J. Appl. Probab. **60,** 797–811 (2023)
*doi:*10.1017/jpr.2022.92

TRANSITION DENSITY OF AN INFINITE-DIMENSIONAL DIFFUSION WITH THE JACK PARAMETER

YOUZHOU ZHOU,* *Xi'an Jiaotong-Liverpool University*

Abstract

From the Poisson–Dirichlet diffusions to the Z-measure diffusions, they all have explicit transition densities. We show that the transition densities of the Z-measure diffusions can also be expressed as a mixture of a sequence of probability measures on the Thoma simplex. The coefficients are the same as the coefficients in the Poisson–Dirichlet diffusions. This fact will be uncovered by a dual process method in a special case where the Z-measure diffusions are established through an up–down chain in the Young graph.

Keywords: Jack graph; transition density; dual process; up–down Markov chain; Kingman coalescent

2020 Mathematics Subject Classification: Primary 60J60
Secondary 60C05

1. Introduction

The Poisson–Dirichlet distribution PD$(0, \theta)$, $\theta > 0$, was proposed in [9] in the simplex $\overline{\nabla}_\infty = \left\{x \in [0, 1]^\infty \mid x_1 \geq x_2 \geq \cdots \geq 0, \ \sum_{i=1}^{\infty} x_i \leq 1\right\}$. The one-parameter Poisson–Dirichlet distribution PD$(0, \theta)$ was generalized to the two-parameter Poisson–Dirichlet distribution PD(α, θ), $\alpha \in (0, 1)$, $\theta + \alpha > 0$, in [15]. They are also the representing measure of the Ewens–Pitman partition structure. The partition structure coined in [10] is an exchangeable partition distribution M of the set $\mathbb{N} = \{1, 2, \ldots, n, \ldots\}$. Due to its exchangeability, the restriction of M to $\{1, 2, \ldots, n\}$ is $M_n(\eta) = \dim(\eta)\varphi(\eta)$, where $\eta = (\eta_1, \ldots, \eta_l)$, $|\eta| =: \sum_{i=1}^{l} \eta_i = n$, is an integer partition of n. Moreover, $\varphi(\eta)$ is the probability of a single partition with cluster sizes η. Due to exchangeability, two partitions have the same probability as long as their cluster sizes are the same. So, $\dim(\eta)$ is the multiplicities of such partitions with cluster size η. Then, the family of distributions $\{M_n, n \geq 1\}$ will satisfy a natural consistency condition and they are uniquely determined by a representing measure in the Kingman simplex $\overline{\nabla}_\infty$ due to its exchangeability and the de Finetti theorem.

Let Γ_n be the totality of integer partitions of n; then $\Gamma = \cup_{n\geq 0}^{\infty}\Gamma_n$ exhausts all integer partitions, and $\Gamma_0 = \emptyset$ is treated as an empty partition. Usually Γ can be the vertex set of a graded branching diagram (Γ, χ), where $\chi(\eta, \omega)$ assigns positive weight to an edge joining η and ω, and 0 otherwise (see Figure 2). Each integer partition is represented by a Young diagram (see Figure 1); we say $\eta \subset \zeta$ if the Young diagram of η is contained in the Young diagram of ζ. There are various kinds of edge weights. In the algebra $\mathcal{A}$ of symmetric functions with variables $(x_1, \ldots, x_n, \ldots)$, there are various kinds of linear bases $\{f_\eta, \eta \in \Gamma\}$ (refer to [12] or the appendix), such as the monomial functions, the Shur functions, and the Jack functions. These

Received 17 October 2021; revision received 24 July 2022.
* Postal address: 111 Renai Road, Suzhou, Jiangsu, China 215 123. Email address: Youzhou.zhou@xjtlu.edu.cn

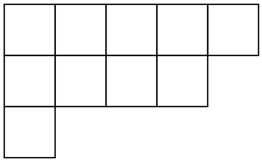

FIGURE 1. Young diagram $\eta = (5, 4, 1)$.

functions usually satisfy the Pieri formula $\left(\sum_{i=1}^{\infty} x_i\right) f_\eta = \sum_{\eta \subset \zeta, |\eta|+1=|\zeta|} \chi(\eta, \zeta) f_\zeta$, where η can be obtained from ζ by removing a box. The edge weights are chosen to be the coefficients in the above Pieri formula. The weights from the monomial functions, the Shur functions, and the Jack functions define the Kingman graph, the Young graph, and the Jack graph respectively. The specialization of the symmetric functions can be regarded as an algebra homomorphism $\Phi\colon \mathcal{A} \to \mathbb{R}$. In particular, a specialization mapping base functions to positive values will determine a positive harmonic function $\varphi(\eta) = \Phi(f_\eta)$ on Γ, satisfying $\varphi(\eta) = \sum_{\eta \subset \zeta, |\eta|+1=|\zeta|} \chi(\eta, \zeta)\varphi(\zeta)$. Under pointwise convergence topology, the space $\mathcal{C}$ of all positive harmonic functions becomes compact and convex. Therefore, we can expect that $\varphi(\eta) = \int_E K(\eta, x)\mu(\mathrm{d}x)$, where E is the Martin boundary of $\mathcal{C}$, $K(\eta, x)$ is the extremity, and $\mu(\mathrm{d}x)$ is a probability measure. For the Kingman graph, this representation is the Kingman's one-to-one correspondence of the partition structures [10] and $K(\eta, x)$ is the continuous extension of the monomial functions. For the Jack graph, the representation is also established in [8], and $K(\eta, x)$ is the extended Jack functions. The Z-partition structure is a partition structure defined on the Jack graph, and its representing measure is called the Z-measure, denoted as $\mathcal{Z}(z, z', \vartheta)$, $z, z' \in \mathbb{C}$, $\vartheta > 0$.

The two-parameter Poisson–Dirichlet diffusion with the stationary distribution $\mathcal{PD}(\alpha, \theta)$ was established through an up–down Markov chain on the Kingman graph in [14].

Similarly, [13] established a reversible diffusion on the Thoma simplex,

$$\Omega = \Big\{(\alpha, \beta) \in [0, 1]^\infty \times [0, 1]^\infty \mid \\ \alpha_1 \geq \alpha_2 \geq \cdots \geq 0,\ \beta_1 \geq \beta_2 \geq \cdots \geq 0,\ \textstyle\sum_{i=1}^\infty \alpha_i + \sum_{j=1}^\infty \beta_j \leq 1\Big\}.$$

Its stationary distribution is the measure $\mathcal{Z}(z, z', \vartheta)$. In this paper, we will use the Z-measure diffusion to refer to the diffusion obtained in [13].

Interestingly, both types of diffusion have similar explicit transition densities, and the spectrums of their generators are also the same, $\left\{0, \frac{1}{2}n(n-1+\theta) \mid n \geq 2\right\}$, where $\theta = zz'/\vartheta > 0$ for the Z-measure diffusion.

Recently, spectral expansion was again applied to the Z-measure diffusion in [11] to derive its explicit transition density. A similar method was previously used to establish the transition density of the two-parameter Poisson–Dirichlet diffusion in [6]. In fact, this method was first adopted in [4]. Surprisingly, by rearranging the density in [6], [17] obtained the expression

$$p(t, x, y) = \widetilde{d}_1^\theta(t) + \sum_{n=2}^\infty d_n^\theta(t) p_n(x, y), \tag{1}$$

where $p_n(x, y)$ is a transition kernel and $\left\{\widetilde{d}_1^\theta(t), d_n^\theta(t), n \geq 2\right\}$, $\theta > -1$, is the distribution of a pure death process related to the Kingman coalescent in [17].

