The Annals of Probability
2012, Vol. 40, No. 5, 1861–1896
DOI: 10.1214/11-AOP674

A PROBABILISTIC INTERPRETATION OF THE MACDONALD POLYNOMIALS

BY PERSI DIACONIS[1] AND ARUN RAM[2]

Stanford University and University of Melbourne

The two-parameter Macdonald polynomials are a central object of algebraic combinatorics and representation theory. We give a Markov chain on partitions of k with eigenfunctions the coefficients of the Macdonald polynomials when expanded in the power sum polynomials. The Markov chain has stationary distribution a new two-parameter family of measures on partitions, the inverse of the Macdonald weight (rescaled). The uniform distribution on cycles of permutations and the Ewens sampling formula are special cases. The Markov chain is a version of the auxiliary variables algorithm of statistical physics. Properties of the Macdonald polynomials allow a sharp analysis of the running time. In natural cases, a bounded number of steps suffice for arbitrarily large k.

1. Introduction. The Macdonald polynomials $P_\lambda(x; q, t)$ are a widely studied family of symmetric polynomials in variables $X = (x_1, x_2, \dots, x_n)$. Let Λ_n^k denote the vector space of homogeneous symmetric polynomials of degree k (with coefficients in $\mathbb{Q}$). The Macdonald inner product is determined by setting the inner product between power sum symmetric functions p_λ as

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\lambda(q, t)$$

with

$$z_\lambda(q, t) = z_\lambda \prod_i \left(\frac{1 - q^{\lambda_i}}{1 - t^{\lambda_i}} \right) \quad \text{and} \quad z_\lambda = \prod_i i^{a_i} a_i! \tag{1.1}$$

for λ a partition of k with a_i parts of size i.

For each q, t, as λ ranges over partitions of k, the $P_\lambda(x; q, t)$ are an orthogonal basis for Λ_n^k. Special values of q, t give classical bases such as Schur functions ($q = t$), Hall–Littlewood functions ($t = 0$) and the Jack symmetric functions (limit as $t \to 1$ with $q^\alpha = t$). An enormous amount of combinatorics, group theory and algebraic geometry is coded into these polynomials. A more careful description and literature review is in Section 2.

Received February 2011; revised March 2011.
[1]Supported in part by NSF Grant 0804324.
[2]Supported in part by NSF Grant 0353038 and ARC Grants DP0986774 and DP087995.
MSC2010 subject classifications. Primary 05E05; secondary 60J10.
Key words and phrases. Macdonald polynomials, random permutations, measures on partitions, auxiliary variables, Markov chain, rates of convergence.

The original definition of Macdonald constructs $P_\lambda(x;q,t)$ as the eigenfunctions of a somewhat mysterious family of operators $D_{q,t}(z)$. This is used to develop their basic properties in [42]. A main result of the present paper is that the Macdonald polynomials can be understood through a natural Markov chain $M(\lambda,\lambda')$ on the partitions of k. For $q,t>1$, this Markov chain has stationary distribution

$$\pi_{q,t}(\lambda)=\frac{Z}{z_\lambda(q,t)}\qquad \text{with } Z=\frac{(q,q)_k}{(t,q)_k},\qquad (x,y)_k=\prod_{i=0}^{k-1}(1-xy^i). \tag{1.2}$$

Here $z_\lambda(q,t)$ is the Macdonald weight (1.1), and Z is a normalizing constant. The coefficients of the Macdonald polynomials expanded in the power sums give the eigenvectors of M, and there is a simple formula for the eigenvalues.

Here is a brief description of M. From a current partition λ, choose some parts to delete: call these λ_J. This leaves $\lambda_{J^c}=\lambda\setminus\lambda_J$. The choice of λ_{J^c} given λ is made with probability

$$w_\lambda(\lambda_{J^c})=\frac{1}{q^k-1}\prod_{i=1}^{k}\binom{a_i(\lambda)}{a_i(\lambda_{J^c})}(q^i-1)^{a_i(\lambda)-a_i(\lambda_{J^c})}. \tag{1.3}$$

It is shown in Section 2.4 that for each λ, $w_\lambda(\cdot)$ is a probability distribution with a simple-to-implement interpretation. Having chosen λ_{J^c}, choose a partition μ of size $|\lambda|-|\lambda_{J^c}|$ with probability

$$\pi_{\infty,t}(\mu)=\frac{t}{t-1}\frac{1}{z_\mu}\prod\left(1-\frac{1}{t^i}\right)^{a_i(\mu)}. \tag{1.4}$$

Adding μ to λ_{J^c} gives a final partition ν. These two steps define the Markov chain $M(\lambda,\nu)$ with stationary distribution $\pi_{q,t}$. It will be shown to be a natural extension of basic algorithms of statistical physics: the Swendsen–Wang and auxiliary variables algorithms. Properties of the Macdonald polynomials give a sharp analysis of the running time for M.

Section 2 gives background on Macdonald polynomials (Section 2.1), Markov chains (Section 2.2) and auxiliary variables algorithms (Section 2.3). The Markov chain M is shown to be a special case of auxiliary variables and hence is reversible with $\pi_{q,t}(\lambda)$ as stationary distribution. Section 2.4 reviews some of the many different measures used on partitions, showing that w_λ and $\pi_{\infty,t}$ above have simple interpretations and efficient sampling algorithms.

The main theorems are in Section 3. The Markov chain M is identified as one term of the Macdonald operators $D_{q,t}(z)$. The coefficients of the Macdonald polynomials in the power sum basis (suitably scaled) are shown to be the eigenfunctions of M with a simple formula for the eigenvalues. Values of the eigenvectors are derived. A heuristic overview of the argument is given (Section 3.2), which may be read now for further motivation.

The main theorem is an extension of earlier work by Hanlon [16, 34], giving a similar interpretation of the coefficients of the family of Jack symmetric functions as eigenfunctions of a natural Markov chain: the Metropolis algorithm on the symmetric group for generating from the Ewens sampling formula. Section 4 develops the connection to the present study.

Section 5 gives an analysis of the convergence of iterates of M to the stationary distribution $\pi_{q,t}$ for natural values of q and t. Starting from (k), it is shown that a bounded number of steps suffice for arbitrary k. Starting from 1^k, order $\log k$ steps are necessary and sufficient for convergence.

2. Background and examples. This section contains needed background on four topics: Macdonald polynomials, Markov chains, auxiliary variables algorithms and measures on partitions and permutations. Each of these has a large literature. We give basic definitions, needed formulas and pointers to literature. Section 2.3 shows that the Markov chain M of the introduction is a special case of the auxiliary variables algorithm. Section 2.4 shows that the steps of the algorithm are easy to run.

2.1. *Macdonald polynomials.* Let Λ_n be the algebra of symmetric polynomials in n variables (coefficients in $\mathbb{Q}$). There are many useful bases of Λ_n: the monomial $\{m_\lambda\}$, power sum $\{p_\lambda\}$, elementary $\{e_\lambda\}$, homogeneous $\{h_\lambda\}$ and Schur functions $\{s_\lambda\}$ are bases whose change of basis formulas contain a lot of basic combinatorics [53], Chapter 7, [42], Chapter I. More esoteric bases such as the Hall–Littlewood functions $\{H_\lambda(q)\}$, zonal polynomials $\{Z_\lambda\}$ and Jack symmetric functions $\{J_\lambda(\alpha)\}$ occur as the spherical functions of natural homogeneous spaces [42]. In all cases, as λ runs over partitions of k, the associated polynomials form a basis of the vector space Λ_n^k, homogeneous symmetric polynomials of degree k.

Macdonald introduced a two-parameter family of bases $P_\lambda(x; q, t)$ which, specializing q, t in various ways, gives essentially all the previous bases. The Macdonald polynomials can be succinctly characterized by using the inner product $\langle p_\lambda, p_\mu\rangle = \delta_{\lambda\mu} z_\lambda(q, t)$ with $z_\lambda(q, t)$ from (1.1). For $q, t > 1$ this is positive definite, and there is a unique family of symmetric functions $P(x; q, t)$ such that $\langle P_\lambda, P_\mu\rangle_{q,t} = 0$ if $\lambda \neq \mu$ and $P_\lambda = \sum_{\mu\leq\lambda} u_{\lambda\mu} m_\mu$ with $u_{\lambda\lambda} = 1$ [42], Chapter VI, (4.7). The properties of P_λ are developed by studying P_λ as the eigenfunctions of a family of operators $D_{q,t}(z)$ from Λ_n to Λ_n.

Define an operator T_{u,x_i} on polynomials by $T_{u,x_i} f(x_1, \dots, x_n) = f(x_1, \dots, ux_i, \dots, x_n)$. Define $D_{q,t}(z)$ and $D^r_{q,t}$ by

$$D_{q,t}(z) = \sum_{r=0}^{n} D^r_{q,t} z^r = \frac{1}{a_\delta} \sum_{w\in S_n} \det(w) x^{w\delta} \prod_{i=1}^{n} \bigl(1 + z t^{(w\delta)_i} T_{q,x_i}\bigr), \tag{2.1}$$

where $\delta = (n-1, n-2, \dots, 0)$, a_δ is the Vandermonde determinant and $x^\gamma = x_1^{\gamma_1} \cdots x_n^{\gamma_n}$ for $\gamma = (\gamma_1, \dots, \gamma_n)$. For any $r = 0, 1, \dots, n$,

$$D^r_{q,t} = \sum_I A_I(x; t) \prod_{i\in I} T_{q,x_i}, \tag{2.2}$$

where the sum is over all r-element subsets I of $\{1, 2, \dots, n\}$ and

$$A_I(x;t) = \frac{1}{a_\delta}\Big(\prod T_{t,x_i}\Big)a_\delta = t^{r(r-1)/2} \prod_{\substack{i\in I,\\ j\notin I}} \frac{tx_i - x_j}{x_i - x_j} \tag{2.3}$$

[42], Chapter VI, $(3.4)_r$.

Macdonald [42], Chapter VI, (4.15), shows that the Macdonald polynomials are eigenfunctions of $D_{q,t}(z)$

$$D_{q,t}(z)P_\lambda(x;q,t) = \prod_{i=1}^{n}(1 + zq^{\lambda_i}t^{n-i})P_\lambda(x;q,t). \tag{2.4}$$

This implies that the operators $D^r_{q,t}$ commute and have the P_λ as eigenfunctions with eigenvalues the rth elementary symmetric function in $\{q^{\lambda_i}t^{n-i}\}$. We will use $D^1_{q,t}$ in our work below. The $D^r_{q,t}$ are self adjoint in the Macdonald inner product $\langle D^r_{q,t} f, g\rangle = \langle f, D^r_{q,t} g\rangle$. This will translate into having $\pi_{q,t}$ as stationary distribution.

The Macdonald polynomials may be expanded in the power sums [42], Chapter VI, (8.19),

$$P_\lambda(x;q,t) = \frac{1}{c_\lambda(q,t)} \sum_\rho \Big[z_\rho^{-1} \prod_i (1 - t^{\rho_i}) X^\lambda_\rho(q,t)\Big] p_\rho(x) \tag{2.5}$$

with [42], Chapter VI, (8.1), $c_\lambda(q,t) = \prod_{s\in\lambda}(1 - q^{a(s)}t^{l(s)+1})$ where the product is over the boxes in the shape of λ, $a(s)$ the arm length and $l(s)$ the leg length of box s. The $X^\lambda_\rho(q,t)$ are closely related to the two-parameter Kostka numbers $K_{\mu\lambda}(q,t)$ via [42], Chapter VI, (8.20),

$$X^\lambda_\rho(q,t) = \sum_\mu \chi^\lambda_\rho K_{\mu\lambda}(q,t), \qquad K_{\mu\lambda}(q,t) = \sum_\rho z_\rho^{-1} \chi^\mu_\rho X^\lambda_\rho(q,t) \tag{2.6}$$

with χ^λ_ρ the characters of the symmetric group for the λth representation at the ρth conjugacy class. These $K_{\mu\lambda}(q,t)$ have been a central object of study in algebraic combinatorics [6, 26, 29–33]. The main result of Section 3 shows that $X^\lambda_\rho(q,t)\prod_i(1 - q^{\rho_i})$ are the eigenfunctions of the Markov chain M.

The Macdonald polynomials used here are associated to the root system A_n. Macdonald [43] has defined analogous functions for the other root systems using similar operators. In a major step forward, Cherednik [15] gives an independent development in all types, using the double affine Hecke algebra. See [42, 44] for a comprehensive treatment. Using this language, Ram and Yip [51] give a "formula" for the Macdonald polynomials in general type. In general type the double affine Hecke is a powerful tool for understanding actions. We believe that our Markov chain can be developed in general type if a suitable analog of the power sum basis is established.

2.2. *Markov chains*. Let $\mathcal{X}$ be a finite set. A Markov chain on $\mathcal{X}$ may be specified by a matrix $M(x,y) \geq 0$, $\sum_y M(x,y) = 1$. The interpretation being that

$M(x, y)$ is the chance of moving from x to y in one step. Then $M^2(x, y) = \sum_z M(x, z)M(z, y)$ is the chance of moving from x to y in two steps, and $M^\ell(x, y)$ is the chance of moving from x to y in ℓ steps. Under mild conditions, always met in our examples, there is a unique stationary distribution $\pi(x) \geq 0$, $\sum_x \pi(x) = 1$. This satisfies $\sum_x \pi(x)M(x, y) = \pi(y)$. Hence, the (row) vector π is a left eigenvector of M with eigenvalue 1. Probabilistically, picking x from π and taking one further step in the chain leads to the chance $\pi(y)$ of being at y.

All of the Markov chains used here are reversible, satisfying the detailed balance condition $\pi(x)M(x, y) = \pi(y)M(y, x)$, for all x, y in $\mathcal{X}$. Set $L^2(\mathcal{X})$ to be $\{f : \mathcal{X} \to \mathbb{R}\}$ with $(f_1, f_2) = \sum_x \pi(x) f_1(x) f_2(x)$. Then M acts as a contraction on $L^2(\mathcal{X})$ by $Mf(x) = \sum_y M(x, y) f(y)$. Reversibility is equivalent to M being self-adjoint. In this case, there is an orthogonal basis of (right) eigenfunctions f_i and real eigenvalues β_i, $1 = \beta_0 \geq \beta_1 \geq \cdots \geq \beta_{|\mathcal{X}|-1} \geq -1$ with $Mf_i = \beta_i f_i$. For reversible chains, if $f_i(x)$ is a left eigenvector, then $f_i(x)/\pi(x)$ is a right eigenvector with the same eigenvalue.

