The Annals of Applied Probability
2015, Vol. 25, No. 3, 1581–1615
DOI: 10.1214/14-AAP1031
© Institute of Mathematical Statistics, 2015

MIXING TIME OF METROPOLIS CHAIN BASED ON RANDOM TRANSPOSITION WALK CONVERGING TO MULTIVARIATE EWENS DISTRIBUTION[1]

BY YUNJIANG JIANG

Stanford University

We prove sharp rates of convergence to the Ewens equilibrium distribution for a family of Metropolis algorithms based on the random transposition shuffle on the symmetric group, with starting point at the identity. The proofs rely heavily on the theory of symmetric Jack polynomials, developed initially by Jack [*Proc. Roy. Soc. Edinburgh Sect. A* **69** (1970/1971) 1–18], Macdonald [*Symmetric Functions and Hall Polynomials* (1995) New York] and Stanley [*Adv. Math.* **77** (1989) 76–115]. This completes the analysis started by Diaconis and Hanlon in [*Contemp. Math.* **138** (1992) 99–117]. In the end we also explore other integrable Markov chains that can be obtained from symmetric function theory.

1. Introduction. There is a well-known bijection between the set of partitions of n and the conjugacy classes of the symmetric group S_n. The partition that a permutation $\sigma \in S_n$ corresponds to is simply given by its cycle structure. In fact this connection is the basis for the classical representation theory of S_n (see, e.g., [9]): the set of irreducible representations of S_n is indexed by the set $\mathcal{P}_n$ of partitions of n. Since S_n is finite, it can also be endowed with a probability space structure. The most natural measure on S_n is thus the uniform measure, with each permutation getting a weight of $1/n!$. Sampling from this uniform measure is important for many statistical applications [5], such as testing independence of n i.i.d. uniform random variables on an ordered set. Its intimate connection with card shuffling models has also generated a wonderful array of mathematics, most notably the determination of their mixing times; see, for instance, [2, 7] and [19].

One of the most natural generalizations of the uniform measure on S_n is a 1-parameter family of so-called multivariate Ewens distributions, named after Warren Ewens, who derived the partition function of this probability measure. It is defined by giving each permutation σ a weight of $\alpha^{\ell(\sigma)}$, $\alpha > 0$, where $\ell(\sigma)$ is the number of cycles in σ; hence it can be viewed as an exponentially tilted family based on the uniform measure. The distribution was first applied to population genetics, in which it describes the distribution of frequencies of alleles in a sample of genes ([13], Chapter 41).

Received May 2013; revised March 2014.

[1]Supported in part by NSF Graduate Fellowship.

MSC2010 subject classifications. Primary 60J05; secondary 05E05.

Key words and phrases. Jack polynomials, Metropolis algorithm, mixing time, random transposition.

Many important properties of the uniform measure on S_n continues to hold for the multivariate Ewens distribution [1]. For instance, the two perfect sampling schemes of the uniform measure, Feller coupling and the Chinese restaurant process, both generalize to the Ewens case. In this article, we describe a much more subtle property of the uniform measure that has been successfully generalized to the α deformed setting. In a nutshell, the characters of the symmetric group S_n form a basis in the Fourier space of class functions on S_n under the uniform measure [5]. When the underlying measure is α-deformed from the uniform, one can only make sense of Fourier transforms of a particular type of class functions, namely the ones supported on transpositions and the identity class. In that case, the basis in the Fourier space becomes the matrix coefficients of the transition from Jack symmetric polynomials basis to the power sum symmetric polynomials basis. These generalize the characters of S_n, which happen to be the transition coefficients from Schur polynomial basis to the power-sum polynomial basis.

This Fourier analytic property was first obtained by Stanley in [20]. Later Hanlon [11] applied it to the study of the Metropolis Markov chain based on random transposition walk on S_n that converges to the multivariate Ewens distribution. Diaconis and Hanlon [21] further initiated the investigation of total variation mixing time of this chain.

In light of the sharp result in [7] for the uniform case ($\alpha = 1$), it is natural to wonder what's the exact mixing time for the Diaconis–Hanlon Metropolis walk ($\alpha \neq 1$). In this paper, we prove a pair of matching upper and lower bound for the mixing time that applies to all $\alpha > 0$, which together imply the cut-off phenomenon. Previously Diaconis and Hanlon [21] outlined a proof of the upper bound in the case $\alpha > 1$ and conjectured that it was tight.

In the Appendices, we include some preliminary attempts to generalize the walk studied here in various directions. These were motivated by questions of Diaconis on whether other nontrivial Markov chains can be constructed from symmetric function theory. First we look at the action of the Sekiguchi–Debiard operator on other classical bases of symmetric polynomials. We also consider higher order operators, as given by the operator valued generating function in [17], page 317. These turn out to give new local move Markov chains converging to MED(θ), albeit without simple group theoretic interpretations. Finally we look at Laplace–Beltrami operators associated with other root systems. Recall the Schur–Weyl duality between the simple Lie groups $SU(n)$ and the finite groups S_n. This leads to an interpretation of the Sekiguchi–Debiard operator (as well as their quantized version given by Macdonald [17]) as associated with root system of type A_n. The appropriate generalizations were first discovered by Heckman and Opdam [12] in the context of Hamiltonian systems of particles on a circle, and later extended to the Macdonald case in [18]; see also [15] and references therein for a 5-parameter generalization.

Jack polynomials, which form the backbone of the argument presented here, turn out to be special cases of Macdonald polynomials of type A_n. Diaconis and Ram [6] interpreted them as eigenfunctions of an auxilliary variable algorithm on the space of partitions, which can be viewed as a quantized version of the local walk studied here.

2. Metropolis walk starting at the identity class. The multivariate Ewens distribution with parameter α is defined on S_n with $P(\sigma)$ proportional to $\alpha^{\ell(\sigma)}$ where $\ell(\sigma)$ is the number of cycles of σ. The normalization constant $z_n(\alpha) = \alpha_{(n)} := \alpha(\alpha+1)\cdots(\alpha+n-1)$.

Consider now the random transposition walk on S_n, defined by picking a pair of numbers $i \neq j$ at random, and multiplying the current state in S_n by the transposition (ij). In this form, the walk is periodic and does not converge. But if we make it lazy, then it converges to the uniform measure on S_n. By metropolizing the nonlazy walk to the multivariate Ewens distribution with parameter α, we create a new Markov chain that converges to MED(α); see [21] for details on the Metropolis algorithm. The walk behaves differently for $\alpha > 1$ and $\alpha < 1$. And as long as $\alpha \neq 1$, it converges, because it always has positive holding probability.

Warning. The α parameter here will be the reciprocal of θ below.

If we start the Metropolis walk described in the introduction from the identity element, then we can view it as a walk either on the symmetric group or on the set of partitions. It is the latter interpretation that allows for sharp analysis with other starting points.

THEOREM 2.1. *For $\theta \in (0,\infty) \setminus \{1\}$, let P_θ be the discrete time Metropolis chain based on random transposition walk starting from* id, *converging to the multivariate Ewens distribution π with parameter θ^{-1} (so identity has the largest mass when $\theta < 1$). Explicitly, let $\lambda(\pi)$ be the cycle structure of the permutation π, λ^t be the transposition of λ as a Ferrers diagram and denote $n(\lambda) = \sum_{i=1}^n \binom{\lambda_i^t}{2}$. Then the transition rule is given by*

$$P_\theta(\pi,\sigma) = \begin{cases} 1 - 1\wedge\theta + \dfrac{n(\lambda(\pi)^t)}{\binom{n}{2}}(1\wedge\theta - 1\wedge\theta^{-1}), & \\ & \textit{if } \sigma = \pi, \\ \dfrac{1}{\binom{n}{2}}(1\wedge\theta), & \textit{if } \sigma = \pi(i,j) \textit{ and } \ell(\sigma) = \ell(\pi) - 1, \\ \dfrac{1}{\binom{n}{2}}(1\wedge\theta^{-1}), & \textit{if } \sigma = \pi(i,j) \textit{ and } \ell(\sigma) = \ell(\pi) + 1, \\ 0, & \textit{otherwise.} \end{cases}$$

Here $1 \le i < j \le n$*. The chain* P_θ *has a total variation cut-off at* $t = \frac{1}{2}(\frac{1}{\theta} \vee 1)n \log n$*. This means*

$$\lim_{c\to\infty} \limsup_{n\to\infty} \| P_{\mathrm{id}}^{t(c)} - \pi \|_{\mathrm{TV}} = 0,$$

$$\lim_{c\to-\infty} \liminf_{n} \| P_{\mathrm{id}}^{t(c)} - \pi \|_{\mathrm{TV}} = 1$$

for $t(c) := \frac{1}{2}(\frac{1}{\theta} \vee 1)n(\log n + c)$.

REMARK 1. (1) Notice the chain has no intrinsic holding: when $\theta = 1$ it corresponds to the completely industrious random transposition walk with probability $\frac{1}{\binom{n}{2}}$ to go to a neighboring permutation. If one inserts a holding of $1/n$, that is, $P \mapsto \frac{1}{n} I + (1 - \frac{1}{n})P$, then the asymptotic cut-off profile stays the same and has no removable discontinuity at $\theta = 1$.

(2) The Metropolis chain defined above can be projected to conjugacy classes of S_n, namely partitions, provided we start at the identity element. The transition matrix takes the following form:

$$P_\theta(\lambda, \mu) = \begin{cases} 1 - 1 \wedge \theta + \dfrac{n(\lambda^t)}{\binom{n}{2}}(1 \wedge \theta - 1 \wedge \theta^{-1}), & \\ & \text{if } \mu = \lambda, \\ \dfrac{\lambda_i \lambda_j}{\binom{n}{2}}(1 \wedge \theta), & \text{if } \mu_k = \lambda_i + \lambda_j, \\ \dfrac{\lambda_k}{\binom{n}{2}}(1 \wedge \theta^{-1}), & \text{if } \mu_i + \mu_j = \lambda_k \text{ and } \mu_i \ne \mu_j, \\ \dfrac{\lambda_k}{2\binom{n}{2}}(1 \wedge \theta^{-1}), & \text{if } \mu_i + \mu_j = \lambda_k \text{ and } \mu_i = \mu_j, \\ 0, & \text{otherwise.} \end{cases}$$

Here $1 \le i < j \le \ell(\mu)$ and $1 \le k \le \ell(\lambda)$. Furthermore in the second line, $\mu \setminus \mu_k = \lambda \setminus \{\lambda_i, \lambda_j\}$. In other words, μ is obtained from λ by joining λ_i and λ_j into a single part μ_k. Similarly, for the third and fourth lines of the formula above, $\mu \setminus \{\mu_i, \mu_j\} = \lambda \setminus \lambda_k$, that is, μ is obtained by breaking a part λ_k in λ into two parts, μ_i and μ_j.

(3) The first order phase transition at $\theta = 1$ for the cut-off value is not surprising, because the Metropolis chain has different forms for $\theta < 1$ and for $\theta > 1$.

(4) θ denotes the inverse of the Ewens sampling parameter the chain P_θ converges to. This choice of convention is justified by the fact that the left eigenfunctions of the chain P_θ are the transition coefficients from the Jack polynomials with parameter θ to the power sum polynomials, as derived in [11].

(5) It will be interesting to see what happens when the θ value in the transition probability is allowed to be state dependent, but satisfying some uniform bound

$c < \theta(\lambda) < c^{-1}$ for c independent of n. My conjecture is that it will always take at least $\frac{1}{2}n\log n$ steps to converge to its stationarity distribution, which is no longer in the Ewens family.

The next four sections will be devoted to the proof of Theorem 2.1.

3. Preliminaries on $\mathcal{L}^2$ mixing time.

LEMMA 3.1. *Given a* reversible *ergodic Markov chain P on a finite state space X, let f_j be the right eigenfunctions*, *normalized so that*

$$\sum_x f_j(x)^2\pi(x) = 1,$$

with corresponding eigenvalues β_j. Then $g_j(x) := f_j(x)\pi(x)$ are left eigenfunctions of P, with the same *eigenvalues*, *satisfying*

$$\sum_x g_j(x)^2\frac{1}{\pi(x)} = 1.$$

Furthermore,

$$\frac{1}{\pi(x)} = \sum_j f_j^2(x), \tag{1}$$

$$\pi(x) = \sum_j g_j^2(x). \tag{2}$$

PROOF. By reversibility, we have $\pi(x)P(x,y) = \pi(y)P(y,x)$. Therefore,

$$\begin{aligned}\beta_j f_j(x) &= \sum_y P(x,y)f_j(y) = \sum_y \frac{\pi(x)}{\pi(x)}P(x,y)f_j(y)\\ &= \sum_y \frac{\pi(y)}{\pi(x)}P(y,x)f_j(y).\end{aligned}$$

Now multiplying both sides by $\pi(x)$, we get

$$\beta_j\pi(x)f_j(x) = \sum_y \pi(y)f_j(y)P(y,x).$$

This proves the first part.

