsetlin decomposition.

Definition 2.2. Let $\lambda \in \widehat{\mathcal{G}_\mathrm{n}}$ and consider the irreducible $\mathcal{G}_\mathrm{n}$-module (the space for the representation of $\mathcal{G}_\mathrm{n}$) V^λ. Since the branching is simple [7, Section 3], the decomposition into irreducible $\mathcal{G}_{n-1}$-modules is given by

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

where the sum is over all $\mu \in \widehat{\mathcal{G}}_{n-1}$, 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 $\mathcal{G}_1$-submodules, i.e.,

$$V^\lambda = \bigoplus_{T} V_T, \tag{2.1}$$

where the sum is over all possible chains $T = \lambda_1 \nearrow \lambda_2 \nearrow \cdots \nearrow \lambda_n$ with $\lambda_i \in \widehat{\mathcal{G}}_i$ and $\lambda_n = \lambda$. We call (2.1) the *Gelfand–Tsetlin* decomposition of V^λ and each V_T in (2.1) a *Gelfand–Tsetlin* subspace of V^λ. We note that if $0 \neq v_T \in V_T$, then $\mathbb{C}[\mathcal{G}_i]v_T = V^{\lambda_i}$ from the definition of V_T. Unlike $\mathcal{G}_i = G_i \wr S_i$, here $\mathcal{G}_i$ denotes the subgroup of $\mathcal{G}_\mathrm{n}$ given as follows

$$\mathcal{G}_i = \{(g_1, \ldots, g_n, \pi) \in \mathcal{G}_\mathrm{n} : \pi(j) = j \text{ for } i+1 \leq j \leq n\},\ 1 \leq i \leq n.$$

The Young–Jucys–Murphy elements act by scalars on the Gelfand–Tsetlin subspaces.

From Lemma 6.2 and Theorem 6.4 of [7], we may parametrize the irreducible representations of $\mathcal{G}_\mathrm{n}$ by elements of $\mathcal{Y}_n(\widehat{G_n})$.

Theorem 2.3 ([7, Theorem 6.5]). *Let $\mu \in \mathcal{Y}_n(\widehat{G_n})$. Then we may index the Gelfand–Tsetlin subspaces of V^μ by standard Young G_n-tableaux of shape μ and write the Gelfand–Tsetlin decomposition as*

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

where each V_T is closed under the action of G_n^n and as a G_n^n-module, is isomorphic to the irreducible G_n^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 on V_T are given by $\frac{|G_n|}{\dim(W^{r_T(i)})} c(b_T(i))$.

Theorem 2.4 ([7, Theorem 6.7]). *Let $\mu \in \mathcal{Y}_n(\widehat{G_n})$. Write the elements of $\widehat{G_n}$ as $\{\sigma_1, \ldots, \sigma_t\}$ 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*

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

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

Without loss of generality, from now on we will assume $\sigma_1 = \mathbb{1}$ (the trivial representation of G_n) and $\mu \left(\in \mathcal{Y}_n(\widehat{G_n})\right)$ as the tuple $(\mu^{(1)}, \ldots, \mu^{(t)})$. We note that for $T \in \mathrm{tab}_{G_n}(n, \mu)$ the dimension of V_T is $d_1^{m_1} \cdots d_t^{m_t}$, notations follow the same meaning as of Theorem 2.4.

Lemma 2.5. *Let G be a finite group and $\sigma \in \widehat{G}$. If W^σ (respectively χ^σ) denotes the irreducible G-module (respectively character) and d_ρ is the dimension of W^σ, then the action of the group algebra element $\sum_{g \in G} g$ on W^σ is given by the following scalar matrix*

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

Here I_{d_σ} is the identity matrix of order $d_\sigma \times d_\sigma$.

Proof. It is clear that $\sum_{g \in G} g$ is in the centre of $\mathbb{C}[G]$. Therefore by Schur's lemma ([11, Proposition 4]), we have $\sum_{g \in G} g = c I_{d_\sigma}$ for some $c \in \mathbb{C}$. The value of c can be obtained by equating the traces of $\sum_{g \in G} g$ and $c I_{d_\sigma}$. □

Theorem 2.6. *For each $\mu = (\mu^{(1)}, \ldots, \mu^{(t)}) \in \mathcal{Y}_n(\widehat{G_n})$, let $\widehat{P}(R)\big|_{V^\mu}$ denotes the restriction of $\widehat{P}(R)$ to the irreducible $\mathcal{G}_n$-module V^μ. Then the eigenvalues of $\widehat{P}(R)\big|_{V^\mu}$ 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 $T \in \mathrm{tab}_{G_n}(n, \mu)$.

