

DEPARTMENT OF
STATISTICS

Up-down ordered Chinese Restaurant Processes: Representations and Asymptotics

Submitted for the degree of DPhil in Statistics

April 3rd 2020

Dane Rogers
Department of Statistics
Merton College
University of Oxford

Abstract

We study composition-valued Markov chains that appear naturally in the framework of up-down Chinese Restaurant Processes (CRPs), introducing first discrete time models as in a similar way to those introduced by Petrov (2009), and also in continuous time via a Poissonisation technique similar to that used in Pal (2011).

As time evolves in the continuous time model, new customers arrive (up-step) and existing customers leave (down-step) at suitable rates derived from the ordered CRP of Pitman and Winkel (2009). We relate such up-down CRPs to the splitting trees of Lambert (2010) inducing spectrally positive Lévy processes. Conversely, we develop a theorem of Ray–Knight type to recover more general up-down CRPs from the heights of Lévy processes with jumps marked by integer-valued paths. We further establish limit theorems for the Lévy process and the integer-valued paths that connects to work by Forman et al. (2018+) on interval partition diffusions.

Fulman 2009 shows that the framework of commutation relations can be used to study the behaviour of discrete time down-up Markov chains, and he applies this to the unordered up-down (α, θ)-CRP from Pitman 2006. We demonstrate that these techniques still yield valid results when applied to the ordered down-up (α, θ)-CRP of Pitman and Winkel 2009, and generalise a theorem of Fulman 2009 on the explicit form and asymptotics of the maximal separation distance from the case $\theta = 1$ to general the case $\theta > 0$.

CONTENTS

1. Introduction 7

1.1 Ordered up-down Chinese Restaurants 7

1.2 Structure of the thesis 9

1.2.1 Chapter 2: Ray-Knight representation of up-down Chinese Restaurant Processes 10

1.2.2 Further directions for Chapter 2 14

1.2.3 Chapter 3: Mixing times and the Ordered Kingman lattice . . 15

2. Ray-Knight representation of up-down Chinese Restaurant Processes . . . 17

2.1 Introduction 17

2.2 Chinese Restaurants and related processes 24

2.2.1 Discrete-time Chinese Restaurants 24

2.2.2 Continuous-time Chinese Restaurants 29

2.3 Results on splitting tree connections 36

2.3.1 Preliminaries on splitting trees 36

2.3.2 The JCCP of the ordered Poissonised up-down CRP 41

2.3.3 Construction of skewer processes 43

2.3.4 The skewer is a Poissonised up-down CRP 47

2.3.5 Generalised Chinese Restaurants 50

2.4 Scaling limits and $(1+\alpha)$-stable Lévy processes 52

2.4.1 Relevant material from the literature 52

2.4.2 Results on convergence to scaling limits 58

3. Mixing times and the Ordered Kingman lattice 65
3.1 Introduction 65
3.2 Main results from Fulman 69
3.2.1 Markov chains on branching diagrams 69
3.2.2 Commutation relations 77
3.2.3 Results on the unordered down-up Chinese restaurant 79
3.3 New results for the ordered lattice 81

Acknowledgements

Firstly, I wish to thank my supervisor Matthias Winkel, for his support, guidance and patience throughout my time in Oxford. I also wish to thank my transfer and confirmation of status examiners Julien Berestycki; Christina Goldschmidt and David Steinsaltz, and to thank my undergraduate project supervisors Andreas Kyprianou and Gunnar Traustason for their supervision while an undergraduate at Bath; I also wish to thank my thesis examiners Christina Goldschmidt and Loïc Chaumont for being willing to examine this thesis. I also wish to thank the EPSRC for providing funding during the first 3.5 years of my course, and also to Merton College for providing an additional doctoral completion bursary of roughly £2,500.

I wish to thank St Hilda's College for offering me a teaching contract during 2017-2019, and also the undergraduate students I taught there as well, as I found it a highly rewarding and entertaining task teaching them. Thanks are also due to my office mates both current and previous for the engaging environment, and the statistics department more broadly for being welcoming and providing copious amounts of cake.

I also wish to thank my maths teacher Ken Davis for being willing to take me on as an additional student for maths A-levels while still home educated. Without him it is entirely possible that I would not have gone to Bath for my undergraduate degree, and it is almost certainly the case that I would have been made to suffer the far higher tuition fees of over £9,000/year; credit is equally due to my parents for being willing to make large sacrifices and move all over the place to support me during my time at Bath. I also wish to thank numerous friends both here and at Bath, for making my time at both universities enjoyable, and also especially those who brought me to Christ while I was an undergraduate at Bath.

Finally, I also wish to thank other members of the Oxford Climate Justice Campaign for working towards a more just university not funded by or supportive of fossil fuel companies, who obviously do not get any thanks for causing the climate

emergency. Acknowledgements and thanks are most certainly not due in any form to Nigel Farage or the two most recent leaders of the Conservative Party for causing me stress near the end of my degree over the prospect of a no-deal Brexit and the potential loss of EU citizenship and free movement rights.

1. INTRODUCTION

1.1 Ordered up-down Chinese Restaurants

One of the main objectives of this thesis is the study of a class of continuous-time Markov chains in the state space

$$\mathcal{C} := \{(n_1, \ldots, n_k) \colon k \geq 0,\ n_1, \ldots, n_k \geq 1\}$$

of integer compositions, which includes, for $k = 0$, the empty vector that we also denote by $\varnothing$. Such Markov chains arise naturally in the framework of the Dubins–Pitman two-parameter Chinese Restaurant Process (CRP) [38], when considered with the additional order structure of Pitman and Winkel [40]. Specifically, we interpret $n_1, \ldots, n_k$ as the numbers of customers at an ordered list of k tables in a restaurant, short *table sizes*. In the most basic model, we allow only the following transitions from $(n_1, \ldots, n_k) \in \mathcal{C}$.

- At rate $n_i - \alpha$, a new customer joins the i-th table, leading to a transition into state $(n_1, \ldots, n_{i-1}, n_i + 1, n_{i+1}, \ldots, n_k)$, $1 \leq i \leq k$.
- At rate α a new customer opens a new table inserted directly to the left of the i-th table, leading to state $(n_1, \ldots, n_{i-1}, 1, n_i, n_{i+1}, \ldots, n_k)$, $1 \leq i \leq k$.
- At rate θ a new customer opens a new table inserted in the right-most position, leading to state $(n_1, \ldots, n_k, 1)$.
- At rate 1 each existing customer at table i, $1 \leq i \leq k$, leaves, leading to either $(n_1, \ldots, n_{i-1}, n_i{-}1, n_{i+1}, \ldots, n_k)$ if $n_i \geq 2$ or $(n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k)$ if $n_i = 1$.

These transition rates give rise to a $\mathcal{C}$-valued continuous-time Markov chain if

$0 \le \alpha \le 1$ and $\theta \ge 0$, which we call a *continuous-time up-down ordered Chinese Restaurant Process with parameters* (α, θ), or, for the purposes of this paper, just an *up-down OCRP*(α, θ).

The first three bullet points increase the number of customers and we call any one of them an *up-step*, while the last bullet point decreases their number, a *down-step*. Conditionally given an up-step, the (induced discrete-time) transition probabilities are those of an ordered CRP [40]. The up-step rates have been related to the usual Dubins–Pitman CRP [38, Section 3.4] without a table order which combines the middle two bullet points to a rate $k\alpha + \theta$ for a new table; see also [28] for a related process. See [27, 21, 41] for other discussions of order structures related to CRPs. It was shown in [40, Proposition 6] that starting from $(1) \in \mathcal{C}$, the distribution p_{n+1} after n consecutive up-steps is the same as a Gnedin–Pitman [21] regenerative composition structure that is known to be weakly sampling consistent, meaning that the push-forward of p_{n+1} under a down-step is also p_n. Specifically writing $N_i = n_i + \cdots + n_k$,

$$p_n(n_k, \ldots, n_1) = \prod_{i=1}^{k} r(N_i, n_i), \quad \text{where } r(n,m) = \binom{n}{m} \frac{(n-m)\alpha + m\theta}{n} \frac{(1-\alpha)_{(m-1)\uparrow}}{(n-m+\theta)_{m\uparrow}},$$

for $1 \le m \le n$, with $x_{y\uparrow} := x(x+1)\cdots(x+y-1)$. This suggests to consider a *discrete-time up-down, or alternatively down-up OCRP*(α, θ) on the subspace $\mathcal{C}_n$ of $\mathcal{C}$ with $\mathcal{C}_n := \{(n_1, \ldots, n_k) \in \mathcal{C} \colon n_1 + \cdots + n_k = n\}$, $n \ge 1$, in which each transition consists of an up-step followed by a down-step or respectively a down-step followed by a up-step. Then p_n is stationary. Similar up-down chains on related state spaces were studied in [7, 17], and for the corresponding discrete-time up-down CRP without the order of tables, Petrov [37] noted that the stationary distribution, is the push-forward of p_n under the map that ranks a vector into decreasing order. Petrov's main result was the existence of a diffusive scaling limit of his up-down chain, when represented in the infinite-dimensional space of decreasing sequences with sum 1.

While Petrov [37] used analytic methods to study how generators act on a certain core of symmetric functions, we develop here probabilistic methods in order to study scaling limits associated with the (continuous-time) up-down or down up OCRP(α, θ). As demonstrated by Pal [36] in a finite-dimensional setting, asymptotic results for continuous-time Markov chains associated with discrete-time Markov chains via a method called "Poissonization" can sometimes be used to deduce asymptotic results for the discrete-time Markov chains via what can be referred to as "de-Poissonization".