A basic theorem of Markov chain theory shows that $M_x^\ell(y) = M^\ell(x, y) \xrightarrow[\infty]{\ell} \pi(y)$. (Again, there are mild conditions, met in our examples.) The distance to stationarity can be measured in L^1 by the total variation distance,

$$\|M_x^\ell - \pi\|_{\mathrm{TV}} = \max_{A \subseteq \mathcal{X}} |M^\ell(x, A) - \pi(A)| = \frac{1}{2} \sum_y |M^\ell(x, y) - \pi(y)|. \tag{2.7}$$

Distance is measured in L^2 by the chi-squared distance

$$\|M_x^\ell - \pi\|_2^2 = \sum_y \frac{(M^\ell(x, y) - \pi(y))^2}{\pi(y)} = \sum_{i=1}^{|\mathcal{X}|-1} \bar{f}_i^2(x) \beta_i^{2\ell}, \tag{2.8}$$

where $\bar{f}_i$ is the eigenvector f_i, normalized to have L^2-norm 1. The Cauchy–Schwarz inequality shows

$$4\|M_x^\ell - \pi\|_{\mathrm{TV}}^2 \leq \|M_x^\ell - \pi\|_2^2. \tag{2.9}$$

Using these bounds calls for getting one's hands on eigenvalues and eigenvectors. This can be hard work, but it has been done in many cases. A central question is this: given M, $\varepsilon > 0$, and a starting state x, how large must ℓ be so that $\|M_x^\ell - \pi\|_{\mathrm{TV}} < \varepsilon$?

Background on the quantitative study of rates of convergence of Markov chains is treated in the textbook of Brémaud [13]. The identities and inequalities that appear above are derived in the very useful treatment by Saloff-Coste [52]. He shows how tools of analysis can be brought to bear. The recent monograph of Levin, Peres and Wilmer [40] is readable by nonspecialists and covers both analytic and probabilistic techniques.

2.3. *Auxiliary variables.* This is a method of constructing a reversible Markov chain with π as stationary distribution. It was invented by Edwards and Sokal [22] as an abstraction of the remarkable Swendsen–Wang algorithm. The Swendsen–Wang algorithm was introduced as a superfast method for simulating from the Ising and Potts models of statistical mechanics. It is a block-spin procedure which changes large pieces of the current state. A good overview of such block spin algorithms is in [45]. The abstraction to auxiliary variables is itself equivalent to several other classes of widely used procedures, data augmentation and the hit-and-run algorithm. For these connections and much further literature, see [3].

To describe auxiliary variables, let $\pi(x) > 0$, $\sum_x \pi(x) = 1$ be a probability distribution on a finite set $\mathcal{X}$. Let I be an auxiliary index set. For each $x \in \mathcal{X}$, let $w_x(i)$ be a probability distribution on I (the chance of moving to i). These define a joint distribution $f(x,i) = \pi(x)w_x(i)$ and a marginal distribution $m(i) = \sum_x f(x,i)$. Let $f(x|i) = f(x,i)/m(i)$ denote the conditional distribution. The final ingredient needed is a Markov matrix $M_i(x,y)$ with $f(x|i)$ as reversing measure $[f(x|i)M_i(x,y) = f(y|i)M_i(y,x)$ for all $x, y]$. This allows for defining

$$M(x,y) = \sum_i w_x(i)M_i(x,y). \tag{2.10}$$

The Markov chain M has the following interpretation: from x, choose $i \in I$ from $w_x(i)$ and then $y \in \mathcal{X}$ from $M_i(x,y)$. The resulting kernel is reversible with respect to π:

$$\begin{aligned}\pi(x)M(x,y) &= \sum_i \pi(x)w_x(i)M_i(x,y) = \sum_i \pi(y)w_y(i)M_i(y,x)\\ &= \pi(y)\sum_i w_y(i)M_i(y,x) = \pi(y)M(y,x).\end{aligned}$$

We now specialize things to $\mathcal{P}_k$, the space of partitions of k. Take $\mathcal{X} = \mathcal{P}_k$, $I = \bigcup_{i=1}^k \mathcal{P}_i$. The stationary distribution is as in (1.2),

$$\pi(\lambda) = \pi_{q,t}(\lambda) = \frac{Z}{z_\lambda(q,t)}. \tag{2.11}$$

From $\lambda \in \mathcal{P}_k$, the algorithm chooses some parts to delete; call these λ_J, leaving parts $\lambda_{J^c} = \lambda \setminus \lambda_J$. Thus if $\lambda = 322111$ and $\lambda_J = 31$, $\lambda_{J^c} = 2211$. We allow $\lambda_J = \lambda$ but demand $\lambda_J \neq \varnothing$. Clearly, λ and λ_J determine λ_{J^c}, and (λ, λ_{J^c}) determines λ_J. We let λ_{J^c} be the auxiliary variable. The choice of λ_{J^c} given λ is made with probability

$$\begin{aligned}w_\lambda(\lambda_{J^c}) &= \frac{1}{q^k-1}\prod_{i=1}^k \binom{a_i(\lambda)}{a_i(\lambda_{J^c})}(q^i-1)^{a_i(\lambda_J)}\\ &= \frac{1}{q^k-1}\prod_{i=1}^k \binom{a_i(\lambda)}{a_i(\lambda_{J^c})}(q^i-1)^{a_i(\lambda)-a_i(\lambda_{J^c})}.\end{aligned} \tag{2.12}$$

Thus, for $\lambda = 1^3 2 3^2$; $\lambda_J = 13$, $\lambda_{J^c} = 1^2 2 3$; $w_\lambda(\lambda_{J^c}) = \frac{1}{q^{11}-1}\binom{3}{2}\binom{1}{1}\binom{2}{1}(q-1)(q^2-1)^0(q^3-1)$. It is shown in Section 2.4 below that $w_\lambda(\lambda_{J^c})$ is a probability distribution with a simple interpretation. Having chosen λ_J with $0 < |\lambda_J| \le k$, the algorithm chooses $\mu \vdash |\lambda_J|$ with probability $\pi_{\infty,t}(\mu)$ given in (1.4). Adding these parts to λ_{J^c} gives ν. More carefully,

$$M_{\lambda_{J^c}}(\lambda,\nu) = \pi_{\infty,t}(\mu) = \frac{t}{(t-1)}\frac{1}{z_\mu}\prod_i\left(1-\frac{1}{t^i}\right)^{a_i(\mu)}. \tag{2.13}$$

Here it is assumed that λ_{J^c} is a part of both λ and ν; the kernel $M_{\lambda_{J^c}}(\lambda,\nu)$ is zero otherwise.

It is shown in Section 2.4 below that $M_{\lambda_{J^c}}$ has a simple interpretation which is easy to sample from. The joint density $f(\lambda,\lambda_{J^c}) = \pi(\lambda)w_\lambda(\lambda_{J^c})$ is proportional to $f(\lambda|\lambda_{J^c})$ and to

$$\frac{\prod_i(1-1/t^i)^{a_i(\lambda)}}{\prod_i i^{a_i(\lambda)}(a_i(\lambda)-a_i(\lambda_{J^c}))!}. \tag{2.14}$$

The normalizing constant depends on λ_{J^c}, but this is fixed in the following. We must now check reversibility of $f(\lambda|\lambda_{J^c})M_{\lambda_{J^c}}(\lambda,\nu)$. For this, compute $f(\lambda|\lambda_{J^c})M_{\lambda_{J^c}}(\lambda,\nu)$ (up to a constant depending on λ_{J^c}) as

$$\frac{\prod_i(1-1/t^i)^{a_i(\lambda)+a_i(\nu)}}{\prod_i i^{a_i(\lambda)+a_i(\nu)}(a_i(\lambda)-a_i(\lambda_{J^c}))!(a_i(\nu)-a_i(\lambda_{J^c}))!}.$$

This is symmetric in λ,ν and so equals $f(\nu|\lambda_{J^c})M_{\lambda_{J^c}}(\nu,\lambda)$. This proves the following:

PROPOSITION 2.1. *With definitions* (2.11)–(2.14), *the kernel on* $\mathcal{P}_k$,

$$M(\lambda,\nu) = \sum_{\lambda_{J^c}} w_\lambda(\lambda_{J^c})M_{\lambda_{J^c}}(\lambda,\nu)$$

generates a reversible Markov chain with $\pi_{q,t}(\lambda)$ *as stationary distribution.*

EXAMPLE 1. With $k=2$, let

$$\pi_{q,t}(2) = \frac{Z}{2}\frac{(t^2-1)}{(q^2-1)}, \qquad \pi_{q,t}(1^2) = \frac{Z}{2}\left(\frac{t-1}{q-1}\right)^2$$

$$\text{for } Z = \frac{(1-q)(1-q^2)}{(1-t)(1-tq)}.$$

From the definitions, with rows and columns labeled (2), 1^2, the transition matrix is

$$\begin{aligned} M &= \begin{pmatrix} \frac{1}{2}\left(1+\frac{1}{t}\right) & \frac{1}{2}\left(1-\frac{1}{t}\right) \\ \frac{q-1}{q+1}\frac{1}{2}\left(1+\frac{1}{t}\right) & \frac{4t+(q-1)(t-1)}{2(q+1)t} \end{pmatrix} \\ &= \frac{1}{2t}\begin{pmatrix} t+1 & t-1 \\ \frac{(q-1)(t+1)}{q+1} & \frac{4t+(q-1)(t-1)}{q+1} \end{pmatrix}. \end{aligned} \tag{2.15}$$

In this $k = 2$ example, it is straightforward to check that $\pi_{q,t}$ sums to 1, the rows of M sum to 1, and that $\pi_{q,t}(\lambda)M(\lambda, \nu) = \pi_{q,t}(\nu)M(\nu, \lambda)$.

2.4. *Measures on partitions and permutations*. The measure $\pi_{q,t}$ of (1.2) has familiar specializations: to the distribution of conjugacy classes of a uniform permutation ($q = t$), and the Ewens sampling measure ($q^\alpha = t \to 1$). After recalling these, the measures $w_{\lambda_{J^c}}(\cdot)$ and $M_{\lambda_{J^c}}(\lambda, \cdot)$ used in the auxiliary variables algorithm are treated. Finally, there is a brief review of the many other, nonuniform distributions used on partitions $\mathcal{P}_k$ and permutations S_k. Along the way, many results on the "shape" of a typical partition drawn from $\pi_{q,t}$ appear.

2.4.1. *Uniform permutations* ($q = t$). If σ is chosen uniformly on S_k, the chance that the cycle type of σ is λ is $1/z_\lambda = \pi_{q,q}(\lambda)$. There is a healthy literature on the structure of random permutations (number of fixed points, cycles of length i, number of cycles, longest and shortest cycles, order, ...). This is reviewed in [25, 49], which also contain extensions to the distribution of conjugacy classes of finite groups of Lie type.

One natural appearance of the measure $1/z_\lambda$ comes from the *coagulation/fragmentation* process. This is a Markov chain on partitions of k introduced by chemists and physicists to study clump sizes. Two parts are chosen with probability proportional to their size. If different parts are chosen, they are combined. If the same part is chosen twice, it is split uniformly into two parts. This Markov chain has stationary distribution $1/z_\lambda$. See [2] for a review of a surprisingly large literature and [18] for recent developments. These authors note that the coagulation/fragmentation process is the random transpositions walk, viewed on conjugacy classes. Using the Metropolis algorithm (as in Section 2.4.6 below) gives a similar process with stationary distribution $\pi_{q,t}$.

Algorithmically, a fast way to pick λ with probability $1/z_\lambda$ is by *uniform stick-breaking*: pick $U_1 \in \{1, \dots, k\}$ uniformly. Pick $U_2 \in \{1, \dots, k - U_1\}$ uniformly. Continue until the first time T that the uniform choice equals its maximum attainable value. The partition with parts $U_1, U_2, \dots, U_T$ equals λ with probability $1/z_\lambda$.

2.4.2. *Ewens and Jack measures*. Set $q = t^\alpha$, and let $t \to 1$. Then $\pi_{q,t}(\lambda)$ converges to

$$\pi_\alpha(\lambda) = \frac{Z}{z_\lambda}\alpha^{-\ell(\lambda)}, \qquad Z = \frac{\alpha^k k!}{\prod_{i=1}^{k-1}(i\alpha + 1)}, \tag{2.16}$$

$\ell(\lambda)$ the number of parts of λ.

In population genetics, setting $\alpha = 1/\theta$, with $\theta > 0$ a "fitness parameter," this measure is called the *Ewens sampling formula*. It has myriad practical appearances through its connection with Kingman's coalescent process, and has generated a large enumerative literature in the combinatorics and probability community [5, 35, 50]. It also makes numerous appearances in the statistics literature through its occurrence in nonparametric Bayesian statistics via Dirichlet random measures and the Dubins–Pitman Chinese restaurant process [27] and [50], Section 3.1.

Algorithmically, a fast way to pick λ with probability $\pi_{1/\theta}(\lambda)$ is by the Chinese restaurant construction. Picture a collection of circular tables. Person 1 sits at the first table. Successive people sit sequentially, by choosing to sit to the right of a (uniformly chosen) previously seated person (probability θ) or at a new table (probability $1-\theta$). When k people have been seated, this generates the cycles of a random permutation with probability $\pi_{1/\theta}$. It would be nice to have a similar construction for the measures $\pi_{q,t}$.

The Macdonald polynomials associated to this weight function are called the *Jack symmetric functions* [42], Chapter VI, Section 1. Hanlon [16, 34] uses properties of Jack polynomials to diagonalize a related Markov chain; see Section 4. When $\alpha = 2$, the Jack polynomials become the zonal-spherical functions of GL_n/O_n. Here, an analysis closely related to the present paper is carried out for a natural Markov chain on perfect matchings and phylogenetic trees [14], Chapter X, [17].

2.4.3. *The measure* w_λ. Fix $\lambda \vdash k$ with ℓ parts and $q > 1$. Define, for $J \subseteq \{1, \dots, \ell\}$, $J \neq \varnothing$,

$$w_\lambda(J) = \frac{1}{q^k - 1} \prod_{i \in J} (q^{\lambda_i} - 1). \tag{2.17}$$

The auxiliary variables algorithm for sampling from $\pi_{q,t}$ involves sampling from $w_\lambda(J)$, and setting $\lambda_J = \{\lambda_i : i \in J\}$; see (1.3) and (2.12). The measure $w_\lambda(J)$ has the following interpretation, which leads to a useful sampling algorithm: consider k places divided into blocks of length λ_i,

$$\underbrace{- - \cdots -}_{\lambda_1} \underbrace{- - \cdots -}_{\lambda_2} \cdots \underbrace{- - \cdots -}_{\lambda_l}, \qquad \lambda_1 + \cdots + \lambda_l = k.$$

Flip a $1/q$ coin for each place. Let, for $1 \le i \le \ell$,

$$X_i = \begin{cases} 1, & \text{if the } i\text{th block is } \textit{not} \text{ all ones,} \\ 0, & \text{otherwise.} \end{cases} \tag{2.18}$$

Thus $P(X_i = 1) = 1 - 1/q^{\lambda_i}$. Let $J = \{i : X_i = 1\}$. So $P\{J = \varnothing\} = 1 - 1/q^k$ and

$$P\{J | J \neq \varnothing\} = \frac{1}{1 - 1/q^k} \prod_{i \in J} \left(1 - \frac{1}{q^{\lambda_i}}\right) \prod_{j \in J^c} \frac{1}{q^{\lambda_j}} = w_\lambda(J). \tag{2.19}$$

This makes it clear that summing $w_\lambda(J)$ over all nonempty subsets of $\{1, \dots, \ell\}$ gives 1.