Proof. We first find the eigenvalues of $X_n + \sum_{g \in G_n} (\mathrm{e}, \ldots, \mathrm{e}, g; \mathrm{id})$. Let $I_{\dim(V_T)}$ denote the identity matrix of order $\dim(V_T) \times \dim(V_T)$. Then from Theorem 2.3 we have

$$V^\mu = \bigoplus_{T \in \mathrm{tab}_{G_n}(n,\mu)} V_T \quad \text{and} \quad X_n\big|_{V_T} = \frac{|G_n|}{\dim(W^{r_T(n)})} c(b_T(n)) I_{\dim(V_T)}. \tag{2.2}$$

Again from Theorem 2.3 and Lemma 2.5 we have

$$\sum_{g\in G_n}(\mathrm{e},\ldots,\mathrm{e},g;\mathrm{id})\big|_{V_T}=\frac{|G_n|}{\dim(W^{r_T(n)})}\langle\chi^{r_T(n)},\chi^{\mathbb{1}}\rangle I_{\dim(V_T)}. \tag{2.3}$$

Also $\widehat{P}(R)=\dfrac{1}{n|G_n|}\displaystyle\sum_{g\in G_n}\left(R\left((\mathrm{e},\ldots,\mathrm{e},g;\mathrm{id})\right)+\sum_{i=1}^{n-1}R\left((\mathrm{e},\ldots,\mathrm{e},g^{-1},\mathrm{e},\ldots,\mathrm{e},g;(i,n))\right)\right)$.

Therefore $n|G_n|\widehat{P}(R)$ is nothing but the action of $X_n+\sum_{g\in G_n}(\mathrm{e},\ldots,\mathrm{e},g;\mathrm{id})$ on $\mathbb{C}[\mathcal{G}_\mathrm{n}]$ from right. Since $\dim(V_T)=d_1^{m_1}\cdots d_t^{m_t}$, the theorem follows from (2.2) and (2.3). □

Remark 2.7. In the regular representation of a finite group, each irreducible representation occurs with multiplicity equal to its dimension [11, Section 2.4]. Therefore, Theorems 2.4 and 2.6 provide the eigenvalues of $\widehat{P}(R)$.

3 Upper bound for total variation distance

In this section, we will prove the theorem giving an upper bound of the total variation distance $||P^{*k}-U_{\mathcal{G}_\mathrm{n}}||_{\mathrm{TV}}$ using the upper bound lemma [1, Lemma 4.2]. Before proving the main result of this section, first we prove some lemmas which will be useful. For any positive integer ℓ, we write $\lambda\vdash\ell$ to denote λ is a partition of ℓ.

Lemma 3.1. *Let ℓ be a positive integer and s be any non-negative real number. For $\lambda\vdash\ell$, if λ_1 denotes the largest part of λ, then*

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

Proof. For any partition ζ of $\ell-\lambda_1$ with largest part $\zeta_1\leq\lambda_1$, we have $f^\lambda\leq\binom{\ell}{\lambda_1}f^\zeta$. Therefore $\sum_{\lambda\vdash\ell}(f^\lambda)^2\left(\frac{\lambda_1-s}{\ell}\right)^{2k}$ is less than or equal to

$$\begin{aligned}\sum_{\lambda_1=1}^{\ell}\sum_{\substack{\zeta\vdash(\ell-\lambda_1)\\ \zeta_1\leq\lambda_1}}\binom{\ell}{\lambda_1}^2(f^\zeta)^2\left(\frac{\lambda_1-s}{\ell}\right)^{2k}&\leq\sum_{\lambda_1=1}^{\ell}\binom{\ell}{\lambda_1}^2\left(\frac{\lambda_1-s}{\ell}\right)^{2k}\sum_{\zeta\vdash(\ell-\lambda_1)}(f^\zeta)^2\\ &=\sum_{u=0}^{\ell-1}\binom{\ell}{u}^2\left(1-\frac{u+s}{\ell}\right)^{2k}u!.\end{aligned}\tag{3.1}$$