In this thesis, there is no direct exploration of the passage between discrete and continuous time. We relate the (continuous-time) up-down OCRP(α, θ) to genealogical trees [20] and their jumping chronological contour processes (JCCPs) [31] that lead to representations as spectrally positive Lévy processes, the jumps of which we mark by integer-valued paths corresponding to table size evolutions. We show that the up-down OCRP(α, θ) and natural generalisations can be recovered from the heights of such a marked Lévy process. This result is reminiscent of the recovery of a geometric Galton–Watson process from the occupation measure (upcrossing counts) of a suitably stopped simple symmetric random walk [29] or indeed, in the scaling limit the recovery of a Feller diffusion (specifically a squared Bessel process of dimension 0) as the local time process of a stopped Brownian motion; see also [5] for an analogous result involving continuous state branching processes. In our framework, we similarly establish scaling limits of the Lévy process and of the integer-valued paths marking its jumps in order to connect to the framework of Forman et al. [15, 12, 13].

1.2 Structure of the thesis

The structure of the thesis is as follows. Each of the two main chapters of the thesis is roughly a self-contained paper, in which no reference will be made to this introduction. In section 1.2.1 we provide specific material on the results of Chapter

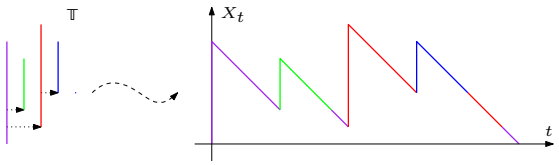


Fig. 1.1: The relationship between a genealogical tree $\mathbb{T}$ and its JCCP X.

2, with section 1.2.2 providing a short discussion of potential future directions for research, and in section 1.2.3 we discuss specific background material on the results of Chapter 3. Research for both papers was done with the supervision of Professor Matthias Winkel, and Chapter 2 is intended to be submitted to a journal in modified form as [45].

1.2.1 Chapter 2: Ray-Knight representation of up-down Chinese Restaurant Processes

Let us first explain how an up-down OCRP$(\alpha, 0)$, which we will denote by $(\mathcal{S}^y)_{y\geq 0}$, induces a genealogy. The rates stated at the beginning of the introduction are such that every table evolves in size as an integer-valued Markov chain with up-rate $m-\alpha$ and down-rate m when in state $m \geq 1$, and is removed when hitting state 0. Hence, tables have "death" times. Furthermore, while a table is "alive", new tables are inserted ("children born") directly to its left at rate α. Note the recursive nature of the model and the independence of table size evolutions. This is the genealogy [20, 31] of a binary homogeneous Crump–Mode–Jagers (CMJ) branching process. Figure 1.1 captures this in a tree $\mathbb{T}$ with a vertical line for each table and horizontal arrows linking each parent table to its children at heights/levels corresponding to birth times.

Travelling around this genealogical tree $\mathbb{T}$ recording heights as in Figure 1.1 yields Lambert's [31] jumping chronological contour process (JCCP) $(X)_{t\geq 0}$: each vertical line yields a jump from a birth level to a death level, exploration is by sliding down at unit speed and recursively jumping up at the birth levels of children. As jump heights are i.i.d. table lifetimes and jumps occur at the rate α of table insertions, this process is a Lévy process starting from an initial table lifetime and stopped when reaching

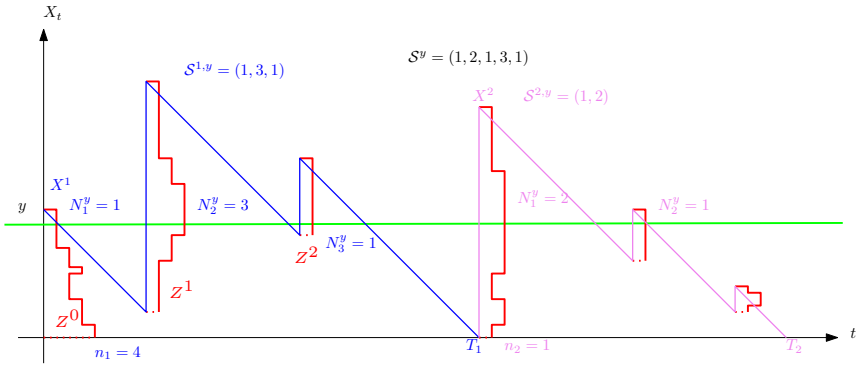


Fig. 1.2: The skewer at level y of the marked JCCP when $\theta = 0$.

0. We further mark each jump U_i of the JCCP, of height $\zeta_i = X_{U_i} - X_{U_i^-}$ by the table size evolution $(Z^i(s))_{0 \le s \le \zeta_i}$, as in Figure 1.2. In a general framework of CMJ processes, this is what Jagers [25, 26] studied: each individual has "characteristics" that vary during its lifetime. Key quantities of interest in a CMJ process are the characteristics $Z^i(y - X_{U_i^-})$ at each time $y \in [U_i^-, U_i]$, or summary statistics such as sums of these characteristics.

Conversely, in a marked JCCP setting we can at each level y extract a composition $\mathcal{S}^y$ by listing from right to left for each jump crossing level y the size given by the marks for that level. In this construction, we refer to $(\mathcal{S}^y)_{y \ge 0}$ as the *skewer process* as we imagine piercing the marked JCCP at level y and pushing together all sizes that we find at this level to form a sequence without gaps. See Figure 1.2. This is a discrete analogue of the skewer process introduced in [12]. We provide a rigorous set-up in Section 2.3.3 and prove carefully the following result, formulated here in the setting of size evolutions on $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$, which we call Q_α-Markov chains. These have Q-matrix Q_α whose only non-zero off-diagonal entries are $q_{m,m+1} = m - \alpha$ and $q_{m,m-1} = m$ for $m \ge 1$, and $q_{0,m} = 0$ for $m \ge 0$.

Theorem 1.2.1. *Let $0 \le \alpha \le 1$ and $n_0 \ge 1$. Let Z^i be independent Q_α-Markov chains with $Z^0(0) = n_0$ and $Z^i(0) = 1$ for $i \ge 1$. Let $\zeta_i = \inf\{s \ge 0 \colon Z^i(s) = 0\}$ be the absorption times. Let $(J_t)_{t \ge 0}$ be an independent Poisson process of rate α,*

$$X_t = -t + \sum_{i=0}^{J_t} \zeta_i, \quad t \geq 0, \qquad \textit{and} \qquad T = \inf\{t \geq 0 \colon X_t = 0\}. \tag{1.2.1}$$

Consider $(X, Z) := ((X_t)_{0 \leq t \leq T}, (Z^i)_{0 \leq i \leq J_T})$*, with* Z^i *marking the* i*-th jump of* X*. Then the skewer process of* (X, Z) *is an up-down OCRP*$(\alpha, 0)$ *starting from* $(n_0) \in \mathcal{C}$*.*

We can also generalise this in three further ways: first to start from any $(n_1, \ldots, n_k) \in \mathcal{C}$, we concatenate k independent copies of (X, Z) replacing n_0 by $n_1, \ldots, n_k$, respectively. Second, to obtain an up-down OCRP(α, θ) for $\theta > 0$, we add a similar construction $((X_t)_{t<0}, (Z^{-i})_{i \geq 1})$ to provide right-most tables, as well as their "children" and further "descendants"; see Figure 1.3 or Theorem 2.3.11 in Section 2.3.3 for details. Third, if we replace Q_α by other Q-matrices on $\mathbb{N}_0$ subject to conditions that ensure appropriate absorption in 0, and appropriately relax the restriction that new tables always start from a single customer, we still obtain a Markovian skewer process, which we refer to as a *generalised up-down OCRP*(α, θ). See Section 2.3.4 for details. We remark that absorption times of continuous-time Markov chains are called "phase-type distributions." The JCCPs associated with phase-type jumps have been studied in other contexts [4].

Returning to the setup of Theorem 1.2.1, we establish distributional scaling limits for the table size evolution and the total number of customers in an up-down OCRP(α, θ), as the initial number of customers tends to infinity, and for the Lévy process $(X_t)_{t \geq 0}$. To formulate the limit theorems, let us discuss squared Bessel processes starting from $a \in (0, \infty)$ with dimension parameter $\delta \in \mathbb{R}$. We follow [22, 44] and consider the unique strong solution of the stochastic differential equation (SDE)

$$dY(s) = \delta \, ds + 2\sqrt{|Y(s)|} \, dW(s), \qquad Y(0) = a, \tag{1.2.2}$$

driven by a standard Brownian motion $(W(s))_{s \geq 0}$. For $\delta \geq 0$ this exists as a non-negative stochastic process for all $s \geq 0$, never reaching 0 when $\delta \geq 2$, reflecting from 0 when $\delta \in (0, 2)$ and absorbed in 0 when $\delta = 0$. We denote this distribution

by $BES_a^2(\delta)$. For $\delta < 0$, let $\zeta = \inf\{s \geq 0 \colon Y(s) = 0\}$ and note that the $BES_a^2(\delta)$ distribution of the process now passes through 0 to $(-\infty, 0]$, and almost surely either returns to reflect from 0 infinitely often or not depending on if $\delta \in (2, 0)$ or $\delta \leq -2$.

Theorem 1.2.2. *For all $0 \leq \alpha \leq 1$, the table size evolution Z_n, i.e. a Q_α-Markov chain, starting from $Z_n(0) = \lfloor nz \rfloor$, satisfies*

$$\left(\frac{Z_n(2ns)}{n}\right)_{s\geq 0} \xrightarrow{D} \texttt{BESQ}_z(-2\alpha) \qquad \textit{under the Skorokhod topology,}$$

where $\texttt{BESQ}_z(-2\alpha)$ is a modified $BES_z^2(-2\alpha)$ process absorbed at the first hitting time of 0. Furthermore, the convergence holds jointly with the convergence of hitting times of 0.

We remark that the convergence of hitting times is not automatic since hitting times are not continuous functionals on the Skorokhod space described in [24].