The last two identities are nothing but a restatement of the fact that the matrix $\Pi(x,y) := \pi(x)P(x,y)/\pi(y)$ is doubly stochastic; that is, each row and column sums to 1. Here is a formal proof. Since $\{f_j\}$ forms a basis, we can decompose the function $z \mapsto 1_x(z)$ in it,

$$1_x(z) = \sum_j c_j f_j,$$

where the coefficients c_j are given by

$$c_j = \langle f_j, 1_x \rangle_{\mathcal{L}^2(\pi)} = \sum_z 1_x(z) f_j(z)\pi(z) = f_j(x)\pi(x).$$

The first equality follows immediately. The second is similar. $\square$

LEMMA 3.2. *Under the same notation as the previous lemma, one can bound the total variation distance to stationarity at time k starting at state x by*

$$4\|P_x^k - \pi\|_{\mathrm{TV}}^2 \le \left\| \frac{P_x^k}{\pi(x)} - 1 \right\|_{\mathcal{L}^2(\pi)}^2 \tag{3}$$

$$= \frac{1}{\pi^2(x)} \sum_j \beta_j^{2k} g_j^2(x) - 1. \tag{4}$$

PROOF. The first inequality (3) follows directly from the Cauchy–Schwarz inequality. To prove the second formula (4), first write

$$\left\| \frac{P_x^k}{\pi} - 1 \right\|_2^2 = \sum_y \pi(y)\left[\left(\frac{P_x^k(y)}{\pi(y)}\right)^2 - 1\right] = \sum_y \frac{(P_x^k(y))^2}{\pi(y)} - \pi(y).$$

Using reversibility again [in the extended form $\pi(x)P^k(x,y) = \pi(y)P^k(y,x)$], we can write

$$\frac{(P_x^k(y))^2}{\pi(y)} = \frac{P_x^k(y)P_y^k(x)}{\pi(x)}.$$

Thus summing over $y \in X$, we get

$$\left\| \frac{P_x^k}{\pi} - 1 \right\|_2^2 = \frac{P^{2k}(x,x)}{\pi(x)} - 1.$$

Next write the function $y \mapsto P_x^{2k}(y)$ as the result of a row vector multiplied by a matrix,

$$P_x^{2k}(y) = \sum_z 1_x(z) P^{2k}(z,y).$$

By the previous lemma, we have

$$1_x(y) = \sum_j c_j g_j(y), \tag{5}$$

where $c_j = \sum_z \frac{1}{\pi(z)} 1_x(z) g_j(z) = \frac{g_j(x)}{\pi(x)}$.

Finally evaluating at $y = x$ in (5), we obtain

$$\frac{P^{2k}(x,x)}{\pi(x)} = \sum_j \frac{g_j(x)^2}{\pi(x)^2} \beta_j^{2k}.$$

$\square$

4. Results from symmetric function theory. First we recall from Hanlon [11] that the eigenvalues for the chain P_θ projected onto conjugacy classes of S_n, with $\theta > 1$, are given by

$$\beta_\lambda = \frac{n(\lambda^t) - \theta^{-1} n(\lambda)}{\binom{n}{2}}. \tag{6}$$

For an independent proof with pointers to literature, see the proof of Theorem 4.2.

Next we derive the eigenvalues of the chain P_θ, for $\theta \in (0, 1)$. Notice this is not the same chain as that studied in [11]. Here the identity element gets the biggest mass, whereas in [11], identity has the smallest mass [Ewens sampling with parameter $\in (0, 1)$]. But the same result of Macdonald can be used here to derive eigenvalues. Indeed, consider the following matrix T_θ defined by

$$T_\theta(\pi, \sigma) = \begin{cases} \dfrac{(\theta - 1)n(\pi)}{\theta\binom{n}{2}}, & \text{if } \sigma = \pi, \\ \dfrac{1}{\binom{n}{2}}, & \text{if } \sigma = \pi(i, j) \text{ and } \ell(\sigma) = \ell(\pi) - 1, \\ \dfrac{1}{\theta\binom{n}{2}}, & \text{if } \sigma = \pi(i, j) \text{ and } \ell(\sigma) = \ell(\pi) + 1, \\ 0, & \text{otherwise,} \end{cases}$$

where $n(\pi) = \sum_i \binom{\pi_i^t}{2}$ and $\{\pi_i\}$ is the partition structure of π. This quantity gives the number of ways to break a part in the partition structure of π into two parts, using multiplication by a transposition.

Hanlon considered the case $\theta \geq 1$ (his α is our θ), here we extend to $\theta \in (0, 1)$, which is no longer a Markov matrix because the diagonal entries are no longer nonnegative. Nevertheless The rows still sum to 1. Then his Theorem 3.5 continues to hold because the proof never uses $\theta \geq 1$. Likewise, Theorem 3.9 holds for $\theta < 1$. To get P_θ, we simply need to rescale T_θ by θ and add a constant multiple cI of identity matrix. c can be obtained by looking at the top eigenvalue. By Theorems 3.5 and 3.9 of [11], θT_θ has eigenvalues $\frac{\theta n(\lambda^t) - n(\lambda)}{\binom{n}{2}}$. Thus we need to add $1 - \theta$ in order for $\beta_{(n)}$ to equal 1.

Combining the two cases, we have the following formula for eigenvalues of P_θ:

$$\beta_\lambda(\theta) = 1 - \theta \wedge 1 + \frac{\theta n(\lambda^t) - n(\lambda)}{(\theta \vee 1)\binom{n}{2}}. \tag{7}$$

Denote $r(\lambda) = \frac{n(\lambda^t) - n(\lambda)}{\binom{n}{2}}$. The following lemma collects a bunch of estimates about β_λ:

LEMMA 4.1. *Let $\succeq$ be the natural partial order on the set of partitions defined as follows: given two partitions represented by Ferrers diagrams λ and λ', say $\lambda \succeq \lambda'$ if λ can be obtained by successive up and right moves of blocks of λ'.*

(1) *$n(\lambda^t)$ is monotone, and $n(\lambda)$ is anti-monotone in the above partial order; that is, for $\lambda \succeq \lambda'$,*

$$n(\lambda^t) > n(\lambda'^t), \qquad n(\lambda) < n(\lambda').$$

(2) *β_λ is monotone with respect to the natural partial order on λ. Thus $\beta_{(n)} \geq \beta_\lambda$ for all $\lambda \vdash n$, $\beta_\lambda \leq \beta_{(\lambda_1, n-\lambda_1)}$ and $\beta_\lambda \geq \beta_{(\lambda_1, 1^{n-\lambda_1})}$.*

(3) *Furthermore, for $\lambda_1 \geq \frac{n}{2}$,*

$$\beta_\lambda \leq 1 - (\theta \wedge 1)\frac{2\lambda_1(n-\lambda_1)}{n(n-1)}, \tag{8}$$

and in general

$$\beta_\lambda \leq 1 - (\theta \wedge 1)\left(1 - \frac{\lambda_1 - 1}{n-1}\right). \tag{9}$$

In particular, if $\beta_\lambda(\theta) \geq 0$, the above two inequalities hold with $|\beta_\lambda|$.

(4) *Finally, if $\beta_\lambda(\theta) < 0$, then $\beta_{\lambda^t}(\theta) > 0$ and $|\beta_\lambda| \leq \beta_{\lambda^t}$.*

PROOF. (1) It suffices to check the first assertion for λ and λ' that differ by one block, that is, $\lambda_i = \lambda_i' + 1$, $\lambda_j = \lambda_j' - 1$, $i < j$. Then

$$\begin{aligned} n(\lambda^t) - n(\lambda'^t) &= \tfrac{1}{2}\big[\lambda_i(\lambda_i - 1) + \lambda_j(\lambda_j - 1) - \lambda_i'(\lambda_i' - 1) - \lambda_j'(\lambda_j' - 1)\big] \\ &= \lambda_i' - \lambda_j \geq 0, \end{aligned}$$

using the fact $\lambda_i \geq \lambda_j$ and $\lambda_i' \geq \lambda_j'$ by definition of Ferrers diagram. The antimonotonicity of $n(\lambda)$ follows by taking transpose.

(2) This follows directly from the previous assertion and formula (7) for β_λ in terms of $n(\lambda)$ and $n(\lambda^t)$.

(3) Equation (8) follows from $\lambda \preceq (\lambda_1, n - \lambda_1)$ and monotonicity, after throwing away the term $-\frac{n(\lambda)}{(\theta \vee 1)\binom{n}{2}}$.

For (9), we again throw away the term $-\frac{n(\lambda)}{\binom{n}{2}}$ in β_λ to obtain

$$\begin{aligned} \beta_\lambda &\leq 1 - \theta \wedge 1 + \frac{\theta n(\lambda^t)}{(\theta \vee 1)\binom{n}{2}} = 1 - (\theta \wedge 1)\left(1 - \frac{n(\lambda^t)}{\binom{n}{2}}\right) \\ &= 1 - (\theta \wedge 1)\left(1 - \frac{\sum_j \lambda_j(\lambda_j - 1)}{n(n-1)}\right) \\ &\leq 1 - (\theta \wedge 1)\left(1 - \frac{(\lambda_1 - 1)\sum_j \lambda_j}{n(n-1)}\right) = 1 - (\theta \wedge 1)\left(1 - \frac{\lambda_1 - 1}{n-1}\right). \end{aligned}$$

(4) Here we consider $\theta \geq 1$ and $\theta < 1$ separately. When $\theta \geq 1$,

$$\beta_\lambda = \frac{1}{\theta\binom{n}{2}}\big(\theta n(\lambda^t) - n(\lambda)\big).$$

If $\beta_\lambda \le 0$, then $\theta n(\lambda^t) - n(\lambda) \le 0$. Since $\theta \ge 1$, $n(\lambda) \ge n(\lambda^t)$. So

$$\theta n(\lambda) - n(\lambda^t) \ge n(\lambda) - \theta n(\lambda^t) \ge 0,$$

which implies $\beta_{\lambda^t} \ge |\beta_\lambda| \ge 0$.

Next let $\theta < 1$. Then we can write

$$\beta_\lambda = 1 - \theta + \theta \frac{n(\lambda^t)}{\binom{n}{2}} - \frac{n(\lambda)}{\binom{n}{2}} = \left[1 - \frac{n(\lambda)}{\binom{n}{2}}\right] - \theta\left[1 - \frac{n(\lambda^t)}{\binom{n}{2}}\right].$$

If $\beta_\lambda \le 0$, then since $\theta < 1$,

$$1 - \frac{n(\lambda^t)}{\binom{n}{2}} \ge 1 - \frac{n(\lambda)}{\binom{n}{2}}.$$

Switching λ and λ^t we again get

$$\beta_{\lambda^t} \ge |\beta_\lambda| \ge 0. \qquad \square$$

We also need some definitions and results from Diaconis and Hanlon [21]:

DEFINITION 1. We collect some notation to be used in the main proof below, some of which will be repeated; they are, for the most part, taken from [21]:

(1) Given a partition $\lambda \vdash n$, and a position $s = (i, j)$ in its Ferrers diagram (i.e., $j \le \lambda_i$), define

$$h^*(s) = h^*_\lambda(s) := (a+1)\theta + \ell,$$
$$h_*(s) = h^\lambda_*(s) := a\theta + (\ell + 1),$$

where $a = \lambda_i - j$ denotes the number of positions in the same row and strictly to the right of s (the arm length), and $\ell = \lambda^t_j - i$ denotes the number of positions in the same column and strictly below s (the leg length).

(2) Define the generalization of hooklength product,

$$j_\lambda = j_\lambda(\theta) := \prod_{s \in \lambda} h_*(s) h^*(s).$$

When $\theta = 1$, this becomes the product of the hooklengths of all the blocks in the diagram of λ.

(3) Define $c_{\lambda,\rho} = c_{\lambda,\rho}(\theta)$ to be the change of basis coefficients from Jack symmetric polynomials $J_\lambda(\theta)$ (not to be confused with j_λ above) to power sum polynomials, that is,

$$J_\lambda(\theta) = \sum_{\rho \vdash n} c_{\lambda,\rho}(\theta) p_\rho. \tag{10}$$

See [20] for extensive development of properties of Jack polynomials. When $\theta = 1$, $J_\lambda(1) = H_\lambda s_\lambda$, where $H_\lambda = j_\lambda(1)$ is the hooklength product, and s_λ is the Schur polynomial indexed by λ.

(4) Denote by $\pi = \pi_\theta$ the Ewens sampling measure with parameter θ^{-1}; recall $\pi_\theta(\sigma) = \theta^{-\ell(\sigma)}/z_n(\theta^{-1})$, where $z_n(\theta^{-1}) = \prod_{i=1}^n(\theta^{-1}+i-1)$ is the Ewens sampling formula. Also let $\Pi = \Pi_{n,\theta} := \pi_\theta(1^n)^{-1} = \prod_{i=1}^n(1+\theta(i-1))$.

Note that when $\theta = 1$, $j_\lambda(1)$ is exactly the square of the product of hook lengths of all positions in λ, which is well known to be $(\frac{n!}{\dim \pi_\lambda})^2$ by the hooklength formula. By Wedderburn's structure theorem (see [8], Chapter 18, Theorem 10), we also have

$$n! = \sum_\lambda \dim \pi_\lambda^2,$$

therefore $\sum_\lambda \frac{1}{j_\lambda(1)} = \frac{1}{n!}$.

THEOREM 4.2. *The left eigenfunctions of the Metropolis chain P_θ defined on partitions, normalized in $\mathcal{L}^2(\mathcal{P}_n, 1/\pi_\theta)$, are given by*

$$g_\lambda(\rho) = \frac{c_{\lambda,\rho}}{(j_\lambda \Pi/(\theta^n n!))^{1/2}} \tag{11}$$

with corresponding eigenvalues stated in (7).

PROOF. We synthesize the arguments found in [11], Definition 3.8 to Definition 3.12 and [21], Theorem 1. The result from [17], Chapter VI, Section 4, shows that the Macdonald polynomials are simultaneous eigenfunctions of the Macdonald operators $D_{q,t}^r$, $r = 0, \dots, n$. Specializing to the limit $q = t^\theta$, $t \to 1$ and after some affine linear transformation, the same results hold for Jack polynomials and the associated Sekiguchi–Debiard operators (37), $D_\theta(X)$. The X^2 coefficient of this operator valued generating function turns out to be the following Laplace–Beltrami-type operator (our notation differs slightly from [17], page 320): let f be a homogeneous polynomial of degree N in n variables, then

$$D_\theta^2 f = \left(-\frac{\theta^2}{2} U_n - \theta V_n + c_n\right) f,$$

where $U_n = \sum_{i=1}^n (x_i \partial_i)^2 - x_i \partial_i = \sum_{i=1}^n x_i^2 \partial_i^2$, $V_n = \frac{1}{2}\sum_{i\neq j} \frac{x_i^2\partial_i - x_j^2 \partial_j}{x_i - x_j}$, and $c_n = \theta^2\binom{N}{2} + \theta N\binom{n}{2} + \frac{1}{4}\binom{n}{3}(3n-1)$; see (40) for a proof.