The simple rejection algorithm for sampling from w_λ is: flip coins as above. If $J \neq \varnothing$, output $\lambda_J = \{\lambda_i : i \in J\}$. If $J = \varnothing$, sample again. The chance of success is $1 - 1/q^k$. Thus, unless q is very close to 1, this is an efficient algorithm.

As q tends to ∞, w_λ converges to point mass at $J = \{1, \dots, k\}$. As q tends to 1, w_λ converges to the measure putting mass λ_i/k on $\{i\}$.

2.4.4. *The measure* $\pi_{\infty,t}$. Generating from the kernel $M_{\lambda_{J^c}}(\lambda, \nu)$ of (2.13) with $r = |\lambda \setminus \lambda_{J^c}|$, requires generating a partition in $\mathcal{P}_r$ from

$$\pi_{\infty,t}(\mu) = \left(\frac{t}{t-1}\right)\frac{1}{z_\mu}\prod_i\left(1 - \frac{1}{t^i}\right)^{a_i(\mu)}.$$

This measure has the following interpretation: pick $\mu^{(1)} \vdash r$ with probability $1/z_{\mu^{(1)}}$. This may be done by picking a random permutation in S_r uniformly and reporting the cycle decomposition, or by the uniform stick-breaking of Section 2.4.1 above. For each part $\mu_j^{(1)}$ of $\mu^{(1)}$, flip a $1/t$ coin $\mu_j^{(1)}$ times. If this comes up tails at least once, and this happens simultaneously for each j, set $\mu = \mu^{(1)}$. If some part of $\mu^{(1)}$ produces all heads, start again and choose $\mu^{(2)} \vdash r$ with probability $1/z_{\mu^{(2)}} \dots$. The chance of failure is $1/t$, independent of r. Thus, unless t is close to 1, this gives a simple, useful algorithm.

The shape of a typical pick from $\pi_{\infty,t}$ is described in the following section. When t tends to infinity, the measure converges to $1/z_\mu$. When t tends to one, the measure converges to point mass at the one part partition (r).

2.4.5. *Multiplicative measures*. For $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_k), \eta_i \geq 0$, define a probability on $\mathcal{P}_k$ (equivalently, S_k) by

$$\pi_{\boldsymbol{\eta}}(\lambda) = \frac{Z}{z_\lambda}\prod_{i=1}^{k}\eta_i^{a_i(\lambda)} \qquad \text{with } Z^{-1} = \sum_{\mu\vdash k}\frac{1}{z_\mu}\prod_i \eta_i^{a_i(\mu)}. \tag{2.20}$$

Such multiplicative measures are classical objects of study. They are considered in [5, 9, 23, 54] and [56], where many useful cases are given. The measures $\pi_{q,t}$ fall into this class with $\eta_i = \frac{(t^i-1)}{(q^i-1)}$. If $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ are two sequences of numbers and $V_\lambda(X)$ is a multiplicative basis of Λ_n^k such as $\{e_\lambda\}, \{p_\lambda\}, \{h_\lambda\}$, setting $\eta_i = V_i(x)V_i(y)$ gives $\pi_{\boldsymbol{\eta}}(\lambda) = \frac{Z}{z_\lambda}V_\lambda(x)V_\lambda(y)$. This is in rough analogy to the Schur measures defined in Section 2.4.7. For the choices $e_\lambda, p_\lambda, h_\lambda$, with x_i, y_j positive numbers, the associated measures are positive. The power sums, with all $x_i = a$, $y_i = b$, gives the Ewens measure with $\alpha = ab$. Setting $x_1 = y_1 = c$ and $x_j = y_j = 0$ otherwise, gives the measure $1/z_\lambda$ after normalization. Multiplicative systems are studied in [42], Chapter VI, Section 1, example.

The asymptotic distribution of the parts of a partition chosen from $\pi_{\boldsymbol{\eta}}$ when k is large can be studied by classical tools of combinatorial enumeration. For fixed values of q, t, these problems fall squarely into the domain of the logarithmic combinatorial structures studied in [5]. A series of further results for more general $\boldsymbol{\eta}$ have been developed by Jiang [37] and Zhao [57]. The following brief survey of their results gives a good picture of typical partitions.

Of course, the theorems vary with the choice of η_i. One convenient condition, which includes the measure $\pi_{q,t}$ for fixed $q, t > 1$, is

$$\sum_{i=1}^{\infty} \left| \frac{(\eta_i - 1)}{i} \right| < \infty. \tag{2.21}$$

THEOREM 2.2. *Suppose* $\eta_i, 1 \le i < \infty$, *satisfy* (2.21). *If* $\lambda \in \mathcal{P}_k$ *is chosen from* $\pi_{\boldsymbol{\eta}}$ *of* (2.20), *then, for k large*:

(2.22) *For any* j, *the distribution of* $(a_1(\lambda), \dots, a_j(\lambda))$ *converges to the distribution of an independent Poisson vector with parameters* η_i / i, $1 \le i \le j$.

(2.23) *The number of parts of* λ *has mean and variance asymptotic to* $\log k$ *and, normalized by its mean and standard deviation, a limiting standard normal distribution.*

(2.24) *The length of the k largest parts of* λ *converge to the Poisson–Dirichlet distribution* [10, 28, 41].

These and other results from [5, 37] show that the parts of a random partition are quite similar to the cycles of a uniformly chosen random permutation, with the small cycles having slightly adjusted parameters. These results are used to give a lower bound on the mixing time of the auxiliary variables Markov chain in Proposition 3.2 below.

2.4.6. *Simulation algorithms.* The distribution $\pi_{q,t}(\lambda)$ can be far from uniform. For example, with $k = 10, q = 4, t = 2, \pi_{4,2}(10) \doteq 0.16, \pi_{4,2}(1^{10}) \doteq 0$. The auxiliary variables algorithm for the measure $\pi_{q,t}$ has been programmed by Jiang [37] and Zhao [57]. It seems to work well over a wide range of q and t. In our experiments, the choice of q and t does not seriously affect the running time, and simulations seem possible for k up to 10^6.

It is natural to try out the simple rejection algorithms of Sections 2.4.3 and 2.4.4 for the measures $\pi_{\boldsymbol{\eta}}$. To begin, suppose that $0 < \eta_i < 1$ for all i. The measure $\pi_{\boldsymbol{\eta}}$ has the following interpretation: pick $\lambda' \in \mathcal{P}_k$ with probability $1/z_{\lambda'}$. As above, for each part of λ' of size i, generate a random variable taking values 1 or 0 with probability $\eta_i, 1 - \eta_i$. If the values for all parts equal 1, set $\lambda = \lambda'$. If not, try again. For more general η_i, divide all η_i by $\eta_* = \max \eta_i$, and generate from η_i / η_*^i. This yields the measure $\pi_{\boldsymbol{\eta}}$ on partitions. Alas, this algorithm performs poorly for η_i and k in ranges of interest. For example, with $\eta_i = \frac{t^i - 1}{q^i - 1}$ for $t = 2, q = 4$, when $k = 10, 11, 12, 13$, the chance of success (empirically) is 1/2,000, 1/4,000, 1/7,000, 1/12,000. We never succeeded in generating a partition for any $k \ge 15$.

TABLE 1
Mixing times to $\pi_{q,t}$ for $q = 4$ and $t = 2$

k	**10**	**20**	**30**	**40**	**50**
Aux	1	1	1	1	1
Met	8	17	26	37	53
$p(k)$	42	627	5,604	37,338	204,226

We also compare with the Metropolis algorithm. This works by simulating permutations from $\pi_{q,t}$ lifted to S_k. From the current permutation σ, propose σ' by making a random transposition [all $\binom{n}{2}$ choices equally likely]. If $\pi_{q,t}(\sigma') \geq \pi_{q,t}(\sigma)$, move to σ'. If $\pi_{q,t}(\sigma')/\pi_{q,t}(\sigma) < 1$, flip a coin with probability $\pi_{q,t}(\sigma')/\pi_{q,t}(\sigma)$ and move to σ' if the coin comes up heads; else stay at σ. For small values of k, Metropolis is competitive with auxiliary variables. Zhao has computed the mixing time for $k = 10, 20, 30, 40, 50$ by a clever sampling algorithm. For $q = 4, t = 2$, Table 1 shows the number of steps required to have total variation distance less than $1/10$ from stationary, starting from the partition (k). Also shown is $p(k)$, the number of partitions of k, to give a feeling for the size of the state space (see Table 1).

The theorems of Section 3 show that auxiliary variables requires a bounded number of steps for arbitrary k. In the computations above, the distance to stationarity after one step of the auxiliary variables is 0.093 (within a 1% error in the last decimal) for $k = 10, \ldots, 50$. For larger k (e.g., $k = 100$), the Metropolis algorithm seemed to need a very large number of steps to move at all. This is consistent with other instances of auxiliary variables, such as the Swendsen–Wang algorithm for the Ising and Potts model (away from the critical temperature; see [11]). Further numerical examples are in [20].

2.4.7. *Other measures on partitions.* This portmanteau section gives pointers to some of the many other measures that have been studied on $\mathcal{P}_k$ and S_k. Often these studies are fascinating, deep, and extensive. All measures studied here seem distinct from $\pi_{q,t}$.

A remarkable two-parameter family of measures on partitions has been introduced by Jim Pitman. For $\theta \geq 0$, $0 \leq \alpha \leq 1$, and $\lambda \vdash k$ with ℓ parts, set

$$P_{\theta,\alpha}(\lambda) = \frac{k!}{z_\lambda} \frac{\theta^{(\alpha,\ell-1)}}{(\theta + 1 - \alpha)^{(1,k-1)}} \prod_{j=1}^{k} \left[(1-\alpha)^{(1,j-1)}\right]^{a_j(\lambda)},$$

where

$$\theta^{(a,m)} = \begin{cases} 1, & \text{if } m = 0, \\ \theta(\theta + a) \cdots \big(\theta + (m-1)a\big), & \text{for } m = 1, 2, 3, \ldots. \end{cases}$$

These measures specialize to $1/z_\lambda$ ($\theta = 1, \alpha = 0$), and the Ewens measure (θ fixed, $\alpha = 0$); see [50], Section 3.2. They arise in a host of probability problems connected to stable stochastic problems of index α. They are also being used in applied probability connected to genetics and Bayesian statistics. They satisfy elegant consistency properties as k varies. For example, deleting a random part gives the corresponding measure on $\mathcal{P}_{k-1}$. For these and many other developments, see the book-length treatments of [8] and [50], Section 3.2.

One widely-studied measure on partitions is the Plancherel measure,

$$p(\lambda) = f(\lambda)^2/k!$$

with $f(\lambda)$ the dimension of the irreducible representation of S_k associated to shape λ. This measure was perhaps first studied in connection with Ulam's problem on the distribution of the length of the longest increasing sequence in a random permutation; see [41, 55]. For extensive developments and references, see [1, 38].

The Schur measures of [12, 46–48] are generalizations of the Plancherel measure. Here the chance of λ is taken as proportional to $s_\lambda(\mathbf{x})s_\lambda(\mathbf{y})$, with s_λ the Schur function and $\mathbf{x}, \mathbf{y}$ collections of real-valued entries. Specializing $\mathbf{x}$ and $\mathbf{y}$ in various ways yields a variety of previously-studied measures. One key property, if the partition is "tilted 135°" to make a v-shape and the local maxima projected onto the x-axis, the resulting points form a determinantal point process with a tractable kernel. This gives a fascinating collection of shape theorems for the original partition.

One final distribution, the uniform distribution on $\mathcal{P}_k$, has also been extensively studied. For example, a uniformly chosen partition has order $\sqrt{6k}/\pi$ parts of size 1, the largest part is of size $(\sqrt{6k}/\pi) \cdot \log(\sqrt{6k}/\pi)$, the number of parts is of size $\sqrt{6k}\log(k/(2\pi))$. A survey with much more refined results is in [24].

The above only scratches the surface. The reader is encouraged to look at [46–48] to see the breadth and depth of the subject as applied to Gromov–Witten theory, algebraic geometry and physics. The measures there seem closely connected to the "Plancherel dual" of our $\pi_{q,t}$. This dual puts mass proportional to $c(\lambda)c'(\lambda)$ on λ, with c, c' the arm-leg length products defined in Section 3.1 below.

3. Main results. This section shows that the auxiliary variables Markov chain M with stationary distribution $\pi_{q,t}(\lambda), \lambda \in \mathcal{P}_k$, is explicitly diagonalizable with eigenfunctions $f_\lambda(\mu)$ essentially the coefficients of the Macdonald polynomials expanded in the power sum basis. The result is stated in Section 3.1. The proof, given in Section 3.3, is somewhat computational. An explanatory overview is in Section 3.2. In Section 5, these eigenvalue/eigenvector results are used to bound rates of convergence of M.

3.1. *Statement of main results.* Fix $q, t > 1$ and $k \geq 2$. Let $M(\lambda, \mu) = \sum_{\lambda_{J^c}} w_\lambda(\lambda_{J^c}) M_{\lambda_{J^c}}(\lambda, \mu)$ be the auxiliary variables Markov chain on $\mathcal{P}_k$. Here, $w_\lambda(\cdot)$ and $M_{\lambda_{J^c}}(\lambda, \mu)$ are defined in (2.12), (2.13) and studied in Sections 2.4.3

and 2.4.4. For a partition λ, let

$$(3.1)\qquad c_\lambda(q,t)=\prod_{s\in\lambda}\bigl(1-q^{a(s)}t^{l(s)+1}\bigr)\quad\text{and}\quad c'_\lambda(q,t)=\prod_{s\in\lambda}\bigl(1-q^{a(s)+1}t^{l(s)}\bigr),$$

where the product is over the boxes in the shape λ, and $a(s)$ is the arm length and $l(s)$ the leg length of box s [42], Chapter VI, (8.1). Write $l(\lambda)$ for the number of parts of the partition λ.