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

In this paper, we rearrange the transition density of the Z-measure diffusion as in [17]. We obtain a similar expression for the Z-measure diffusion,

$$q(t, \sigma, \omega) = \widetilde{d}_1^{\theta}(t) + \sum_{n=2}^{\infty} d_n^{\theta}(t)\mathcal{K}_n(\sigma, \omega), \tag{2}$$

where $\mathcal{K}_n(\sigma, \omega)$ is also a transition kernel and $\{\widetilde{d}_1^{\theta}(t), d_n^{\theta}(t), n \geq 2\}$, $\theta > 0$, are the same as the coefficients in (1).

The next natural question would be why their transition densities have the same coefficients. This question has been resolved for the two-parameter Poisson–Dirichlet diffusion in [7] by a dual process method. In order to understand why the Z-diffusion also has similar coefficients in its density expression (2), we also apply the dual process method to the Z-measure diffusion Y_t. But we can only find the dual process of the Z-measure diffusion when $\vartheta = 1$ because we rely on [3, (5)]. When $\vartheta \neq 1$, a similar equation is not known. For $\vartheta = 1$, the dual process $\mathcal{D}_t$ is also a partition-valued jump process characterized by the generator

$$\mathcal{L}_1^{\mathrm{J}} f(\eta) = -\frac{n(n-1+\theta)}{2} f(\eta) + \frac{n(n-1+\theta)}{2} \sum_{\zeta \subset \eta, |\eta| = |\zeta|+1} p_1^{\downarrow}(\eta, \zeta) f(\zeta), \qquad \mathcal{L}_1^{\mathrm{J}} 1 = 0,$$

where $p_1^{\downarrow}(\eta, \zeta)$ is the transition probability of the down Markov chain in the Young graph discussed in [3], and f is the continuous function on the Thoma simplex. The dual relation is defined through a bivariate function $F(\eta, \omega) = s_\eta^o(\omega)/\mathbb{E}_{z,z',1} s_\eta^o$, which is the normalized kernel in the representation of the Young graph. The notation $\mathbb{E}_{z,z',1}$ should always be interpreted as the expectation with respect to the Z-measure $\mathcal{Z}(z, z', 1)$. We show that duality relation reads as $\mathbb{E}_\omega F(\eta, Y_t) = \mathbb{E}_\eta F(\mathcal{D}_t, \omega)$. Because the distribution of $\mathcal{D}_t$ is easier to calculate, then the expression in (2) can be obtained. Now the radial process $|\mathcal{D}_t|$ of the dual process $\mathcal{D}_t$ will be exactly the same as that in the two-parameter Poisson–Dirichlet diffusion. This will eventually determine the coefficients in the density expression (2). Although when $\vartheta \neq 1$ we cannot find the dual process, the result when $\vartheta = 1$ encourages us to conjecture that the dual process of the Z-measure diffusion should also be a partition-valued jump process whose jump rate is $\frac{1}{2} n(n-1+\theta)$ and its embedded chain is the down Markov chain in [13].

The plan of the paper is as follows. In Section 2 we introduce the branching diagrams and the up and down Markov chains. In Section 3, we talk about the Z-measure diffusion and its transition density. In Section 4 we use the dual process method to derive the transition density of the Z-measure diffusion when $\vartheta = 1$. In the last section, a few conjectures will be presented.

2. Branching diagram and up–down Markov chain

2.1. Branching diagram

For $n, l \in \mathbb{Z}_+$, $\eta = (\eta_1, \ldots, \eta_l) \in \mathbb{N}^l$ is called an integer partition of n if $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_l > 0$ and $|\eta| := \sum_{i=1}^{l} \eta_i = n$. Define $\alpha_i(\eta) = \#\{1 \leq j \leq l \mid \eta_j = i\}$, $1 \leq i \leq n$; then $(\alpha_1(\eta), \ldots, \alpha_n(\eta))$ is a different representation of the partition η. Denote by Γ_n the set of all integer partitions of n; then $\Gamma = \cup_{n \geq 0} \Gamma_n$ is the set of all integer partitions, where $\Gamma_0 = \{\emptyset\}$ and $\emptyset$ is an empty partition. An integer partition $\eta = (\eta_1, \ldots, \eta_l)$ may also be represented by a Young diagram (Figure 1) defined by attaching boxes at position (i, j), where $1 \leq i \leq l$, $1 \leq j \leq \eta_i$. Here the row number increases from top to bottom and the column number increases from left to right.

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

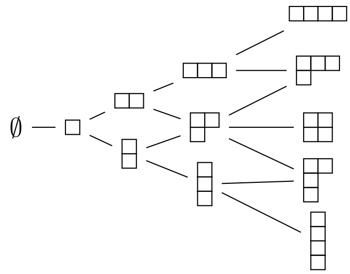

FIGURE 2. Branching diagram.

In this paper, we will interchangeably use η to represent either an integer partition or its Young diagram. We define the diagonal line of the Young diagram η as the set of boxes $\{(i, i) \mid i \le \eta_i, \eta'_i\}$, and r as the length of the diagonal line of η. The transposition of η with respect to its diagonal line will give us a new partition η' called the conjugate of η. We say $\eta \subset \nu$ if the Young diagram of η is contained in the Young diagram of ν. Then, $\subset$ is a partial ordering in Γ. For each box $b = (i, j)$ in the partition η, we define its arm length as $a(b) = \eta_i - j$ and its leg length as $l(b) = \eta'_j - i$. A partition η may also be represented by its Frobenius coordinates $\widetilde{\eta} = \big(a(1, 1) + \frac{1}{2}, \ldots, a(r, r) + \frac{1}{2}; l(1, 1) + \frac{1}{2}, \ldots, l(r, r) + \frac{1}{2}\big)$.

A branching diagram is a graded graph (Γ, χ) (Figure 2), where $\Gamma = \cup_{n\ge 0}\Gamma_n$ is the vertex set. There is an edge joining $\eta \in \Gamma_n$ and $\nu \in \Gamma_{n+1}$ if and only if $\eta \subset \nu$, and their edge weight is $\chi(\eta, \nu)$.

When the edge weights are determined by the coefficients in the Pieri formula of the monomial symmetric functions, we will have a Kingman graph. Its edge weight is defined as $\chi^{\mathrm{K}}(\eta, \zeta) = \alpha_{\zeta_i}(\zeta)$ if $\zeta - \eta = (i, j)$, i.e. ζ can be obtained from η by attaching a box at (i, j). If we choose the edge weight to be the coefficients in the Pieri formula of the Jack functions, then we end up with the Jack graph, whose edge weight, denoted as $\chi^{\mathrm{J}}_{\vartheta}(\eta, \zeta)$, will be

$$\chi^{\mathrm{J}}_{\vartheta}(\eta, \zeta) = \prod_b \frac{(a(b) + (l(b) + 2)\vartheta)(a(b) + 1 + l(b)\vartheta)}{(a(b) + 1 + (l(b) + 1)\vartheta)(a(b) + (l(b) + 1)\vartheta)},$$

where b runs over all boxes in the jth column of η if $\zeta - \eta = (i, j)$. When $\vartheta = 1$, then $\chi^{\mathrm{J}}_1 = 1$ and the Jack graph reduces to the Young graph.