After an affine transform, we arrive at the following cleaner operator:

$$L_\theta^2 := \frac{1}{\binom{N}{2}}\left(\frac{1}{2} U_n + \frac{1}{\theta}(V_n - (n-1)N)\right), \tag{12}$$

which readily admits a Markov chain interpretation, when acting on power sum polynomials. Combining (39), (38) and (40), we have

$$L_\theta^2 p_\lambda = \frac{p_\lambda}{\binom{N}{2}}\left(\sum_{s<t} \lambda_s \lambda_t \frac{p_{\lambda_s+\lambda_t}}{p_{\lambda_s} p_{\lambda_t}} + (1-\theta^{-1}) n(\lambda') + \theta^{-1} \sum_s \frac{\lambda_s}{2} \sum_{r=1}^{\lambda_s - 1} \frac{p_r p_{\lambda_s - r}}{p_{\lambda_s}}\right).$$

Observe that for $\theta > 1$, the first and third terms above correspond to joining two cycles into one and splitting a cycle into two cycles, respectively, whereas the middle term gives the holding probability at λ. In other words, the probability of going from λ to μ in one step under the Jack–Metropolis walk is given by the p_μ coefficient of $L_\theta^2 p_\lambda$. This translates to $L_\theta^2 p_\lambda = \sum_{\mu \vdash N} T_\theta(\lambda, \mu) p_\mu$.

Next we show that $L_\theta^2 J_\lambda = \beta_\lambda J_\lambda$, with β given by (6) (i.e., when $\theta > 1$); the general case follows by an appropriate affine transform. In [17], page 317, it is shown that for the Macdonald operator-valued generating function $D_n(X; q, t)$, eigenvalues are given by $\beta_\lambda(X; q, t) = \prod_{i=1}^n (1 + X t^{n-i} q^{\lambda_i})$. Now using Example 3(c) on page 320, one can derive the eigenvalues for $D_n(X; \alpha)$, by considering the limiting operator $\lim_{t \to 1} (t-1)^{-n} Y^n D_n(Y^{-1}; q, t)$ where $Y = (t-1)X - 1$. Extracting the X^{n-2} term gives β_λ, which is stated in Example 3(b) of page 327 as $\alpha \binom{N}{2} e_\lambda(\alpha)$ (α is the same as θ in our notation), since $\square_n^\alpha = \alpha \binom{N}{2} L_\alpha^2$.

Finally we prove the formula for the left eigenfunctions. Define the inner product $\langle \cdot, \cdot \rangle_\theta$ by $\langle p_\lambda, p_\mu \rangle_\theta = \delta_{\lambda\mu} z_\lambda \theta^{\ell(\lambda)}$. In [20] (see also Lemma 3.11 of [11]), it is shown that $\langle J_\lambda, J_\lambda \rangle_\theta = j_\lambda(\theta)$ as defined before. Here recall the normalization of J_λ is fixed by requiring that in the monomial symmetric function basis, its m_{1^N} coefficient be 1. Therefore expressing J_λ in terms of p_μ's, we have

$$\sum_{\rho \vdash N} c_{\lambda,\rho}^2 z_\rho \theta^{\ell(\rho)} = j_\lambda.$$

On the other hand, the normalization constant for the MED(θ^{-1}) distribution is $z_n(\theta^{-1}) = \theta^{-1}(\theta^{-1} + 1) \cdots (\theta^{-1} + n - 1) = \Pi \theta^{-n}$, hence $\pi_\theta(\rho) = \theta^{-\ell(\rho)} \frac{n!}{z_\rho} z_n(\theta)^{-1}$, and with $g_\lambda(\rho)$ given in (11) we have

$$\sum_{\rho \vdash N} g_\lambda(\rho)^2 \pi_\theta(\rho)^{-1} = \sum_\rho \frac{c_{\lambda,\rho}^2}{j_\lambda \Pi / (\theta^n n!)} \Pi \theta^{\ell(\rho) - n} \frac{z_\rho}{n!} = \sum_\rho \frac{c_{\lambda,\rho}^2 \theta^{\ell(\rho)} z_\rho}{j_\lambda} = 1,$$

by the previous equation. This shows g_λ are indeed left eigenfunctions by Lemma 3.1. $\square$

COROLLARY 4.3. *The right eigenfunctions of P_θ are proportional to*

$$f_\lambda(\rho) = g_\lambda(\rho) \theta^{\ell(\rho)} z_\rho.$$

PROOF. Since $\pi_\theta(\rho) \propto \theta^{-\ell(\rho)} \frac{n!}{z_\rho}$, this follows from Lemma 3.1 and the previous theorem. $\square$

LEMMA 4.4. *For any $\lambda \vdash n$,*

$$c_{\lambda, 1^n} = 1.$$

PROOF. This follows from the following formula in [20]:

$$J_\lambda(1^n;\theta) = \prod_{(i,j)\in\lambda} \big(n-(i-1)+\theta(j-1)\big),$$

true for all $n \in \mathbb{N}$, by reading coefficients of powers of n; See [21], Section 4, Theorem 1. □

LEMMA 4.5. *j_λ admits the following inductive bound on the parts of λ:*

$$j_\lambda \geq \lambda_1!^2\theta^{2\lambda_1-1}\lambda_1^{\theta^{-1}-1}e^{-\pi^2/12\theta^2}j_{(\lambda_2,\dots,\lambda_n)}.$$

Note that the constant $e^{-\pi^2/12\theta^2}$ is not important.

PROOF OF LEMMA 4.5. From the definition, we have

$$\begin{aligned}
j_\lambda &\geq \left[\prod_{i=1}^{\lambda_1}(i\theta)\prod_{i=1}^{\lambda_1-1}(i\theta+1)\right]j_{(\lambda_2,\dots,\lambda_n)} \\
&= \left[\prod_{i=1}^{\lambda_1}(i\theta)\prod_{i=1}^{\lambda_1-1}(i\theta)\big(1+(i\theta)^{-1}\big)\right]j_{(\lambda_2,\dots,\lambda_n)} \\
&\geq \lambda_1!\theta^{\lambda_1}(\lambda_1-1)!\theta^{\lambda_1-1}\exp\left(\frac{1}{\theta}\log\lambda_1 - \frac{1}{\theta^2}\frac{\pi^2}{12}\right)j_{(\lambda_2,\dots,\lambda_n)} \\
&= \lambda_1!\theta^{\lambda_1}(\lambda_1-1)!\theta^{\lambda_1-1}\lambda_1^{1/\theta}\exp\left(-\frac{1}{\theta^2}\frac{\pi^2}{12}\right)j_{(\lambda_2,\dots,\lambda_n)},
\end{aligned} \tag{13}$$

where we used the fact that $1+x \geq e^{x-x^2/2}$ for $x \geq 0$, applied to $x = (i\theta)^{-1}$, and the zeta sum,

$$\sum_i \frac{1}{2i^2} \leq \frac{\pi^2}{12}.$$

□

5. Mixing time upper bound. By Theorem 4.2, under the same notation there, *four times* the total variation distance of $P_{1^n}^k$ from π can be bounded by

$$\|P_x^k - \pi\|_2^2 \leq \frac{1}{\pi^2(x)}\sum_\lambda \beta_\lambda^{2k}g_\lambda^2(1^n) - 1,$$

where we use the sloppy (but standard) notation $\|P_x^k - \pi\|_2$ to mean $\|\frac{P_x^k}{\pi} - 1\|_{\mathcal{L}^2(\pi)}$.

For $\lambda = (n)$, corresponding to the trivial representation on S_n, and starting point $x = (1^n)$, the summand exactly cancels -1: $\beta_{(n)} = 1$, $j_{(n)} = \Pi\theta^n n!$ and $c_{(n),1^n} = 1$ (by Lemma 4.4), whereas $\pi(1^n) = \Pi^{-1}$, so

$$\frac{1}{\pi(1^n)^2}\beta_{(n)}^{2k}g_{(n)}^2(1^n) = 1.$$

Thus using the explicit formula for g_λ, we immediately have

$$\| P_x^k - \pi \|_2^2 = \theta^n n! \Pi_n \sum_{\lambda \vdash n, \lambda \neq (n)} \frac{\beta_\lambda^{2k}}{j_\lambda}. \tag{14}$$

We now break the sum according to the sign of β_λ,

$$\sum_{\lambda \vdash n, \lambda \neq (n)} \frac{\beta_\lambda^{2k}}{j_\lambda} = \sum_{\beta_\lambda \geq 0}^{*} \frac{\beta_\lambda^{2k}}{j_\lambda} + \sum_{\beta_\lambda < 0} \frac{\beta_\lambda^{2k}}{j_\lambda}, \tag{15}$$

where $\sum^*$ denotes summation skipping the top eigenvalue indexed by $\lambda = (n)$. Next we can rewrite the first summand on the right according to the size of λ_1, and obtain the following bound:

$$\begin{aligned} \theta^n n! \Pi_n \sum_{\beta_\lambda \geq 0}^{*} \frac{\beta_\lambda^{2k}}{j_\lambda} &= \sum_{s=n-1}^{1} \sum_{\lambda : \lambda_1 = s, \beta_\lambda \geq 0} \theta^n n! \Pi_n \frac{\beta_\lambda^{2k}}{j_\lambda} \\ &\leq \sum_{s=n-1}^{1} \max\{\beta_\lambda^{2k} : \beta_\lambda \geq 0, \lambda_1 = s\} \sum_{\lambda : \lambda_1 = s} \frac{\theta^n n! \Pi_n}{j_\lambda}. \end{aligned} \tag{16}$$

Splitting $\Pi = \Pi_{n,\theta}$ into two subproducts, and using Lemma 4.5, we have

$$\frac{\Pi n! \theta^n}{j_\lambda} \leq \frac{(\Pi_n/(\Pi_{n-\lambda_1}))(n!/((n-\lambda_1)!))\theta^{\lambda_1}}{\theta^{2\lambda_1 - 1} \lambda_1!^2 \lambda_1^{\theta^{-1}-1} e^{-\pi^2/12\theta^2}} \frac{\Pi_{n-\lambda_1}(n-\lambda_1)! \theta^{n-\lambda_1}}{j_{(\lambda_2, \ldots, \lambda_n)}},$$

where the second quotient factor happens to be $\frac{1}{\pi(1^{n-\lambda_1})^2} g^2_{(\lambda_2,\ldots,\lambda_n)}(1^{n-\lambda_1})$; see Theorem 4.2 and Lemma 4.4. Also denote the first factor by q_{n,λ_1}.

By (2), the definition of $\Pi_{n-s} := \pi(1^{n-s})^{-1}$ (see Definition 1), and the fact $\Pi_n = \theta^n (n-1)! e^{\theta^{-1} \sum_{i=2}^{n} 1/(i-1)}$, we can bound the summand of the right-hand side of (16) for a fixed s as

$$\begin{aligned} \sum_{\lambda : \lambda_1 = s} \frac{\Pi_n n! \theta^n}{j_\lambda} &\leq q_{n,s} \sum_{\mu \vdash n-s} \frac{1}{\pi(1^{n-s})^2} g_\mu^2(1^{n-s}) = q_{n,s} \frac{1}{\pi(1^{n-s})} \\ &\leq \left(\frac{n}{s}\right)^{\theta^{-1}-1} \theta^{n-s+1} \left(\frac{n!}{s!}\right)^2 \frac{e^{\pi^2/12\theta^2}}{(n-s)!}. \end{aligned}$$

We will now reduce the $\mathcal{L}^2$ bound (14) to bounding the following quantity:

$$b_{n,+} := \sum_{s=n-1}^{1} \beta_{s,+}^{2k} \left(\frac{n}{s}\right)^{\theta^{-1}-1} \theta^{n-s+1} \left(\frac{n!}{s!}\right)^2 \frac{e^{\pi^2/12\theta^2}}{(n-s)!}, \tag{17}$$

where $\beta_{s,+} := \max\{\beta_\lambda : \beta_\lambda \geq 0, \lambda_1 = s\}$.

For the second summand of (15), we obtain

$$
\begin{aligned}
\sum_{\beta_\lambda<0} \frac{\beta_\lambda^{2k}}{j_\lambda} \le{}& \sum_{s=n-1}^{1} \max\{\beta_\lambda^{2k} : \beta_\lambda < 0, \lambda_1^t = s\} \sum_{\lambda : \lambda_1^t = s} \frac{\theta^n n! \Pi_n}{j_\lambda} \\
&+ \beta_{1^n}^{2k} \frac{\theta^n n! \Pi_n}{j_{1^n}}.
\end{aligned} \tag{18}
$$

Using the explicit formula (7) for β_λ, we get

$$
\beta_{1^n} = 1 - (\theta \wedge 1) - (\theta^{-1} \wedge 1) \in (-1, 1),
$$

for $\theta \neq 1$. On the other hand,

$$
\Pi_n n! \theta^n / j_{1^n} = \prod_{i=1}^{n-1} (1 + i\theta) n! \theta^n \Big/ \prod_{i=1}^{n} (i\theta)\bigl(1 + (i-1)\theta\bigr) = 1,
$$

since $j_{1^n} = \prod_{i=1}^n (i\theta)(1 + (i-1)\theta)$ by definition. Thus

$$
\beta_{1^n}^{\Omega(n)} \Pi_n n! \theta^n / j_{1^n} = o(1),
$$

which has negligible contribution in (14).