THEOREM 3.1. (1) *The Markov chain $M(\lambda,\nu)$ is reversible and ergodic with stationary distribution $\pi_{q,t}(\lambda)$ defined in* (1.2). *This distribution is properly normalized.*

(2) *The eigenvalues of M are $\{\beta_\lambda\}_{\lambda\in\mathcal{P}_k}$ given by*

$$\beta_\lambda=\frac{t}{q^k-1}\sum_{i=1}^{\ell(\lambda)}(q^{\lambda_i}-1)t^{-i}.$$

Thus, $\beta_k=1$, $\beta_{k-1,1}=\frac{t}{q^k-1}(\frac{q^{k-1}-1}{t}+\frac{q-1}{t^2}),\dots.$

(3) *The corresponding right eigenfunctions are*

$$f_\lambda(\rho)=X^\lambda_\rho(q,t)\prod_{i=1}^{\ell(\rho)}(1-q^{\rho_i})$$

with $X^\lambda_\rho(q,t)$ the coefficients occurring in the following expansion of the Macdonald polynomials in terms of the power sums [42], *Chapter* VI, (8.19):

$$(3.2)\qquad P_\lambda(x;q,t)=\frac{1}{c_\lambda(q,t)}\sum_\rho\Biggl[z_\rho^{-1}X^\lambda_\rho(q,t)\prod_{i=1}^{\ell(\rho)}(1-t^{\rho_i})\Biggr]p_\rho(x).$$

(4) *The $f_\lambda(\rho)$ are orthogonal in $L^2(\pi_{q,t})$ with*

$$\langle f_\lambda,f_\mu\rangle=\delta_{\lambda\mu}c_\lambda(q,t)c'_\lambda(q,t)\frac{(q,q)_k}{(t,q)_k}.$$

EXAMPLE 2. When $k=2$, the matrix M with rows and columns indexed by 2, 1^2, is as in (2.15). Macdonald [42], page 359, gives tables of $K(\lambda,\mu)$ for $2\le k\le 6$. For $k=2$, $K(\lambda,\mu)$ is $\begin{pmatrix}1&q\\t&1\end{pmatrix}$. The character matrix is $\begin{pmatrix}1&1\\-1&1\end{pmatrix}$, and the product is $\begin{pmatrix}1-q&1+q\\t-1&t+1\end{pmatrix}$. From Theorem 3.1(3), the rows of this matrix, multiplied coordinatewise by $(1-q^2)$, $(1-q)^2$, give the right eigenvectors

$$f_{(2)}(2)=f_{(2)}(1^2)=(1-q)^2(1+q),$$
$$f_{(1^2)}(2)=(t-1)(1-q^2)\quad\text{and}\quad f_{(1^2)}(1^2)=(t+1)(1-q)^2.$$

Then $f_{(2)}(\rho)$ is a constant function, and $f_{(1^2)}(\rho)$ satisfies $\sum_\rho M(\lambda,\rho)f_{(1^2)}(\rho)=\beta_{(1^2)}f_{(1^2)}(\lambda)$, with $\beta_{(1^2)}=\frac{1+t^{-1}}{1+q}$.

Further useful formulas, used in Section 5, are [42], Chapter VI, Section 8, Example 8,

$$
\begin{aligned}
X_\rho^{(k)}(q,t) &= (q,q)_k \prod_{i=1}^{\ell(\rho)} (1-t^{\rho_i}),\\
X_\rho^{(1^k)}(q,t) &= (-1)^{|\rho|-\ell(\rho)}(t,t)_k \prod_{i=1}^{\ell(\rho)} (1-t^{\rho_i})^{-1},\\
X_{(k)}^\lambda(q,t) &= \prod_{\substack{(i,j)\in\lambda\\ (i,j)\neq(1,1)}} (t^{i-1}-q^{j-1}),\\
X_{(1^k)}^\lambda(q,t) &= \frac{c'_\lambda(q,t)}{(1-t)^k}\sum_T \varphi_T(q,t)
\end{aligned}
\tag{3.3}
$$

with the sum over standard tableaux T of shape λ, and $\varphi_T(q,t)$ from [42], Chapter VI, page 341, (1), and [42], Chapter VI, (7.11).

3.2. *Overview of the argument.* Macdonald [42], Chapter VI, defines the Macdonald polynomials as the eigenfunctions of the operator $D_{q,t}^1 : \Lambda_n \to \Lambda_n$ from (2.2). As described in Section 2.1 above, $D_{q,t}^1$ is self-adjoint for the Macdonald inner product and sends Λ_n^k into itself [42], Chapter VI, (4.15). For $\lambda \vdash k, k \le n$,

$$
D_{q,t}^1 P_\lambda(x;q,t) = \bar\beta_\lambda P_\lambda(x;q,t) \qquad \text{with } \bar\beta_\lambda = \sum_{i=1}^{\ell(\lambda)} q^{\lambda_i} t^{k-i}.
\tag{3.4}
$$

The Markov chain M is related to an affine rescaling of the operator $D_{q,t}^1$, which Macdonald [42], Chapter VI, (4.1), calls E_n. We work directly with $D_{q,t}^1$ to give direct access to Macdonald's formulas. The affine rescaling is carried out at the end of Section 3.3 below.

The *integral form* of Macdonald polynomials [42], Chapter VI, Section 8, is

$$
J_\lambda(x;q,t) = c_\lambda(q,t) P_\lambda(x;q,t)
$$

for c_λ defined in (3.1). Of course, the J_λ are also eigenfunctions of $D_{q,t}^1$. The J_λ may be expressed in terms of the shifted power sums via [42], Chapter VI, (8.19).

$$
\begin{aligned}
J_\lambda(x;q,t) &= \sum_\rho z_\rho^{-1} X_\rho^\lambda(q,t) p_\rho(x;t),\\
p_\rho(x;t) &= p_\rho(x) \prod_{i=1}^{\ell(\rho)} (1-t^{\rho_i}).
\end{aligned}
\tag{3.5}
$$

This gives our equation (3.2) above. In Proposition 3.2 below, we compute the action of $D^1_{q,t}$ on the power sum basis: for λ with ℓ parts,

$$
\begin{aligned}
D^1_{q,t} p_\lambda &\stackrel{\text{def}}{=} \sum_\mu \bar{M}(\lambda,\mu) p_\mu \\
&= [n]p_\lambda + \frac{t^n}{t-1} \sum_{J\subseteq\{1,\dots,\ell\}} p_{\lambda_{J^c}} \prod_{k\in J}(q^{\lambda_k}-1) \sum_{\mu\vdash|\lambda_J|} \prod_m (1-t^{-\mu_m}) \frac{p_\mu}{z_\mu}.
\end{aligned} \tag{3.6}
$$

On the right, the coefficient of $p_{\lambda_{J^c}} p_\mu$ is essentially the Markov chain M; we use $\bar{M}$ for this unnormalized version. Indeed, we first computed (3.6) and then recognized the operator as a special case of the auxiliary variables operator.

Equations (3.4)–(3.6) show that simply scaled versions of X^λ_ρ are eigenvectors of the matrix $\bar{M}$ as follows. From (3.4), (3.5),

$$
\begin{aligned}
\bar{\beta}_\lambda P_\lambda(x;q,t) = D^1_{q,t} P_\lambda(x;q,t) &= \frac{1}{c_\lambda} D^1_{q,t}(J_\lambda) \\
&= \frac{1}{c_\lambda} D^1_{q,t}\left(\sum_\rho X^\lambda_\rho \frac{1}{z_\rho} p_\rho(x;t)\right) \\
&= \frac{1}{c_\lambda} \sum_\rho X^\lambda_\rho \frac{\prod(1-t^{\rho_i})}{z_\rho} D^1_{q,t} p_\rho(x) \\
&= \frac{1}{c_\lambda} \sum_\rho \frac{\prod(1-t^{\rho_i})}{z_\rho} X^\lambda_\rho \sum_\mu \bar{M}(\rho,\mu) p_\mu(x) \\
&= \frac{1}{c_\lambda} \sum_\mu p_\mu \sum_\rho \frac{\prod(1-t^{\rho_i})}{z_\rho} X^\lambda_\rho \bar{M}(\rho,\mu).
\end{aligned} \tag{3.7}
$$

Also, from (3.4) and (3.5),

$$
\bar{\beta}_\lambda P_\lambda(x;q,t) = \frac{\bar{\beta}_\lambda}{c_\lambda} J_\lambda(x;q,t) = \frac{\bar{\beta}_\lambda}{c_\lambda} \sum_\mu X^\lambda_\mu \frac{1}{z_\mu} \prod_i (1-t^{\mu_i}) p_\mu(x). \tag{3.8}
$$

Equating coefficients of $p_\mu(x)$ on both sides of (3.7), (3.8), gives

$$
\frac{\bar{\beta}_\lambda}{c_\lambda} X^\lambda_\mu \frac{1}{z_\mu} \prod_i (1-t^{\mu_i}) = \frac{1}{c_\lambda} \sum_\rho \frac{X^\lambda_\rho}{z_\rho} \prod_i (1-t^{\rho_i}) \bar{M}(\rho,\mu). \tag{3.9}
$$

This shows that $h_\lambda(\mu) = \frac{X^\lambda_\mu \prod_i(1-t^{\mu_i})}{z_\mu}$ is a left eigenfunction for $\bar{M}$ with eigenvalue $\bar{\beta}_\lambda$. It follows from reversibility ($\pi_{q,t}(\rho)\bar{M}(\rho,\mu) = \pi_{q,t}(\mu)\bar{M}(\mu,\rho)$) that $h_\lambda(\mu)/\pi_{q,t}(\mu)$ is a right eigenfunction for $\bar{M}$. Since $\pi_{q,t}(\mu) = Zz^{-1}_\mu(q,t)$, simple manipulations give the formulae of part (3) of Theorem 3.1.

As explained in Section 2.1 above, the Macdonald polynomials diagonalize a family of operators $D^r_{q,t}$, $0 \le r \le n$. The argument above applies to all of these.

In essence, the method consists of interpreting equations such as (3.5) as linear combinations of partitions, equating p_λ with λ.

3.3. *Proof of Theorem* 3.1. As in Section 2.1 above, let $D_{q,t}(z) = \sum_{r=0}^{n} D_{q,t}^r \times z^r$. Let $[n] = \sum_{i=1}^{n} t^{n-i}$. The main result identifies $D_{q,t}^1$, operating on the power sums, as an affine transformation of the auxiliary variables Markov chain. The following proposition is the first step, providing the expansion of $D_{q,t}^1$ acting on power sums. A related computation is in [7], Appendix B, Proposition 2.

PROPOSITION 3.2. (a) *If f is homogeneous*, *then*

$$D_{q,t}^0 f = f, \qquad D_{q,t}^n f = q^{\deg(f)} f$$

and

$$D_{q,t}^{n-1} f = t^{\deg(f)+n(n-1)/2} q^{\deg(f)} D_{q^{-1},t^{-1}}^1 f.$$

(b) *If $\lambda = (\lambda_1, \dots, \lambda_\ell)$ is a partition*, *then*

$$\begin{aligned} D_{q,t}^1 p_\lambda = [n] p_\lambda + \sum_{\substack{J \subseteq \{1,\dots,\ell\} \\ J \neq \varnothing}} p_{\lambda_{J^c}} \Biggl(\prod_{k \in J} (q^{\lambda_k} - 1) \Biggr) \\ \times \frac{t^n}{t-1} \sum_{\mu \vdash |\lambda_J|} \Biggl(\prod_{m=1}^{\ell(\mu)} (1 - t^{-\mu_m}) \Biggr) \frac{1}{z_\mu} p_\mu. \end{aligned} \tag{3.10}$$

PROOF. (a) If f is homogeneous, then

$$D_{q,t}^n f = \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| = n}} A_I(x;t) \prod_{i \in I} T_{q,x_i} f = T_{q,x_1} T_{q,x_2} \cdots T_{q,x_n} f = q^{\deg(f)} f. \tag{3.11}$$

By definition,

$$A_I(x;t) = \frac{1}{a_\delta} \Biggl(\prod_{i \in I} T_{t,x_i} \Biggr) a_\delta = t^{r(r-1)/2} \prod_{\substack{i \in I \\ j \notin I}} \frac{t x_i - x_j}{x_i - x_j}. \tag{3.12}$$

Letting $x^\gamma = x_1^{\gamma_1} \cdots x_n^{\gamma_n}$ for $\gamma = (\gamma_1, \dots, \gamma_n)$,

$$\begin{aligned} T_{q,x_1} T_{q,x_2} \cdots \hat{T}_{q,x_j} \cdots T_{q,x_n} x^\gamma &= q^{\gamma_1 + \cdots + \gamma_n - \gamma_j} x_1^{\gamma_1} \cdots x_n^{\gamma_n} \\ &= q^{\deg(x^\gamma)} q^{-\gamma_j} x^\gamma \\ &= q^{\deg(x^\gamma)} T_{q^{-1},x_j} x^\gamma, \end{aligned}$$

and it follows that

$$T_{q,x_1} T_{q,x_2} \cdots \hat{T}_{q,x_j} \cdots T_{q,x_n} f = q^{\deg(f)} T_{q^{-1},x_j} f, \tag{3.13}$$

if f is homogeneous. Thus,

$$\begin{aligned}
D_{q,t}^{n-1} f &= \sum_{\substack{I\subseteq\{1,\dots,n\}\\ |I|=n-1}} A_I(x;t)\Big(\prod_{i\in I} T_{q,x_i}\Big) f \\
&= \sum_{j=1}^{n} A_{\{j\}^c}(x;t) T_{q,x_1}\cdots \hat{T}_{q,x_j}\cdots T_{q,x_n} f \\
&= \sum_{j=1}^{n} \frac{1}{a_\delta} T_{t,x_1}\cdots \hat{T}_{t,x_j}\cdots T_{t,x_n} a_\delta T_{q,x_1}\cdots \hat{T}_{q,x_j}\cdots T_{q,x_n} f \\
&= \sum_{j=1}^{n} \frac{1}{a_\delta} t^{\deg(f)+\deg(a_\delta)} T_{t^{-1},x_j} a_\delta q^{\deg(f)} T_{q^{-1},x_j} f \\
&= t^{\deg(f)+n(n-1)/2} q^{\deg(f)} \sum_{j=1}^{n} A_j(x;t^{-1}) T_{q^{-1},x_j} f \\
&= t^{\deg(f)+n(n-1)/2} q^{\deg(f)} D^1_{q^{-1},t^{-1}} f.
\end{aligned}$$

Hence,

$$D_{q,t}^{n-1} f = t^{\deg(f)+n(n-1)/2} q^{\deg(f)} D^1_{q^{-1},t^{-1}} f. \tag{3.14}$$

(b) By [42], Chapter VI, (3.7), (3.8),

$$\begin{aligned}
D_{1,t}(z) m_\lambda &= \sum_{\beta\in S_n\lambda} \Big(\prod_{i=1}^{n} (1+zt^{n-i})\Big) s_\beta = \Big(\prod_{i=1}^{n} (1+zt^{n-i})\Big) \sum_{\beta\in S_n\lambda} s_\beta \\
&= \Big(\prod_{i=1}^{n} (1+zt^{n-i})\Big) m_\lambda = \sum_{r=0}^{n} t^{r(r-1)/2} \begin{bmatrix} n \\ r \end{bmatrix} z^r m_\lambda,
\end{aligned}$$

where m_λ denotes the monomial symmetric function. Thus, since $D_{q,t}(z) = \sum_{r=0}^{n} D_{q,t}^r z^r$ and

$$D_{1,t}^r = \sum_{\substack{I\subseteq\{1,\dots,n\}\\ |I|=r}} A_I(x;t) \prod_{i\in I} T_{1,x_i} = \sum_{\substack{I\subseteq\{1,\dots,n\}\\ |I|=r}} A_I(x;t),$$

it follows that

$$\sum_{j=1}^{n} A_j(x;t) f = D^1_{1,t} f = [n] f \tag{3.15}$$

for a symmetric function f. By [42], Chapter VI, Section 3, Example 2,

$$(t-1) \sum_{i=1}^{n} A_i(x;t) x_i^r = t^n g_r(x;0,t^{-1}) - \delta_{0r}, \tag{3.16}$$

where, from [42], Chapter VI, (2.9),

$$g_r(x;q,t)=\sum_{\lambda\vdash n} z_\lambda(q,t)^{-1}p_\lambda(x)$$

with $z_\lambda(q,t)$ as in (1.1).