Equality in (3.1) is obtained by writing $u=\ell-\lambda_1$. Using $1-x\leq e^{-x}$ for all $x\geq0$ and $\binom{\ell}{u}\leq\frac{\ell^u}{u!}$, the expression in the right hand side of (3.1) is less than or equal to

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

□

Corollary 3.2. *Following the notations of Lemma 3.1, we have*

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

Lemma 3.3. *Following the notations from Theorems 2.3 and 2.4 we have*

$$\sum_{T\in\mathsf{tab}_{G_n}(n,\mu)}\left(\frac{c(b_T(n))+\langle\chi^{r_T(n)},\chi^{\mathbb{1}}\rangle}{n\dim(W^{r_T(n)})}\right)^{2k}$$
$$<\binom{n}{m_1,\dots,m_t}f^{\mu^{(1)}}\cdots f^{\mu^{(t)}}\sum_{j=1}^{t}\left(\frac{\mu_1^{(j)}-1+\langle\chi^{\sigma_j},\chi^{\mathbb{1}}\rangle}{nd_j}\right)^{2k}.$$

Here $\mu_1^{(i)}$ denotes the largest part of the partition $\mu^{(i)}$ for each $1\leq i\leq t$.

Proof. The set $\mathsf{tab}_{G_n}(n,\mu)$ is a disjoint union of the sets $\mathcal{T}_1,\dots,\mathcal{T}_t$, where

$$\mathcal{T}_i=\{(T_1,\dots,T_t)\in\mathsf{tab}_{G_n}(n,\mu)\mid b_T(n)\text{ is in }T_i\}$$

for each $1\leq i\leq t$. Therefore

$$\sum_{T\in\mathsf{tab}_{G_n}(n,\mu)}\left(\frac{c(b_T(n))+\langle\chi^{r_T(n)},\chi^{\mathbb{1}}\rangle}{n\dim(W^{r_T(n)})}\right)^{2k}=\sum_{i=1}^{t}\sum_{T\in\mathcal{T}_i}\left(\frac{c(b_T(n))+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{nd_i}\right)^{2k}$$
$$=\sum_{i=1}^{t}\left(\frac{1}{d_i}\right)^{2k}\binom{n-1}{m_1,\dots,m_i-1,\dots,m_t}\frac{f^{\mu^{(1)}}\cdots f^{\mu^{(t)}}}{f^{\mu^{(i)}}}\sum_{T_i\in\mathsf{tab}(\mu^{(i)})}\left(\frac{c(b_T(m_i))+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{n}\right)^{2k}$$
$$<\sum_{i=1}^{t}\left(\frac{1}{d_i}\right)^{2k}\binom{n}{m_1,\dots,m_t}\frac{f^{\mu^{(1)}}\cdots f^{\mu^{(t)}}}{f^{\mu^{(i)}}}\sum_{T_i\in\mathsf{tab}(\mu^{(i)})}\left(\frac{\mu_1^{(i)}-1+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{n}\right)^{2k}. \tag{3.2}$$

The result follows from $\sum_{T_i\in\mathsf{tab}(\mu^{(i)})}\left(\frac{\mu_1^{(i)}-1+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{n}\right)^{2k}=f^{\mu^{(i)}}\left(\frac{\mu_1^{(i)}-1+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{n}\right)^{2k}$ and (3.2). □

Proposition 3.4. *For the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ driven by P, we have*

$$4\,||P^{*k}-U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}^2<\left(e^{n^2e^{-\frac{2k}{n}}}-1\right)+e\left(e^{n^2(|G_n|-1)e^{-\frac{2k}{n}}}-1\right)$$
$$+(t-1)\left(e^{-\frac{2k}{n}}e^{n^2e^{-\frac{2k}{n}}}+\frac{e}{n^2}\left(e^{n^2(|G_n|-1)e^{-\frac{2k}{n}}}-1\right)\right),$$

for all $k\geq n\log n$.