For the remaining terms in (18), first observe that

$$
j_{\lambda^t} \ge \prod_{i=1}^{\lambda_1} i \prod_{i=1}^{\lambda_1 - 1} (i + \theta) j_{(\lambda_2, \dots, \lambda_n)^t}.
$$

Hence when $\theta < 1$,

$$
j_{\lambda^t} \ge \prod_{i=1}^{\lambda_1} (i\theta) \prod_{i=1}^{\lambda_1 - 1} (i\theta + 1) j_{(\lambda_2, \dots, \lambda_n)^t}
$$

and using the bound $|\beta_\lambda| \le \beta_{\lambda^t}$, for $\beta_\lambda < 0$, we get

$$
\sum_{\beta_\lambda < 0} \frac{\beta_\lambda^{2k}}{j_\lambda} \le \sum_{\beta_\lambda \ge 0}^{*} \frac{\beta_\lambda^{2k}}{j_\lambda} + o(1).
$$

If $\theta > 1$, the j_{λ^t} is comparable to j_λ within an exponential factor

$$
j_{\lambda^t} \ge j_\lambda \theta^{2n}.
$$

Furthermore by the explicit formula of β_λ, we have

$$
\begin{aligned}
-\beta_\lambda(\theta) &= -\left(\frac{n(\lambda^t)}{\binom{n}{2}} - \theta^{-1} \frac{n(\lambda)}{\binom{n}{2}} \right) \le \theta^{-1} \frac{n(\lambda)}{\binom{n}{2}} - \frac{n(\lambda^t)}{\binom{n}{2}} \\
&\le \theta^{-1} \left(\frac{n(\lambda) - n(\lambda^t)}{\binom{n}{2}} \right) \le -\theta^{-1} \beta_\lambda(1).
\end{aligned}
$$

So since $k := \frac{1}{2(\theta \wedge 1)} n(c + \log n) = \Theta(n \log n)$, $\theta^{-2k}\theta^{2n} = o(1)$. Thus we can still compare the negative β_λ sum to the positive one,

$$\sum_{\beta_\lambda < 0} \frac{\beta_\lambda^{2k}}{j_\lambda} \le b_{n,+} + o(1).$$

It remains to bound $2b_{n,+}$. First note that the factor 2 in front is immaterial, since for $\lambda \neq (n)$,

$$\beta_\lambda^{cn} < \beta_{(n-1,1)}^{cn} \le e^{-\Omega_\theta(c)},$$

thanks to the monotonicity of β_λ's. So by increasing c in $k = \frac{1}{2(\theta \wedge 1)} n(c + \log n)$, we can decrease $b_{n,+}$ by a factor of 2. The factor $c_\theta = e^{\pi^2/2\theta^2}$ can be ignored similarly. We can also get rid of the factor $(\frac{n}{s})^{\theta^{-1}-1}$ in (17) as follows.

For $s \ge n/2$, $(\frac{n}{s})^{\theta^{-1}-1} = \mathcal{O}_\theta(1)$, so again increasing c annihilates it. For $s < n/2$, recall the second bound on β_λ (9), which implies that for $\lambda_1 < n/2$, β_λ is bounded away from 1 uniformly in n. Now in the definition of $b_{n,+}$, β_λ is assumed to be nonnegative (alternatively, β_{1^n} is bounded uniformly away from -1), hence raising β_λ to the power $\Omega(n)$ easily cancels any power of n, that is,

$$n^{\theta^{-1}-1} \beta_\lambda^{cn} = o(1).$$

So together, by increasing c, we can reduce the problem to bounding the following quantity:

$$B_{n,+} := \sum_{s=n-1}^{1} \frac{\theta^{n-s+1}}{(n-s)!} \left(\frac{n!}{s!}\right)^2 \beta_{s,+}^{2k}.$$

The only estimates we rely on now are (8) and (9) from Lemma 4.1; the idea will be similar to [7]; see also [5]. First note that it suffices to show

$$\frac{\theta^{n-s+1}}{(n-s)!} \left(\frac{n!}{s!}\right)^2 \beta_{s,+}^{2k} = \mathcal{O}(1), \tag{19}$$

uniformly for all $s \in [1, n-1]$ and c sufficiently large. Indeed using (9), we have

$$\beta_{s,+} \le e^{-x-x^2/2} \le e^{-x}$$

for $x = (\theta \wedge 1)\frac{n-s}{n}$. Therefore

$$\mathcal{O}(1) \sum_{s=n-1}^{1} \beta_{s,+}^{(cn)/(2(1\wedge\theta))} \le \mathcal{O}(1) \sum_{t=1}^{n} e^{-tc} = o_c(1),$$

by geometric summation; in fact, using (8) we can get a better bound, but that's not necessary.

Next recall (8) as well as the estimates (no Stirling formula needed)

$$\frac{n!}{s!} \le e^{\int_s^n \log x\,dx + \log n - \log s} = e^{n\log n - s\log s - (n-s) + \log n - \log s},$$

$$(n-s)! \ge e^{\int_1^{n-s} \log x\,dx} = e^{(n-s)\log(n-s) - (n-s-1)}.$$

Taking logarithm, and letting $s = \alpha n$, we can bound the left-hand side of (19) by

$$\begin{aligned}
&\log\left[\frac{\theta^{n-s+1}}{(n-s)!}\left(\frac{n!}{s!}\right)^2 \beta_{s,+}^{2k}\right] \\
&\quad \le (1-2\alpha)(1-\alpha)n\log n - (1-\alpha)n\log(1-\alpha) \qquad (20)\\
&\qquad - 2\alpha n\log\alpha - 2\log\alpha + \bigl(C_1(\theta) - 2c\alpha\bigr)(1-\alpha)n + C_2(\theta),
\end{aligned}$$

where $C_1(\theta)$, $C_2(\theta)$ are constants that depend only on θ.

For $\alpha \ge \alpha_0 \in (1/2, 1)$, the right-hand side of (20) can be further simplified to

$$(1-\alpha)n\bigl[(1-2\alpha)\log n - \log(1-\alpha) + C_1'(\theta) - c\bigr] + C_2'(\theta).$$

Since $(1-\alpha)n \ge 1$, and we can choose c as large as we want, it suffices to show the expression inside the square brackets above is $\mathcal{O}(1)$. But when $\alpha = 1 - 1/n$ or $1/2$, this is clearly true. Furthermore, the derivative

$$\frac{d}{d\alpha}\bigl[(1-2\alpha)\log n - \log(1-\alpha)\bigr] = -2\log n + \frac{1}{1-\alpha}$$

is monotone increasing, showing that the function $\alpha \mapsto (1-2\alpha)\log n - \log(1-\alpha)$ is convex, and its value for any $\alpha \in [1/2, 1-1/n]$ is bounded above by the values at the boundary points.

Next let $\alpha < \alpha_0$. Using the second bound for β_λ, (9), we have

$$\beta_\lambda \le e^{-(\theta\wedge 1)(1-(s/n))}.$$

Then the logarithm of the left-hand side of (19) is bounded by

$$\begin{aligned}
&(1-\alpha)n\log n + (1-\alpha)\bigl(\log\theta - \log(1-\alpha) + 1\bigr)n - 2\alpha n\log\alpha \\
&\quad + 2\log\alpha - (1-\alpha)n\log n - (1-\alpha)cn \qquad (21)\\
&\quad \le (1-\alpha)\bigl(C(\theta) - c\bigr)n + 2(1-\alpha n)\log\alpha.
\end{aligned}$$

Clearly $(1-\alpha n)\log\alpha = \mathcal{O}(n)$ for $\alpha \in [\frac{1}{n}, \alpha_0]$. So for sufficiently large c, the right-hand side above goes to $-\infty$. Together this shows (19) is true for all $s \in [\frac{1}{n}, 1-\frac{1}{n}]$, and concludes the upper bound for the mixing time.

6. Mixing time lower bound.

We rely heavily on results from [20]. Again we collect some notation needed in the analysis below:

DEFINITION 2. For $\lambda, \rho \vdash n$, let:

- H_λ be the product of all hook-lengths of the Ferrers diagram for λ;
- $z_\rho := \prod_{i=1}^n i^{m_i} m_i!$ and $m_i = m_i(\rho)$ is the number of parts in ρ of length i;

- $\chi_\lambda(\rho)$ be the character of S_n indexed by λ evaluated at an element of conjugacy class ρ; alternatively, they can be defined by the system

$$p_\rho = \sum_\lambda \chi_\lambda(\rho) s_\lambda,$$

where s_ρ are the Schur polynomials.

Warning. Note the Schur polynomials are not direct specializations of the Jack polynomials; they differ by a factor

$$s_\lambda = H_\lambda^{-1} J_\lambda(1). \tag{22}$$

Thus the matrix $\chi_\lambda(\rho)$ is the inverse of $H_\lambda^{-1} c_{\lambda,\rho}(1)$.

LEMMA 6.1 ([17], Chapter I, equation (7.5); see also [20], equation (50)). *The relation between $c_{\lambda,\rho}(1)$ and $\chi_\lambda(\rho)$ is given by*

$$c_{\lambda,\rho}(1) = H_\lambda z_\rho^{-1} \chi_\lambda(\rho).$$

COROLLARY 6.2. *The inverse matrix to $\chi_\lambda(\rho)$ is given by $\chi_\lambda(\rho) z_\rho^{-1}$, that is,*

$$s_\lambda = \sum_\rho \chi_\lambda(\rho) z_\rho^{-1} p_\rho.$$

PROOF. By relation (22) and the lemma above, we have

$$\begin{aligned} s_\lambda &= H_\lambda^{-1} J_\lambda(1) \\ &= H_\lambda^{-1} \sum_\rho H_\lambda z_\rho^{-1} \chi_\lambda(\rho) p_\rho. \end{aligned}$$

Comparing the coefficients with (10) in Definition 1 yields the result. □

As in the $\theta = 1$ case studied by Diaconis and Shahshahani, the strategy will be to use a certain eigenfunction f of the chain as test function and compare the probabilities of the event $\{f < \eta\}$ for some $\eta \in \mathbb{R}$ under the stationary distribution and the distribution at time slightly before the mixing time, which in our case is $k(c) := \frac{1}{2(\theta \wedge 1)} n(\log n - c)$.

In the case where $\theta = 1$, $\rho \mapsto \chi_{(n-1,1)}(\rho) = m_1(\rho) - 1$ is the desired eigenfunction. So it is natural to guess that a suitable affine transformation of the fixed-point (aka 1-cycle) counting function $\rho \mapsto m_1(\rho)$ is the desired eigenfunction.

Lemma 6.1 shows that the following normalized version of $c_{\lambda,\rho}$ is the right analogue of characters of the symmetric group

$$d_{\lambda,\rho}(\theta) := c_{\lambda,\rho}(\theta) z_\rho \theta^{-(n-\ell(\rho))} H_\lambda^{-1}. \tag{23}$$

Thus our candidate test function will be $d_\lambda(\rho) = d_{\lambda,\rho}$.

It is straightforward to compute $\mathbb{P}_\infty(d_\lambda < \eta)$ where $\mathbb{P}_\infty$ denotes the stationary measure; the number of cycles are asymptotically independent and Poisson distributed. To estimate $\mathbb{P}_k(d_\lambda < \eta)$, one uses the second moment method. The first moment of $d_{(n-1,1)}$ is easily computed since it is proportional to the right eigenfunctions of the chain P_θ; see Corollary 4.3. For second moments, we need to decompose $d^2_{(n-1,1)}$ as linear combinations of other d_λ's. This is accomplished by first expressing $d_{(n-1,1)}$ and other d_λ's in terms of powers of m_i's, the number of i-cycles (not to be confused with monomial symmetric functions), then deducing the relationship by solving the appropriate system of linear equations. The analysis below will be an elaboration of this strategy.

First we need:

PROPOSITION 6.3 ([20], Proposition 7.5). *Let $d_{\lambda,\rho} = d_{\lambda,\rho}(\theta)$ be defined as above. Then*

$$\begin{aligned}&\theta^{k+1}(k+n)d_{(n,1^k),\rho}(\theta)\\&\quad= \sum_{j=0}^{k}(-1)^{k-j}\big(j+(n+k-j)\theta\big)\sum_{\nu\vdash j}\left[\prod_i \binom{m_i(\rho)}{m_i(\nu)}\right](-1)^{j-\ell(\nu)}\theta^{\ell(\nu)}.\end{aligned}$$

Note that the partitions in the proposition above is for $n+k$, rather than n.

From this, we easily obtain

$$d_{(n-1,1),\rho}(\theta) = -\frac{1}{\theta} + \frac{1+(n-1)\theta}{\theta n} m_1(\rho) \tag{24}$$

and

$$\begin{aligned}d_{(n-2,1^2),\rho}(\theta) &= \frac{1}{\theta^2} - \left(\frac{1+(n-1)\theta}{n\theta^2} + \frac{2+(n-2)\theta}{2\theta n}\right)m_1(\rho)\\&\quad + \frac{2+(n-2)\theta}{2\theta n}m_1(\rho)^2 - \frac{2+(n-2)\theta}{\theta^2 n}m_2(\rho).\end{aligned}\tag{25}$$

Note in particular,

$$\chi_{(n-1,1)}(\rho) = m_1(\rho) - 1, \tag{26}$$

$$\chi_{(n-2,1^2)}(\rho) = 1 - \tfrac{3}{2}m_1(\rho) + \tfrac{1}{2}m_1(\rho)^2 - m_2(\rho), \tag{27}$$

as expected.