We define the weight of a path t: $\eta = \eta(0) \subset \cdots \subset \eta(k) = \zeta$ as $\prod_{i=0}^{k-1} \chi(\eta(i), \eta(i + 1))$. Then we define the total weight between η and ζ in the Kingman graph as $\dim^{\mathrm{K}}(\eta, \zeta) = \sum_{\eta=\eta(0)\subset\cdots\subset\eta(k)=\zeta} \prod_{i=0}^{k-1} \chi(\eta(i), \eta(i + 1))$. Similarly, the total weight between η and ζ in the Jack graph will be $\dim^{\mathrm{J}}_{\vartheta}(\eta, \zeta) = \sum_{\eta=\eta(0)\subset\cdots\subset\eta(k)=\zeta} \prod_{i=0}^{k-1} \chi(\eta(i), \eta(i + 1))$. In particular, when $\eta = \emptyset$, we regard $\dim^{\mathrm{K}}(\nu) = \dim^{\mathrm{K}}(\emptyset, \nu)$ as the total weight of ν in the Kingman graph and $\dim^{\mathrm{J}}_{\vartheta}(\nu) = \dim^{\mathrm{J}}_{\vartheta}(\emptyset, \nu)$ as the total weight of ν in the Jack graph. Naturally, we have

$$\dim^{\mathrm{K}}(\zeta) = \sum_{\eta\subset\zeta, |\zeta|=|\eta|+1} \chi^{\mathrm{K}}(\eta, \zeta) \dim^{\mathrm{K}}(\eta), \tag{3}$$

$$\dim^{\mathrm{J}}_{\vartheta}(\zeta) = \sum_{\eta\subset\zeta, |\zeta|=|\eta|+1} \chi^{\mathrm{J}}(\eta, \zeta) \dim^{\mathrm{J}}_{\vartheta}(\eta).$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

In the Kingman graph, the total weight of η is $\dim^{\mathrm{K}}(\eta) = n!/(\eta_1! \cdots \eta_l!)$ if $\eta = (\eta_1, \ldots, \eta_l) \in \Gamma_n$. In the Jack graph, the total of η is $\dim^{\mathrm{J}}_\vartheta(\eta) = n!/(H(\eta; \vartheta)H'(\eta; \vartheta))$, where $H(\eta; \vartheta) = \prod_{b\in\eta}(a(b) + 1 + l(b)\eta)$, $H'(\eta; \vartheta) = \prod_{b\in\eta}(a(b) + (1 + l(b))\eta)$. When $\vartheta = 1$, the total weight of η is $\dim^{\mathrm{J}}_1(\eta) = n!/H^2(\eta; 1)$ in the Young graph.

2.2. Down–up Markov chain

Due to (3), we can easily construct a down Markov chain $\{D^{\mathrm{K}}_n, n \geq 1\}$ in the Kingman graph with the transition probability

$$p^{\downarrow,\mathrm{K}}(\zeta, \eta) = \frac{\chi^{\mathrm{K}}(\eta, \zeta)\dim^{\mathrm{K}}(\eta)}{\dim^{\mathrm{K}}(\zeta)} = \frac{\alpha_{\zeta_i}(\zeta)\zeta_i}{n}, \qquad \zeta - \eta = (i, j),\ \zeta \in \Gamma_n.$$

This down Markov chain is the embedded chain of the dual process in [7]. Similarly, we can also construct a down Markov chain $\{D^{\mathrm{J},\vartheta}_n, n \geq 1\}$ in the Jack graph with the transition probability

$$p^{\downarrow,\mathrm{J}}_\vartheta(\zeta, \eta) = \frac{\chi^{\mathrm{J}}_\vartheta(\eta, \zeta)\dim^{\mathrm{J}}_\vartheta(\eta)}{\dim^{\mathrm{J}}_\vartheta(\zeta)}, \qquad \zeta \in \Gamma_n,\ \eta \in \Gamma_{n-1}.$$

When $\vartheta = 1$, this down Markov chain is the embedded chain of the dual process of the Z-measure diffusion that we are going to discuss in Section 4.

As you may see, the down Markov chain depends only on the edge weights of the graph. But we can also construct an up Markov chain if we have exchangeable partition structures in the branching diagram. The Ewens–Pitman partition structure can be used to construct an up Markov chain $\{U^{\mathrm{K},\theta,\alpha}_n, n \geq 1\}$ in the Kingman graph [14]. Then we can define an up–down chain $\{UD^{\mathrm{K},\theta,\alpha,n}_m, m \geq 1\}$ on Γ_n, updating itself as

$$\begin{aligned}\mathbb{P}\big(UD^{\mathrm{K},\alpha,\theta,n}_{m+1} = \zeta \mid UD^{\mathrm{K},\alpha,\theta,n}_m = \eta\big) \\ = \sum_{|\nu|=n+1} \mathbb{P}\big(U^{\mathrm{K},\theta,\alpha}_{n+1} = \nu \mid U^{\mathrm{K},\theta,\alpha}_n = \eta\big)\mathbb{P}\big(D^{\mathrm{K}}_n = \zeta \mid D^{\mathrm{K}}_{n+1} = \nu\big). \qquad (4)\end{aligned}$$

Naturally, the partition distribution $M^{\mathrm{K},\theta,\alpha}_n$ will serve as the stationary distribution of this up–down chain. The usual space and time scaling yields the two-parameter Poisson–Dirichlet diffusion in [14]. The Ewens–Pitman partition structure can be replicated by the Blackwell–MacQueen urn model, and it has found many applications in classification problems through Bayesian statistics [1].

For the Jack graph there is also a special Z partition structure [2]

$$M^{\mathrm{J},z,z',\vartheta}_n(\eta) = \dim^{\mathrm{J}}_\vartheta(\eta)\frac{(z)_{\eta,\vartheta}(z')_{\eta,\vartheta}}{\theta_{(n)}H'(\eta; \vartheta)}, \qquad \theta = \frac{zz'}{\vartheta},\ \vartheta > 0, \qquad (5)$$

where either (i) $z \in \mathbb{C} - (\mathbb{Z}_{\leq 0} + \vartheta\mathbb{Z}_{\geq 0})$ and $z' = \bar{z}$ (principal case), or (ii) ϑ is rational number and z, z' are real numbers belonging to an interval between two consecutive lattice points in $\mathbb{Z} + \vartheta\mathbb{Z}$ (complementary case). Here, $(z)_{\eta;\vartheta} = \prod_{(i,j)\in\eta}(z + (j-1) - (i-1)\vartheta)$. The representing measure $\mathcal{Z}(z, z', \vartheta)$ of the partition structure (5) is the Z-measure. Similarly, we can construct an up Markov chain $\{U^{\mathrm{J},z,z',\vartheta}_n, n \geq 1\}$ [13]. By mimicking the update in (4), we can also construct an up–down Markov chain $\{UD^{\mathrm{J},z,z',\vartheta,n}_m, m \geq 1\}$ on Γ_n. Then the usual space and time scaling in Section 3 yields the Z-diffusion in [13]. To the best of the author's knowledge, no quick replication of the Z partition structure $M^{\mathrm{J},z,z',\vartheta}_n(\eta)$ has been identified. So whether the Z partition structure can be applied to classification problems is still open.

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

3. The *Z*-measure diffusion and its transition density

3.1. The *Z*-measure diffusion

The diffusion approximation of the up–down Markov chain $\{UD_m^{\mathrm{J},z,z',\vartheta,n}, m \geq 1\}$ on Γ_n can be carried out by the space scaling

$$\pi : \eta \to \frac{\widetilde{\eta}}{n} = \left(\frac{a(1,1)+\frac{1}{2}}{n}, \ldots, \frac{a(r,r)+\frac{1}{2}}{n}; \frac{l(1,1)+\frac{1}{2}}{n}, \ldots, \frac{l(r,r)+\frac{1}{2}}{n}\right),$$

where $\widetilde{\eta}$ is the Frobenius coordinates of η. As $n \to \infty$, [13] showed that $\pi(UD_{[n^2t]}^{\mathrm{J},z,z',\vartheta,n})$ converges to the Z-measure diffusion Y_t on the Thoma simplex Ω with the pre-generator

$$A_{z,z',\vartheta} = \frac{1}{2}\sum_{i,j\geq 2} ij(\varphi^o_{i+j-1} - \varphi^o_i\varphi^o_j)\frac{\partial^2}{\partial\varphi^o_i\partial\varphi^o_j} + \frac{\vartheta}{2}\sum_{i,j\geq 1}(i+j+1)\varphi^o_i\varphi^o_j\frac{\partial}{\partial\varphi^o_{i+j+1}}$$
$$+ \frac{1}{2}\sum_{i\geq 2}\left[(1-\vartheta)i(i-1)\varphi^o_{i-1} + (z+z')i\varphi^o_{i-1} - i(i-1)\varphi^o_i - i\frac{zz'}{\vartheta}\varphi^o_i\right]\frac{\partial}{\partial\varphi^o_i}.$$