Proof. Using the upper bound lemma [1, Lemma 4.2], we have

$$4\,||P^{*k}-U_{\mathcal{G}_n}||_{\mathrm{TV}}^2 \leq \sum_{\mu\in\mathcal{Y}_n(\widehat{G_n}):\,\mu(\mathbb{1})\neq(n)} \dim(V^{\mu})\,\mathrm{trace}\left(\left(\widehat{P}(R)\big|_{V^{\mu}}\right)^{2k}\right). \tag{3.3}$$

Here we recall that $\widehat{G_n}=\{\sigma_1,\dots,\sigma_t\}$ 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 $\mu=(\mu^{(1)},\dots,\mu^{(t)})$, where $\mu^{(i)}\vdash m_i$ and $\sum_{i=1}^{t}m_i=n$. First we partition the set $\mathcal{Y}_n(\widehat{G_n})$ into two disjoint subsets $\mathcal{A}_1,\mathcal{A}_2$ as follows:

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

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

It can be easily seen that $\mathcal{B}_i$'s are disjoint. Therefore by using Theorem 2.6 and Remark 2.7, the inequality (3.3) become

$$\begin{aligned}4\,||P^{*k}-U_{\mathcal{G}_n}||_{\mathrm{TV}}^2 \leq &\sum_{\substack{\mu\in\mathcal{B}_1\\ \mu(\mathbb{1})\neq(n)}}\dim(V^{\mu})\sum_{T\in\mathrm{tab}_{G_n}(n,\mu)}\left(\frac{c(b_T(n))+1}{nd_1}\right)^{2k}d_1^n\\ &+\sum_{i=2}^{t}\sum_{\mu\in\mathcal{B}_i}\dim(V^{\mu})\sum_{T\in\mathrm{tab}_{G_n}(n,\mu)}\left(\frac{c(b_T(n))}{nd_i}\right)^{2k}d_i^n\\ &+\sum_{\mu\in\mathcal{A}_2}\dim(V^{\mu})\sum_{T\in\mathrm{tab}_{G_n}(n,\mu)}\left(\frac{c(b_T(n))+\langle\chi^{r_T(n)},\chi^{\mathbb{1}}\rangle}{n\dim(W^{r_T(n)})}\right)^{2k}d_1^{m_1}\cdots d_t^{m_t}.\end{aligned} \tag{3.4}$$

Given a partition ξ of any positive integer, now on the largest part of ξ will be denoted by ξ_1. The first two terms in the right hand side of inequality (3.4) is equal to

$$\begin{aligned}&\sum_{\substack{\lambda\vdash n\\ \lambda_1\neq n}}f^{\lambda}d_1^n\sum_{T\in\mathrm{tab}(\lambda)}\left(\frac{c(b_T(n))+1}{nd_1}\right)^{2k}d_1^n+\sum_{i=2}^{t}\sum_{\lambda\vdash n}f^{\lambda}d_i^n\sum_{T\in\mathrm{tab}(\lambda)}\left(\frac{c(b_T(n))}{nd_i}\right)^{2k}d_i^n\\ \leq\ &\frac{1}{d_1^{2k-2n}}\sum_{\substack{\lambda\vdash n\\ \lambda_1\neq n}}f^{\lambda}\sum_{T\in\mathrm{tab}(\lambda)}\left(\frac{\lambda_1}{n}\right)^{2k}+\sum_{i=2}^{t}\frac{1}{d_i^{2k-2n}}\sum_{\lambda\vdash n}f^{\lambda}\sum_{T\in\mathrm{tab}(\lambda)}\left(\frac{\lambda_1-1}{n}\right)^{2k}\\ \leq\ &\sum_{\substack{\lambda\vdash n\\ \lambda_1\neq n}}(f^{\lambda})^2\left(\frac{\lambda_1}{n}\right)^{2k}+\sum_{i=2}^{t}\sum_{\lambda\vdash n}(f^{\lambda})^2\left(\frac{\lambda_1-1}{n}\right)^{2k}\\ <\ &\left(e^{n^2e^{-\frac{2k}{n}}}-1\right)+(t-1)e^{-\frac{2k}{n}}e^{n^2e^{-\frac{2k}{n}}}.\end{aligned} \tag{3.5}$$

The inequality in (3.5) follows from Lemma 3.1 and Corollary 3.2. Using Lemma 3.3, the third term in the right hand side of (3.4) is less than

$$\sum_{\mu\in\mathcal{A}_2}\binom{n}{m_1,\dots,m_t}^2 (f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2 d_1^{2m_1}\dots d_t^{2m_t}\sum_{j=1}^{t}\left(\frac{\mu_1^{(j)}-1+\langle\chi^{\sigma_i},\chi^{\mathbb{1}}\rangle}{nd_j}\right)^{2k}. \tag{3.6}$$