Using the Schur–Weyl relation

$$\chi^2_{n-1,1} = \chi_n + \chi_{n-1,1} + \chi_{n-2,2} + \chi_{n-2,1^2},$$

and we also obtain

$$\chi_{(n-2,2)} = \tfrac{1}{2}m_1^2 - \tfrac{3}{2}m_1 + m_2.$$

To get $J_{(n-2,2)}$, we need the conjecture right after Proposition 7.2 as well as Corollary 3.5 from [20]. The conjecture has been proved in [14]. Notice the parameter α is the same as our parameter θ.

PROPOSITION 6.4 ([20], Proposition 7.2). *The Jack polynomials corresponding to the partition* $(2^i, 1^j)$ *have the following expansion in terms of the monomial symmetric basis* m_λ:

$$J_{(2^i,1^j)} = \sum_{r=0}^{i} (i)_r (\theta + i + j)_r \big(2(i-r) + j\big)! m_{(2^r, 1^{2(i-r)+j})},$$

where $(i)_r := i(i-1)\cdots(i-r+1)$.

PROPOSITION 6.5 ([14], Theorem 1.1). *In terms of Schur polynomial basis, we have*

$$J_{(2^i,1^j)} = \sum_{r=0}^{i} (i)_r (\theta + i + j)_r (i - r - \theta)_{i-r} (i + j - r)! s_{(2^r, 1^{2(i-r)+j})}.$$

The next result relates Jack polynomials corresponding to conjugate partitions, when expressed in terms of the power sum basis.

PROPOSITION 6.6 ([20], Corollary 3.5). *Let* $J_\lambda = \sum_\mu c_{\lambda\mu}(\theta) p_\mu$, *then*

$$J_{\lambda^t} = \sum_\mu (-\theta)^{n-\ell(\mu)} c_{\lambda\mu}\left(\frac{1}{\theta}\right) p_\mu.$$

We also recall from Corollary 6.2 that $s_\lambda = \sum_\rho \chi_\lambda(\rho) z_\rho^{-1} p_\rho$, where $\chi_\lambda(\rho)$ is the character of λ evaluated at ρ.

Combining the previous results, we easily get

$$\begin{aligned}
&J_{n-2,2}(\theta)\\
&\quad = \sum_\rho (-\theta)^{n-\ell(\rho)} \bigg[\left(2 - \frac{1}{\theta}\right)\left(1 - \frac{1}{\theta}\right)(n-2)\chi_\rho^{1^n}\\
&\qquad\qquad + 2\left(n - 2 + \frac{1}{\theta}\right)\left(1 - \frac{1}{\theta}\right)(n-3)!\chi_{2^1,1^{n-2}}(\rho)\\
&\qquad\qquad + 2\left(n - 2 + \frac{1}{\theta}\right)\left(n - 3 + \frac{1}{\theta}\right)(n-4)!\chi_{2^2,1^{n-4}}(\rho)\bigg] z_\rho^{-1} p_\rho,
\end{aligned}$$

where $\chi_\lambda(\rho)$ is the irreducible character λ evaluated at ρ. Therefore we can read off the coefficients

$$\begin{aligned}
&c_{(n-2,2),\rho}(\theta)\\
&\quad = \frac{(-\theta)^{n-\ell(\rho)}}{z_\rho}\bigg(\left(2 - \frac{1}{\theta}\right)\left(1 - \frac{1}{\theta}\right)(n-2)!\chi_{1^n}(\rho)\\
&\qquad\qquad + 2\left(n - 2 + \frac{1}{\theta}\right)\left(1 - \frac{1}{\theta}\right)(n-3)!\chi_{2^1,1^{n-2}}(\rho)\\
&\qquad\qquad + 2\left(n - 2 + \frac{1}{\theta}\right)\left(n - 3 + \frac{1}{\theta}\right)(n-4)!\chi_{2^2,1^{n-4}}(\rho)\bigg),
\end{aligned} \tag{28}$$

and using the relation $\chi_{\lambda^t}(\rho) = \chi_\lambda(\rho)\,\mathrm{sgn}(\rho)$, as well as the formula for χ_n, $\chi_{n-1,1}$, $\chi_{n-2,2}$ and $\chi_{n-2,1^2}$ derived above, we also get

$$\begin{aligned} d_{(n-2,2),\rho}(\theta) &= m_1 \frac{n-2+(1/\theta)}{(n-1)(n-2)}\left[(n-3)\left(1-\frac{1}{\theta}\right) - \frac{3}{2}\left(n-3+\frac{1}{\theta}\right)\right] \\ &\quad + m_1^2 \frac{(n-2+(1/\theta))(n-3+(1/\theta))}{2(n-1)(n-2)} \\ &\quad + m_2 \frac{(n-2+(1/\theta))(n-3+(1/\theta))}{(n-1)(n-2)}. \end{aligned} \tag{29}$$

Using $i=0$ and $j=n$, one gets

$$J_{(n)}(\theta) = \sum_\rho \theta^{n-\ell(\rho)} n! z_\rho^{-1} p_\rho.$$

Hence

$$c_{(n),\rho}(\theta) = \theta^{n-\ell(\rho)} n! z_\rho^{-1}, \tag{30}$$

and

$$d_{(n),\rho}(\theta) = 1. \tag{31}$$

To mimic the case of $\theta = 1$, we need to express $d_{(n-1,1),\rho}(\theta)^2$ in terms of the other d_λ's. First using (24), we get

$$d_{(n-1,1),\rho}(\theta)^2 = \frac{1}{\theta^2} + \left(\frac{1}{\theta n} + \frac{n-1}{n}\right)^2 m_1^2 - \frac{2}{\theta}\left(\frac{1}{\theta n} + \frac{n-1}{n}\right) m_1.$$

Using m_1, m_1^2, m_2 and 1 as a basis, we can write

$$d_{(n-1,1),\rho}(\theta)^2 = u + v d_{(n-1,1),\rho}(\theta) + w d_{(n-2,1^2),\rho}(\theta) + x d_{(n-2,2),\rho}(\theta)^2, \tag{32}$$

for some indeterminates u, v, w, x.

Comparing coefficients of the m_i^k's, we get the following four equations:

$$u - \frac{v}{\theta} + \frac{w}{\theta^2} = \frac{1}{\theta^2}, \tag{33}$$

$$\begin{aligned} &v\left(\frac{1}{\theta n} + \frac{n-1}{n}\right) - w\left(\frac{1+(n-1)\theta}{n\theta^2} + \frac{2+(n-2)\theta}{2\theta n}\right) \\ &\quad + x \frac{n-2+(1/\theta)}{(n-1)(n-2)}\left[(n-3)\left(1-\frac{1}{\theta}\right) - \frac{3}{2}\left(n-3+\frac{1}{\theta}\right)\right] \\ &\quad = -\frac{2}{\theta}\left(\frac{1}{\theta n} + \frac{n-1}{n}\right), \end{aligned} \tag{34}$$

$$w\frac{2+(n-2)\theta}{2\theta n} + x\frac{(n-2+(1/\theta))(n-3+(1/\theta))}{2(n-1)(n-2)} = \left(\frac{1}{\theta n} + \frac{n-1}{n}\right)^2, \tag{35}$$

$$-w\frac{2+(n-2)\theta}{\theta^2 n} + x\frac{(n-2+(1/\theta))(n-3+(1/\theta))}{(n-1)(n-2)} = 0. \tag{36}$$

Solving, we get

$$\begin{aligned}
u &= \frac{(n^2-4n+3)\theta^3+(n-1)\theta^2+(n-1)^2\theta+n-3}{\theta^2(\theta+1)(\theta(n-3)+1)n} \\
&= \frac{1}{\theta^2(\theta+1)}+\frac{1}{\theta+1}+\mathcal{O}\left(\frac{1}{n}\right), \\
v &= \frac{(n^3-6n^2+11n-6)\theta^3+2(2n^2-7n+4)\theta^2+(3n+2)\theta-4}{\theta n((n^2-5n+6)\theta^2+(3n-8)\theta+2)} \\
&= 1+\mathcal{O}\left(\frac{1}{n}\right), \\
w &= \frac{2(1+\theta(n-1))^2}{(1+\theta)(2+\theta(n-2))n} = \frac{2\theta}{1+\theta}+\mathcal{O}\left(\frac{1}{n}\right), \\
x &= \frac{2(1+\theta(n-1))^2(n-1)(n-2)}{(1+\theta)n^2(1+\theta(2n-5)+\theta^2(n-2)(n-3))} = \frac{2}{1+\theta}+\mathcal{O}\left(\frac{1}{n}\right).
\end{aligned}$$

Notice $x+w=2+O(\frac{1}{n})$.

Also by (24), (25), (29) and (31), we get

$$\begin{aligned}
d_{n,1^n} &= 1, \\
d_{(n-1,1),1^n} &= n-1, \\
d_{(n-2,1^2),1^n} &= \frac{(n-1)(n-2)}{2}, \\
d_{(n-2,2),1^n} &= \frac{n(n-3)}{2},
\end{aligned}$$

which are independent of θ because of the normalization chosen for Jack polynomials.

Finally we recall $c_\lambda(\theta)$ are eigenfunctions of P_θ, hence so are $d_\lambda(\theta)$, with eigenvalue β_λ. We list the relevant eigenvalues here:

$$\begin{aligned}
\beta_{(n)} &= 1, \\
\beta_{(n-1,1)} &= 1-(\theta\wedge 1)+\frac{\theta\binom{n-1}{2}-1}{(\theta\vee 1)\binom{n}{2}} = 1-(1\wedge\theta)\frac{2}{n}+\mathcal{O}\left(\frac{1}{n^2}\right), \\
\beta_{(n-2,1^2)} &= 1-(\theta\wedge 1)+\frac{\theta\binom{n-2}{2}-3}{(\theta\vee 1)\binom{n}{2}} = 1-(\theta\wedge 1)\frac{4}{n}+\mathcal{O}\left(\frac{1}{n^2}\right), \\
\beta_{(n-2,2)} &= 1-(\theta\wedge 1)+\frac{\theta\binom{n-2}{2}-2}{(\theta\vee 1)\binom{n}{2}} = 1-(\theta\wedge 1)\frac{4}{n}+\mathcal{O}\left(\frac{1}{n^2}\right).
\end{aligned}$$

Notice

$$\lim_{n\to\infty}\frac{(1-\beta_{(n-1,1)})}{1-\beta_{(n-2,1^2)}}=\lim_{n\to\infty}\frac{(1-\beta_{(n-1,1)})}{1-\beta_{(n-2,2)}}=2.$$

Now we are fully equipped to prove the lower bound. First observe $\mathcal{L}_\infty(d_{(n-1,1)}(\theta))\prec \mathrm{Poi}(\theta^{-1})+1$, which comes from Feller coupling. Here $d_{(n-1,1)}$ stands for the random variable $d_{(n-1,1),\rho}$ where ρ has Ewens sampling distribution with parameter θ^{-1}, as indicated by the subscript ∞. Therefore,

$$\lim_{\eta\to\infty}\mathbb{P}_\infty\big(d_{(n-1,1)}(\theta)\le\eta\big)=1.$$

Furthermore,

$$\mathbb{P}_k\big(d_{(n-1,1)}(\theta)\le\eta\big)\le\frac{\mathrm{var}_k(d_{(n-1,1)})}{(\eta-\mathbb{E}_k(d_{(n-1,1)}))^2}.$$

Let $k=\frac{1}{2}(\theta^{-1}\vee 1)n(\log n-c)$ for any $c>0$.

Since d_λ are eigenfunctions, we can compute the mean and variance at time k,

$$\begin{aligned}
\mathbb{E}_k d_{(n-1,1)}&=(n-1)\left(1-(\theta\wedge 1)\frac{2}{n}\right)^k+O(1)=e^c+\mathcal{O}(1),\\
\mathrm{var}_k\, d_{(n-1,1)}&=\mathbb{E}_k d_{(n-1,1)}^2-(\mathbb{E}_k d_{(n-1,1)})^2\\
&=u+(n-1)ve^{-(\theta\wedge 1)((2k)/n)(1+\mathcal{O}(1/n))}\\
&\quad+\left(\frac{(n-1)(n-2)}{2}w+\frac{n(n-3)}{2}x\right)e^{-(\theta\wedge 1)((4k)/n)(1+O(1/n))}\\
&\quad-(n-1)^2e^{-(\theta\wedge 1)((4k)/n)(1+\mathcal{O}(1/n))}\\
&\le\frac{1}{1+\theta}+\frac{1}{\theta^2(1+\theta)}+(n-1)e^{-(\theta\wedge 1)((2k)/n)}+\mathcal{O}\left(\frac{1}{n}\right)\le\mathcal{O}(e^c).
\end{aligned}$$

Therefore if we let $\eta=\frac{1}{2}e^c$, then

$$\lim_{c\to\infty}\liminf_n\mathbb{P}_{k(c)}[d_{(n-1,1)}<\eta]\le\lim_{c\to\infty}\liminf_n\mathcal{O}(1)\frac{e^c}{((1/2)e^c+\mathcal{O}(1))^2}=0.$$

Thus

$$\begin{aligned}
&\lim_{c\to\infty}\lim_{n\to\infty}\|\delta_{1^n}P^{k(c)}-\pi\|_{\mathrm{TV}}\\
&\qquad\ge\lim_{c\to\infty}\liminf_n\left|\mathbb{P}_{k(c)}\left[d_{(n-1,1)}<\frac{1}{2}e^c\right]-\mathbb{P}_\infty\left[d_{(n-1,1)}<\frac{1}{2}e^c\right]\right|=1.
\end{aligned}$$

REMARK 2. Wilson's method gives a suboptimal lower bound, once we know the "geometric" information that $d_{(n-1,1),\rho}=-\frac{1}{\theta}+\frac{1+(n-1)\theta}{\theta n}m_1(\rho)$: let X_1 be the random variable distributed as $\delta_x P$. We have

$$R := \sup_{x \in S_n} \mathbb{E}(d_{(n-1,1),X_1} - d_{(n-1,1),x})^2 \le 1,$$

$$\log \frac{1}{\beta_{(n-1,1)}} = (\theta \wedge 1)\frac{2}{n} + O\left(\frac{1}{n^2}\right)$$

and $d_{(n-1,1),1^n} = n-1$. Hence by Wilson [16],

$$\begin{aligned} t_{\mathrm{mix}}(\varepsilon) &\ge \frac{1}{2\log(1/\beta_{(n-1,1)})}\left[\log\left[\frac{(1-\beta_{(n-1,1)})d^2_{(n-1,1),1^n}}{2R}\right] + \log\frac{1-\varepsilon}{\varepsilon}\right] \\ &\ge \frac{n}{4(1\wedge\theta)}\log n + \log \varepsilon^{-1} + \mathcal{O}(1). \end{aligned}$$

This misses a factor of 2 from the lower bound obtained by second moment method. The discrepancy is possibly due to the nonlocal nature of the random transposition walk.