Theorem 1.2.3. *For all $0 \leq \alpha \leq 1$ and $\theta \geq 0$, the evolution M_n of the total number of customers in an up-down OCRP(α, θ), i.e. a continuous-time Markov chain whose non-zero off-diagonal entries are $q_{m,m+1} = m + \theta$ and $q_{m,m-1} = m$, $m \geq 0$, starting from $M_n(0) = \lfloor na \rfloor$, satisfies*

$$\left(\frac{M_n(2ns)}{n}\right)_{s\geq 0} \xrightarrow{D} \texttt{BESQ}_a(2\theta) \qquad \textit{under the Skorokhod topology.}$$

Theorem 1.2.4. *For all $0 < \alpha < 1$, the Lévy process $(X_t)_{t\geq 0}$ of (1.2.1) satisfies*

$$\left(\frac{X_{2n^{1+\alpha}t}}{2n}\right)_{t\geq 0} \xrightarrow{D} \texttt{Stable}(1+\alpha) \qquad \textit{under the Skorokhod topology,}$$

where $\texttt{Stable}(1+\alpha)$ is the distribution of a spectrally positive stable Lévy process with Laplace exponent $\psi(\lambda) = \lambda^{1+\alpha}/2^\alpha\Gamma(1+\alpha)$.

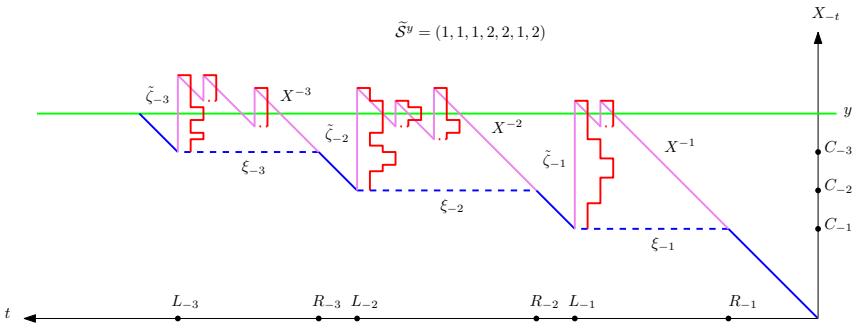


Fig. 1.3: The skewer at level y of the modified JCCP when $\theta > 0$.

1.2.2 *Further directions for Chapter 2*

Theorems 1.2.2 and 1.2.4 connect to work by Forman et al. [15, 12, 13], where the starting point is a σ-finite $\texttt{BESQ}(-2\alpha)$ excursion measure Θ due to Pitman and Yor [42]. Specifically, a $\texttt{Stable}(1+\alpha)$ Lévy process $(\overline{X}_t)_{t\geq 0}$ is constructed from some $\overline{Z}^0 \sim \texttt{BESQ}_z(-2\alpha)$ and a Poisson Random Measure $\overline{N} := \sum_{t>0:\, \overline{Z}^t \not\equiv 0} \delta_{t,\overline{Z}^t}$ with intensity measure Leb $\times\, \Theta$ by using the excursion lengths $\overline{\zeta}_t := \inf\{s \geq 0 \colon \overline{Z}^t_s = 0\}$ as jump heights in a compensating limit

$$\overline{X}_t = \lim_{\varepsilon \downarrow 0} \left(\sum_{0\leq r\leq t:\, \overline{\zeta}_r > \varepsilon} \overline{\zeta}_r \;-\; t \frac{1+\alpha}{2^\alpha \Gamma(1-\alpha)\Gamma(1+\alpha)} \varepsilon^{-\alpha} \right).$$

Furthermore, an interval partition diffusion is extracted from $(\overline{X}, \overline{N})$ using a continuous analogue of the skewer process described above. This interval partition diffusion is called a type-1 evolution. Type-0 evolutions are obtained by specifying a semi-group that has further intervals added on the left-hand side in a way that achieves some left-right symmetry. It is expected that further generalisations will be explored in [11].

Conjecture 1.2.5. *Let $(C^n_j)_{j\geq 0}$ be a $\mathcal{C}_n$-valued discrete-time up-down OCRP(α,θ), for each $n \geq 1$. If $n^{-1}C^n_0$ converges, as $n \to \infty$, then $(n^{-1}C^n_{[n^2 y]})_{y\geq 0}$ has a diffusive scaling limit, when suitably represented on a space of interval partitions. For $0 < \theta = \alpha < 1$ and $0 = \theta < \alpha < 1$, these are the de-Poissonised type-0 and type-1 evolutions of [13].*

Conjecture 1.2.6. *Let $(\mathcal{S}^y_n)_{y\geq 0}$ be a $\mathcal{C}$-valued continuous-time up-down OCRP(α,θ),*

for each $n \geq 1$. If $n^{-1}\mathcal{S}_n^0$ converges, as $n \to \infty$, then $(n^{-1}\mathcal{S}_n^{ny})_{y\geq 0}$ has a diffusive scaling limit, when suitably represented on a space of interval partitions. For $0 < \theta = \alpha < 1$ and $0 = \theta < \alpha < 1$, these are the type-0 and type-1 evolutions of [13].

1.2.3 Chapter 3: Mixing times and the Ordered Kingman lattice

In this chapter, we will adapt and generalise Fulman's results [17] on the maximal separation distance after r steps of a down-up unordered (α, θ)-Chinese restaurant Markov chain on the Kingman lattice to the ordered Chinese restaurant, and furthermore extend his results on asymptotics and exact forms of the maximal separation distance from the case $\theta = 1$ to the case $\theta > 0$.

Let K be the space $\{(\lambda_1, \lambda_2, \ldots, \lambda_l, 0, 0, \ldots) : \lambda_i \in \mathbb{N} \text{ if } 1 \leq i \leq l\}$ known as the *Ordered Kingman lattice* with subspace $K_n := \{\lambda \in K : \sum_{i=1}^{l} \lambda_i = n\}$, where $K_0 := \{(0, 0, 0 \ldots)\}$. Note that in Fulman, the Kingman lattice has decreasing elements, this corresponds to the standard two-parameter Dubins-Pitman CRP [38] with the structure of [40].

If P is the transition matrix of a Markov chain X on a finite set I with stationary distribution π, then define the *separation distance* $s(r, x)$ of X started from $x \in I$ and *maximal separation distance* $s^*(r)$ by

$$s(r, x) := \max_{y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right),$$
$$s^*(r) := \max_{x \in I} s(r, x) = \max_{x,y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right).$$

Note that from proposition 2.2.6, we are able to use the fact that an up-down ordered (α, θ)-CRP chain has a stationary distribution along with the consistency condition to conclude that the maximal separation distance is well-defined.

We can study the transitions of these Markov chains as follows via a pair of linear maps. For any set S, $\mathbb{C}S$ denotes the complex vector space with basis S. We define two linear maps $D_n := \mathbb{C}K_n \to \mathbb{C}K_{n-1}$ and $U_n := \mathbb{C}K_n \to \mathbb{C}K_{n+1}$ that capture the transitions in the following way.

Let $p_D(\lambda, \rho)$ and $p_U(\lambda, \tau)$ be the respective probabilities of transitioning from λ to ρ or τ in upward or downward steps respectively, where we assume that $\rho \in K_{n-1}, \lambda \in K_n$ and $\tau \in K_{n+1}$ and define

$$D_n(\lambda) := \sum_{\rho \in K_{n-1}} p_D(\lambda, \rho)\rho \quad \text{and} \quad U_n(\lambda) := \sum_{\tau \in K_{n+1}} p_U(\lambda, \tau)\tau.$$

These maps are such that the coefficients are transition probabilities, and moreover the coefficients of $\mu \in K_n$ in $U_{n-1}D_n(\lambda)$ or $D_{n+1}U_n(\lambda)$ are the probabilities of transitioning from λ to μ after a down-up or up-down step respectively.

Our first main result confirms that the maps satisfy the following commutation relation, so that Fulman's [17] theory applies and is adaptable.

Theorem 1.2.7. *Suppose that $\theta > 0$, $0 \leq \alpha < 1$. For $I_n : \mathbb{C}K_n \to \mathbb{C}K_n$ the identity map, the functions U_n and D_n still satisfy the relation*

$$D_{n+1}U_n = a_n U_{n-1}D_n + b_n I_n$$

with $a_n = \frac{n(\theta+n-1)}{(n+1)(\theta+n)}$ and $b_n = 1 - a_n$ as before.

Our second main result is the following statement about the maximal separation distance. This is a generalisation from the case $\theta = 1$ to the case $\theta > 0$ of [17, Theorem 7.5], and additionally holds in both our framework of the ordered CRP and also Fulman's unordered case.

Theorem 1.2.8. *If $s^*(r)$ is the maximal separation distance after $r \geq 1$ iterations of the down-up chain on the Kingman lattice or ordered Kingman lattice and $0 \leq \alpha < 1$, then if $\theta > 0$,*

$$s^*(r) = \sum_{i=2}^{n} \left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^r \frac{(-1)^i(i-1)(\theta+2i-1)}{\theta+i-1} \binom{n}{i} \frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)}.$$

Moreover for any fixed $c > 0$,

$$\lim_{n\to\infty} s^*(cn(\theta+n-1)) = \sum_{i=2}^{\infty} e^{-ci(\theta+i-1)} \frac{(-1)^i(i-1)(\theta+2i-1)}{i!(\theta+i-1)} \prod_{j=1}^{i}(\theta+j).$$

2. RAY-KNIGHT REPRESENTATION OF UP-DOWN CHINESE RESTAURANT PROCESSES

Abstract

The standard two-parameter (α, θ)-Chinese Restaurant Process is well-studied, and an ordered version was introduced by Pitman and Winkel 2009. We introduce up-down chains, which can be Poissonised in the sense of Pal 2011 and which also yield splitting trees in the sense of Lambert 2010. These yield a spectrally positive Lévy process, and a theorem of Ray-Knight type allows us to recover a more general Ordered Chinese Restaurant up-down chain from the heights of the Lévy process, marked by suitable $\mathbb{N}_0$-valued Markov chains. We further show that scaling the tables yields $BES^2(-2\alpha)$ diffusions, and scaling the Lévy process yields a $(1+\alpha)$-stable Lévy process, connecting to work by Forman, Rizzolo, Pal and Winkel 2017.