We now deal with (3.6) by considering two separate cases namely $j=1$ and $1<j\leq t$. The partial sum corresponding to $j=1$ in (3.6) equal to

$$\begin{aligned}
&\sum_{m_1=0}^{n-1}\sum_{\substack{(m_2,\dots,m_t)\\ \sum m_k=n-m_1\\ 0\leq m_k\leq n-1}}\sum_{\substack{\mu^{(i)}\vdash m_i\\ 1\leq i\leq t}}\binom{n}{m_1}^2\binom{n-m_1}{m_2,\dots,m_t}^2 (f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2 d_1^{2m_1}\dots d_t^{2m_t}\left(\frac{\mu_1^{(1)}}{nd_1}\right)^{2k}\\
&<\sum_{m_1=0}^{n-1}\sum_{\substack{(m_2,\dots,m_t)\\ \sum m_k=n-m_1\\ m_k\geq 0}}\sum_{\substack{\mu^{(i)}\vdash m_i\\ 1\leq i\leq t}}\binom{n}{m_1}^2\binom{n-m_1}{m_2,\dots,m_t}^2 (f^{\mu^{(1)}})^2\cdots(f^{\mu^{(t)}})^2 d_1^{2m_1}\dots d_t^{2m_t}\left(\frac{\mu_1^{(1)}}{nd_1}\right)^{2k}\\
&=\sum_{m_1=0}^{n-1}(d_2^2+\cdots+d_t^2)^{n-m_1}\binom{n}{m_1}^2(n-m_1)!\left(\frac{1}{d_1}\right)^{2k-2m_1}\left(\frac{m_1}{n}\right)^{2k}\sum_{\mu^{(1)}\vdash m_1}(f^{\mu^{(1)}})^2\left(\frac{\mu_1^{(1)}}{m_1}\right)^{2k}\\
&<\sum_{m_1=0}^{n-1}(d_2^2+\cdots+d_t^2)^{n-m_1}\binom{n}{m_1}^2(n-m_1)!\left(\frac{1}{d_1}\right)^{2k-2m_1}\left(\frac{m_1}{n}\right)^{2k}e^{m_1^2e^{-\frac{2k}{m_1}}}.
\end{aligned}\tag{3.7}$$

The inequality in (3.7) follows from Lemma 3.1. As $k\geq n\log n$, we have $k\geq m_1\log m_1$ and $k\geq n$. Thus using $1-x\leq e^{-x}$ for all $x\geq 0$ and writing $n-m_1$ by u, the expression in (3.7) is less than or equal to

$$\begin{aligned}
&e\sum_{u=1}^{n}\left(\frac{d_2^2+\cdots+d_t^2}{d_1^2}\right)^{u}\left(\frac{1}{d_1}\right)^{2k-2n}\binom{n}{u}^2u!\left(1-\frac{u}{n}\right)^{2k}\\
&\leq e\sum_{u=1}^{n}\frac{1}{u!}\left(n^2\left(\frac{|G_n|}{d_1^2}-1\right)e^{-\frac{2k}{n}}\right)^{u}<e\left(e^{\left(n^2\left(\frac{|G_n|}{d_1^2}-1\right)e^{-\frac{2k}{n}}\right)}-1\right).
\end{aligned}\tag{3.8}$$

By carrying out similar process as of $j=1$, the partial sum corresponding to $1<j\leq t$ in (3.6) turns out to be

$$\frac{e}{n^2}\left(e^{\left(n^2\left(\frac{|G_n|}{d_j^2}-1\right)e^{-\frac{2k}{n}}\right)}-1\right). \tag{3.9}$$

Using $\frac{1}{d_j}\leq 1$ for all $1\leq j\leq t$, the proposition follows from (3.4), (3.5), (3.8) and (3.9). $\square$

Theorem 3.5. *Recall that, $|\widehat{G_n}|$ denotes the number of irreducible representations of G_n. For the random walk on $\mathcal{G}_{\mathrm{n}}$ driven by P, we have the following:*

1. *Let $|\widehat{G_n}| > e+1$ and $c > \frac{1}{2}$. Then for $k \geq n\log n + \frac{1}{2}n\log(|G_n|-1) + cn\log(|\widehat{G_n}|-1)$, we have*
$$||P^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} < \sqrt{\frac{e+1}{2}}\, e^{-c} + o(1).$$
For $|\widehat{G_n}| = 2,3$ take $k \geq n\log n + \frac{1}{2}n\log(|G_n|-1) + cn$.