APPENDIX A: SEKIGUCHI–DEBIARD OPERATOR OVER OTHER BASES

Having seen the probabilistic interpretation of the second order differential operator (12) expressed in the power sum symmetric basis p_λ, it is natural to consider the following:

QUESTION 1. *Are there other bases of the symmetric polynomials Λ_n over which L^2_θ has natural probabilistic interpretation?*

Here we examine the remaining four fundamental bases: monomial, elementary and complete. The action of L^2_θ on the monomial basis m_λ is well known to be strictly upper triangular ([17], page 317), when the rows and columns of the Markov matrix are arranged in a total order compatible with the natural partial order on the set of partitions $\mathcal{P}_n$ of n: $\mu < \lambda$ if $\mu_1 + \cdots + \mu_r \le \lambda_1 + \cdots + \lambda_r$ for all r. In particular, if L^2_θ does define a Markov matrix (meaning the entries are nonnegative), it has a single absorbing state at (1^n).

Next consider the action of L^2_θ on e_λ, the elementary symmetric polynomials. This has been studied in detail in [3]. Here we give a quick development avoiding lengthy computations. The action of U is easy to describe. For any simple elementary polynomial e_r, the operators $x_i \partial_i$ and $(x_i\partial_i)^2$ simply collect all the terms in e_r that contains the factor x_i. So after summing over $i \in [n]$, this results in a constant multiple of the identity. Thus to get a nontrivial action, one must consider a composite $e_{r_1,r_2} := e_{r_1}e_{r_2}$. In this case, one can show that for $r_1 \le r_2$,

$$U(e_{r_1,r_2}) = 3(1+r_1)e_{r_1,r_2} - \sum_{j=0}^{r_1-1} 2(r_1 + r_2 - 2j)e_{r_1+r_2-j,j}.$$

Thus U is strictly lower triangular with respect to the partial order $\preceq$. It turns out that the action of V on the e_λ is diagonal: first of all V satisfies a product rule on $e_\lambda = \prod_{j=1}^{\ell(\lambda)} e_{\lambda_j}$; by pairing up $j \neq k$, one also sees that Ve_r consists of monomials with no repeated factors, hence by symmetry must be a multiple of e_r. Thus the following linear combination yields a legitimate Markov matrix:

$$M_e(c_1, c_2) := c_1 I + c_2\left(\frac{\theta}{2}U + V\right).$$

Notice that we need to add the multiple of I to make sure that the diagonal entries of M_e are nonnegative. Also observe that the Jack parameter θ needs to be nonpositive for the off-diagonal entries to be nonnegative. It is clear from the description of U and V that this Markov chain is absorbing at (n), because the next step either stays in the current state or goes to a state corresponding to a partition bigger than that of the current state.

Next we look at the complete symmetric polynomials h_λ. First consider the action of $\langle X, \nabla\rangle$ on h_r, one of the generators. Since $h_r = s_{(r)}$ is a degree r homogeneous polynomial, the action of $\langle X, D\rangle = \sum_{i=1}^n x_i \partial_i$ is simply multiplication by r; that is, any homogeneous polynomials are eigenfunctions $\langle X, D\rangle$. However, the operator $(\langle X, D\rangle)^2$ acts nontrivially on h_r,

$$\sum_{i=1}^n (x_i \partial_i)^2 h_4 = -2h_{1^4} + 10h_{2,1,1} - 8h_{2^2} - 12h_{3,1} + 28h_4.$$

For partitions of more than one part, the computation gets unwieldy, and I have not tried to express $U(h_{r_1}h_{r_2})$ and $V(h_{r_1}h_{r_2})$ in terms of h_λ explicitly because of the following numerical observation: for $\lambda = (3, 2, 1)$, we have

$$Uh_\lambda = 2h_{2,1^4} - 8h_{2^2,1^2} - 2h_{3,1^3} + 14h_{3,2,1} + 6h_{3^2} + 6h_{4,1^2} + 8h_{4,2} + 10h_{5,1},$$
$$Vh_\lambda = -h_{2,1^4} + 4h_{2^2,1^2} + h_{3,1^3} + 32h_{3,2,1}.$$

The only linear combination of the above two expressions that yields nonnegative coefficients is $\frac{1}{2}U + V$, which corresponds to $\theta = 1$. But in that case, the Markov chain is again absorbing at (6). So we arrive at the following result:

PROPOSITION A.1. *The operator L_θ^2 gives a Markov matrix under the complete symmetric polynomial basis h_λ for all n if and only if $\theta = 1$. In this case, the Markov chain never goes toward partitions of fewer or equal parts, hence is absorbing at (n).*

PROOF. When $\theta = 1$, h_λ are dual to e_λ with respect to the Jacobi–Trudy identity. Hence the walk defined by D_1^2 on h_λ can be obtained from the upper-triangular walk on e_λ under the map $\lambda \mapsto \lambda^t$; in particular the walk is absorbing at (r). For $\theta \neq 1$, the numerical example above suffices to show the associated walk is not positive. $\square$

For $\theta \neq 1$, the resulting signed Markov matrix seems to have nontrivial left eigenvector corresponding to the eigenvalue 1. I have not checked if this corresponds to some nice stationary distribution on $\mathcal{P}_n$; presumably it will define a signed measure.

APPENDIX B: HIGHER ORDER SEKIGUCHI–DEBIARD OPERATORS

Throughout this section N will denote the number of underlying variables in the symmetric functions, and n will denote the weight of partitions, consistent with previous sections. It is possible to study higher order differential operators on Λ_N from the Sekiguchi–Debiard operator-valued generating function (see [17], page 328),

$$\begin{aligned} D_\theta(X) &:= a_\delta(x)^{-1} \sum_{w \in S_N} \varepsilon(w) x^{w\delta} \prod_{j=1}^{N} \big(X + (w\delta)_j + \theta x_j \partial_j\big) \\ &= \sum_{k=0}^{N} D_\theta^k X^{N-k}. \end{aligned} \tag{37}$$

Since the seminal work of Diaconis and Ram [6] interpreting D_θ^2 above as the generator of an auxilliary variable Markov chain, it has been tempting to consider the following:

QUESTION 2. *Does any of the higher order D_θ^k's admit natural probabilistic interpretation?*

Below we give complete analysis of D_θ^3, and show that the answer is not nearly as nice as for D_θ^2. To begin, it suffices to understand the following two-parameter family of operators:

$$\begin{aligned} D(\lambda, \mu; h) &= a_\delta(x)^{-1} \sum_{w \in S_N} \varepsilon(w) \prod_{i=1}^{\ell} \left(\sum_{j=1}^{N} (w\delta)_j^{\lambda_i} (x_j \partial_j)^{\mu_i} \right) \sum_{j=1}^{N} (x_j \partial_j)^h \\ &= a_\delta(x)^{-1} \sum_{j_1, \dots, j_\ell} \prod_{i=1}^{\ell} (x_{j_i} \partial_{j_i})^{\lambda_i} a_\delta(x) (x_{j_i} \partial_{j_i})^{\mu_i} \sum_{j=1}^{N} (x_j \partial_j)^h, \end{aligned}$$

where λ, μ are positive integer compositions and $\ell = \ell(\mu) = \ell(\lambda)$. Indeed, it is easy to see that [denoting by $(j_1, \dots, j_k)$ all distinct indices]

$$D_\theta^k = a_\delta(x)^{-1} \sum_{w \in S_N} \sum_{(j_1, \dots, j_k)} \sum_{u=0}^{k} \frac{\theta^u}{u!(k-u)!} \prod_{i=1}^{k-u} (w\delta)_{j_i} \prod_{i=k-u+1}^{k} (x_{j_i} \partial_{j_i}),$$

which can be expressed as a linear combination of $D(\lambda, \mu, h)$'s with $|\lambda| + |\mu| + h = k$ and $h \neq 1$; the factors of the form $\sum_{j=1}^{N} (w\delta)_j^k$ or $\sum_{j=1}^{N} (x_j \partial_j)^k$ all evaluate to constant by symmetry, and the operator $\sum_{j=1}^{N} (x_j \partial_j)$ acts on p_λ by constant multiplication.

For $k=3$, we thus have three operators to consider: $D((1),(2);0)$, $D((2),(1);0)$ and $D(\varnothing,\varnothing;3)$. We compute the action of each on p_λ below. We need the following three computations:

- $a_\delta(x)^{-1}x_i\partial_i a_\delta(x)=\sum_{j\neq i}\frac{x_i}{x_i-x_j}$;
- $$\begin{aligned}(x_i\partial_i)^2p_\lambda&=x_i\partial_i\sum_{s=1}^{\ell(\lambda)}\frac{p_\lambda}{p_{\lambda_s}}(x_i\partial_i)p_{\lambda_s}\\&=\sum_{s=1}^{\ell(\lambda)}\frac{p_\lambda}{p_{\lambda_s}}(x_i\partial_i)^2p_{\lambda_s}+\sum_{t\neq s}\frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}}\big[(x_i\partial_i)p_{\lambda_s}\big]\big[(x_i\partial_i)p_{\lambda_t}\big]\\&=\sum_{s=1}^{\ell(\lambda)}\lambda_s^2x_i^{\lambda_s}\frac{p_\lambda}{p_{\lambda_s}}+\sum_{s\neq t}\lambda_s\lambda_t x_i^{\lambda_s+\lambda_t}\frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}};\end{aligned}$$
- $$\begin{aligned}\sum_{i=1}^{N}\sum_{j\neq i}\frac{x_i^{r+1}}{x_i-x_j}&=\frac{1}{2}\sum_{i\neq j}\frac{x_i^{r+1}-x_j^{r+1}}{x_i-x_j}=\frac{1}{2}\sum_{i\neq j}\sum_{u=0}^{r}x_i^u x_j^{r-u}\\&=\frac{1}{2}\sum_{\substack{r_1+r_2=r:\\ r_i\geq 1}}p_{r_1}p_{r_2}+\frac{2N-r-1}{2}p_r.\end{aligned}$$

From this we have

$$\begin{aligned}&D\big((1),(2);0\big)p_\lambda\\&\quad=a_\delta(x)^{-1}\sum_{i=1}^{N}\big[x_i\partial_i a_\delta(x)\big](x_i\partial_i)^2p_\lambda\\&\quad=p_\lambda\Bigg(\sum_{s=1}^{\ell(\lambda)}\frac{\lambda_s^2}{2}\Bigg[(2N-\lambda_s-1)+\sum_{u=1}^{\lambda_s-1}\frac{p_u p_{\lambda_s-u}}{p_{\lambda_s}}\Bigg]\\&\qquad\quad+\sum_{s\neq t}\frac{\lambda_s\lambda_t}{2}\Bigg[(2N-\lambda_s-\lambda_t-1)\frac{p_{\lambda_s+\lambda_t}}{p_{\lambda_s}p_{\lambda_t}}+\sum_{u=1}^{\lambda_s+\lambda_t-1}\frac{p_u p_{\lambda_s+\lambda_t-u}}{p_{\lambda_s}p_{\lambda_t}}\Bigg]\Bigg).\end{aligned}$$

It is worth pointing out that the sum $\sum_{u=1}^{\lambda_s+\lambda_t-1}\frac{p_u p_{\lambda_s+\lambda_t-u}}{p_{\lambda_s}p_{\lambda_t}}$ contains a constant term (when $u\in\{\lambda_s,\lambda_t\}$).

Next we consider $D((2),(1);0)$. Again we collect some computations below:

- $x_i\partial_i p_\lambda=\sum_{s=1}^{\ell(\lambda)}\lambda_s x_i^{\lambda_s}\frac{p_\lambda}{p_{\lambda_s}}$.

- $a_\delta(x)^{-1}(x_i\partial_i)^2 a_\delta(x)$

$$\begin{aligned}
&= (x_i\partial_i)^2 \log a_\delta(x) + x_i^2\big(\partial_i \log a_\delta(x)\big)^2 \\
&= \sum_{j\neq i} \frac{-x_i x_j}{(x_i-x_j)^2} + \sum_{j\neq i}\sum_{k\neq i,j} \frac{x_i^2}{(x_i-x_j)(x_i-x_k)} + \sum_{j\neq i} \frac{x_i^2}{(x_i-x_j)^2} \\
&= \sum_{j\neq i} \frac{x_i}{x_i-x_j} + \sum_{\substack{j\neq k:\\ j,k\neq i}} \frac{x_i^2}{(x_i-x_j)(x_i-x_k)}.
\end{aligned}$$

- $\sum_{i=1}^{N} \sum_{\substack{j\neq k:\\ j,k\neq i}} \frac{x_i^{r+2}}{(x_i-x_j)(x_i-x_k)}$

$$\begin{aligned}
&= \sum_{\substack{T\subset[N]:\\ |T|=3}} \sum_{i\in T} \sum_{\substack{j\neq k:\\ j,k\in T\setminus\{i\}}} \frac{x_i^{r+2}}{(x_i-x_j)(x_i-x_k)} \\
&= 2\sum_{\substack{T\subset[N]:\\ |T|=3}} s_r(T) = 2\sum_{\substack{T\subset[N]:\\ |T|=3}} \sum_{\lambda\vdash r} m_\lambda(T) \\
&= 2\sum_{\substack{T\subset[N]:\\ |T|=3}} \sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)\le 3}} m_\lambda(T) \\
&= 2\left[\binom{N-3}{0} \sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=3}} m_\lambda + \binom{N-2}{1} \sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=2}} m_\lambda + \binom{N-1}{2} \sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=1}} m_\lambda\right],
\end{aligned}$$

where $s_\lambda(T)$ denotes the Schur polynomial over the variables indexed by T, and similarly for m_λ.