The core of $A_{z,z',\vartheta}$ is spanned by $\{\varphi^o_j, j \geq 1\}$, where φ^o_j is the image of $\varphi_j(x) = \sum_{i=1}^{\infty} x_i^j$ under the special algebra homomorphisms $\Phi_\omega\colon \mathcal{A} \to \mathbb{R}$, $\omega = (\alpha, \beta) \in \Omega$, $\Phi_\omega(\varphi_n) = \sum_{i=1}^{\infty}\alpha_i^n + (-\vartheta)^{n-1}\sum_{j=1}^{\infty}\beta_j^n$, $n \geq 2$, $\Phi_\omega(\varphi_1) = 1$. Since the Jack functions $J_\zeta(x; \vartheta)$ can be written as linear combinations of $\varphi_{\eta_1}\cdots\varphi_{\eta_l}$, where $\eta = (\eta_1, \ldots, \eta_l) \subset \zeta$, then $J^o_\zeta(\omega; \vartheta) := \Phi_\omega(J_\zeta(x; \vartheta))$.

In particular, when $\vartheta = 1$ we have the extended Shur function $s^o_\zeta(\omega) := \Phi_\omega(J_\zeta(x; 1))$. The Z-measure diffusion reduces to the diffusion in [3].

3.2. Transition density of the *Z*-measure diffusion Y_t

By spectral expansion of $A_{z,z',\vartheta}$ in the Hilbert space $L^2(\Omega, \mathcal{Z}(z, z', \vartheta))$, the following explicit transition density of Y_t was obtained in [11].

Proposition 1. *The transition density of Y_t is*

$$q(t, \sigma, \omega) = 1 + \sum_{m=2}^{\infty} \mathrm{e}^{-t\lambda_m} G_m(\sigma, \omega), \tag{6}$$

where $\theta = zz'/\vartheta$, $\lambda_m = \frac{1}{2}m(m-1+\theta)$, and

$$G_m(\sigma, \omega) = \sum_{n=0}^{m}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{m-1}}{m!}\mathcal{K}_n(\sigma, \omega),$$

$$\mathcal{K}_n(\sigma, \omega) = \sum_{|\eta|=n}\frac{j_\eta(\sigma; \vartheta)j_\eta(\omega; \vartheta)}{\mathbb{E}_{z,z',\vartheta}j_\eta}, \qquad j_\eta(\omega; \vartheta) = \dim^{\mathrm{J}}_\vartheta(\eta)J^o_\eta(\omega; \vartheta).$$

In particular, when $n = 0, 1$, $\mathcal{K}_n(\sigma, \omega) = 1$.

In [11], it was shown that

$$K^o_n(\sigma, \omega) = \sum_{|\eta|=n}\frac{H'(\eta; \vartheta)}{H(\eta; \vartheta)}\frac{J^o(\sigma; \vartheta)J^o(\omega; \vartheta)}{(z)_{\eta,\vartheta}(z')_{\eta,\vartheta}}.$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

By the following Proposition 2, [2] showed that the representing measure of the Z partition structure is the Z-measure $\mathcal{Z}(z, z', \vartheta)$.

Proposition 2. *For a partition structure $\{M_n, n \geq 1\}$ on the Jack graph, there is a unique probability measure μ on the Thoma simplex Ω such that $M_n(\eta) = \int_\Omega j_\eta(\omega; \vartheta)\mu(\mathrm{d}\omega)$. Moreover, $\mu_n(\mathrm{d}\omega) = \sum_{\eta \in \Gamma_n} M_n(\eta)\delta_{\widetilde{\eta}/n}(\mathrm{d}\omega)$ will converge weakly to μ. Here, $\widetilde{\eta} = \big(a(1, 1) + \frac{1}{2}, \ldots, a(r, r) + \frac{1}{2}; l(1, 1) + \frac{1}{2}, \ldots, l(r, r) + \frac{1}{2}\big)$ is the Frobenius coordinate of η.*

Therefore, similar to the Ewens sampling formula, we have

$$\mathbb{E}_{z,z',\vartheta} j_\eta = M_n^{z,z',\vartheta}(\eta) \tag{7}$$

and $K_n^o(\sigma, \omega) = \mathcal{K}_n(\sigma, \omega)/(n!(\theta)_n)$. Then, we can easily see that (6) is equivalent to the density in [11].

In this paper, we will rearrange the right-hand side of (6) to yield a new representation.

Theorem 1. *The transition density of Y_t is*

$$q(t, \sigma, \omega) = d_0^\theta(t) + d_1^\theta(t) + \sum_{n=2}^{\infty} d_n^\theta(t)\mathcal{K}_n(\sigma, \omega), \tag{8}$$

where

$$d_0^\theta(t) = 1 - \sum_{m=1}^{\infty} \frac{2m - 1 + \theta}{m!}(-1)^{m-1}\theta_{(m-1)}\mathrm{e}^{-\lambda_m t},$$

$$d_n^\theta(t) = \sum_{m=1}^{\infty} \frac{2m - 1 + \theta}{m!}(-1)^{m-n}\binom{m}{n}(n + \theta)_{(m-1)}\mathrm{e}^{-\lambda_m t}, \qquad n \geq 1.$$

Due to the estimation of G_m in [11], the proof of Theorem 1 is the same as the proof of [17, Theorem 2.1].

Proof. By [11, Proposition 16], we know there exist positive constants c, d such that $\sup_{\sigma,\omega\in\Omega} |G_m(\sigma, \omega)| \leq cm^{dm}$. Then, we can show that the series in (6) is uniformly convergent (see [17] or [11]). We can then switch the order of summation in (6):

$$\begin{aligned} q(t, \sigma, \omega) = 1 + \sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\Bigg[&\sum_{n=2}^{m} (-1)^{m-n}\binom{m}{n}\frac{(\theta + 2m - 1)(\theta + n)_{m-1}}{m!}\mathcal{K}_n(\sigma, \omega) \\ &+ \frac{2m + \theta - 1}{m!}(-1)^{m-1}(\theta + 1)_{(m-1)} m\mathcal{K}_1(\sigma, \omega) \\ &+ \frac{2m + \theta - 1}{m!}(-1)^{m}(\theta)_{(m-1)}\mathcal{K}_0(\sigma, \omega)\Bigg]. \end{aligned}$$

Since $\mathcal{K}_1(\sigma, \omega) = \mathcal{K}_0(\sigma, \omega) = 1$, we have

$$q(t, \sigma, \omega) = 1 + \sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t} \sum_{n=2}^{m} (-1)^{m-n}\binom{m}{n}\frac{(\theta + 2m - 1)(\theta + n)_{m-1}}{m!}\mathcal{K}_n(\sigma, \omega)$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

$$
\begin{aligned}
&+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\left[\frac{2m+\theta-1}{m!}(-1)^{m-1}(\theta+1)_{(m-1)}m \right.\\
&\qquad\qquad \left. +\frac{2m+\theta-1}{m!}(-1)^{m}(\theta)_{(m-1)}\right]\\
&=1+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\sum_{n=2}^{m}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{(m-1)}}{m!}\mathcal{K}_n(\sigma,\omega)\\
&+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\left[\frac{2m+\theta-1}{m!}(-1)^{m-1}(\theta+1)_{(m-1)}m \right.\\
&\qquad\qquad \left. +\frac{2m+\theta-1}{m!}(-1)^{m}(\theta)_{(m-1)}\right]\\
&=1+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\sum_{n=2}^{m}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{(m-1)}}{m!}\mathcal{K}_n(\sigma,\omega)\\
&+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\frac{2m+\theta-1}{m!}(-1)^{m-1}[(\theta+1)_{(m-1)}m-(\theta)_{(m-1)}].
\end{aligned}
$$