2.1 Introduction

We consider unordered (α, θ)-Chinese Restaurant Processes (CRP) [38] as discrete time stochastic processes taking values in the space

$$\mathcal{C} := \{(n_1, \ldots, n_k) : k \geq 1,\ n_i \geq 1,\ \forall i\colon 1 \leq i \leq k\} \cup \{\varnothing\}$$

of compositions. They have the following explicit description for any $0 \leq \alpha \leq 1$ and $\theta > -\alpha$ or $\alpha < 0$ and $\theta = -m\alpha$ for m some positive integer.

From state $(n_1, \ldots, n_k) \in \mathcal{C}_n := \{(n_1, \ldots, n_k) \in \mathcal{C} : n_1 + \cdots + n_k = n\}$, an (α, θ)-CRP moves into

- state $(n_1, \ldots, n_{i-1}, n_i + 1, n_{i+1}, \ldots, n_k) \in \mathcal{C}_{n+1}$ with probability proportional to $n_i - \alpha$,
- state $(n_1, \ldots, n_k, 1) \in \mathcal{C}_{n+1}$ with probability proportional to $k\alpha + \theta$.

For simplicity, assume that all Chinese Restaurants start from the initial state $(1) \in \mathcal{C}_1$.

These compositions can be viewed as arising naturally from a sequence $(\Pi_n)_{n\geq 1}$ of partitions corresponding to labelled customers at tables in a restaurant; at the n-th timestep a customer labelled $n+1$ arrives and forms a new table or sits at an existing table with n_i customers according to the probabilities described above. It is well known [39, p.150] that such a partition is exchangeable and the exchangeable partition probability function $p_{\alpha,\theta}(\#A_1, \ldots, \#A_k) := \mathbb{P}(\Pi_n = (A_1, \ldots, A_k))$ is given by

$$p_{\alpha,\theta}(n_1, \ldots, n_k) = \frac{(\alpha+\theta)_{(k-1)\uparrow\alpha} \prod_{i=1}^{k} (1-\alpha)_{(n_i-1)\uparrow 1}}{(1+\theta)_{(n-1)\uparrow 1}} \text{ for } (n_1, \ldots, n_k) \in \mathcal{C}_n,$$

where $x_{0\uparrow\alpha} := 1$ and $x_{y\uparrow\alpha} := (x + (y-1)\alpha) x_{(y-1)\uparrow\alpha}$, for $y \in \mathbb{N}$.

An up-down chain here is a modification of the CRP where after a new customer arrives, a customer instantaneously leaves at the same time, chosen uniformly at random, and with existing customers relabelled if necessary as described in Section 2. We refer to the resulting process on $\mathcal{C}_n$ as an (α, θ)-up-down chain. Certain classes of up-down and down-up chains were studied in [7], [17] and [19], and Petrov [37] showed that the above function $p_{\alpha,\theta}$ induces the stationary distribution of an up-down chain on the subset $\mathcal{C}_n^{\downarrow}$ of $\mathcal{C}_n$ ordered by decreasing elements. Specifically, the stationary distribution $P(n_1, \ldots, n_k)$ on $\mathcal{C}_n^{\downarrow}$ is given by

$$P(n_1, \ldots, n_k) = \Big(n! / (\prod_{i=1}^{k} n_i! \cdot \prod_{j=1}^{\infty} [(n_1, \ldots, n_k); j]!) \Big) p_{\alpha,\theta}(n_1, \ldots, n_k),$$

where $[(n_1, \ldots, n_k); j] := \#\{i : n_i = j\}$.

In this paper, we study a variation of the CRP introduced by Pitman and Winkel [40], known as the Ordered Chinese Restaurant Process. If we specify instead that when a new table of size 1 is inserted it is inserted to the right of all other tables with weight $\theta \geq 0$ and to the left of any individual table with weight α, the construction is called an Ordered (α, θ)-Chinese Restaurant Process.

The notion of an Ordered up-down CRP is defined in much the same way and is a discrete-time analogue of the down-up chain described in [14]. Explicitly, we add our new customer according to the rules of the Ordered CRP and then uniformly select a customer for immediate removal.

We show that the stationary distribution $p_n(n_1, \ldots, n_k)$ of the Ordered up-down CRP is the distribution of an Ordered CRP on $\mathcal{C}_n$, and describe p_n explicitly. Moreover, we note that these compositions are exchangeable if and only if $\alpha = 1$ or $\alpha = \theta$.

We discuss the notion of Poissonisation introduced by Pal [36] and described more explicitly in [35, Section 1.1] and Forman, Pal, Rizzolo and Winkel [14, Section 1.1]. Replacing the relative probabilities $n_i - \alpha$, α and θ in the definition of the ordered CRP by rates and supposing that each customer leaves independently at rate 1, the number of customers at each table evolves according to a version of a continuous-time Markov chain Z on $\mathbb{N}_0$ with transition rates $q_{n,n-1} = n, q_{n,n+1} = n - \alpha, q_{i,j} = 0$ if $|i - j| > 1$ for $n \geq 1$, $q_{0,j} = 0$ for all $j \geq 0$. Such a process is called a Poissonised Ordered (α, θ)-up-down CRP, or sometimes a Poissonised up-down CRP for short.

Unlike in the discrete setting, the Poissonised Ordered up-down CRP is transient when $\theta > 0$ and consequently no stationary distribution exists; when $\theta = 0$ the up-down chain is absorbed at state $\varnothing \in \mathcal{C}_0$ almost surely.

The Poissonised Ordered CRP can be related to the splitting trees studied by Lambert [31], and introduced previously by Geiger and Kersting [20]. Specifically, our genealogy is such that we view the children of a table of initial size n_i as all

those singleton tables which were ever inserted to the left of the size n_i block.

This paper follows the work of Forman-Pal-Rizzolo-Winkel [14], which constructs continuum analogues to up-down chains on CRPs with $(\alpha, \theta) = \left(\frac{1}{2}, \frac{1}{2}\right)$ or $(\alpha, \theta) = \left(\frac{1}{2}, 0\right)$. In particular, the main results of the present work are generalisations of results suggested informally and without proof in [14, Chapter 1.1]. The convergence results presented here were proved in the $\left(\frac{1}{2}, 0\right)$-case in unpublished work by those authors [16]. In [14] and [35] it is noted that the Poissonised ordered CRP can be studied by way of a sequence of splitting trees.

Each of these trees is individually in bijection with a realisation of a Lévy process $(X_t)_{t\geq 0}$ of the form

$$X_t = -t + \sum_{i=0}^{J_t} \zeta_i,$$

where $(J_t)_{t\geq 0}$ is an independent Poisson process of rate α, Z^i are independent versions of Z started from $Z^0(0) = n_0$, $Z^i(0) = 1$ if $i \geq 1$ and $\zeta_i := \inf\{s \geq 0 : Z^i(s) = 0\}$, representing the evolution of a table size and its lifetime.

We make the following claim. Let T_j be the points of the Poisson process $(J_t)_{t\geq 0}$ and write $B_j = X_{T_{j-}}$, $T := \inf\{t \geq 0 : X_t = 0\}$. To identify the jumps which cross level y define

$$I_1^y := \inf\{i \geq 0 \colon B_i \leq y < B_i + \zeta_i\},\ I_{j+1}^y := \inf\{i \geq I_j^y + 1 : B_i \leq y < B_i + \zeta_i\} \text{ for } j \geq 1,$$
$$K^y := \sup\{0\} \cup \{j \geq 1 \colon T_{I_j^y} \leq T\} \text{ and } N_j^y := Z^{I_j^y}(y - B_{I_j^y}),\ \text{for } 1 \leq j \leq K^y.$$

Note that K^y counts the numbers of jumps which cross y and the N_j^y identify the numbers of customers in each table at level y. Using terminology similar to Forman et al. [14], we show that the skewer $(\mathcal{S}^y)_{y\geq 0}$ given by $\mathcal{S}^y := (N_{K^y}^y, \ldots, N_1^y)$ evolves as a Poissonised Ordered (α, θ)-up-down CRP started from $(n_0) \in \mathcal{C}$. The back to front indexing accounts for the fact that insertions at rate α are located to the left of current tables, while the Lévy process $(X_t)_{t\geq 0}$ is a contour process of planar splitting trees with new branches drawn to the right.

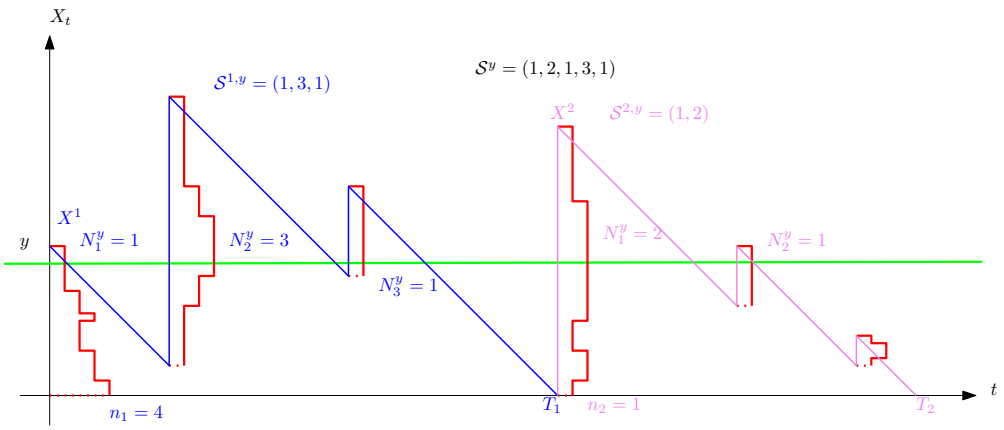


Fig. 2.1: How to find an $(\alpha, 0)$-skewer process from the JCCP.