- $\sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=2}} m_\lambda = \sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=2}} \sum_{i\neq j} x_i^{\lambda_1} x_j^{\lambda_2} I(\lambda_1\neq\lambda_2) + \frac{1}{2} x_i^{\lambda_1} x_j^{\lambda_2} I(\lambda_1=\lambda_2)$

$$= \frac{1}{2} \sum_{\substack{r_1+r_2=r:\\ r_i\ge 1}} (p_{r_1}p_{r_2} - p_r) = \frac{1}{2} \sum_{\substack{r_1+r_2=r:\\ r_i\ge 1}} p_{r_1}p_{r_2} - \frac{r-1}{2} p_r.$$

- $\sum_{\substack{\lambda\vdash r:\\ \ell(\lambda)=3}} m_\lambda = \frac{1}{6}\Bigg[\sum_{\substack{r_1+r_2+r_3=r:\\ r_i\ge 1}} p_{r_1}p_{r_2}p_{r_3}$

$$- 3\sum_{\substack{r_1+r_2=r:\\ r_i\ge 1}} (r_1-1)p_{r_1}p_{r_2} + (r-1)(r-2)p_r\Bigg]$$

$$= \frac{1}{6} \sum_{\substack{r_1+r_2+r_3=r:\\ r_i\geq 1}} p_{r_1}p_{r_2}p_{r_3} - \frac{r-2}{4} \sum_{\substack{r_1+r_2=r:\\ r_i\geq 1}} p_{r_1}p_{r_2} + \frac{(r-1)(r-2)}{6} p_r.$$

Putting everything together we have

$$\begin{aligned} &D\big((2),(1);0\big)p_\lambda \\ &= p_\lambda \sum_{s=1}^{\ell(\lambda)} \lambda_s \Bigg[\bigg(N - \frac{1+\lambda_s}{2}\bigg) \sum_{\substack{r_1+r_2=\lambda_s:\\ r_1,r_2\geq 1}} p_{r_1}p_{r_2} + \frac{1}{3} \sum_{\substack{r_1+r_2+r_3=\lambda_s:\\ r_i\geq 1}} p_{r_1}p_{r_2}p_{r_3} \\ &\qquad + \bigg((N-1)(N-\lambda_s) + \frac{(2\lambda_s-3)(\lambda_s-1)}{6}\bigg)p_{\lambda_s}\Bigg]. \end{aligned}$$

Finally we compute the action of $D(\varnothing,\varnothing;3)$,

$$\begin{aligned} D(\varnothing,\varnothing;3)p_\lambda &= \sum_{i=1}^N (x_i\partial_i)^2 \sum_{s=1}^{\ell(\lambda)} \lambda_s x_i^{\lambda_s} \frac{p_\lambda}{p_{\lambda_s}} \\ &= \sum_{i=1}^N (x_i\partial_i) \Bigg[\sum_{s=1}^{\ell(\lambda)} \Bigg[\lambda_s^2 x_i^{\lambda_s} \frac{p_\lambda}{p_{\lambda_s}} + \sum_{t\neq s} \lambda_s\lambda_t x_i^{\lambda_s+\lambda_t} \frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}}\Bigg]\Bigg] \\ &= \sum_{i=1}^N \sum_{s=1}^{\ell(\lambda)} \Bigg[\lambda_s^3 x_i^{\lambda_s} \frac{p_\lambda}{p_{\lambda_s}} + \sum_{t\neq s} \Bigg[\lambda_s\lambda_t(\lambda_s+\lambda_t) x_i^{\lambda_s+\lambda_t} \frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}} \\ &\qquad + \sum_{u\neq s,t} \lambda_s\lambda_t\lambda_u x_i^{\lambda_s+\lambda_t+\lambda_u} \frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}p_{\lambda_u}}\Bigg] \\ &\qquad + \sum_{t\neq s} \lambda_s^2\lambda_t x_i^{\lambda_s+\lambda_t} \frac{p_\lambda}{p_{\lambda_s}p_{\lambda_t}}\Bigg] \\ &= p_\lambda \Bigg[\sum_{s=1}^{\ell(\lambda)} \lambda_s^3 + 3\sum_{t\neq s} \lambda_s^2\lambda_t \frac{p_{\lambda_s+\lambda_t}}{p_{\lambda_s}p_{\lambda_t}} + \sum_{(s,t,u)} \lambda_s\lambda_t\lambda_u \frac{p_{\lambda_s+\lambda_t+\lambda_u}}{p_{\lambda_s}p_{\lambda_t}p_{\lambda_u}}\Bigg], \end{aligned}$$

where $\sum_{(s,t,u)}$ denotes summation over all distinct triples.

Next we compute the operators $D((1),(1);0)$ and $D(\varnothing,\varnothing;2)$.

$$\begin{aligned} D\big((1),(1);0\big)p_\lambda &= a_\delta(x)^{-1} \sum_{w\in S_N} \varepsilon(w)x^{w\delta} \sum_{i=1}^N (w\delta)_i(x_i\partial_i)p_\lambda \\ &= p_\lambda \sum_{s=1}^{\ell(\lambda)} \lambda_s \Bigg(\frac{1}{2} \sum_{\substack{r_1+r_2=\lambda_s\\ r_i\geq 1}} \frac{p_{r_1}p_{r_2}}{p_{\lambda_s}} + \frac{2N-\lambda_s-1}{2}\Bigg) \end{aligned} \tag{38}$$

and

$$
\begin{aligned}
D(\varnothing,\varnothing;2)p_\lambda &= a_\delta(x)^{-1}\sum_{w\in S_N}\varepsilon(w)x^{w\delta}\sum_{i=1}^N (x_i\partial_i)^2 p_\lambda \\
&= p_\lambda\left(\sum_{s=1}^{\ell(\lambda)}\lambda_s^2+\sum_{s\neq t}\lambda_s\lambda_t\frac{p_{\lambda_s+\lambda_t}}{p_{\lambda_s}p_{\lambda_t}}\right).
\end{aligned}
\tag{39}
$$

We can now compute the action of D_θ^2 on power sum polynomials; see [17], Example VI.3.3(e).

$$
\begin{aligned}
D_\theta^2 p_\lambda &= a_\lambda(x)^{-1}\sum_{w\in S_N}\varepsilon(w)x^{w\delta}\sum_{i\neq j}\biggl(\frac{1}{2}(w\delta)_i(w\delta)_j+\theta(w\delta)_i(x_j\partial_j) \\
&\qquad\qquad + \frac{\theta^2}{2}(x_i\partial_i)(x_j\partial_j)\biggr)p_\lambda \\
&= a_\delta(x)^{-1}\sum_w \varepsilon(w)x^{w\delta}\biggl(\frac{1}{2}\biggl[\Bigl(\sum_i (w\delta)_i\Bigr)^2-\sum_i (w\delta)_i^2\biggr] \\
&\qquad + \theta\biggl[\Bigl(\sum_i (w\delta)_i\Bigr)\Bigl(\sum_i x_i\partial_i\Bigr)-\sum_i (w\delta)_i(x_i\partial_i)\biggr] \\
&\qquad + \frac{\theta^2}{2}\biggl[\Bigl(\sum_i x_i\partial_i\Bigr)^2-\sum_i (x_i\partial_i)^2\biggr]\biggr) \\
&= \frac{1}{2}\biggl[\Bigl(\frac{N(N-1)}{2}\Bigr)^2-\frac{(N-1)N(2N-1)}{6}\biggr] \\
&\qquad + \theta\biggl[\frac{N(N-1)n}{2}-D\bigl((1),(1);0\bigr)\biggr]+\frac{\theta^2}{2}\bigl[n^2-D(\varnothing,\varnothing;2)\bigr]p_\lambda \\
&= \biggl(\frac{1}{2}\biggl[\theta^2n^2+\theta nN(N-1)+\frac{N(N-1)(N-2)(3N-1)}{12}\biggr] \\
&\qquad - \theta D\bigl((1),(1);0\bigr)-\frac{\theta^2}{2}D(\varnothing,\varnothing;2)\biggr)p_\lambda.
\end{aligned}
\tag{40}
$$

Similarly we can compute D_θ^3 using the inclusion-exclusion principle,

$$
\begin{aligned}
D_\theta^3 p_\lambda &= a_\delta(x)^{-1} \\
&\quad\times\sum_w \varepsilon(w)x^{w\delta}\sum_{(i,j,k)}\biggl[\frac{1}{6}(w\delta)_i(w\delta)_j(w\delta)_k+\frac{\theta}{3}(w\delta)_i(w\delta)_j x_k\partial_k \\
&\qquad + \frac{\theta^2}{3}(w\delta)_i(x_j\partial_j)(x_k\partial_k)+\frac{\theta^3}{6}(x_i\partial_i)(x_j\partial_j)(x_k\partial_k)\biggr]
\end{aligned}
$$

$$= \Big(c_N(3,\theta) + \frac{\theta^3}{6}\big[2D(\varnothing,\varnothing;3) - 3nD(\varnothing,\varnothing;2)\big] \\ + \frac{\theta^2}{3}\Big[2D\big((1),(2);0\big) - 2nD\big((1),(1);0\big) - \binom{N}{2}D(\varnothing,\varnothing;2)\Big] \\ + \frac{2\theta}{3}\Big[D\big((2),(1);0\big) - \binom{N}{2}D\big((1),(1);0\big)\Big]\Big)p_\lambda,$$

where $c_N(3,\theta) = \frac{1}{6}[\binom{N}{2}^3 - \frac{3}{4}\binom{N}{2}\binom{2N}{3} + 2\binom{N}{2}^2] + \frac{\theta}{2}[\binom{N}{2}^2 n - \frac{1}{4}\binom{2N}{3}n] + \frac{\theta^2}{3}\binom{N}{2}n^2 + \frac{\theta^3 n^3}{6}$. Unfortunately I cannot extract any natural interpretation of Markov chains from the right-hand side. This is not so surprising since the composite walk P_θ^k for $k \geq 2$ does not correspond to some affine transformation of the Metropolization of P_1^k with respect to the measure $\mathrm{MED}(\theta)$. Nonetheless this gives a new Markov chain that converges to the multivariate Ewens distribution with parameter θ^{-1}, since the operators D_θ^r are simultaneously diagonalized and the left eigenfunction corresponding to the eigenvalue 1 is simply the stationary distribution.

I also computed a numerical example using the symmetric reduction function in mathematica and the SF package in maple. We take the power sum polynomial p_λ with $\lambda = (3, 1^2)$:

$$D_\theta^2 p_3 p_1^2 = -3\theta p_2 p_1^3 + (33\theta + 35 + 7\theta^2)p_3 p_1^2 - 6\theta^2 p_4 p_1 - p_2 p_3 \theta^2,$$

$$D_\theta^3 p_3 p_1^2 = (2/3)\theta p_1^5 + (-4\theta^2 - 8\theta)p_2 p_1^3 \\ + (22\theta^2 + (73/6)\theta^3 + 50 + (307/3)\theta)p_3 p_1^2 + 4\theta^2 p_1 p_2^2 \\ + (4\theta^3 - 20\theta^2)p_4 p_1 + (-(4/3)\theta^3 - 2\theta^2)p_3 p_2 + 6\theta^3 p_5,$$

$$D_\theta^2 \circ D_\theta^2 p_3 p_1^2 = 3\theta^2 p_1^5 + (-21\theta - 4\theta^3 - 19\theta^2)p_2 p_1^3 \\ + (505\theta^3 + 1225 + 2310\theta + 4\theta^4 + 1579\theta^2)p_3 p_1^2 + 24\theta^3 p_1 p_2^2 \\ + (-41\theta^3 - 6\theta^4 - 42\theta^2)p_4 p_1 \\ + (-6\theta^3 - \theta^4 - 7\theta^2)p_3 p_2 + 30 p_\theta^4.$$

This example shows that D_θ^2, D_θ^3, $D_\theta^2 \circ D_\theta^2$ and id are independent operators on Λ_N. Notice also that $D_\theta^3 p_\lambda$ has positivity issues: the partitions of Cayley distance 2 from the starting partition λ are always positive, whereas the ones that differ from λ by one transposition might become negative. So in order to make D_θ^3 into a Markov matrix, one needs to add a sufficiently negative multiple of D_θ^2. We have not tried to compute the optimal multiple here since we are unable to glean any nice pattern from the numerical example above; in particular, the coefficients cannot be made into simple powers of θ. Observe that $D_\theta^2 \circ D_\theta^2 p_\lambda$ also has positivity problem, but it is much easier to fix, since one can simply add a multiple of the identity to D_θ^2 as in the case treated by [21].