Note that $(\theta+1)_{(m-1)}m-(\theta)_{(m-1)}=0$ for $m=1$. Therefore,

$$
\begin{aligned}
q(t,\sigma,\omega)&=1+\sum_{m=2}^{\infty} \mathrm{e}^{-\lambda_m t}\sum_{n=2}^{m}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{(m-1)}}{m!}\mathcal{K}_n(\sigma,\omega)\\
&+\sum_{m=1}^{\infty} \mathrm{e}^{-\lambda_m t}\frac{2m+\theta-1}{m!}(-1)^{(m-1)}[(\theta+1)_{(m-1)}m-(\theta)_{(m-1)}]\\
&=1-\sum_{m=1}^{\infty} \mathrm{e}^{-\lambda_m t}\frac{2m+\theta-1}{m!}(-1)^{(m-1)}(\theta)_{(m-1)}\\
&+\sum_{m=1}^{\infty} \mathrm{e}^{-\lambda_m t}\frac{2m+\theta-1}{m!}(-1)^{(m-1)}(\theta+1)_{(m-1)}m\\
&+\sum_{m=2}^{\infty} e^{-\lambda_m t}\sum_{n=2}^{m}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{(m-1)}}{m!}\mathcal{K}_n(\sigma,\omega)\\
&=d_0^{\theta}(t)+d_1^{\theta}(t)\\
&+\sum_{n=2}^{\infty}\mathcal{K}_n(\sigma,\omega)\sum_{m=n}^{\infty} \mathrm{e}^{-\lambda_m t}(-1)^{m-n}\binom{m}{n}\frac{(\theta+2m-1)(\theta+n)_{(m-1)}}{m!}\\
&=\widetilde{d_1^{\theta}}(t)+\sum_{n=2}^{\infty} d_n^{\theta}(t)\mathcal{K}_n(\sigma,\omega).
\end{aligned}
$$

□

Corollary 1. *The diffusion Y_t satisfies the ergodic inequality*

$$\sup_{\omega\in\Omega} \|\mathbb{P}_\omega(Y_t \in \cdot) - \mathcal{Z}(z, z', \vartheta)(\cdot)\|_{\text{Var}} \le \tfrac{1}{2}(\theta+1)(\theta+2)\mathrm{e}^{-(\theta+1)t}, \qquad \theta = \frac{zz'}{\vartheta} > 0.$$

This inequality can be easily derived from an inequality of tail probabilities (see [16]): $\sum_{n=2}^{\infty} d_n^\theta(t) \le \frac{1}{2}(\theta+1)(\theta+2)\mathrm{e}^{-(\theta+1)t}$.

4. Dual process of the Z-measure diffusion when $\vartheta = 1$

The transition density of the Z-measure diffusion looks very similar to the transition density of the two-parameter Poisson–Dirichlet diffusion. Given their huge differences, it is surprising that their transition densities are both mixtures of distributions with exactly the same coefficients. To figure out why they have the same coefficients, we adopt the dual process method used in [7]. The duality between Y_t and its dual process $\mathcal{D}_t$ is defined through a bivariate function $F(\eta, \omega)$, where $\eta \in \Gamma$ and $\omega \in \Omega$. The bivariate function $F(\eta, \omega)$ is usually chosen to be the normalized kernel $F(\eta, \omega) = j_\eta(\omega; \vartheta)/\mathbb{E}_{z,z',\vartheta} j_\eta$ in Proposition 2. Then, Y_t and its dual $\mathcal{D}_t$ satisfy $\mathbb{E}^*_\eta F(\mathcal{D}_t, \omega) = \mathbb{E}_\omega F(\eta, Y_t)$, where $\mathbb{E}^*_\eta$ is the expectation with respect to the distribution of $\mathcal{D}_t$. In this paper, we use this dual equation to derive the transition density (8) directly when $\vartheta = 1$. This derivation clearly explains why the coefficients $d_n^\theta(t)$ show up in the transition density (8). However, when $\vartheta \neq 1$, we fail to verify the dual process of Y_t because we don't know whether the Z-measure diffusion is similar to [13, (5)].

4.1. Dual process

In this section we consider test functions $g_\eta(\omega) = j_\eta(\omega; 1)/\mathbb{E}_{z,z',\vartheta} j_\eta$, so $\mathbb{E}_{z,z',\vartheta}\, g(\eta) = 1$. When $\vartheta = 1$, $J_\eta(x; 1)$ is just the Shur function $s_\eta(x)$, so $j_\eta(\omega; 1)$ is $\dim_1^{\mathrm{J}}(\eta)\Phi_\omega(s_\eta) = \dim_1^{\mathrm{J}}(\eta) s^o_\eta(\omega)$. Moreover, by (7), we know that

$$\mathbb{E}_{z,z',1} \dim_1^{\mathrm{J}}(\eta) s^o_\eta = \frac{n!}{H^2(\eta; 1)} \frac{(z)_{\eta,1}(z')_{\eta,1}}{(\theta)_{(n)}}.$$

We define $F(\eta, \omega) = g_\eta(\omega)$, $\eta \in \Gamma$, $\omega \in \Omega$. So,

$$F(\eta, \omega) = g_\eta(\omega) = \frac{H(\eta; 1)\theta_{(n)} s^o_\eta(\omega)}{(z)_{\eta,1}(z')_{\eta,1}}. \tag{9}$$

Now consider a jump process $\mathcal{D}_t$ defined by

$$\mathcal{L}_1^{\mathrm{J}} f(\eta) = -\frac{n(n-1+\theta)}{2} f(\eta) + \frac{n(n-1+\theta)}{2} \sum_{\zeta\subset\eta, |\eta|=|\zeta|+1} \frac{\dim_1^{\mathrm{J}}(\zeta)}{\dim_1^{\mathrm{J}}(\eta)} f(\zeta),$$

where $\dim_1^{\mathrm{J}}(\zeta)/\dim_1^{\mathrm{J}}(\eta) = p_1^{\downarrow, J}(\eta, \zeta)$ is the transition probability of the down Markov chain $D_n^{\mathrm{J},1}$ in the Young graph. Moreover, $\mathcal{L}_1^{\mathrm{J}} 1 = 0$ and. because $s^o_{(1)} = 1$, $\mathcal{D}_t$ will be absorbed at state $\eta = (1)$. The radial process $|\mathcal{D}_t|$ of $\mathcal{D}_t$ is only determined by the jump rates $\frac{1}{2} n(n-1+\theta)$, $n \ge 1$. Therefore, $|\mathcal{D}_t|$ is exactly the radial process in the Kingman coalescent by collapsing states 0 and 1 as a new state 1. The distribution of $|\mathcal{D}_t|$ is obtained in [16, 17]. So as long as the jump rates are the same for the dual process, the coefficients in their transition density, if it exists, will be the same.