2. *For any $\epsilon \in (0,1)$, if $|\widehat{G_n}| = O(n^2)$ then $k_n = \lfloor(1+\epsilon)\left(n\log n + \frac{1}{2}n\log(|G_n|-1)\right)\rfloor$ implies*
$$\lim_{n\to\infty} ||P^{*k_n} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} = 0.$$

Proof. This follows from Proposition 3.4 and straightforward calculations. □

4 Lower bound for total variation distance

In this section, our focus will be on a lower bound of the total variation distance $||P^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}}$. We will give the outline without details. The details will appear in a later publication [5]. Also we prove Theorem 1.1 in this section. To start, we define an auxiliary representation ρ of $\mathcal{G}_{\mathrm{n}}$ and a random variable X on $\mathcal{G}_{\mathrm{n}}$.

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 $\rho : \mathcal{G}_{\mathrm{n}} \longrightarrow GL(V)$ on the basis elements of V by

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

The random variable X counts the number of fixed points of the action of ρ, i.e. X is the character of ρ. Let $E_k(X)$ be the expectation and $V_k(X)$ the variance of X with respect to the probability measure P^{*k} on $\mathcal{G}_{\mathrm{n}}$. $E_U(X)$ denotes the expectation of X with respect to the uniform distribution on $\mathcal{G}_{\mathrm{n}}$. It can be seen that $E_U(X) = 1$.

Theorem 4.1. *If $|G_n| = o(n^\delta)$ for every $\delta > 0$, then $\lim_{n\to\infty} ||P^{*k_n} - U_{\mathcal{G}_n}||_{\mathrm{TV}} = 1$ for any $\epsilon \in (0,1)$ and $k_n = \lfloor(1-\epsilon)\left(n\log n + \frac{1}{2}n\log(|G_n|-1)\right)\rfloor$.*

Proof. Using Chebychev's and Markov's inequality first we obtain,

$$||P^{*k} - U_{\mathcal{G}_{\mathrm{n}}}||_{\mathrm{TV}} \geq 1 - \frac{4V_k(X)}{(E_k(X))^2} - \frac{2}{E_k(X)}. \tag{4.1}$$

The theorem now follows by putting the values of $E_k(X)$ and $V_k(X)$ in (4.1). □

Proof of Theorem 1.1. The first part of Theorem 3.5 implies that the mixing time for this shuffle is of order $n \log n + \frac{1}{2} n \log(|G_n| - 1)$ and the second part of Theorem 3.5 and Theorem 4.1 prove that the warp-transpose top with random shuffle on $\mathcal{G}_{\mathrm{n}}$ satisfies the cutoff phenomenon. □

Acknowledgements

I would like to thank my advisor Arvind Ayyer for all the insightful discussions during the preparation of this paper. I would also like to thank the anonymous reviewer of FPSAC 2019 for the suggestion to extend the work from the $S_2 \wr S_n$ to $G \wr S_n$. I would like to acknowledge support in part by a UGC Centre for Advanced Study grant.

References

[1] P. Diaconis. "Applications of non-commutative Fourier analysis to probability problems". *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*. Springer, 1988, pp. 51–100.

[2] P. Diaconis. *Group representations in probability and statistics*. Vol. 11. Institute of Mathematical Statistics Lecture Notes—Monograph Series. 1988, pp. vi+198.

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

[4] P. Diaconis and M. Shahshahani. "Generating a random permutation with random transpositions". *Z. Wahrsch. Verw. Gebiete* **57**.2 (1981), pp. 159–179. Link.

[5] S. Ghosh. "Total variation cutoff phenomenon for the warp-transpose top with random shuffle on $G \wr S_n$". In preparation.

[6] D. Levin, Y. Peres, and E. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009, pp. xviii+371.

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

[8] A. Prasad. *Representation theory: a combinatorial viewpoint*. Vol. 147. Cambridge University Press, 2015.

[9] B. Sagan. *The symmetric group*. Second. Vol. 203. Graduate Texts in Mathematics. Representations, combinatorial algorithms, and symmetric functions. Springer-Verlag, New York, 2001, pp. xvi+238.

[10] L. Saloff-Coste. "Random walks on finite groups". *Probability on discrete structures*. Vol. 110. Encyclopaedia Math. Sci. Springer, Berlin, 2004, pp. 263–346.

[11] J.-P. Serre. *Linear representations of finite groups*. Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, New York-Heidelberg, 1977, pp. x+170.