Let $\lambda=(\lambda_1,\dots,\lambda_\ell)$ be a partition, and

$$\text{for } J\subseteq\{1,\dots,\ell\}\qquad \text{let } \lambda_J=(\lambda_{j_1},\dots,\lambda_{j_k})\qquad \text{if } J=\{j_1,\dots,j_k\}.$$

Then, using

$$T_{q,x_i}p_r=q^rx_i^r-x_i^r+p_r=(q^r-1)x_i^r+p_r, \tag{3.17}$$

(3.16) and (3.15),

$$\begin{aligned}
D^1_{q,t}p_\lambda &= \sum_{j=1}^n A_j(x;t)T_{q,x_j}p_{\lambda_1}\cdots p_{\lambda_\ell}\\
&= \sum_{j=1}^n A_j(x;t)\bigl((q^{\lambda_1}-1)x_j^{\lambda_1}+p_{\lambda_1}\bigr)\cdots\bigl((q^{\lambda_\ell}-1)x_j^{\lambda_\ell}+p_{\lambda_\ell}\bigr)\\
&= \sum_{j=1}^n A_j(x;t)\sum_{J\subseteq\{1,\dots,\ell\}}\Bigl(\prod_{k\in J}(q^{\lambda_k}-1)\Bigr)x_j^{|\lambda_J|}\prod_{s\notin J}p_{\lambda_s}\\
&= \sum_{J\subseteq\{1,\dots,\ell\}}\prod_{s\notin J}p_{\lambda_s}\Bigl(\prod_{k\in J}(q^{\lambda_k}-1)\Bigr)\sum_{j=1}^n A_j(x;t)x_j^{|\lambda_J|}\\
&= \sum_{j=1} A_j(x;t)p_\lambda+\sum_{\substack{J\subseteq\{1,\dots,\ell\}\\ J\neq\varnothing}} p_{\lambda_{J^c}}\Bigl(\prod_{k\in J}(q^{\lambda_k}-1)\Bigr)\frac{t^n}{t-1}g_{|\lambda_J|}(x;0,t^{-1})\\
&= [n]p_\lambda+\sum_{\substack{J\subseteq\{1,\dots,\ell\}\\ J\neq\varnothing}} p_{\lambda_{J^c}}\Bigl(\prod_{k\in J}(q^{\lambda_k}-1)\Bigr)\frac{t^n}{t-1}\sum_{\mu\vdash|\lambda_J|}\frac{1}{z_\mu(0;t^{-1})}p_\mu\\
&= [n]p_\lambda+\sum_{\substack{J\subseteq\{1,\dots,\ell\}\\ J\neq\varnothing}} p_{\lambda_{J^c}}\Bigl(\prod_{k\in J}(q^{\lambda_k}-1)\Bigr)\\
&\qquad\qquad\times\frac{t^n}{t-1}\sum_{\mu\vdash|\lambda_J|}\Bigl(\prod_{m=1}^{\ell(\mu)}(1-t^{-\mu_m})\Bigr)\frac{1}{z_\mu}p_\mu. \qquad\square
\end{aligned}$$

Let us show that the measure $\pi_{q,t}(\lambda)$ is properly normalized and compute the normalization of the eigenvectors.

LEMMA 3.3. *Let $\pi_{q,t}(\lambda)$ be as in (1.2) and let $f_\lambda(\rho) = X^\lambda_\rho \prod_{i=1}^{\ell(\rho)}(1-q^{\rho_i})$ be as in Theorem 3.1(3). Then*

$$\sum_{\lambda\vdash k}\pi_{q,t}(\lambda) = 1 \quad \textit{and} \quad \sum_{\rho\vdash k} f_\lambda^2(\rho)\pi_{q,t}(\lambda) = \frac{(q,q)_k}{(t,q)_k}c_\lambda c'_\lambda.$$

PROOF. From [42], Chapter VI, (2.9), (4.9), the Macdonald polynomial $P_{(k)}(x;q,t)$ can be written

$$P_{(k)} = \frac{(q,q)_k}{(t,q)_k}\cdot g_k = \frac{(q,q)_k}{(t,q)_k}\sum_{\lambda\vdash k} z_\lambda(q,t)^{-1}p_\lambda.$$

From [42], Chapter VI, (4.11), (6.19),

$$\langle P_\lambda, P_\lambda\rangle = c'_\lambda/c_\lambda,$$

and it follows that

$$\sum_{\lambda\vdash k}\pi_{q,t}(\lambda) = \frac{(q,q)_k}{(t,q)_k}\sum_{\lambda\vdash k} z_\lambda(q,t)^{-1} = \frac{(t,q)_k}{(q,q)_k}\langle P_{(k)}, P_{(k)}\rangle = \frac{(t,q)_k}{(q,q)_k}\frac{c'_{(k)}}{c_{(k)}} = 1.$$

To get the normalization of $f_\lambda(\rho) = X^\lambda_\rho \prod_{i=1}^{\ell(\rho)}(1-q^{\rho_i})$ in Theorem 3.1(4), use (3.5) and

$$\begin{aligned}
c_\lambda c'_\lambda &= (c_\lambda)^2\langle P_\lambda, P_\lambda\rangle = \langle J_\lambda, J_\lambda\rangle \\
&= \sum_{\rho\vdash k} z_\rho^{-2}\left(X^\lambda_\rho(q,t)\prod_{i=1}^{\ell(\rho)}(1-t^{\rho_i})\right)^2\langle p_\rho, p_\rho\rangle \\
&= \sum_{\rho\vdash k} z_\rho^{-1}\left(X^\lambda_\rho(q,t)\prod(1-t^{\rho_i})\right)^2\prod_{i=1}^{\ell(\rho)}\frac{(1-q^{\rho_i})}{(1-t^{\rho_i})} \\
&= \sum_{\rho\vdash k} f_\lambda^2(\rho) z_\rho^{-1}(q,t) = \frac{(t,q)_k}{(q,q)_k}\sum_{\rho\vdash k} f_\lambda^2(\rho)\pi_{q,t}(\lambda).
\end{aligned}$$

□

We next show that an affine renormalization of the discrete version $\bar{M}$ (3.5) of the Macdonald operator equals the auxiliary variables Markov chain of Section 2.3. Along with Macdonald [42], Chapter VI, (4.1), define

$$E_k = t^{-k}D^1_{q,t} - \sum_{i=1}^{k} t^{-i} \qquad \text{and let } \tilde{E}_k = \frac{t}{q^k-1}E_k,$$

operating on Λ^k_n. From (3.3), the eigenvalues of E_k are $\beta_\lambda = \sum_{i=1}^{\ell(\lambda)}(q^{\lambda_i}-1)t^{-i}$. Noting that $\beta_{(k)} = \frac{q^k-1}{t}$, the operator $\tilde{E}_k$ is a normalization of E_k with top eigenvalue 1. From Proposition 3.2(b), for λ a partition with ℓ parts,

$$\tilde{E}_k p_\lambda = \frac{1}{(1-t^{-1})(q^k-1)} \sum_{\substack{J\subseteq\{1,\ldots,\ell\}\\ J\neq\varnothing}} \prod_{k\in J}(q^{\lambda_k}-1)p_{\lambda_{J^c}} \sum_{\mu\vdash|\lambda_J|}\prod_{i=1}^{\ell(\mu)}(1-t^{-k})\frac{p_\mu}{z_\mu}.$$

Using p_λ as a surrogate for λ as in Section 3.2, the coefficient of $\nu = \lambda_{J^c}\mu$ is exactly $M(\lambda,\nu)$ of Section 2.3.

This completes the proof of Theorem 3.1.

EXAMPLE 3. When $k=2$, from the definitions

$$\tilde{E}_2 p_2 = \left(\frac{1-t^{-1}}{2}\right)p_{1^2} + \left(\frac{1+t^{-1}}{2}\right)p_2,$$

$$\tilde{E}_2 p_{1^2} = \frac{1}{2(q+1)}\big((q+3-qt^{-1}+t^{-1})p_{1^2} + (q-1)(1+t^{-1})p_2\big).$$

Thus, on partitions of 2, the matrix of $\tilde{E}_2$ is the matrix of (2.15), derived there from the probabilistic description.

4. Jack polynomials and Hanlon's walk. The Jack polynomials are a one-parameter family of bases for the symmetric polynomials, orthogonal for the weight $\langle p_\lambda, p_\mu\rangle_\alpha = \alpha^{\ell(\lambda)} z_\lambda \delta_{\lambda\mu}$. They are an important precursor to the full two-parameter Macdonald polynomial theory, containing several classical bases: the limits $\alpha=0, \alpha=\infty$, suitably interpreted, give the $\{e_\lambda\}, \{m_\lambda\}$ bases; $\alpha=1$ gives Schur functions; $\alpha=2$ gives zonal polynomials for GL_n/O_n; $\alpha=\frac{1}{2}$ gives zonal polynomials for $GL_n(\mathbb{H})/U_n(\mathbb{H})$ where $\mathbb{H}$ is the quaternions (see [42], Chapter VII). A good deal of the combinatorial theory for Macdonald polynomials was first developed in the Jack case. Further, the Jack theory has been developed in more detail [34, 39, 53] and [42], Chapter VI, Section 10.

Hanlon [34] managed to interpret the differential operators defining the Jack polynomials as the transition matrix of a Markov chain on partitions with stationary distribution $\pi_\alpha(\lambda) = Z\alpha^{-\ell(\lambda)}/z_\lambda$, described in Section 2.4.2 above. In later work [16], this Markov chain was recognized as the Metropolis algorithm for generating π_α from the proposal of random transpositions. This gives one of the few cases where this important algorithm can be fully diagonalized. See [36] for a different perspective.

Our original aim was to extend Hanlon's findings, adding a second "sufficient statistic" to $\ell(\lambda)$, and discovering a Metropolis-type Markov chain with the Macdonald coefficients as eigenfunctions. It did not work out this way. The auxiliary variables Markov chain makes more vigorous moves than transpositions, and there is no Metropolis step. Nevertheless, as shown below, Hanlon's chain follows from interpreting a limiting case of D_α^1, one of Macdonald's D_α^r operators. We believe that all of the operators D^r_α should have interesting interpretations. In this section, we indicate how to derive Hanlon's chain from the Macdonald operator perspective.

Overview. There are several closely-related operators used to develop the Jack theory. Macdonald [42], Chapter VI, Section 3, Example 3, uses $D_\alpha(u)$ and D^r_α, defined by

$$D_\alpha(u) = \sum_{r=0}^{n} D^r_\alpha u^{n-r} = \frac{1}{a_\delta}\sum_{w\in S_n}\det(w)x^{w\delta}\prod_{i=1}^{n}\left(u+(w\delta)_i+\alpha x_i\frac{\partial}{\partial x_i}\right), \tag{4.1}$$

where $\delta = (n-1, n-2, \ldots, 1, 0)$, a_δ is the Vandermonde determinant, and $x^\gamma = x_1^{\gamma_1}\cdots x_n^{\gamma_n}$ for $\gamma = (\gamma_1, \ldots, \gamma_n)$. He shows [42], Chapter VI, Section 3, Example 3c, that

$$D_\alpha(u) = \lim_{t\to 1}\frac{z^n}{(t-1)^n}D_{t^\alpha,t}(z^{-1}) \qquad \text{if } z = (t-1)u-1, \tag{4.2}$$

so that the Jack operators are a limiting case of Macdonald polynomials. Macdonald [42], Chapter VI, Section 4, Example 2b, shows that the Jack polynomials J^α_λ are eigenfunctions of $D_\alpha(u)$ with eigenvalues $\beta_\lambda(\alpha) = \prod_{i=1}^{n}(u+n-i+\alpha\lambda_i)$.

Stanley [53], proof of Theorem 3.1, and Hanlon [34], (3.5), use $D(\alpha)$ defined as follows. Let

$$\partial_i = \frac{\partial}{\partial x_i}, \qquad U_n = \frac{1}{2}\sum_{i=1}^{n}x_i^2\partial_i^2, \qquad V_n = \sum_{i\neq j}\frac{x_i^2}{x_i-x_j}\partial_i \quad \text{and} \tag{4.3}$$

$$D(\alpha) = \alpha U_n + V_n. \tag{4.4}$$

Hanlon computes the action of $D(\alpha)$ on the power sums in the form [see (4.7)]

$$D(\alpha)p_\lambda = (n-1)rp_\lambda + \alpha\binom{r}{2}\sum_{\mu}\ell_{\mu\lambda}(\alpha)p_\mu, \tag{4.5}$$

where n is the number of variables, and λ is a partition of r.

The matrix $\ell_{\mu\lambda}(\alpha)$ can be interpreted as the transition matrix of the following Markov chain on the symmetric group S_r. For $w \in S_r$, set $c(w) = \#$ cycles. If the chain is currently at w_1, pick a transposition (i, j) uniformly; set $w_2 = w_1(i, j)$. If $c(w_2) = c(w_1)+1$, move to w_2. If $c(w_2) = c(w_1)-1$, move to w_2 with probability $1/\alpha$; else stay at w_1. This Markov chain has transition matrix

$$H_\alpha(w_1, w_2) = \begin{cases} \dfrac{1}{\binom{r}{2}}, & \text{if } w_2 = w_1(i, j) \text{ and } c(w_2) = c(w_1)+1,\\[2ex] \dfrac{1}{\alpha\binom{r}{2}}, & \text{if } w_2 = w_1(i, j) \text{ and } c(w_2) = c(w_1)-1,\\[2ex] \dfrac{n(w_1)(1-\alpha^{-1})}{\binom{r}{2}}, & \text{if } w_1 = w_2,\end{cases}$$

where $n(w_1) = \sum_i (i-1)\lambda_i$ for w_1 of cycle type λ. Hanlon notes that this chain only depends on the conjugacy class of w_1, and the induced process on conjugacy classes is still a Markov chain for which the transition matrix is the matrix of $\ell_{\mu\lambda}(\alpha)$ of (4.5).

The Jack polynomial theory now gives the eigenvalues of the Markov chain $H_\alpha(w_1, w_2)$, and shows that the corresponding eigenvectors are the coefficients when the Jack polynomials are expanded in the power sum basis. The formulas available for Jack polynomials then allow for a careful analysis of rates of convergence to stationarity; see [16, 37].

We may see this from the present perspective as follows.

PROPOSITION 4.1. *Let $D_\alpha(u)$ and $D(\alpha)$ be defined by* (4.1), (4.4).