Theorem 2. *When $\vartheta = 1$, the Z-measure diffusion Y_t and $\mathcal{D}_t$ satisfy the duality*

$$\mathbb{E}^*_\eta F(\mathcal{D}_t, \omega) = \mathbb{E}_\omega F(\eta, Y_t). \tag{10}$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

Proof. By [3, Lemma 5.4], we know that

$$A_{z,z',1}s_{\eta}^{o}(\omega)=-\frac{n(n-1+\theta)}{2}s_{\eta}^{o}(\omega)+\frac{1}{2}\sum_{\zeta\subset\eta,|\eta|=|\zeta|+1}\frac{(z)_{\eta,1}(z')_{\eta,1}}{(z)_{\zeta,1}(z')_{\zeta,1}}s_{\zeta}^{0}(\omega). \tag{11}$$

From (9), we know that

$$s_{\eta}^{o}(\omega)=g_{\eta}(\omega)\frac{(z)_{\eta,1}(z')_{\eta,1}}{H(\eta;1)\theta_{(n)}},\qquad s_{\zeta}^{o}(\omega)=g_{\zeta}(\omega)\frac{(z)_{\zeta,1}(z')_{\zeta,1}}{H(\zeta;1)\theta_{(n-1)}}.$$

Replacing s_{η}^{o} and s_{ζ}^{o} in (11) yields

$$A_{z,z',1}g_{\eta}(\omega)=-\frac{n(n-1+\theta)}{2}g_{\eta}(\omega)+\frac{1}{2}\sum_{\zeta\subset\eta,|\eta|=|\zeta|+1}\frac{H(\eta;1)\theta_{(n)}}{H(\zeta;1)\theta_{(n-1)}}g_{\zeta}(\omega).$$

Because

$$\frac{H(\eta;1)}{H(\zeta;1)}=n\frac{\dim_1^{\mathrm{J}}(\zeta)}{\dim_1^{\mathrm{J}}(\eta)},\qquad \frac{\theta_{(n)}}{\theta_{(n-1)}}=n-1+\theta,$$

we have

$$A_{z,z',1}g_{\eta}(\omega)=-\frac{n(n-1+\theta)}{2}g_{\eta}(\omega)+\frac{n(n-1+\theta)}{2}\sum_{\zeta\subset\eta,|\eta|=|\zeta|+1}\frac{\dim_1^{\mathrm{J}}(\zeta)}{\dim_1^{\mathrm{J}}(\eta)}g_{\zeta}(\omega).$$

Therefore, $A_{z,z',1}F(\eta,\omega)=\mathcal{L}_1^{\mathrm{J}}F(\eta,\omega)$. By [5, Theorem 4.4.1], we can show that the Z-measure diffusion Y_t and $\mathcal{D}_t$ satisfy the dual equation (10). □

Proposition 3. *The dual process $\mathcal{D}_t$ has the transition probability $\mathbb{P}(\mathcal{D}_t=\eta\mid\mathcal{D}_0=\nu)=d_{mn}^{\theta}(t)\mathcal{H}(\eta,\nu)$, $\eta\in\Gamma_n$, $\nu\in\Gamma_m$, $n\le m$. Here,*

$$\mathcal{H}(\eta,\nu)=\frac{\dim_1^{\mathrm{J}}(\eta)\dim_1^{\mathrm{J}}(\eta,\nu)}{\dim_1^{\mathrm{J}}(\nu)},$$

$$d_{mn}^{\theta}(t)=\sum_{k=n}^{m}\mathrm{e}^{-\frac{1}{2}k(k+\theta-1)t}(-1)^{k-n}\frac{(2k+\theta-1)(n+\theta)_{(k-1)}}{n!(k-n)!}\frac{m_{[k]}}{(\theta+m)_{(k)}}.$$

Proof. By the definition of $|\mathcal{D}_t|$ and [17, Proposition 2.1], we know that $\mathbb{P}(|\mathcal{D}_t|=n\mid|\mathcal{D}_0|=m)=d_{mn}^{\theta}(t)$. Then, $\mathbb{P}_{\nu}(\mathcal{D}_t=\eta)=\mathbb{P}(\mathcal{D}_t=\eta\mid\mathcal{D}_0=\nu)=\mathbb{P}(\mathcal{D}_t=\eta\mid|\mathcal{D}_t|=n,\mathcal{D}_0=\nu)\mathbb{P}(|\mathcal{D}_t|=n\mid\mathcal{D}_0=\nu)=d_{mn}^{\theta}(t)\mathbb{P}(\mathcal{D}_t=\eta\mid|\mathcal{D}_t|=n,\mathcal{D}_0=\nu)$. For $\nu\in\Gamma_m$, $\eta\in\Gamma_n$, there are paths of length $k=m-n$ joining ν and η. These paths are the realizations of the embedded down Markov chain $D_n^{\mathrm{J},1}$. Thus,

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

$$\mathbb{P}(\mathcal{D}_t=\eta \mid |\mathcal{D}_t|=n, \mathcal{D}_0=\nu) = \sum_{\eta=\eta(k)\subset\cdots\subset\eta(0)=\nu} \prod_{i=0}^{k-1} p_1^{\downarrow,J}(\eta(i), \eta(i+1))$$

$$= \sum_{\eta=\eta(k)\subset\cdots\subset\eta(0)=\nu} \prod_{i=0}^{k-1} \frac{\dim_1^{\mathrm{J}}(\eta(i+1))\chi_1^{\mathrm{J}}(\eta(i), \eta(i+1))}{\dim_1^{\mathrm{J}}(\eta(i))}$$

$$= \frac{\dim_1^{\mathrm{J}}(\eta(k))}{\dim_1^{\mathrm{J}}(\eta(0))} \sum_{\eta=\eta(k)\subset\cdots\subset\eta(0)=\nu} \prod_{i=0}^{k-1} \chi_1^{\mathrm{J}}(\eta(i), \eta(i+1))$$

$$= \frac{\dim_1^{\mathrm{J}}(\eta)}{\dim_1^{\mathrm{J}}(\nu)} \dim_1^{\mathrm{J}}(\eta, \nu). \qquad \square$$

Proposition 4. *For $j_\eta(\omega; 1)$, $\eta \in \Gamma_m$, we have*

$$\mathbb{E}_{z,z',1}[j_\eta(Y_t; 1) \mid Y_0=\omega] = (\mathbb{E}_{z,z',1} j_\eta)\left[d_{m1}^\theta(t) + \sum_{n=2}^m d_{mn}^\theta(t) \sum_{\zeta\subset\eta, |\zeta|=n} \mathcal{H}(\eta, \zeta)\frac{j_\zeta(\omega; 1)}{\mathbb{E}_{z,z',1} j_\zeta}\right]. \quad (12)$$

Proof. By the duality equation (10), we know that

$$\mathbb{E}_\omega g_\eta(Y_t) = \mathbb{E}_\eta^* g_{\mathcal{D}_t}(\omega) = \sum_{\zeta\subset\eta} \mathbb{P}_\eta(\mathcal{D}_t=\zeta) g_\zeta(\omega) = \sum_{n=1}^m d_{mn}^\theta(t) \sum_{|\zeta|=n, \zeta\subset\eta} \mathcal{H}(\zeta, \eta) g_\zeta(\omega).$$

Then, we have (12) if we replace

$$g_\eta(\omega) = \frac{j_\eta(\omega; 1)}{\mathbb{E}_{z,z',1} j_\eta}, \qquad g_\zeta(\omega) = \frac{j_\zeta(\omega; 1)}{\mathbb{E}_{z,z',1} j_\zeta}. \qquad \square$$