APPENDIX C: COMPOSITIONS OF D_θ^2 FOR DIFFERENT θ VALUES

In general the eigenvalues and eigenfunctions of a Markov chain can be highly intractable, due to the need to solve for high degree polynomials. For instance, the Metropolis chain based on 3-cycle shuffle on $\mathcal{P}_n$ already requires taking square roots for $n = 4$:

$$M_4^{(3)}(\theta) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1/8 & 3/4 & 0 & 1/8 \\ 0 & 1 & 0 & 0 & 0 \\ \theta^2 & 0 & 0 & 1-\theta^2 & 0 \\ 0 & \theta^2 & 0 & 0 & 1-\theta^2 \end{pmatrix}.$$

The eigenvalues are

$$\{1, 1, -\theta^2, \tfrac{1}{16}(1 - 8\theta^2 - \sqrt{193 - 208\theta^2 + 64\theta^4}),$$
$$\tfrac{1}{16}(1 - 8\theta^2 + \sqrt{193 - 208\theta^2 + 64\theta^4})\}.$$

The following result was discovered in numerical experiments:

PROPOSITION C.1. *For any Laurent polynomial p in m variables, $p(D_{\theta_1}^2, \dots, D_{\theta_m}^2)$ gives rise to a Markov chain on the set of partitions $\mathcal{P}_n$, with eigenvalues, and left and right eigenvectors given by rational functions of $\theta_1, \dots, \theta_m$. In particular, the stationary distribution is given by rational functions of θ_i's also.*

PROOF. When expressed in the monomial symmetric basis, D_θ^2 is unipotent (upper triangular with 1's on the diagonal), with respect to any total ordering on $\mathcal{P}_n$ compatible with the natural partial ordering $\preceq$, whereby $\lambda \preceq \mu$ if $\lambda_1 + \cdots + \lambda_r \le \mu_1 + \cdots + \mu_r$ for all r; see [17], page 317, equation (3.7). Thus fixing this total-ordering, any Laurent polynomial of D_{θ_i}'s is clearly still unipotent. The eigenvalues are simply the diagonal entries, and the eigenvectors can be computed using simple row reduction, which also results in rational components. □

The above result is clearly also true for D_θ^k in general and even Macdonald operators. Thus in principle, one can compute the $\mathcal{L}^2$ mixing time of Markov chain of the form $D_{\theta_1}^2 D_{\theta_2}^2$, whose stationary distribution can be quite complicated: for $n = 3$, the stationary probabilities are

$$\left\{\frac{\theta_2(\theta_1(3-4\theta_2) + \theta_1^2(-1+\theta_2) + \theta_2)}{8 + \theta_2 + \theta_1^2(-1+\theta_2)\theta_2 - \theta_1(-1+9\theta_2+\theta_2^2)},\right.$$
$$\frac{3(-(-3+\theta_2)\theta_2 + \theta_1(1-4\theta_2+\theta_2^2))}{8 + \theta_2 + \theta_1^2(-1+\theta_2)\theta_2 - \theta_1(-1+9\theta_2+\theta_2^2)},$$
$$\left.\frac{-2\theta_1 + 2(-2+\theta_2)^2}{8 + \theta_2 + \theta_1^2(-1+\theta_2)\theta_2 - \theta_1(-1+9\theta_2+\theta_2^2)}\right\}.$$

APPENDIX D: EXTENSIONS TO OTHER ROOT SYSTEMS

The appropriate generalization of the Laplace–Beltrami operator to root systems other than A_N is given by the Heckman–Opdam operator (see [12] and [3]),

$$L_N(\kappa, R) = \Delta + \kappa V_N := \sum_{i=1}^{N} \partial_{t_i}^2 + \sum_{\alpha \in R_+} \kappa_\alpha \coth(\alpha/2)\partial_\alpha,$$

where R denotes an arbitrary root system, R_+ a designated set of positive roots and κ_α is called a multiplication function, invariant under the action of the Weyl group on R_+. For more on root systems and Weyl groups, consult the first 3 chapters of [10] as well as Chapters 19 and 20 of [4].

Fascinated by the success of the A_N root system, Diaconis raised the following:

QUESTION 3. *Are there other root systems beside those of type A_N whose associated Heckman–Opdam operators give rise to nontrivial Markov chains with algebraically tractible spectral decomposition?*

We study root system D_N in detail here; B_N and C_N are similar. These come from the irreducible decomposition of the adjoint representation of the maximal torus in the compact Lie groups $SO(2N, \mathbb{R})$. The positive roots can be chosen as the set $\{x_i x_j^{-1}, x_i x_j : 1 \le i < j \le N\}$ on the maximal torus; in the associated Cartan subalgebra (the Lie subalgebra corresponding to the maximal torus), they become $\{t_i - t_j, t_i + t_j : 1 \le i < j \le N\}$. The appropriate analogue of the power sum polynomials appears to be the following power sum symmetric Laurent polynomials:

$$p_a = \sum_{i=1}^{N} \cosh(a t_i) = \sum_{i=1}^{N} [x_i^a + x_i^{-a}]/2,$$

where $x_i = e^{t_i}$. And as in the case of A_N, $p_\lambda = \prod_{i=1}^{\ell(\lambda)} p_{\lambda_i}$. By direct computation we have

$$\Delta p_\lambda = p_\lambda \left\{ \sum_{i=1}^{\ell(\lambda)} \lambda_i^2 + \sum_{1 \le i < j \le \ell(\lambda)} \lambda_i \lambda_j \left[\frac{p_{\lambda_i + \lambda_j}}{p_{\lambda_i} p_{\lambda_j}} - \frac{p_{\lambda_i - \lambda_j}}{p_{\lambda_i} p_{\lambda_j}} \right] \right\},$$

$$\begin{aligned}\sum_{\theta \in R_+} \coth(\theta/2)\partial_\alpha p_a &= 2a \sum_{i \ne j} \sum_{\ell=0}^{a-1} \cosh(\ell t_i) \cosh\bigl((a-\ell)t_j\bigr) \\ &= \bigl(2an - a^2 - a\bigr) p_a + 2a \sum_{\ell=1}^{a-1} p_\ell p_{a-\ell} - a \sum_{\ell=1}^{a-1} p_{a-2\ell},\end{aligned}$$

$$\sum_{\ell=1}^{a-1} p_{a-2\ell} = \begin{cases} 2 \displaystyle\sum_{\ell=1}^{(a-1)/2} p_{a-2\ell}, & \text{if } a \text{ is odd,} \\ N + 2 \displaystyle\sum_{\ell=1}^{(a-2)/2} p_{a-2\ell}, & \text{if } a \text{ is even} \end{cases}$$

$$= 2 \sum_{\ell=1}^{\lfloor a/2 \rfloor} p_{a-2\ell},$$

if we define $p_0 := N/2$.

Therefore for $n = |\lambda| = \sum_i \lambda_i$,

$$\sum_{\alpha \in R_+} \coth(\alpha/2)\partial_\alpha p_\lambda$$
$$= \left((2N-1)n - \sum_{i=1}^{\ell(\lambda)} \lambda_i^2\right) p_\lambda + \sum_{i=1}^{\ell(\lambda)} \frac{p_\lambda}{p_{\lambda_i}} 2\lambda_i \left[\sum_{\ell=1}^{\lambda_i - 1} p_\ell p_{\lambda_i - \ell} - \sum_{\ell=1}^{\lfloor \lambda_i/2 \rfloor} p_{\lambda_i - 2\ell}\right].$$

Restricting to partitions of the top grading, n, clearly the transition coefficients are affine transformation of those in the A_N case, and hence nothing new is obtained this way. There are several pathological features regarding the action of $L_N(D, \kappa) := \Delta + \sum_{\alpha \in R_+} \kappa_\alpha \coth(\alpha/2)\partial_\alpha$ on the power sum analogues of symmetric Laurent polynomials (see the toy example below):

(1) there are no easy ways to make the entries all positive;
(2) the row sums are not the same.

Thus it remains difficult to interpret the full transition matrix as a Markov kernel. For root system D_N, there is only one Weyl orbit, hence $\kappa_\alpha \equiv \kappa$. The Heckman–Opdam functions have rational transition coefficients to this power sum Laurent basis, as illustrated by the following numerical example ($\mathcal{P}_k$ denotes the set of partitions of k):

$$M_\kappa(N)|_{\mathcal{P}_3 \cup \mathcal{P}_1}$$
$$:= \begin{pmatrix} 9 + \kappa(-12 + 6N) & 12\kappa & 0 & -6\kappa \\ 2 & 5 + \kappa(-8 + 6N) & 4\kappa & -2 - 2\kappa N \\ 0 & 3 & 3 + \kappa(-6 + 6N) & -3 \\ 0 & 0 & 0 & 1 + \kappa(-2 + 2N) \end{pmatrix},$$

where the columns and rows are indexed by (3), (21), (1^3), (1). The eigenvalues are very clean,

$$3 + 6\kappa(-2 + N), \qquad 9 + 6\kappa(-1 + N),$$
$$5 + \kappa(-8 + 6N), \qquad 1 + 2\kappa(-1 + N).$$

The left eigenvectors are rational functions of κ, which we display as rows of the following matrix:

$$\begin{pmatrix} \frac{-5+\theta+2N}{3(-1+N)} & -\frac{-5+\theta+2N}{-1+N} & \frac{2(-5+\theta+2N)}{3(-1+N)} & 1 \\ 0 & 0 & 0 & 1 \\ -\frac{2\theta(-1+2\theta+N)}{3(1+2\theta+N)} & \frac{-2(-1+2\theta+N)}{1+2\theta+N} & -\frac{4(-1+2\theta+N)}{3\theta(1+2\theta+N)} & 1 \\ \frac{\theta(-3+2\theta+2N)}{\theta+2\theta^2-2N+2\theta N} & -\frac{2(-1+\theta)(-3+2\theta+2N)}{\theta+2\theta^2-2N+2\theta N} & \frac{12-8\theta-8N}{\theta+2\theta^2-2N+2\theta N} & 1 \end{pmatrix}.$$

Here $\theta = \kappa^{-1}$ corresponds to the parameter in the A_N case.

We have also tried to adjust the diagonal entries to make the row sum equal to 1; the resulting matrix however does not have rational eigenvalues in the entries.

Acknowledgments. I thank Persi Diaconis for introducing me to this problem as well as providing guidance throughout the writing. I also thank Calvin (Yuncheng) Lin for suggesting that the Lie algebra formed by D_α and natural Lie bracket is solvable and hence can be simultaneously uni-upper-triangularized by a result of Serre. Several of the results in the Appendix are computed and verified using Mathematica 6.0 and Maple 13 together with the package SF developed by John Stembridge.

REFERENCES

[1] ARRATIA, R., BARBOUR, A. D. and TAVARÉ, S. (1992). Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2** 519–535. MR1177897

[2] BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313. MR1161056

[3] BEERENDS, R. J. (1991). Chebyshev polynomials in several variables and the radial part of the Laplace–Beltrami operator. *Trans. Amer. Math. Soc.* **328** 779–814. MR1019520

[4] BUMP, D. (2004). *Lie Groups*. *Graduate Texts in Mathematics* **225**. Springer, New York. MR2062813

[5] DIACONIS, P. (1988). *Group Representations in Probability and Statistics*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **11**. IMS, Hayward, CA. MR0964069

[6] DIACONIS, P. and RAM, A. (2012). A probabilistic interpretation of the Macdonald polynomials. *Ann. Probab.* **40** 1861–1896. MR3025704

[7] DIACONIS, P. and SHAHSHAHANI, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR0626813

[8] DUMMIT, D. S. and FOOTE, R. M. (1991). *Abstract Algebra*. Prentice Hall, Englewood Cliffs, NJ. MR1138725

[9] FULTON, W. and HARRIS, J. (1991). *Representation Theory*: *A First Course*. *Graduate Texts in Mathematics* **129**. Springer, New York. MR1153249

[10] GOODMAN, R. and WALLACH, N. R. (1998). *Representations and Invariants of the Classical Groups*. *Encyclopedia of Mathematics and Its Applications* **68**. Cambridge Univ. Press, Cambridge. MR1606831

[11] HANLON, P. (1992). A Markov chain on the symmetric group and Jack symmetric functions. *Discrete Math*. **99** 123–140. MR1158785

[12] HECKMAN, G. J. and OPDAM, E. M. (1987). Root systems and hypergeometric functions. I. *Compos*. *Math*. **64** 329–352. MR0918416

[13] JOHNSON, N. L., KOTZ, S. and BALAKRISHNAN, N. (1997). *Discrete Multivariate Distributions*. Wiley, New York. MR1429617

[14] KOIKE, K. (1993). On a conjecture of Stanley on Jack symmetric functions. *Discrete Math*. **115** 211–216. MR1217630

[15] KOORNWINDER, T. H. (1994). Special functions associated with root systems: Recent progress. In *From Universal Morphisms to Megabytes*: *A Baayen Space Odyssey* 391–404. Math. Centrum, Centrum Wisk. Inform., Amsterdam. MR1490602

[16] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR2466937

[17] MACDONALD, I. G. (1995). *Symmetric Functions and Hall Polynomials*, 2nd ed. *Oxford Mathematical Monographs*. Oxford Univ. Press, New York. MR1354144

[18] MACDONALD, I. G. (2000/2001). Orthogonal polynomials associated with root systems. *Sém*. *Lothar. Combin*. **45** Art. B45a, 40 pp. (electronic). MR1817334

[19] MORRIS, B. (2009). Improved mixing time bounds for the Thorp shuffle and L-reversal chain. *Ann*. *Probab.* **37** 453–477. MR2510013

[20] STANLEY, R. P. (1989). Some combinatorial properties of Jack symmetric functions. *Adv. Math*. **77** 76–115. MR1014073

[21] DIACONIS, P. and HANLON, P. (1992). Eigen-analysis for some examples of the Metropolis algorithm. *Contemp*. *Math*. **138** 99–117.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: jyj@math.stanford.edu