(a) *Let D_α^t be the coefficient of u^{n-t} in $D_\alpha(u)$* (*see* [42], *Chapter* VI, *Section* 3, *Example* 3d). *If f is a homogeneous polynomial in $x_1, \dots, x_n$ of degree r, then*

$$\begin{aligned} &D_\alpha^0 f = f, \qquad D_\alpha^1(f) = \bigl(\alpha r + \tfrac{1}{2}n(n-1)\bigr) f \quad \textit{and} \\ &D_\alpha^2 f = (-\alpha^2 U_n - \alpha V_n + c_n) f, \end{aligned} \tag{4.6}$$

where

$$c_n = \tfrac{1}{2}\alpha^2 r(r-1) + \tfrac{1}{2}\alpha r n(n-1) + \tfrac{1}{24}n(n-1)(n-2)(3n-1).$$

(b) *From* [53], *proof of Theorem* 3.1,

$$\begin{aligned} D(\alpha) p_\lambda = \frac{1}{2} p_\lambda \Biggl(&\sum_{k=1}^{s} \alpha \lambda_k(\lambda_k - 1) + \alpha \sum_{\substack{j,k=1 \\ j \neq k}}^{s} \frac{\lambda_j \lambda_k p_{\lambda_j + \lambda_k}}{p_{\lambda_j} p_{\lambda_k}} \\ &+ \sum_{k=1}^{s} \lambda_k (2n - \lambda_k - 1) + \sum_{k=1}^{s} \frac{\lambda_k}{p_{\lambda_k}} \sum_{m=1}^{\lambda_k - 1} p_{\lambda_k - m} p_m \Biggr). \end{aligned} \tag{4.7}$$

From part (a), up to affine rescaling, D_α^2 is the Stanley–Hanlon operator. From part (b), this operates on the power sums in precisely the way that the Metropolis algorithm operates. Indeed, multiplying a permutation w by a transposition (i, j) changes the number of cycles by one; the change takes place by fusing two cycles [the first term in (4.7)] or by breaking one of the cycles in w into parts [the second term in (4.7)]. The final term constitutes the "holding" probability from the Metropolis algorithm. The proof of Proposition 4.1 is a lengthy but straightforward computation. See [20] for the details.

5. Rates of convergence. This section uses the eigenvectors and eigenvalues derived above to give rates of convergence for the auxiliary variables Markov chain. Section 5.1 states the main results: starting from the partition (k) a bounded number of steps suffice for convergence, independent of k. Section 5.2 contains an overview of the argument and needed lemmas. Section 5.3 gives the proof of Theorem 5.1, and Section 5.4 develops the analysis starting from (1^k), showing that $\log_q(k)$ steps are needed.

5.1. *Statement of main results.* Fix $q, t > 1$ and $k \geq 2$. Let $\mathcal{P}_k$ be the partitions of k, $\pi_{q,t}(\lambda) = Z/z_\lambda(q,t)$ the stationary distribution defined in (1.2), and $M(\lambda, \nu)$ the auxiliary variables Markov chain defined in Proposition 2.1. The total variation distance $\|M^\ell_{(k)} - \pi_{q,t}\|_{\mathrm{TV}}$ used below is defined in (2.7).

THEOREM 5.1. *Consider the auxiliary variables Markov chain on partitions of $k \geq 4$. Then, for all $\ell \geq 2$,*

$$4\|M^\ell_{(k)} - \pi_{q,t}\|^2_{\mathrm{TV}} \leq \frac{1}{(1-q^{-1})^{3/2}(1-q^{-2})^2}\left(\frac{1}{q} + \frac{1}{tq^{k/2}}\right)^{2\ell} + k\left(\frac{t}{t-1}\right)\left(\frac{2}{q^{k/4}}\right)^{2\ell}. \tag{5.1}$$

For example, if $q = 4$, $t = 2$ and $k = 10$ the bound becomes $1.76(0.26)^{2\ell} + 20(1/512)^{2\ell}$. Thus, when $\ell = 2$ the total variation distance is at most 0.05 in this example.

5.2. *Outline of proof and basic lemmas.* Let $\{f_\lambda, \beta_\lambda\}_{\lambda \vdash k}$ be the eigenfunctions and eigenvalues of M given in Theorem 3.1. From Section 2.2, for any starting state ρ,

$$4\|M^\ell_\rho - \pi_{q,t}\|^2_{\mathrm{TV}} \leq \sum_\lambda \frac{(M^\ell(\rho,\lambda) - \pi_{q,t}(\lambda))^2}{\pi_{q,t}(\lambda)} = \sum_{\lambda \neq (k)} \bar{f}^2_\lambda(\rho)\beta^{2\ell}_\lambda \tag{5.2}$$

with $\bar{f}_\lambda$ right eigenfunctions normalized to have norm one. At the end of this subsection we prove the following:

$$\sum_\lambda \bar{f}^2_\lambda(\rho) = \frac{1}{\pi_{q,t}(\rho)} \qquad \text{for any } \rho \in \mathcal{P}_k, \tag{5.3}$$

$$\left(\frac{1-t^{-k}}{1-t^{-1}}\right)\frac{1}{k\pi_{q,t}(k)} \text{ is an increasing sequence bounded by } (1-q^{-1})^{-1/2}, \tag{5.4}$$

(5.5) β_λ is monotone increasing in the usual partial order (moving up boxes); in particular, $\beta_{k-1,1}$ is the second largest eigenvalue and all $\beta_\lambda > 0$,

$$\beta_{k-r,r} \sim \frac{2}{q^r}. \tag{5.6}$$

Using these results, consider the sum on the right-hand side of (5.2), for λ with largest part λ_1 less than $k-r$. Using monotonicity, (5.5) and the bound (5.3),

$$\sum_{\lambda : \lambda_1 \leq k-r} \bar{f}_\lambda^2(k)\beta_\lambda^{2\ell} \leq \left(\frac{2}{q^r}\right)^\ell \pi_{q,t}^{-1}(k) \leq \frac{t}{t-1}\left(\frac{2}{q^r}\right)^\ell k. \tag{5.7}$$

By taking $r = k/4$ gives the second term on the right-hand side of (5.1).

Using monotonicity again,

$$\sum_{\substack{\lambda \neq (k) \\ \lambda_1 > k-j^*}} \bar{f}_\lambda^2(k)\beta_\lambda^{2\ell} \leq \sum_{r=1}^{j^*} \beta_{(k-r,r)}^{2\ell} \sum_{\gamma \vdash r} \bar{f}_{(k-r,\gamma)}^2(k). \tag{5.8}$$

The argument proceeds by looking carefully at $\bar{f}_\lambda^2$ and showing

$$\bar{f}_{(k-r,\gamma)}^2(k) \leq c\bar{f}_\gamma^2(r) \tag{5.9}$$

for a constant c. In (5.9) and throughout this section, $c = c(q,t)$ denotes a positive constant which depends only on q and t, but not on k. Its value may change from line to line. Using (5.3) on $\mathcal{P}_r$ shows $\sum_{\lambda' \vdash r} \bar{f}_{\lambda'}^2(r) = \pi_{q,t}^{-1}(r) \sim cr$. Using this and (5.6) in (5.8) gives an upper bound

$$\sum_{\substack{\lambda \neq (k) \\ \lambda_1 \geq k-j^*}} \bar{f}_\lambda^2(\rho)\beta_\lambda^{2\ell} \leq c\sum_{r=1}^{j^*}\left(\frac{2}{q^r}\right)^\ell r. \tag{5.10}$$

This completes the outline for starting state (k).

This section concludes by proving the preliminary results announced above.

LEMMA 5.2. *For any $\rho \in \mathcal{P}_k$, the normalized eigenfunctions $\bar{f}_\lambda(\rho)$ satisfy*

$$\sum_{\lambda \vdash k} \bar{f}_\lambda(\rho)^2 = \frac{1}{\pi_{q,t}(\rho)}.$$

PROOF. The $\{\bar{f}_\lambda\}$ are orthonormal in $L^2(\pi_{q,t})$. Fix $\rho \in \mathcal{P}_k$, and let $\delta_\rho(\nu) = \delta_{\rho\nu}$ be the measure concentrated at ρ. Expand the function $g(\nu) = \delta_\rho(\nu)/\pi_{q,t}(\rho)$ in this basis: $g(\nu) = \sum_\lambda \langle g|\bar{f}_\lambda\rangle \bar{f}_\lambda(\nu)$. Using the Plancherel identity, $\sum g(\nu)^2 \times \pi_{q,t}(\nu) = \sum_\lambda \langle g|\bar{f}_\lambda\rangle^2$. Here, the left-hand side equals $\pi_{q,t}^{-1}(\rho)$ and $\langle g|\bar{f}_\lambda\rangle = \sum_\nu g(\nu)\bar{f}_\lambda(\nu)\pi_{q,t}(\nu) = \bar{f}_\lambda(\rho)$. So the right-hand side is the needed sum of squares. $\square$

The asymptotics in (5.4) follow from the following lemma.

LEMMA 5.3. *For $q,t>1$, the sequence*

$$P_k = \left(\frac{1-t^{-k}}{1-t^{-1}}\right)\frac{1}{k\pi_{q,t}(k)} = \frac{(1-q^{-k})}{(1-t^{-1})}\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}$$
$$= \prod_{j=1}^{k-1}\frac{1-t^{-1}q^{-j}}{1-q^{-j}}$$

is increasing and bounded by $\frac{1}{\sqrt{1-q^{-1}}}$.

PROOF. The equalities follow from the definitions of $\pi_{q,t}(\lambda)$, $(t,q)_k$ and $(q,q)_k$. Since $\frac{1-t^{-1}q^{-k}}{1-q^{-k}} > 1$, the sequence is increasing. The bound follows from

$$\begin{aligned}
\prod_{j=1}^{\infty}\frac{1-t^{-1}q^{-j}}{1-q^{-j}} &= \exp\left(\sum_{j=1}^{\infty}\log(1-t^{-1}q^{-j}) - \log(1-q^{-j})\right)\\
&= \exp\left(\sum_{j=1}^{\infty}\sum_{n=1}^{\infty}\left(\frac{q^{-jn}}{n} - \frac{t^{-n}q^{-jn}}{n}\right)\right)\\
&= \exp\left(\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\frac{q^{-jn}(1-t^{-n})}{n}\right)\\
&= \exp\left(\sum_{n=1}^{\infty}\frac{(1-t^{-n})}{n}\frac{q^{-n}}{1-q^{-n}}\right)\\
&= \exp\left(\sum_{n=1}^{\infty}\frac{(1-t^{-n})}{q^n-1}\frac{1}{n}\right) \le \exp\left(\sum_{n=1}^{\infty}\frac{1}{2q^n n}\right)\\
&= \exp\left(-\frac{1}{2}\log(1-q^{-1})\right) = \frac{1}{\sqrt{1-q^{-1}}}.
\end{aligned}$$

□

REMARK. The function $P_\infty = \lim_{k\to\infty} P_k$ is an analytic function of q,t for $|q|,|t|>1$, thoroughly studied in the classical theory of partitions [4], Section 2.2.

For the next lemma, recall the usual dominance partial order on $\mathcal{P}_k : \lambda \ge \mu$ if $\lambda_1 + \cdots + \lambda_i \ge \mu_1 + \cdots + \mu_i$ for all i [42], Chapter I, Section 1. This amounts to "moving up boxes" in the diagram for μ. Thus (k) is largest, (1^k) smallest. When $k=6$, $(5,1) > (4,2) > (3,3)$, but $(3,3)$ and $(4,1,1)$ are not comparable. The following result shows that the eigenvalues β_λ are monotone in this order. A similar monotonicity holds for the random transpositions chain [21], the Ewens sampling chain [16] and the Hecke algebra deformation chain [19].

LEMMA 5.4. *For $q, t > 1$, the eigenvalues*

$$\beta_\lambda = \frac{t}{q^k - 1} \sum_{j=1}^{\ell(\lambda)} (q^{\lambda_j} - 1) t^{-j}$$

are monotone in λ.

PROOF. Consider first a partition λ, $i < j$, with $a = \lambda_i \geq \lambda_j = b$, where moving one box from row i to row j is allowed. It must be shown that $q^{a+1} t^{-i} + q^{b-1} t^{-j} > q^a t^{-i} + q^b t^{-j}$. Equivalently,

$$q^{a+1} + q^{b-1} t^{-(j-i)} > q^a + q^b t^{-(j-i)}$$

or

$$q^{a+1} t^{j-i} + q^{b-1} > q^a t^{j-i} + q^b$$

or

$$q^a t^{j-i} (q - 1) > q^{b-1} (q - 1).$$

Since $t^{j-i} > 1$ and $q^{a-b+1} > 1$, this always holds. $\square$

By elementary manipulations, $\frac{q^a - 1}{q^b - 1} < \frac{q^a}{q^b} = \frac{1}{q^{b-a}}$ for $1 < a < b$, so that

$$\begin{aligned} \beta_{(k-r,r)} &= \frac{t}{q^k - 1}\left(\frac{q^{k-r} - 1}{t} + \frac{q^r - 1}{t^2}\right) \leq \frac{1}{q^r} + \frac{1}{t q^{k-r}} \\ &= \frac{1}{q^r}\left(1 + \frac{1}{t q^{k-2r}}\right), \end{aligned} \tag{5.11}$$

which establishes (5.6).

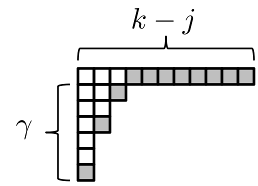
5.3. *Proof of Theorem* 5.1. From Theorem 3.1, the normalized eigenvectors are given by

$$\bar{f}_\lambda(k)^2 = \frac{(X^\lambda_{(k)}(q^k - 1))^2}{c_\lambda c'_\lambda} \cdot \frac{(t, q)_k}{(q, q)_k} \qquad \text{where } X^\lambda_{(k)} = \prod_{\substack{(i,j) \in \lambda \\ (i,j) \neq (1,1)}} (t^{i-1} - q^{j-1}), \tag{5.12}$$

and c_λ and c'_λ are given by (3.1).

LEMMA 5.5. *For* $\lambda = (k - r, \gamma)$, *with* $\gamma \vdash r$ *and* $r \le k/2$,

$$\bar{f}_\lambda(k)^2 \le \bar{f}_\gamma(r)^2 \frac{(1-q^{-k})^2}{(1-q^{-r})^2} \frac{q^k}{t^k} \frac{(t,q)_k}{(q,q)_k} \frac{t^r}{q^r} \frac{(q,q)_r}{(t,q)_r},$$



PROOF. Let $\lambda = (k-r, \gamma)$ with $\gamma \vdash r$ and $r \le k/2$. Let U be the boxes in the first row of λ, and let L be the shaded boxes in the figure above.