Indeed, it is possible to extract a more general (α, θ) up-down Chinese Restaurant Process started from any state $(n_1, \ldots, n_k) \in \mathcal{C}$ via a similar procedure. For the k initial tables, consider k independent excursions of X_t above level 0 with the j-th excursion corresponding to the j-th table and with n_0 replaced by n_{k+1-j} in such an excursion. For the tables arriving at rate θ, consider a variation of X_t on the time-set $(-\infty, 0)$ such that excursions of X occur according to a Poisson process of rate θ and our process otherwise slides up at unit speed. Note that as demonstrated by Figures 2.1 and 2.2, the result described above is a theorem of Ray-Knight type, where the precise notation in the pictures is as introduced in Section 3.3. We extend the result to a more general setting, describing the skewer process as belonging to a new class of processes, which are shown to be a generalisation of the Ordered Poissonised (α, θ)-up-down CRP.

The Poissonised (α, θ)-up-down chain is related to Bessel-squared processes as introduced by Revuz and Yor [44] and extended to negative dimensions by Göing-Jaeschke and Yor [22]. Denote by $BES_z^2(\delta)$ the distribution of the process which is the unique strong solution of the SDE

$$dY_t = \delta dt + 2\sqrt{|Y_t|}dW_t,\ Y_0 = y,$$

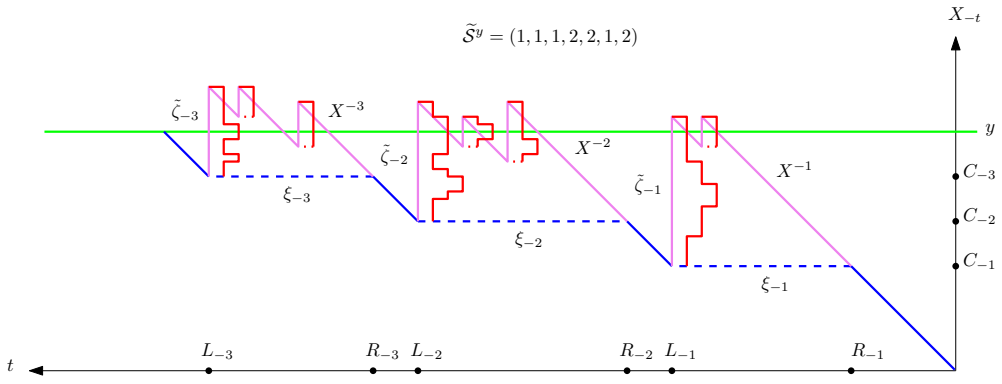


Fig. 2.2: How to account for immigration in finding the (α, θ)-skewer process from the modified JCCP.

for $(W_t)_{t\geq 0}$ a standard Brownian motion on $\mathbb{R}$. Under the Skorokhod topology on the space $\mathbb{D}([0,\infty),\mathbb{R})$ of càdlàg functions $f\colon [0,\infty) \to \mathbb{R}$, we prove that

$$\frac{Z_n(2nt)}{n} \xrightarrow{D} BES_z^2(-2\alpha)(t \wedge \zeta_z^{-2\alpha}) \text{ and } \frac{M_n(2n\cdot)}{n} \xrightarrow{D} BES_y^2(2\theta) \text{ in } \mathbb{D}([0,\infty),\mathbb{R}),$$

where Z_n is a version of the chain Z governing the evolution of a table of initial size $\lfloor nz \rfloor$, $M_n(t)$ is the number of customers in an Ordered Poissonised (α,θ)-up-down CRP with $\lfloor ny \rfloor$ initial customers and $\zeta_z^{-2\alpha}$ is the absorption time at 0 of the $BES_z^2(-2\alpha)$.

We show that there are connections to $(1+\alpha)$-stable Lévy processes. Let Q_z^δ be the law of a $BES_z^2(\delta)$ and ζ the first time at which $0 < z$ is hit. Then as $z \downarrow 0$, $z^{-(1+\alpha)} Q_z^{-2\alpha}(\zeta \in ds) \to \Pi(ds)$ vaguely, where $\Pi(ds) := \frac{s^{-(2+\alpha)}}{2^{1+\alpha}\Gamma(1+\alpha)} ds$ is the Lévy measure of a $(1+\alpha)$-stable Lévy process.

Also, let $X'_t := X_t - X_0$ denote some copy of the Lévy process described previously above, but instead started from 0. Then as $n \to \infty$, $\left((2n)^{-1} X'_{2n^{1+\alpha}t}\right)_{t\geq 0}$ converges in distribution on $\mathbb{D}([0,\infty),\mathbb{R})$ to the $(1+\alpha)$-stable Lévy process with Laplace exponent $\psi(\lambda) = \frac{\lambda^{1+\alpha}}{2^\alpha \Gamma(1+\alpha)}$ and no negative jumps. We also prove that as $n \to \infty$,

$$2\alpha n^{1+\alpha} \mathbb{P}\Big(\frac{\tau(Z)}{2n} > \varepsilon \Big| Z(0) = 1 \Big) \to \Pi(\varepsilon, \infty),$$

where $\tau(Z) := \inf\{t > 0 : Z(t) \leq 0\}$ and Π is the Lévy measure of the $(1+\alpha)$-stable process with Laplace exponent ψ above. It can be shown that

$$2\alpha n^{1+\alpha}\mathbb{P}\Big(\big(n^{-1}Z_{2nt}\big)_{t\geq 0} \in \cdot \Big| Z_0 = 1\Big) \to \Theta(\cdot)$$

vaguely, where Θ is a $BES^2(-2\alpha)$ excursion measure with $\Theta(\tau \in \cdot) = \Pi(\cdot)$. This paper is motivated by constructions of diffusive evolutions on spaces of interval partitions by Forman, Pal, Rizzolo and Winkel, [15, 12, 13] which are continuum analogues of the up-down Chinese restaurant processes that we study here. More specifically, this paper leads to the following conjecture regarding the evolution of such diffusions.

Conjecture 2.1.1. *Let $(\mathcal{S}_n^y)_{y\geq 0}$ be a $\mathcal{C}$-valued continuous-time up-down $OCRP(\alpha, \theta)$, for each $n \geq 1$. If $n^{-1}\mathcal{S}_n^0$ converges, as $n \to \infty$, then $(n^{-1}\mathcal{S}_n^{ny})_{y\geq 0}$ has a diffusive scaling limit, when suitably represented on a space of interval partitions. For $0 < \theta = \alpha < 1$ and $0 = \theta < \alpha < 1$, these are the type-0 and type-1 evolutions of [13].*

It is worth noting that if we consider evolution of the size of an individual block of the interval partitions in the type-0 and type-1 evolutions, it evolves as a BES^2 process with a suitable parameter, with new blocks of the interval coming into existence according to the level sets of a $(1 + \alpha)$-stable Lévy process without negative jumps, and the scaling limits proved here provide evidence towards the conjecture. This is related to two long standing conjectures by [2] and [10] respectively. Specifically, Aldous conjectured the existance of a scaling limit for a simple up-down Markov chain on uniform trees as a diffusive evolution of the Brownian CRT. Feng and Sun conjectured the existence of certain measure-valued diffusions whose stationary distributions are measures with Poisson-Dirichlet distributed atom sizes, also known as Pitman-Yor [47] or three-parameter Dirichlet processes [8] in the Bayesian nonparametrics literature. The paper is structured as follows.

In Section 2.2 we describe the discrete time and Poissonised Chinese Restaurant

Process and up-down chain via partitions, and discuss the existence or otherwise of stationary distributions.

In Section 2.3, we recall some preliminaries on splitting trees and contour processes from Lambert, explain the connection to our Poissonised up-down chain, and explain how to construct a generalised up-down Chinese Restaurant process via a theorem of Ray-Knight type.

In Section 2.4, we discuss the convergences of various scaling limits to Bessel-Squared and $(1+\alpha)$-stable processes, and explain the connections to the work by Forman, Rizzolo, Pal and Winkel after dealing with some necessary technical preliminaries.

Acknowledgement: The proof of the result on scaling limits is adapted from unpublished work [16] by Noah Forman, Matthias Winkel, Douglas Rizzolo and Soumik Pal, from which several ideas arose and which focused on the special case when $\alpha = \frac{1}{2}$.

2.2 Chinese Restaurants and related processes

2.2.1 Discrete-time Chinese Restaurants

The (α, θ)-Chinese Restaurant Process is an exchangeable partition-valued Markov chain studied by [39]. We recall this construction from [38]. Let $\mathcal{P}_n$ denote the set of partitions of $[n]$ and let $\mathcal{C}_n$ denote the set of compositions of n for $n \geq 1$, where $\mathcal{C}_0 := \{\varnothing\}$, $\mathcal{P}_0 = \{\varnothing\}$ and $[n] := \{1, \ldots, n\}$.

Refer to each block of a partition as a table and each element of one of these blocks as a customer. Define $\Pi_1 := \{\{1\}\}$. Whenever we have customers $1, \ldots, n$ seated at $k \geq 1$ non-empty tables, with n_i customers seated at table $1 \leq i \leq k$,

- place customer $n+1$ at occupied table i with probability proportional to $n_i - \alpha$,
- place customer $n+1$ at a new empty table with probability proportional to $\theta + k\alpha$.

For the rules of probability to be satisfied, it is necessary that either $0 \leq \alpha \leq 1$ and $\theta > -\alpha$ or $\alpha < 0$ and $\theta = -m\alpha$ for m some positive integer, by [39].