4.2. Proof of Theorem 1 through the dual process method

Next we use Proposition 4 to deduce the transition density (8). By Proposition 2, we know that $\mu_m(d\omega) = \sum_{|\eta|=m} \mathbb{E}[j_\eta(Y_t; 1) \mid Y_0=\omega]\delta_{\widetilde{\eta}/m}(d\omega)$ will converge weakly to the distribution of Y_t, i.e. the transition probability $q(t, \omega, \cdot) = \mathbb{P}_\omega(Y_t \in \cdot)$. By (12), we know that

$$\mu_n(\mathrm{d}\omega) = d_{m1}^\theta(t)\mu_{1,m}(\mathrm{d}\omega) + \sum_{n=2}^m d_{mn}^\theta(t) \sum_{|\zeta|=n} \mu_{n,m}(\mathrm{d}\omega)\frac{j_\zeta(\omega; 1)}{\mathbb{E}_{z,z',1} j_\zeta},$$

where

$$\mu_{1,m}(\mathrm{d}\omega) = \sum_{|\eta|=m} [\mathbb{E}_{z,z',1} j_\eta]\delta_{\widetilde{\eta}/m}(\mathrm{d}\omega),$$

$$\mu_{n,m}(\mathrm{d}\omega) = \sum_{|\eta|=m, \zeta\subset\eta} \mathcal{H}(\eta, \zeta)[\mathbb{E}_{z,z',1} j_\eta]\delta_{\widetilde{\eta}/m}(\mathrm{d}\omega), \qquad n \geq 2.$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

As $m \to \infty$, we need to show that $\mu_{1,m}(\mathrm{d}\omega) \to \mathcal{Z}(z, z', \vartheta)(\mathrm{d}\omega)$ and $\mu_{n,m}(\mathrm{d}\omega) \to j_\zeta(\omega; 1)\mathcal{Z}(z, z', \vartheta)(\mathrm{d}\omega)$. The first claim can be directly obtained from Proposition 2; now we are going to show the second one. For any $f \in C(\Omega)$, we have

$$\begin{aligned}\int_\Omega f(\omega)\mu_{n,m}(\mathrm{d}\omega) &= \sum_{|\eta|=m, \zeta\subset\eta} f\left(\frac{\widetilde{\eta}}{m}\right)\mathcal{H}(\eta, \zeta)(\mathbb{E}_{z,z',1}j_\eta) \\ &= \sum_{|\eta|=m, \zeta\subset\eta} f\left(\frac{\widetilde{\eta}}{m}\right) j_\zeta\left(\frac{\widetilde{\eta}}{m};1\right)(\mathbb{E}_{z,z',1}j_\eta) + O\left(\frac{1}{\sqrt{m}}\right) \\ &\to \int_\Omega f(\omega) j_\zeta(\omega; 1)\mathcal{Z}(z, z', \vartheta)(\mathrm{d}\omega),\end{aligned}$$

where the second equality is due to [8, Theorems 6.1 and 7.1]. Since $\lim_{m\to\infty} d^\theta_{mn}(t) = d^\theta_n(t)$, we have derived the representation in Theorem 1.

5. Further discussion

5.1. Conjectures on diffusions with given stationary measures

The conclusions in this paper indicate that the dual process is only determined by the weights in the branching diagram, whether it be the Kingman graph or the Jack graph. The dual processes are determined by the generator

$$\mathcal{L}f(\eta) = -\frac{n(n-1+\theta)}{2} f(\eta) + \frac{n(n-1+\theta)}{2} \sum_{\zeta\subset\eta, |\eta|=|\zeta|+1} p^\downarrow(\eta, \zeta) f(\zeta).$$

The diffusions, however, depend on the partition structures on the Kingman graph or the Jack graph. Due to Kingman's representation theorem and the representation theorem in Proposition 2, the partition structures are uniquely determined by their representing measures, which will be the stationary distribution of the diffusions constructed through the up–down Markov chains.

More generally, for a given probability measure μ on the Kingman simplex $\overline{\nabla}_\infty$, we can consider the generator $\mathcal{B}_\mu$ defined on an algebra $\mathcal{A}^o_{\mathrm{K}}$ spanned by $\{1, \varphi_k, k \geq 2\}$ as follows:

$$\mathcal{B}^{\mathrm{K}}_\mu g_\eta(x) = -\frac{n(n-1+\theta)}{2} g_\eta(x) + \frac{n(n-1+\theta)}{2} \sum_{\zeta\subset\eta, |\eta|=|\zeta|+1} \frac{\dim^{\mathrm{K}}(\zeta)\chi^{\mathrm{K}}(\zeta, \eta)}{\dim^{\mathrm{K}}(\eta)} g_\zeta(x), \quad (13)$$

where $\{g_\eta(x) = m_\eta(x)/(\mathbb{E}_\mu m_\eta) \mid \eta \in \Gamma\}$ is a linear base and $\mathbb{E}_\mu$ is the expectation with respect to the probability distribution μ. It will uniquely determine the operation of $\mathcal{B}^{\mathrm{K}}_\mu$ on $\mathcal{A}^o_{\mathrm{K}}$.

Conjecture 1. *For a given probability measure μ on the Kingman simplex $\overline{\nabla}_\infty$, the generator defined in (13) will determine a reversible diffusion W^{K}_t with the stationary distribution μ. Its transition density is $p^{\mathrm{K}}(t, x, y) = d^\theta_0(t) + d^\theta_1(t) + \sum_{n=2}^\infty d^\theta_n(t) p_n(x, y)$, where*

$$p_n(x, y) = \sum_{|\eta|=n} \frac{\dim^{\mathrm{K}}(\eta) m^o_\eta(x) \dim^{\mathrm{K}}(\eta) m^o_\eta(y)}{\mathbb{E}_\mu[\dim^{\mathrm{K}}(\eta) m^o_\eta]}.$$

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

Moreover, for a given probability measure μ in the Thoma simplex Ω, we can also consider the generator $\mathcal{B}^{\mathrm{J}}_{\mu}$ defined on an algebra $\mathcal{A}^{o}_{\mathrm{J}}$ spanned by $\{1, \varphi^{o}_{k}, k \geq 2\}$ as follows:

$$\mathcal{B}^{\mathrm{J}}_{\mu} g_{\eta}(x) = -\frac{n(n-1+\theta)}{2} g_{\eta}(x) + \frac{n(n-1+\theta)}{2} \sum_{\zeta \subset \eta, |\eta|=|\zeta|+1} \frac{\dim^{\mathrm{J}}_{\vartheta}(\zeta) \chi^{\mathrm{J}}_{\vartheta}(\zeta, \eta)}{\dim^{\mathrm{J}}_{\vartheta}(\eta)} g_{\zeta}(x), \quad (14)$$

where $\{g_{\eta}(x) = j_{\eta}(x; \vartheta)/(\mathbb{E}_{\mu} j_{\eta}) \mid \eta \in \Gamma\}$ is a linear base. It will uniquely determine the operation of $\mathcal{B}^{\mathrm{J}}_{\mu}$ on $\mathcal{A}^{o}_{\mathrm{J}}$.

Conjecture 2. *For a given probability measure μ on the Thoma simplex Ω, the generator defined in (14) will determine a reversible diffusion W^{J}_{t} with the stationary distribution μ. Its transition density is $p^{\mathrm{J}}(t, x, y) = d^{\theta}_{0}(t) + d^{\theta}_{1}(t) + \sum_{n=2}^{\infty} d^{\theta}_{n}(t)\mathcal{K}_{n}(x, y)$, where*

$$\mathcal{K}_{n}(x, y) = \sum_{|\eta|=n} \frac{j_{\eta}(x; \vartheta) j_{\eta}(y; \vartheta)}{\mathbb{E}_{\mu}[j_{\eta}]}.$$

Appendix A. Symmetric functions

In this appendix we discuss a few facts about the symmetric functions. Please refer to [12] for further details. Symmetric functions are defined as inverse limits of symmetric polynomials. For $n \in \mathbb{Z}_{+}$, let Λ_{n} be the ring of symmetric polynomials of variables $x_{1}, \ldots, x_{n}$. Define $\rho_{n+1,n} \colon \Lambda_{n+1} \to \Lambda_{n}$ as follows:

$$\rho_{n+1,n}(f(x_{1}, \ldots, x_{n}, x_{n+1})) = f(x_{1}, \ldots, x_{n}, 0). \quad (15)$$

Then Λ is the inverse limit of Λ_{n}, $n \geq 1$.