For a box s in λ, let $i(s)$ be the row number and $j(s)$ the column number of s. Then

$$\begin{aligned}
\left(\frac{X^\lambda_{(k)}}{X^\gamma_{(r)}}\right)^2 &= \frac{\prod_{(i,j)\in\lambda,(i,j)\ne(1,1)}(t^{i-1}-q^{j-1})^2}{\prod_{(i,j)\in\gamma,(i,j)\ne(1,1)}(t^{i-1}-q^{j-1})^2} \\
&= \prod_{s\in L}\bigl(t^{i(s)-1}-q^{j(s)-1}\bigr)^2 = \prod_{s\in U}\bigl(t^{l(s)}-q^{j(s)-1}\bigr)^2 \\
&= \prod_{m=1}^{\gamma_1}(t^{\gamma'_m}-q^{m-1})^2 \prod_{m=\gamma_1+1}^{k-r}(1-q^{m-1})^2,
\end{aligned}$$

where γ'_m is the length of the mth column of γ. Next,

$$\begin{aligned}
\frac{c_\lambda c'_\lambda}{c_\gamma c'_\gamma} &= \frac{\prod_{s\in\lambda}(1-q^{a(s)}t^{l(s)+1})(1-q^{a(s)+1}t^{l(s)})}{\prod_{s\in\gamma}(1-q^{a(s)}t^{l(s)+1})(1-q^{a(s)+1}t^{l(s)})} \\
&= \prod_{s\in U}\bigl(1-q^{a(s)}t^{l(s)+1}\bigr)\bigl(1-q^{a(s)+1}t^{l(s)}\bigr) \\
&= \prod_{m=1}^{\gamma_1}(1-q^{k-r-m}t^{\gamma'_m+1})(1-q^{k-r-m+1}t^{\gamma'_m}) \\
&\quad\times \prod_{m=\gamma_1+1}^{k-r}(1-q^{k-r-m}t)(1-q^{k-r-m+1}) \\
&= q^{-2(k-r)(k-r-1)} \prod_{m=1}^{\gamma_1}(t^{\gamma'_m+1}q^{k-r-1}-q^{m-1})(t^{\gamma'_m}q^{k-r}-q^{m-1}) \\
&\quad\times \prod_{m=\gamma_1+1}^{k-r}(tq^{k-r-1}-q^{m-1})(q^{k-r}-q^{m-1}).
\end{aligned}$$

Thus,

$$\left(\frac{X^{\lambda}_{(k)}}{X^{\gamma}_{(r)}}\right)^2 \frac{c_\gamma c'_\gamma}{c_\lambda c'_\lambda} = q^{2(k-r)(k-r-1)} \prod_{m=1}^{\gamma_1} \frac{(t^{\gamma'_m} - q^{m-1})}{(t^{\gamma'_m+1} q^{k-r-1} - q^{m-1})} \frac{(t^{\gamma'_m} - q^{m-1})}{(t^{\gamma'_m} q^{k-r} - q^{m-1})}$$
$$\times \prod_{m=\gamma_1+1}^{k-r} \frac{(1-q^{m-1})}{(tq^{k-r-1} - q^{m-1})} \frac{(1-q^{m-1})}{(q^{k-r} - q^{m-1})}.$$

Since $k-r-1 \geq m-1$ and $t > 1$, then $t^{\gamma'_m+1} q^{k-r-1} - q^{m-1} > 0$, so that $q^{-(k-r-1)} t^{-1} < 1$ implies

$$\frac{(t^{\gamma'_m} - q^{m-1})}{(t^{\gamma'_m+1} q^{k-r-1} - q^{m-1})} < q^{-(k-r-1)} t^{-1}.$$

Similarly, since $k-r > m-1$ and $t > 1$, then $t^{\gamma'_m} q^{k-r} - q^{m-1} > 0$, so that $q^{-(k-r)} < 1$ implies

$$\frac{(t^{\gamma'_m} - q^{m-1})}{(t^{\gamma'_m} q^{k-r} - q^{m-1})} < q^{-(k-r)}.$$

Similarly, $t^{-1} q^{-(k-r-1)}$ and $q^{-(k-r)} < 1$ imply

$$\frac{(1-q^{m-1})}{(tq^{k-r-1} - q^{m-1})} < t^{-1} q^{-(k-r-1)} \quad \text{and} \quad \frac{(1-q^{m-1})}{(q^{k-r} - q^{m-1})} < q^{-(k-r)}.$$

So

$$\begin{aligned}\left(\frac{X^{\lambda}_{(k)}}{X^{\gamma}_{(r)}}\right)^2 \frac{c_\gamma c'_\gamma}{c_\lambda c'_\lambda} &\leq q^{2(k-r)(k-r-1)} \prod_{m=1}^{\gamma_1} \big(q^{-(k-r-1)} t^{-1}\big)\big(q^{-(k-r)}\big) \\ &\quad \times \prod_{m=\gamma_1+1}^{k-r} \big(t^{-1} q^{-(k-r-1)}\big)\big(q^{-(k-r)}\big) \\ &= q^{2(k-r)(k-r-1)} t^{-(k-r)} q^{-(k-r)^2} q^{-(k-r-1)(k-r)} \\ &= q^{-(k-r)} t^{-(k-r)}.\end{aligned}$$

Thus,

$$\begin{aligned}\frac{\bar{f}_\lambda(k)^2}{\bar{f}_\gamma(r)^2} &= \left(\frac{X^{\lambda}_{(k)}}{X^{\gamma}_{(r)}}\right)^2 \frac{c_\gamma c'_\gamma}{c_\lambda c'_\lambda} \frac{(q^k-1)^2}{(q^r-1)^2} \frac{(t,q)_k}{(q,q)_k} \frac{(q,q)_r}{(t,q)_r} \\ &\leq \frac{1}{q^{k-r} t^{k-r}} \frac{(q^k-1)^2}{(q^r-1)^2} \frac{(t,q)_k}{(q,q)_k} \frac{(q,q)_r}{(t,q)_r}.\end{aligned}$$ □

We may now bound the upper bound sum on the right-hand side of (5.2). Fix $j^* \leq k/2$. Using monotonicity (Lemma 5.4), Lemmas 5.2, 5.3 and the definition of $\pi_{q,t}(r)$ from (1.2),

$$\begin{aligned}
\sum_{\substack{\lambda\neq(k)\\ \lambda_1\geq k-j^*}} \bar{f}_\lambda(k)^2\beta_\lambda^{2\ell} &= \sum_{r=1}^{j^*}\sum_{\lambda=(k-r,\gamma)} \beta_{(k-r,\gamma)}^{2\ell}\bar{f}_\lambda(k)^2\\
&\leq \sum_{r=1}^{j^*}\sum_{\lambda=(k-r,\gamma)} \beta_{(k-r,r)}^{2\ell}\bar{f}_\lambda(k)^2\\
&\leq \sum_{r=1}^{j^*}\beta_{(k-r,r)}^{2\ell}\sum_{\gamma\vdash r}\bar{f}_\gamma(r)^2\frac{(1-q^{-k})^2}{(1-q^{-r})^2}\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}\frac{t^r}{q^r}\frac{(q,q)_r}{(t,q)_r}\\
&\leq \sum_{r=1}^{j^*}\beta_{(k-r,r)}^{2\ell}\frac{1}{\pi_{q,t}(r)}\frac{(1-q^{-k})^2}{(1-q^{-r})^2}\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}\frac{t^r}{q^r}\frac{(q,q)_r}{(t,q)_r}\\
&\leq \sum_{r=1}^{j^*}\beta_{(k-r,r)}^{2\ell} r\frac{q^r}{t^r}\frac{(t,q)_r}{(q,q)_r}\frac{(1-q^{-r})}{(1-t^{-r})}\frac{(1-q^{-k})^2}{(1-q^{-r})^2}\\
&\qquad\times\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}\frac{t^r}{q^r}\frac{(q,q)_r}{(t,q)_r}\\
&\leq (1-q^{-k})^2\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}\sum_{r=1}^{j^*} r\beta_{(k-r,r)}^{2\ell}\frac{1}{(1-q^{-r})(1-t^{-r})}\\
&\leq \frac{(1-q^{-k})^2}{(1-q^{-1})(1-t^{-1})}\frac{q^k}{t^k}\frac{(t,q)_k}{(q,q)_k}\sum_{r=1}^{j^*} r\beta_{(k-r,r)}^{2\ell}.
\end{aligned}$$

Using (5.11) and Lemma 5.3 gives

$$\begin{aligned}
\sum_{\substack{\lambda\neq(k)\\ \lambda_1\geq k-j^*}} \bar{f}_\lambda(k)^2\beta_\lambda^{2\ell} &\leq \frac{(1-q^{-k})}{(1-q^{-1})}\left(\prod_{j=1}^{k-1}\frac{1-t^{-1}q^{-j}}{1-q^{-j}}\right)\sum_{r=1}^{j^*}\frac{r}{q^{2r\ell}}\left(1+\frac{1}{tq^{k-2r}}\right)^{2\ell}\\
&\leq \frac{(1-q^{-k})}{(1-q^{-1})}\left(\prod_{j=1}^{\infty}\frac{1-t^{-1}q^{-j}}{1-q^{-j}}\right)\left(1+\frac{1}{tq^{k-2j^*}}\right)^{2\ell}\sum_{r=1}^{j^*}\frac{r}{q^{2r\ell}}\\
&\leq \frac{(1-q^{-k})}{(1-q^{-1})^{3/2}}\left(1+\frac{1}{tq^{k-2j^*}}\right)^{2\ell}\frac{1}{q^{2\ell}}\left(1-\frac{1}{q^{2\ell}}\right)^{-2}\\
&\leq \frac{(1-q^{-k})}{(1-q^{-1})^{3/2}}\frac{1}{(1-q^{-2})^2}\left(\frac{1}{q}+\frac{1}{tq^{k-2j^*+1}}\right)^{2\ell}
\end{aligned}$$

by Lemma 5.2. Choose j^* (of order $k/4$) so that $k-2j^*+1=k/2$. Then

$$\sum_{\substack{\lambda\neq(k)\\ \lambda_1\geq k-j^*}} \bar{f}_\lambda(k)^2\beta_\lambda^{2\ell} \leq \frac{1}{(1-q^{-1})^{3/2}(1-q^{-2})^2}\left(\frac{1}{q}+\frac{1}{tq^{k/2}}\right)^{2\ell}$$

with a as in the statement of Theorem 5.1.

Now use

$$\begin{aligned}\sum_{\substack{\lambda\\ \lambda_1<3k/4}} \bar{f}_\lambda(k)^2\beta_\lambda^{2\ell} &\leq \sum_{\substack{\lambda\\ \lambda_1<3k/4}} \bar{f}_\lambda(k)^2\beta_{(3k/4,k/4)}^{2\ell}\\ &\leq \sum_{\lambda} \bar{f}_\lambda(k)^2\beta_{(3k/4,k/4)}^{2\ell} \leq \frac{1}{\pi_{q,t}(k)}\beta_{(3k/4,k/4)}^{2\ell}\\ &= \frac{t^k}{q^k}\frac{(q,q)_k}{(t,q)_k}\frac{(1-t^{-k})}{(1-q^{-k})}k\beta_{(3k/4,k/4)}^{2\ell}\\ &\leq k\frac{(1-t^{-k})}{(1-t^{-1})}\left(\prod_{j=1}^{k-1}\frac{1-q^{-j}}{1-t^{-1}q^{-j}}\right)\left(\frac{1}{q^{k/4}}\left(1+\frac{1}{tq^{k/2}}\right)\right)^{2\ell}\end{aligned}$$

so that

$$\sum_{\substack{\lambda\\ \lambda_1<3k/4}} \bar{f}_\lambda(k)^2\beta_\lambda^{2\ell} \leq k\frac{t}{t-1}\left(\frac{2}{q^{k/4}}\right)^{2\ell}. \tag{5.13}$$

This completes the proof of Theorem 5.1.

5.4. *Bounds starting at* (1^k). We have not worked as seriously at bounding the chain starting from the partition (1^k). The following results show that $\log_q(k)$ steps are required, and offer evidence for the conjecture that $\log_q(k)+\theta$ steps suffice [where the distance to stationarity tends to zero with θ, so there is a sharp cutoff at $\log_q(k)$].

The L^2 or chi-square distance on the right-hand side of (5.2) has first term $\beta_{k-1,1}^{2\ell}\bar{f}_{k-1,1}^2(1^k)$.

LEMMA 5.6. *For fixed* $q,t>1$, *as* k *tends to infinity*,

$$\bar{f}_{(k-1,1)}(1^k)^2 = \frac{(X_{(1^k)}^{(k-1,1)}(q-1)^k)^2}{c_{(k-1,1)}c'_{(k-1,1)}(q,q)_k/(t,q)_k} \sim \left(\frac{1-q^{-1}}{1-t^{-1}}\right)k^2.$$

PROOF. From (3.3) and the definition of $\varphi_T(q,t)$ from [42], Chapter VI, page 341, (1), and [42], Chapter VI, (7.11),

$$X_{(1^k)}^{(k-1,1)} = \frac{c'_{(k-1,1)}(q,t)}{(1-t)^k}\sum_T \varphi_T(q,t) = \frac{c'_{(k-1,1)}}{(1-t)^k}\left(\frac{(1-t)^k}{(1-q)^k}p\right) = \frac{c'_{(k-1,1)}}{(1-q)^k}p$$

with

$$p=\left(\frac{(1-t^2)/(1-qt)}{(1-t)/(1-q)}+\frac{(1-qt^2)/(1-q^2t)}{(1-qt)/(1-q^2)}\right.$$
$$\left.+\frac{(1-q^2t^2)/(1-q^3t)}{(1-q^2t)/(1-q^3)}+\cdots+\frac{(1-q^{k-2}t^2)/(1-q^{k-1}t)}{(1-q^{k-2}t)/(1-q^{k-1})}\right).$$

Using the definition of $c_{(k-1,1)}$ and $c'_{(k-1,1)}$ from (3.1), and the definition of $(t,q)_k$ and $(q,q)_k$ from (1.2),

$$\begin{aligned}
&\bar{f}_{(k-1,1)}(1^k)^2\\
&\quad=\frac{(X_{(1^k)}^{(k-1,1)}(1-q)^k)^2(t,q)_k}{c_{(k-1,1)}c'_{(k-1,1)}(q,q)_k}=\frac{c'_{(k-1,1)}p^2}{c_{(k-1,1)}}\frac{(t,q)_k}{(q,q)_k}\\
&\quad=\frac{(t,q)_k}{(q,q)_k}\frac{(1-q)(1-tq^{k-1})(1-q)(1-q^2)\cdots(1-q^{k-2})}{(1-t)(1-t^2q^{k-2})(1-t)(1-tq)\cdots(1-tq^{k-3})}p^2\\
&\quad=\frac{(1-tq^{k-2})(1-tq^{k-1})}{(1-q^{k-1})(1-q^k)}\frac{(1-q)(1-tq^{k-1})}{(1-t)(1-t^2q^{k-2})}p^2\\
&\quad=\frac{(1-t^{-1}q^{-(k-2)})(1-t^{-1}q^{-(k-1)})}{(1-q^{-(k-1)})(1-q^{-k})}\frac{(1-q^{-1})(1-t^{-1}q^{-(k-1)})}{(1-t^{-1})(1-t^{-2}q^{-(k-2)})}p^2,
\end{aligned}$$

and the result follows, since $p\sim k$ for k large. □

COROLLARY 5.7. *There is a constant c such that, for all $k,\ell\geq 2$,*

$$\chi^2_{(1^k)}(\ell)=\sum_{\lambda}\frac{(M^\ell((1^k),\lambda)-\pi_{q,t}(\lambda))^2}{\pi_{q,t}(\lambda)}\geq c\left(\frac{1-q^{-1}}{1-t^{-1}}\right)\frac{k^2}{q^{2\ell}}.$$

PROOF. Using only the lead term in the expression for $\chi^2_{(1^k)}(\ell)$ in (5.4) gives the lower bound $\beta^{2\ell}_{(k-1,1)}\bar{f}^2_{(k-1,1)}(1^k)$. The formula for $\beta_{(k-1,1)}$ in Theorem 3.1(2) gives $\beta_{(k-1,1)}\geq\frac{1}{q}$, and the result then follows from Lemma 5.6. □

The corollary shows that if $\ell=\log_q(k)+\theta$, $\chi^2_{1^k}(\ell)\geq\frac{c}{q^{2\theta}}$. Thus, more than $\log_q(k)$ steps are required to drive the chi-square distance to zero. In many examples, the asymptotics of the lead term in the bound (5.2) sharply controls the behavior of total variation and chi-square convergence. We conjecture this is the case here, and that there is a sharp cut-off at $\log_q(k)$.