Definition 2.2.1 (Exchangeable partitions and compositions, [38])**.** We say that a random partition Π_n of $[n]$ is *exchangeable* if for any partition $\{A_1, \ldots, A_k\}$ of $[n]$, $\mathbb{P}(\Pi_n = \{A_1, \ldots, A_k\}) = p(|A_1|, \ldots, |A_k|)$, where $p(n_1, \ldots, n_k)$ is a symmetric function of compositions of $[n]$ known as the *exchangeable partition probability function* (EPPF) and $|A|$ is the number of elements in the set A. A composition C_n of $[n]$ is *exchangeable* if for any permutation σ of $[k]$,

$$\mathbb{P}(C_n = (n_{\sigma(1)}, \ldots, n_{\sigma(k)})) = \mathbb{P}(C_n = (n_1, \ldots, n_k)).$$

The Chinese restaurant construction given above does not specify the location of new tables, and will henceforth be referred to as an unordered (α, θ)-CRP. Pitman and Winkel [40] provide an ordering which arises naturally from an (α, θ)-tree growth process as introduced by [43] in the case $\theta = \alpha = 1/2$. For this new Chinese Restaurant construction to satisfy the laws of probability, it is necessary and sufficient that $0 \leq \alpha \leq 1$ and $\theta \geq 0$.

Definition 2.2.2 (Ordered (α, θ)-Chinese Restaurant Process)**.** Consider an (α, θ)-CRP where the location of any inserted single tables is determined as follows.

- Insert to the right of all tables with probability proportional to θ,
- insert directly to the left of table i with probability proportional to α.

With the same insertion rule for occupied tables as described before Definition 2.2.1, the resulting partitions are called an *ordered* (α, θ)*-CRP* and are illustrated in Figure 2.3.

The so-called up-down chain is a discrete-time Markov chain on $\mathcal{P}_n$ that can be defined as follows. For a single transition, a customer $n+1$ is seated according to the rules above and a second customer l is then chosen uniformly at random and removed

immediately, along with his table if this is left empty. Customers $l+1, l+2 \ldots, n+1$ are then relabelled as $l, l+1, \ldots, n$ so that we still have a partition of $[n]$.

In what follows, we no longer consider customers individually. Instead, we consider only the relative locations of tables and the numbers of customers at each table. Specifically, if our process is in a state $(A_1, \ldots, A_k) \in \mathcal{P}_n$, we are only interested in the composition $(|A_1|, \ldots, |A_k|) \in \mathcal{C}_n$. This is referred to as the *unlabelled Chinese Restaurant* in contrast to the partition-valued *labelled Chinese Restaurant*. A Chinese Restaurant without qualification is assumed to be unlabelled.

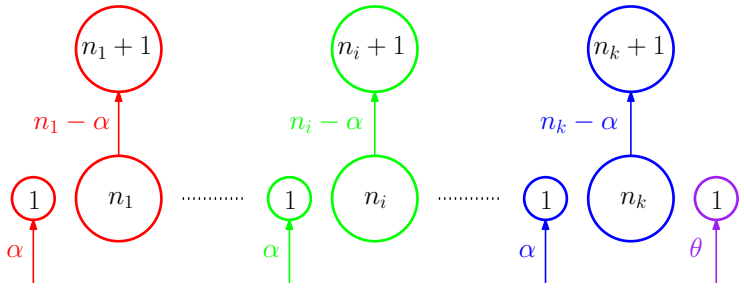


Fig. 2.3: Insertion rules for the ordered unlabelled (α, θ)-CRP.

As in the case of the unordered CRP, it is possible to find the stationary distribution of the up-down chain. We use some definitions given in [21] and [40].

Definition 2.2.3 (Regenerative composition structures)**.** A *composition sequence* is a Markovian sequence $(C_n)_{n\in\mathbb{N}}$ of compositions with $C_n \in \mathcal{C}_n$. It is *regenerative* if conditional on the first part of C_n having size m, the remaining block sizes form a composition of $n-m$ distributed as C_{n-m}. It is *weakly sampling consistent* if the distribution of C_{n+1} with a block size chosen with probability proportional to its size and reduced by 1 is C_n, and in such a case is called a *composition structure*. The *decrement matrix* $r(n, m)$ is the probability that the first block of C_n has size m.

Note that ordered (α, θ)-CRP compositions fit into the above framework by the following result from Pitman and Winkel [40].

Lemma 2.2.4 ([40], Proposition 6 (i),(ii))**.** *The composition sequence $(C_n)_{n\in\mathbb{N}}$ arising from an ordered Chinese restaurant process started from $\varnothing$ is regenerative, and*

also weakly sampling consistent. Moreover, the decrement matrix is given by

$$r(n,m) = \binom{n}{m} \frac{(n-m)\alpha + m\theta}{n} \frac{(1-\alpha)_{(m-1)\uparrow}}{(n-m+\theta)_{m\uparrow}}, \text{ if } 1 \leq m \leq n,$$

where $x_{y\uparrow}$ is a shorthand for $x_{y\uparrow 1}$.

Together with the following lemma, we will be able to identify the distribution of the ordered (α, θ)-CRP.

Lemma 2.2.5 ([21], Proposition 3.1)**.** *A composition structure $(C_n)_{n\in\mathbb{N}}$ with decrement matrix $r(n,m)$ is regenerative if and only if for any $n \in \mathbb{N}$, $(n_1, \ldots, n_k) \in \mathcal{C}_n$,*

$$p_n(n_1, \ldots, n_k) := \mathbb{P}\big(C_n = (n_1, \ldots, n_k)\big) = \prod_{i=1}^{k} r(N_i, n_i), \qquad (**)$$

where $N_i := \sum_{j=i}^{k} n_j$ for $i \in [k]$.

We now put this all together. To the best of our knowledge, the result below does not appear in the existing literature, although it follows in a very straightforward way from [21] and [40].

Corollary 2.2.6. *Let $\alpha \in [0,1]$, $\theta \geq 0$, and let $(n_1, \ldots, n_k) \in \mathcal{C}_n$, $n \geq 1$. The distribution p_n of an ordered (α, θ)-CRP composition on $\mathcal{C}_n$, $n \geq 1$ started from $(1) \in \mathcal{C}_1$ is given by*

$$p_n(n_1, \ldots, n_k) = \left(\prod_{i=1}^{k-1} \binom{N_i - 1}{n_i - 1} \Big(\frac{N_{i+1}\alpha + n_i\theta}{n_i} \Big) \right) \frac{\prod_{i=1}^{k} (1-\alpha)_{(n_i-1)\uparrow}}{(\theta+1)_{(n-1)\uparrow}}, \text{ if } \alpha \neq 1,$$

$$p_n(n_1, \ldots, n_k) = \prod_{i=1}^{k} \delta_{n_i 1} \text{ if } \alpha = 1, \text{ where } \delta_{ij} \text{ is the Kronecker delta.}$$

Proof. The case when $\alpha = 1$ is trivial, so suppose that $0 \leq \alpha < 1$. Note that an ordered sequence of (α, θ)-CRP compositions is weakly sampling consistent as noted in Lemma 2.2.4. The distribution of an ordered (α, θ)-CRP composition started from $(1) \in \mathcal{C}_1$ is found by noting that these compositions on $\mathcal{C}_n$ are regenerative and by

Lemma 2.2.4. From Lemma 2.2.5, the distribution p_n has the form

$$p_n(n_1,\dots,n_k) = \prod_{i=1}^{k} r(N_i,n_i) = \prod_{i=1}^{k}\left(\binom{N_i}{n_i}\frac{(N_i-n_i)\alpha+n_i\theta}{N_i}\frac{(1-\alpha)_{(n_i-1)\uparrow}}{(N_i-n_i+\theta)_{n_i\uparrow}}\right)$$
$$= \left(\prod_{i=1}^{k}\frac{(N_i-1)!}{n_i!(N_i-n_i)!}(N_{i+1}\alpha+n_i\theta)\right)\frac{\prod_{i=1}^{k}(1-\alpha)_{(n_i-1)\uparrow}}{\prod_{i=1}^{k}(N_{i+1}+\theta)_{n_i\uparrow}},$$

where $N_{k+1} := 0$. Writing $\prod_{i=1}^{k}(N_{i+1}+\theta)_{n_i\uparrow} = \theta_{n\uparrow}$, splitting off the $i=k$ term from the first product along with a factor of θ from $\theta_{n\uparrow}$ and simplifying,

$$p_n(n_1,\dots,n_k) = \left(\prod_{i=1}^{k-1}\binom{N_i-1}{n_i-1}\left(\frac{N_{i+1}\alpha+n_i\theta}{n_i}\right)\right)\frac{\prod_{i=1}^{k}(1-\alpha)_{(n_i-1)\uparrow}}{(\theta+1)_{(n-1)\uparrow}}$$

as claimed. □

Corollary 2.2.7. *The distribution of compositions in Corollary 2.2.6 above also coincides with the stationary distribution of the up-down ordered (α,θ)-CRP chain on $\mathcal{C}_n$.*

The proof follows directly via an inductive argument and the weak sampling consistency of Lemma 2.2.4, and we choose not to expand upon it here. While the unordered CRP is always exchangeable, we see that this is no longer the case for the ordered CRP. The result below states precise conditions for exchangeability of the (α,θ)-CRP compositions started from $(1)\in\mathcal{C}_1$.