A.1. Monomial symmetric functions

Consider the symmetric polynomials $m^{n}_{\eta}(x_{1}, \ldots, x_{n}) = \sum_{\eta=(a_{(1)},\ldots,a_{(n)})} x_{1}^{a_{1}} \cdots x_{n}^{a_{n}}$, where $(a_{(1)}, \ldots, a_{(n)})$ is the descending arrangement of $(a_{1}, \ldots, a_{n})$. We can see that (15) is also true for $m^{n}_{\eta}(x_{1}, \ldots, x_{n})$. We can then define $m_{\eta}(x)$ as an inverse limit of $\{m^{n}_{\eta} \mid n \geq 1\}$; this is called a *monomial symmetric function*. The evaluation of $m_{\eta}(x)$ in the Kingman simplex $\overline{\nabla}_{\infty}$ can be performed through continuous extension of $m_{\eta}|_{\nabla_{\infty}}$. Moreover, $m_{\eta}(x)$ satisfies the Pieri formula $m_{\eta}(x) = \sum_{\eta \subset \nu, |\eta|+1=|\nu|} \chi^{\mathrm{K}}(\eta, \nu) m_{\nu}(x)$, and $1 = \sum_{|\eta|=n} \dim^{\mathrm{K}}(\eta) m_{\eta}(x)$.

A.2. Shur functions

Denote a symmetric group as S_{n}. Consider the symmetric polynomials

$$s^{n}_{\eta}(x_{1}, \ldots, x_{n}) = \frac{a^{n}_{\eta+\delta}(x_{1}, \ldots, x_{n})}{a_{\delta}(x_{1}, \ldots, x_{n})},$$

where $\eta = (\eta_{1}, \ldots, \eta_{n})$ and $\delta = (n-1, n-2, \ldots, 1, 0)$ are integer partitions, and $a_{\eta}(x_{1}, \ldots, x_{n}) = \sum_{\sigma \in S_{n}} \mathrm{Sgn}(\sigma) x^{\eta_{1}}_{\sigma(1)} \cdots x^{\eta_{n}}_{\sigma(n)}$. We can also show that $\{s^{n}_{\eta} \mid n \geq 1\}$ satisfy (15). Then, the Shur function $s_{\eta}(x)$ is defined to be the inverse limit of $\{s^{n}_{\eta} \mid n \geq 1\}$. Moreover, the Shur functions also satisfy the Pieri formula, $s_{\eta}(x) = \sum_{\eta \subset \nu, |\eta|+1=|\nu|} \chi^{\mathrm{J}}_{1}(\eta, \nu) s_{\nu}(x)$, and $1 = \sum_{|\eta|=n} \dim^{\mathrm{J}}_{1}(\eta) s_{\eta}(x)$, where $\dim^{\mathrm{J}}_{1} = n!/H(\eta; 1)$.

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press

A.3. Jack functions

Jack polynomials $J^n_\lambda(x_1, \ldots, x_n; \vartheta)$ are defined to be the eigenfunctions of the Sekiguchi operators:

$$D(u; \vartheta) = \frac{1}{\prod_{i<j}(x_i - x_j)} \det\left[x_i^{n-j}\left(x_i \frac{\partial}{\partial x_i} + (n-j)\vartheta + u\right)\right]_{1\le i,j\le n},$$

$$D(u; \vartheta)J^n_\lambda(x; \vartheta) = \left[\prod_{i=1}^{n}(\lambda_i + (n-i)\vartheta + u)\right]J^n_\lambda(x; \vartheta).$$

We can also show that $\{J^n_\eta(x; \vartheta) \mid n \ge 1\}$ satisfy (15). Then the Jack symmetric function $J_\lambda(x_1, \ldots, x_n, \ldots; \vartheta)$ is defined to be the inverse limit of $\{J^n_\eta(x;\vartheta) \mid n \ge 1\}$. Moreover, the Jack functions satisfy the Pieri formula, $J_\eta(x;\vartheta) = \sum_{\eta\subset\nu,|\eta|+1=|\nu|} \chi^{\mathrm{J}}_\vartheta(\eta, \nu)J_\nu(x;\vartheta)$, and $\left(\sum_{i=1}^{\infty} x_i\right)^n = \sum_{|\eta|=n} \dim^{\mathrm{J}}_\vartheta(\eta)J_\eta(x; \vartheta)$. In particular, when $\vartheta = 1$, $J_\eta(x;1) = s_\eta(x)$.

Acknowledgements

The author would like to thank Stewart N. Ethier for introducing him to [11]. Moreover, gratitude should also go to Robert C. Griffiths for many stimulating discussions on dual processes. Last, but not least, the author would like to thank the two referees for their insightful comments.

Funding information

This research is supported by NSFC: 11701570.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BLEI, D. M., NG, A. Y. AND JORDAN, M. I. (2003). Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022.

[2] BORODIN, A. AND OLSHANSKI, G. (2005). Z-measures on partitions and their scaling limits. *Europ. J. Combin.* **26**, 795–834.

[3] BORODIN, A. AND OLSHANSKI, G. (2009). Infinite-dimensional diffusions as limits of random walks on partitions. *Prob. Theory Relat. Fields* **144**, 281–318.

[4] ETHIER, S. N. (1992). Eigenstructure of the infinitely-many-neutral-alleles diffusion model. *J. Appl. Prob.* **29**, 487–498.

[5] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Process: Characterization and Convergence*. Wiley, New York.

[6] FENG, S., SUN, W., WANG, F. Y. AND XU, F. (2011). Functional inequalities for the two-parameter extension of the infinitely-many-neutral-alleles diffusion. *J. Funct. Anal.* **260**, 39–413.

[7] GRIFFITHS, R., SPANÓ, D., RUGGIERO, M. AND ZHOU, Y. Dual process in the two-parameter Poisson–Dirichlet diffusion. Preprint, arXiv:2102.08520.

[8] KEROV, S., OKOUNKOV, A. AND OLSHANSKI, G. (1988). The boundary of Young graph with Jack edge multiplicities. *Int. Math. Res. Not.* **1998**, 173–199.

[9] KINGMAN, J. F. C. (1975). The random discrete distributions. *J. R. Statist. Soc.* B **37**, 1–22.

[10] KINGMAN, J. F. C. (1978). The representation of partition structures. *J. London Math. Soc.* **18**, 374–380.

[11] KOROTKIKH, S. Y. (2020). Transition functions of diffusion processes on the Thoma simplex. *Funct. Anal. Appl.* **54**, 118–134.

[12] MACDONALD, I. (1995). *Symmetric Functions and Hall Polynomials*. Cambridge University Press.

[13] OLSHANSKI, G. (2009). Anisotropic Young diagrams and infinite-dimensional diffusion processes with the Jack parameter. *Int. Math. Res. Not.* **2010**, 1102–1166.
[14] PETROV, L. (2009). A two-parameter family of infinite-dimensional diffusions in the Kingman simplex. *Funct. Anal. Appl.* **43**, 45–66.
[15] PITMAN, J. (1995). Exchangeable and partially exchangeable random partitions. *Prob. Theory Relat. Fields* **102**, 145–158.
[16] TAVARÉ, S. (1984). Line-of-descent and genealogical processes, and their application in population genetics models. *Theoret. Pop. Biol.* **26**, 119–164.
[17] ZHOU, Y. (2015). Ergodic inequality of a two-parameter infinitely-many-alleles diffusion model. *J. Appl. Prob.* **52**, 238–246.

https://doi.org/10.1017/jpr.2022.92 Published online by Cambridge University Press