It is easy to give a total variation lower bound:

PROPOSITION 5.8. *For the auxiliary variables chain $M(\lambda,\lambda')$, after ℓ steps with $\ell=\log_q(k)+\theta$, for k large and $\theta<-\frac{t-1}{q-1}$,*

$$\|M^\ell_{(1^k)}-\pi_{q,t}\|_{\mathrm{TV}}\geq e^{-(t-1)/(q-1)}-e^{-1/q^\theta}+o(1).$$

PROOF. Consider the Markov chain starting from $\lambda = (1^k)$. At each stage, the algorithm chooses some parts of the current partition to discard, with probability given by (1.3). From the detailed description given in Section 2.4.3, the chance of a specific singleton being eliminated is $1/q$. Of course, in the replacement stage (1.4) this (and more singletons) may reappear. Let T be the first time that all of the original singletons have been removed at least once; this T depends on the history of the entire Markov chain. Then T is distributed as the maximum of k independent geometric random variables $\{X_i\}_{i=1}^k$ with $P(X_i > \ell) = 1/q^\ell$ (here X_i is the first time that the ith singleton is removed).

Let $A = \{\lambda \in \mathcal{P}_k : a_1(\lambda) > 0\}$. From the definition

$$\|M_{(1^k)}^\ell - \pi_{q,t}\|_{\mathrm{TV}} = \max_{B \subseteq \mathcal{P}_k} |M^\ell((1^k), B) - \pi_{q,t}(B)| \geq |M^\ell((1^k), A) - \pi_{q,t}(A)|$$

and

$$\begin{aligned} M^\ell((1^k), A) &\geq P\{T > \ell\} = 1 - P\{T \leq \ell\} \\ &= 1 - P\{\max X_i \leq \ell\} \\ &= 1 - P(X_1 \leq \ell)^k \\ &= 1 - e^{k \log(1 - P(X_1 > \ell))} \\ &= 1 - e^{k \log(1 - 1/q^\ell)} \\ &\sim 1 - e^{-k/q^\ell} \\ &= 1 - e^{-1/q^\theta}. \end{aligned}$$

From the limiting results in Section 2.4.5, under $\pi_{q,t}$, $a_1(\lambda)$ has an approximate Poisson $(\frac{t-1}{q-1})$ distribution. Thus, $\pi_{q,t}(A) \sim 1 - e^{-(t-1)/(q-1)}$. The result follows. $\square$

Acknowledgments. We thank John Jiang and James Zhao for their extensive help in understanding the measures $\pi_{q,t}$. We thank Alexei Borodin for telling us about multiplicative measures and Martha Yip for help and company in working out exercises from [42] over the past several years. We thank Cindy Kirby for her expert help.

REFERENCES

[1] ALDOUS, D. and DIACONIS, P. (1999). Longest increasing subsequences: From patience sorting to the Baik–Deift–Johansson theorem. *Bull. Amer. Math. Soc.* (*N.S.*) **36** 413–432. MR1694204

[2] ALDOUS, D. J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** 3–48. MR1673235

[3] ANDERSEN, H. C. and DIACONIS, P. (2007). Hit and run as a unifying device. *J. Soc. Fr. Stat. & Rev. Stat. Appl.* **148** 5–28. MR2502361

[4] ANDREWS, G. E. (1998). *The Theory of Partitions*. Cambridge Univ. Press, Cambridge. MR1634067

[5] ARRATIA, R., BARBOUR, A. D. and TAVARÉ, S. (2003). *Logarithmic Combinatorial Structures*: *A Probabilistic Approach*. Eur. Math. Soc., Zürich. MR2032426

[6] ASSAF, S. H. (2007). Dual equivalence graphs, ribbon tableaux and Macdonald polynomials. Ph.D. thesis, Dept. Mathematics, Univ. California, Berkeley.

[7] AWATA, H., KUBO, H., ODAKE, S. and SHIRAISHI, J. (1996). Quantum $\mathscr{W}_N$ algebras and Macdonald polynomials. *Comm*. *Math*. *Phys*. **179** 401–416. MR1400746

[8] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge Univ. Press, Cambridge. MR2253162

[9] BETZ, V., UELTSCHI, D. and VELENIK, Y. (2011). Random permutations with cycle weights. *Ann*. *Appl*. *Probab*. **21** 312–331. MR2759204

[10] BILLINGSLEY, P. (1972). On the distribution of large prime divisors. *Period*. *Math*. *Hungar*. **2** 283–289. MR0335462

[11] BORGS, C., CHAYES, J. T., FRIEZE, A., KIM, J. H., TETALI, P., VIGODA, E. and VU, V. H. (1999). Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In 40*th Annual Symposium on Foundations of Computer Science* (*New York*, 1999) 218–229. IEEE Comput. Soc., Los Alamitos, CA. MR1917562

[12] BORODIN, A., OKOUNKOV, A. and OLSHANSKI, G. (2000). Asymptotics of Plancherel measures for symmetric groups. *J*. *Amer*. *Math*. *Soc*. **13** 481–515 (electronic). MR1758751

[13] BRÉMAUD, P. (1999). *Markov Chains*: *Gibbs Fields*, *Monte Carlo Simulation*, *and Queues*. *Texts in Applied Mathematics* **31**. Springer, New York. MR1689633

[14] CECCHERINI-SILBERSTEIN, T., SCARABOTTI, F. and TOLLI, F. (2008). *Harmonic Analysis on Finite Groups*: *Representation Theory*, *Gelfand Pairs and Markov Chains*. *Cambridge Studies in Advanced Mathematics* **108**. Cambridge Univ. Press, Cambridge. MR2389056

[15] CHEREDNIK, I. (1992). Double affine Hecke algebras, Knizhnik–Zamolodchikov equations, and Macdonald's operators. *Int*. *Math*. *Res*. *Not*. *IMRN* **9** 171–180. MR1185831

[16] DIACONIS, P. and HANLON, P. (1992). Eigen-analysis for some examples of the Metropolis algorithm. In *Hypergeometric Functions on Domains of Positivity*, *Jack Polynomials*, *and Applications* (*Tampa*, *FL*, 1991). *Contemporary Mathematics* **138** 99–117. Amer. Math. Soc., Providence, RI. MR1199122

[17] DIACONIS, P. and HOLMES, S. P. (2002). Random walks on trees and matchings. *Electron*. *J*. *Probab*. **7** 17 pp. (electronic). MR1887626

[18] DIACONIS, P., MAYER-WOLF, E., ZEITOUNI, O. and ZERNER, M. P. W. (2004). The Poisson–Dirichlet law is the unique invariant distribution for uniform split-merge transformations. *Ann*. *Probab*. **32** 915–938. MR2044670

[19] DIACONIS, P. and RAM, A. (2000). Analysis of systematic scan Metropolis algorithms using Iwahori–Hecke algebra techniques. *Michigan Math*. *J*. **48** 157–190. Dedicated to William Fulton on the occasion of his 60th birthday. MR1786485

[20] DIACONIS, P. and RAM, A. (2010). A probabilistic interpretation of the Macdonald polynomials. Available at arXiv:1007.4779.

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

[22] EDWARDS, R. G. and SOKAL, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Phys*. *Rev*. *D* (3) **38** 2009–2012. MR0965465

[23] ERCOLANI, N. M. and UELTSCHI, D. (2011). Cycle structure of random permutations with cycle weights. Available at arXiv:1102.4796.

[24] FRISTEDT, B. (1993). The structure of random partitions of large integers. *Trans*. *Amer*. *Math*. *Soc*. **337** 703–735. MR1094553

[25] FULMAN, J. (2002). Random matrix theory over finite fields. *Bull. Amer. Math. Soc.* (*N.S.*) **39** 51–85. MR1864086

[26] GARSIA, A. and REMMEL, J. B. (2005). Breakthroughs in the theory of Macdonald polynomials. *Proc. Natl. Acad. Sci. USA* **102** 3891–3894 (electronic). MR2139721

[27] GHOSH, J. K. and RAMAMOORTHI, R. V. (2003). *Bayesian Nonparametrics*. Springer, New York. MR1992245

[28] GONTCHAROFF, V. (1944). Du domaine de l'analyse combinatoire. *Bull. Acad. Sci. USSR Sér. Math.* [*Izvestia Akad. Nauk SSSR*] **8** 3–48. MR0010922

[29] GORDON, I. (2003). On the quotient ring by diagonal invariants. *Invent. Math.* **153** 503–518. MR2000467

[30] HAGLUND, J., HAIMAN, M. and LOEHR, N. (2005). A combinatorial formula for Macdonald polynomials. *J. Amer. Math. Soc.* **18** 735–761 (electronic). MR2138143

[31] HAGLUND, J., HAIMAN, M. and LOEHR, N. (2005). Combinatorial theory of Macdonald polynomials. I. Proof of Haglund's formula. *Proc. Natl. Acad. Sci. USA* **102** 2690–2696 (electronic). MR2141666

[32] HAGLUND, J., HAIMAN, M. and LOEHR, N. (2008). A combinatorial formula for nonsymmetric Macdonald polynomials. *Amer. J. Math.* **130** 359–383. MR2405160

[33] HAIMAN, M. (2006). Cherednik algebras, Macdonald polynomials and combinatorics. In *International Congress of Mathematicians*, *Vol. III* 843–872. Eur. Math. Soc., Zürich. MR2275709

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

[35] HOPPE, F. M. (1987). The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.* **25** 123–159. MR0896430

[36] HORA, A. and OBATA, N. (2007). *Quantum Probability and Spectral Analysis of Graphs*. Springer, Berlin. With a foreword by Luigi Accardi. MR2316893

[37] JIANG, J. (2011). Multiplicative measures on partitions, asymptotic theory. Preprint, Dept. Mathematics, Stanford Univ.

[38] KEROV, S. V. (2003). *Asymptotic Representation Theory of the Symmetric Group and Its Applications in Analysis*. *Translations of Mathematical Monographs* **219**. Amer. Math. Soc., Providence, RI. Translated from the Russian manuscript by N. V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski. MR1984868

[39] KNOP, F. and SAHI, S. (1997). A recursion and a combinatorial formula for Jack polynomials. *Invent. Math.* **128** 9–22. MR1437493

[40] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. With a chapter by James G. Propp and David B. Wilson. MR2466937

[41] LOGAN, B. F. and SHEPP, L. A. (1977). A variational problem for random Young tableaux. *Adv. Math.* **26** 206–222. MR1417317

[42] MACDONALD, I. G. (1995). *Symmetric Functions and Hall Polynomials*, 2nd ed. The Clarendon Press Oxford Univ. Press, New York. With contributions by A. Zelevinsky, Oxford Science Publications. MR1354144

[43] MACDONALD, I. G. (2000/01). Orthogonal polynomials associated with root systems. *Sém. Lothar. Combin.* **45** Art. B45a, 40 pp. (electronic). MR1817334

[44] MACDONALD, I. G. (2003). *Affine Hecke Algebras and Orthogonal Polynomials*. *Cambridge Tracts in Mathematics* **157**. Cambridge Univ. Press, Cambridge. MR1976581

[45] NEWMAN, M. E. J. and BARKEMA, G. T. (1999). *Monte Carlo Methods in Statistical Physics*. The Clarendon Press Oxford Univ. Press, New York. MR1691513

[46] OKOUNKOV, A. (2001). Infinite wedge and random partitions. *Selecta Math.* (*N.S.*) **7** 57–81. MR1856553

[47] OKOUNKOV, A. (2002). Symmetric functions and random partitions. In *Symmetric Functions 2001: Surveys of Developments and Perspectives*. *NATO Sci. Ser. II Math. Phys. Chem.* **74** 223–252. Kluwer Academic, Dordrecht. MR2059364

[48] OKOUNKOV, A. (2005). The uses of random partitions. In *XIVth International Congress on Mathematical Physics* 379–403. World Sci. Publ., Hackensack, NJ. MR2227852

[49] OLSHANSKI, G. (2011). Random permutations and related topics. In *The Oxford Handbook on Random Matrix Theory* (G. Akermann, J. Baik and P. Di Francesco, eds.). Oxford Univ. Press. To appear. Available at http://www.bookdepository.co.uk/Oxford-Handbook-Random-Matrix-Theory-Gernot-Akemann/9780199574001?b=-3&t=-26#Bibliographicdata-26.

[50] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer, Berlin. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR2245368

[51] RAM, A. and YIP, M. (2011). A combinatorial formula for Macdonald polynomials. *Adv. Math.* **226** 309–331. MR2735761

[52] SALOFF-COSTE, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1996). *Lecture Notes in Math.* **1665** 301–413. Springer, Berlin. MR1490046

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

[54] VERSHIK, A. M. (1996). Statistical mechanics of combinatorial partitions, and their limit configurations. *Funktsional. Anal. i Prilozhen.* **30** 19–39, 96. MR1402079

[55] VERŠIK, A. M. and KEROV, S. V. (1977). Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. *Dokl. Akad. Nauk SSSR* **233** 1024–1027. MR0480398

[56] YAKUBOVICH, Y. (2009). Ergodicity of multiplicative statistics. Available at arXiv:0901.4655.

[57] ZHAO, J. T. (2011). Universality results for measures on partitions. Preprint, Dept. Mathematics, Stanford Univ.

DEPARTMENTS OF MATHEMATICS
AND STATISTICS
STANFORD UNIVERSITY
SEQUOIA HALL, 390 SERRA MALL
STANFORD, CALIFORNIA 94305-4065
USA
E-MAIL: ckirby@stanford.edu

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF MELBOURNE
PARKVILLE, VICTORIA 3010
AUSTRALIA