Corollary 2.2.8. *Every composition of the ordered $(1,\theta)$-CRP started from $(1)\in\mathcal{C}_1$ is exchangeable, and the composition sequences generated by (α,θ)-CRPs started from $(1)\in\mathcal{C}_1$ with $0\le\alpha<1$ are exchangeable if and only if $\theta=\alpha$. Moreover, in the $\theta=\alpha<1$ case,*

$$p_n(n_1,\dots,n_k) = \binom{n}{n_1,\dots,n_k}\frac{\prod_{i=1}^{k}(1-\alpha)_{(n_i-1)\uparrow}}{(\alpha+1)_{(n-1)\uparrow}}\alpha^{k-1}.$$

Proof. The $(1,\theta)$-CRP is trivially exchangeable from the form of p_n given in Corol-

lary 2.2.6 above, so suppose that $\theta = \alpha < 1$. Then

$$\binom{N_i - 1}{n_i - 1}\left(\frac{N_{i+1}\alpha + n_i\theta}{n_i}\right) = \alpha\binom{N_i}{n_i},$$

and the form of $p_n(n_1, \ldots, n_k)$ follows from the combinatorial identity

$$\prod_{i=1}^{k-1}\binom{N_i}{n_i} = \prod_{i=1}^{k-1}\binom{\sum_{j=i}^{k} n_j}{n_i} = \binom{n}{n_1, \ldots, n_k}.$$

It is clear from the form of p_n that ordered (α, α)-CRP compositions with $0 \leq \alpha < 1$ are exchangeable. Conversely, for the compositions defined by the ordered (α, θ)-CRP with $0 \leq \alpha < 1$ to be exchangeable, we must have $p_3(2, 1) = p_3(1, 2)$. Using the forms of $p_3(2, 1)$ and $p_3(1, 2)$ from Corollary 2.2.6 once more,

$$\frac{(1-\alpha)(\alpha + 2\theta)}{(1+\theta)(2+\theta)} = \frac{(1-\alpha)(2\alpha + \theta)}{(1+\theta)(2+\theta)},$$

so that $\alpha = \theta$ as claimed. □

2.2.2 Continuous-time Chinese Restaurants

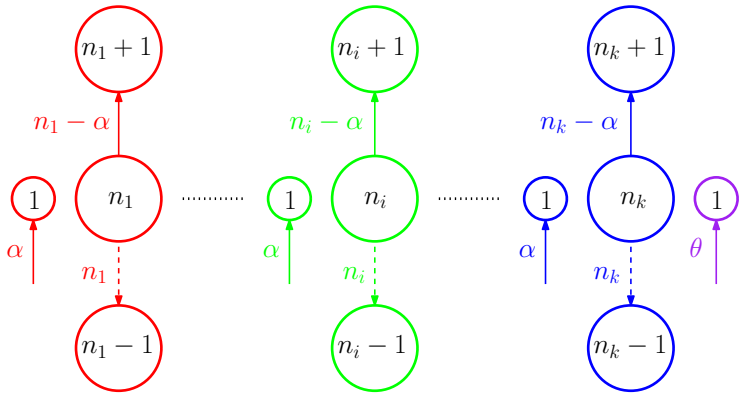


Fig. 2.4: Insertion rules for the ordered unlabelled up-down (α, θ)-CRP.

We study a continuous-time version of the up-down chain. This differs from the discrete-time version in that we do not have a customer leave immediately when a new one arrives. Instead, each customer stays for an exponential time of rate 1

independently of all other customers. Specifically, the number of customers is no longer constant, so this version of the chain is defined on the whole of $\mathcal{C} = \bigcup_{n\geq 0} \mathcal{C}_n$ and not just on some $\mathcal{C}_n$. This modification is known as Poissonisation, and is adapted from Pal [35]. We begin by describing the labelled chain on $\mathcal{P} := \bigcup_{n\geq 0} \mathcal{P}_n$.

Definition 2.2.9 (Poissonised Chinese Restaurant process)**.** Let $0 \leq \alpha \leq 1$, $\theta \geq 0$. The *Labelled Ordered Poissonised* (α, θ)*-Chinese Restaurant process* is a $\mathcal{P}$-valued continuous-time Markov chain, where from state $(A_1, \ldots, A_k) \in \mathcal{P}_n$, $n \geq 0$, we independently move into

- $(A_1, \ldots, A_i \cup \{n+1\}, \ldots, A_k)$ at rate $\#A_i - \alpha$,
- $(A_1, \ldots, A_k, \{n+1\})$ at rate θ,
- $(A_1, \ldots, A_{i-1}, \{n+1\}, A_i, \ldots, A_k)$ at rate α, the string $(A_1, \ldots, A_{i-1})$ is considered empty.

for all $i \in [k]$, with the convention here and throughout that if $i = 1$, The induced $\mathcal{C}$-valued process of compositions is called a *Poissonised Ordered* (α, θ)*-CRP*. The *Labelled Ordered Poissonised up-down* (α, θ)*-CRP* is defined similarly to the Chinese restaurant, but from any state $(A_1, \ldots, A_k) \in \mathcal{P}_n$, $n \geq 1$, we also consider the following additional independent transitions at rate 1. For each customer $l \in [n]$, if i is the unique index such that $l \in A_i$ we consider the following partition of $\{1, \ldots, l-1, l+1, \ldots, n\}$:

- $(A_1, \ldots A_{i-1}, A_i \setminus \{l\}, A_{i+1}, \ldots, A_k)$ if $\#A_i \geq 2$ and $l \in A_i$,
- $(A_1, \ldots A_{i-1}, A_{i+1}, \ldots, A_k)$ if $A_i = \{l\}$ and $n \geq 2$,
- $\varnothing \in \mathcal{P}_0$ if $n = 1$.

In order to arrive at a partition in $\mathcal{P}$, relabel customers $l+1, l+2, \ldots, n$ as $l, l+1, \ldots, n-1$. The compositions which track only the number of customers at

each table are said to come from an *Unlabelled Ordered Poissonised up-down* (α, θ)*-CRP*, which is illustrated in Figure 2.2.2 above. The associated *mass chain* $(M_t)_{t\geq 0}$ is defined to be the number of customers in the up-down chain at time t.

A consequence of the Poissonisation is that a stationary distribution exists if and only if $\theta = 0$, and in this case it is trivial. To prove this, we consider the mass chain $(M_t)_{t\geq 0}$, and observe that $(M_t)_{t\geq 0}$ is an $\mathbb{N}_0$-valued continuous-time chain with transition rates $q_{n+1,n} = n + 1$, $q_{n,n+1} = n + \theta$ for $n \in \mathbb{N}_0$, and all other $q_{i,j}$ with $i \neq j$ zero.

Proposition 2.2.10. *The ordered $\mathcal{C}$-valued Poissonised up-down chain and the associated mass chain $(M_t)_{t\geq 0}$ both have a stationary distribution if and only if $\theta = 0$. In this case, the stationary distribution of the Poissonised up-down chain is a point mass at $\varnothing \in \mathcal{C}_0$, and the stationary distribution of the mass chain is a point mass at 0.*

Proof. Let π be a potential stationary distribution for $M = (M_t)_{t\geq 0}$ and suppose that Q is the Q-matrix, then the equation $\pi Q = 0$ yields

$$-\theta\pi_0 + \pi_1 = 0, \quad (n+1)\pi_{n+1} - (2n+\theta)\pi_n + (n-1+\theta)\pi_{n-1} = 0 \quad \text{for any } n \in \mathbb{N}.$$

If $\theta = 0$ then a trivial induction yields $\pi_0 = 1$, $\pi_i = 0$ if $i \geq 1$, as claimed. If $\theta > 0$, then it is easily proved by induction that $\pi_i = \frac{\theta_{i\uparrow}}{i!}\pi_0$ for any $i \in \mathbb{N}_0$.

To be able to normalise this we need $\sum_{i=0}^{\infty} \frac{\theta_{i\uparrow}}{i!} < \infty$. Now observe that

$$\sum_{i=0}^{n} \frac{\theta_{i\uparrow}}{i!} = \sum_{i=0}^{n} \binom{\theta + i - 1}{i} = \binom{n+\theta}{n},$$

where $\binom{x}{y} := \Gamma(x+1)\big(\Gamma(x-y+1)\Gamma(y+1)\big)^{-1}$ for $x \geq 0, 0 \leq y \leq x$. By applying

Stirling's formula $\Gamma(z) \sim \sqrt{2\pi} z^{z-1/2} e^{-z}$ as $z \to \infty$ to $\binom{n+\theta}{n}$,

$$\begin{aligned}\binom{n+\theta}{n} &\sim \frac{(n+\theta+1)^{n+\theta+1/2} e^{-n-\theta-1}}{(n+1)^{n+1/2} e^{-n-1} \Gamma(\theta+1)} \\ &= \frac{e^{-\theta}(n+\theta+1)^{\theta}}{\Gamma(\theta+1)} \Big(\frac{n+\theta+1}{n+1}\Big)^{n+1/2}.\end{aligned}$$

To verify divergence of the series $\sum_{i=0}^{\infty} \frac{\theta_{i\uparrow}}{i!}$, we will let $n \to \infty$ in the expression above. Note that as $n \to \infty$, the factor $e^{-\theta}(n+\theta+1)^{\theta}\big(\Gamma(\theta+1)\big)^{-1}$ diverges to ∞. Writing $m = n+1$,

$$\Big(\frac{n+\theta+1}{n+1}\Big)^{n+1/2} = \Big(\frac{m+\theta}{m}\Big)^{m-1/2} = \Big(1+\frac{\theta}{m}\Big)^{m} \cdot \sqrt{\Big(1+\frac{\theta}{m}\Big)^{-1}} \to e^{\theta} \text{ as } m \to \infty.$$

It follows that $\sum_{i=0}^{\infty} \frac{\theta_{i\uparrow}}{i!} = \infty$, so the mass chain has no stationary distribution if $\theta > 0$.

Denote the Poissonised chain in this proof by $(U_t)_{t\geq 0}$. To prove the claims made about the case $\theta = 0$, note that $U_t = \varnothing$ if and only if $M_t = 0$. As 0 is an absorbing state of M, $\varnothing$ is an absorbing state of U and the stationary distribution is a point mass at $\varnothing \in \mathcal{C}_0$ as claimed.

Now suppose for a contradiction that U has a stationary distribution π on $\mathcal{C}$ with $\pi(n_1, \ldots, n_k) > 0$ for some composition $(n_1, \ldots, n_k) \in \mathcal{C}_n$, $n \geq 1$. Then $\Big(\sum_{(n_1,\ldots,n_k)\in\mathcal{C}_n} \pi(n_1, \ldots, n_k)\Big)_{n\geq 0}$ would be a stationary distribution for M. This proves that no stationary distribution exists in the case $\theta > 0$. □

To study the long term behaviour of the Poissonised Ordered CRP, we will consider a scaling limit in Theorem 2.4.16. Although we do not yet embark on this line of investigation, we study how long it takes single tables in the CRP to die out. Let $(Z_t)_{t\geq 0}$ denote a version of the continuous-time Markov chain on $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$ governing the evolution of a table size in the Poissonised up-down CRP, and observe that $(Z_t)_{t\geq 0}$ is a birth-death chain with birth rate $n - \alpha$ and death rate n. Let $\zeta := \inf\{s \geq 0 : Z_s = 0\}$ denote the time taken for such a chain to hit 0, define by $\mathbb{P}_r$ the probability measure on the Skorokhod space $\mathbb{D}_r([0, \infty), \mathbb{R})$ of càdlàg

paths $f\colon [0,\infty) \to \mathbb{R}$ started from $r \geq 0$ governing the evolution of a version of Z started from r, and let $\mathbb{E}_r$ denote the corresponding expectation. The proposition below will be used later to prove the result on scaling limits in Theorem 2.4.19.

Proposition 2.2.11. *Under $\mathbb{P}_1$, ζ has Laplace transform given by*

$$\mathbb{E}_1(e^{-\lambda\zeta}) = 1 - \frac{\lambda}{\alpha} + \frac{\lambda^{1+\alpha}}{\alpha e^{\lambda}\Gamma(1+\alpha,\lambda)} \qquad \text{if } 0 < \alpha \leq 1,$$

where the incomplete Gamma function $\Gamma(a,z)$ *is defined for $z > 0, a \in \mathbb{R}$ by*

$$\Gamma(a,z) := \int_z^\infty e^{-t}t^{a-1}dt.$$

In particular, under $\mathbb{P}_1$, ζ is exponential with rate 1 if $\alpha = 1$. Moreover, $\mathbb{E}_1(\zeta) = \alpha^{-1}$ if $0 < \alpha \leq 1$.

Apart from some changes to the rates, this largely follows in the same way as [35, Lemma 12], which corresponds to the case $\alpha = 1/2$.

Proof. Define $\sigma_r := \inf\{t \geq 0\colon Z_t = r\}$ with the convention that $\inf \varnothing = \infty$, suppose that $U_r(t)$ is the density of σ_r at t under $\mathbb{P}_{r+1}$ and $V_r(t) := \mathbb{P}_r(Z_t = r, Z_s \geq r \text{ for all } s \leq t)$. Note that by an application of the strong Markov property immediately after the first jump,

$$V_r(t) = e^{-(2r-\alpha)t} + \int_0^t (r-\alpha)e^{-(2r-\alpha)u}\int_0^{t-u} U_r(s)V_r(t-s-u)dsdu. \qquad (***)$$

For $\lambda > 0$, we can define Laplace transforms by

$$\begin{aligned} F_r(\lambda) &:= \int_0^\infty U_r(t)e^{-\lambda t}dt, \\ H_r(\lambda) &:= \int_0^\infty V_r(t)e^{-\lambda t}dt, \\ G_r(\lambda) &:= \int_0^\infty e^{-(2r-\alpha)t}e^{-\lambda t}dt = (2r-\alpha+\lambda)^{-1}. \end{aligned}$$

Now

$$(r-\alpha)F_r(\lambda)H_r(\lambda)G_r(\lambda) = (r-\alpha)\int_0^\infty \left[\int_0^t \int_0^{t-s} e^{-(2r-\alpha)t} U_r(s)V_r(t-s-u)\right] du ds e^{-\lambda t} dt,$$

which by (***) equals $\int_0^\infty \left(V_r(t) - e^{-(2r-\alpha)t}\right)e^{-\lambda t}dt = H_r(\lambda) - G_r(\lambda)$. Rearranging this and using the expression for $G_r(\lambda)$, it follows that

$$H_r(\lambda) = \frac{1}{2r - \alpha + \lambda - (r-\alpha)F_r(\lambda)}.$$

Now for all $\varepsilon > 0$,

$$\begin{aligned}\int_t^{t+\varepsilon} U_r(s)ds &= \mathbb{P}_{r+1}(t < \sigma_r \leq t+\varepsilon) \\ &= \sum_{k=1}^\infty \mathbb{P}_{r+1}(Z_t = r+k, Z_s \geq r+1 \text{ for all } 0 \leq s \leq t)\mathbb{P}_{r+k}(\sigma_r \leq \varepsilon) \\ &= V_{r+1}(t)\mathbb{P}_{r+1}(\sigma_r \leq \varepsilon) + o(\varepsilon^2) \qquad \text{as } \varepsilon \downarrow 0.\end{aligned}$$

Upon dividing by ε and letting $\varepsilon \downarrow 0$, we see that $U_r(t) = (r+1)V_{r+1}(t)$. Hence for all $r \geq 0$,

$$F_r(\lambda) = (r+1)H_{r+1}(\lambda) = \frac{r+1}{2r+2-\alpha+\lambda-(r+1-\alpha)F_{r+1}(\lambda)}. \qquad (****)$$

Now define $K_{-2}(\lambda) := e^z z^{-a}\Gamma(a,z)$ and $K_{-1}(\lambda)$ implicitly by

$$K_{-2}(\lambda) = \frac{1}{1+z-a-(1-a)K_{-1}(\lambda)}.$$

Then

$$K_{-1}(\lambda) = \frac{(1+z-a)e^z z^{-a}\Gamma(a,z) - 1}{(1-a)e^z z^{-a}\Gamma(a,z)} = \frac{1+z-a}{1-a} - \frac{1}{(1-a)e^z z^{-a}\Gamma(a,z)}.$$

However, it is known from [32, p.278, 3.3.3] that for any $a \in \mathbb{R},\ z > 0$,

$$K_{-2}(\lambda) = e^z z^{-a}\Gamma(a,z) = \cfrac{1}{1+z-a-\cfrac{1(1-a)}{3+z-a-\cfrac{2(2-a)}{5+z-a-\cfrac{3(3-a)}{7+z-a-\cdots}}}} = \frac{1}{1+z-a-(1-a)K_{-1}(\lambda)}.$$

Upon using the recursion given for $F_r(\lambda)$ in (****), we also see that

$$F_0(\lambda) = \cfrac{1}{2+\lambda-\alpha-\cfrac{2(1-\alpha)}{4+\lambda-\alpha-\cfrac{3(2-\alpha)}{6+\lambda-\alpha-\cfrac{4(3-\alpha)}{8+\lambda-\alpha-\cdots}}}}.$$

Substituting in the continued fraction expression for $K_{-2}(\lambda)$ into the expression for $K_{-1}(\lambda)$, it follows that

$$K_{-1}(\lambda) = \cfrac{1}{3+z-a-\cfrac{2(2-a)}{5+z-a-\cfrac{3(3-a)}{7+z-a-\cfrac{4(4-a)}{9+z-a-\cdots}}}}.$$

In the case when $\alpha \neq 0$, this can be identified with $F_0(\lambda)$ by setting $a = 1+\alpha, z = \lambda$, from which it follows that

$$F_0(\lambda) = K_{-1}(\lambda) = \frac{\lambda-\alpha}{-\alpha} - \frac{1}{-\alpha e^{\lambda}\lambda^{-(1+\alpha)}\Gamma(1+\alpha,\lambda)} = 1 - \frac{\lambda}{\alpha} + \frac{\lambda^{1+\alpha}}{\alpha e^{\lambda}\Gamma(1+\alpha,\lambda)}.$$

Since $F_0(\lambda)$ is the Laplace transform of ζ under $\mathbb{P}_1$ by the definition of $U_0(\lambda)$, the proposition is proved upon observing that $\mathbb{E}_1(\zeta) = -F_0'(0)$. □

2.3 Results on splitting tree connections

2.3.1 Preliminaries on splitting trees

The ordered Poissonised Chinese restaurant connects to the splitting trees studied by Lambert [31] and introduced previously by Geiger and Kersting [20]. Our ideas are similar to those of [35, Section 6], see also [14, Section 1.1]. Specifically, we will show that it is possible to view the ordered Poissonised up-down chain as a splitting tree with lifespan measure $\Lambda(dr) := \mathbb{P}(\inf\{t \geq 0 : Z_t = 0\} \in dr)$ for Z the chain governing the evolution of table sizes as described before Proposition 2.2.11, with initial size $Z_0 = 1$. Letting $\mathbb{N} = \{1, 2, 3, \ldots\}$, we recall some relevant notation from Lambert [31].

Set $\mathcal{U} := \cup_{n=0}^{\infty} \mathbb{N}^n$, where $\mathbb{N}^0 := \{\varnothing\}$, and for $u, v \in \mathcal{U}$ let $uv := (u_1, \ldots, u_n, v_1, \ldots, v_m)$, where $\varnothing u := u =: u\varnothing$. We begin by introducing discrete trees as a genealogical model. The initial ancestor of a population is encoded as $\varnothing$, and for $u \in \mathcal{U}, j \in \mathbb{N}$, uj denotes the j-th child of u.

A *discrete tree* or *Ulam-Harris tree* is a set $\mathcal{T} \subset \mathcal{U}$ such that

(i) $\varnothing \in \mathcal{T}$,

(ii) if $uj \in \mathcal{T}$ then $u \in \mathcal{T}$,

(iii) For any $u \in \mathcal{T}$ there is a unique integer $K_u \leq \infty$ such that $uj \in \mathcal{T}$ iff $j \in \{1, \ldots, K_u\}$.

Such trees encompass a wide range of models, but do not take account of when births and deaths occur and are therefore insufficient as a model for studying the ordered Poissonised up-down chain. To remedy this, we work with *chronological trees* as introduced by [20] and studied extensively by [31]. Each individual u is born at a time $\beta(u)$ and dies at a time $\omega(u)$ as illustrated in Figure 2.3.1 below. Set

$\mathbb{U} := \mathcal{U} \times [0, \infty)$, $\rho := (\varnothing, 0)$, and denote for $x := (u, \sigma) \in \mathbb{U}$ the *first and second projections* p_1 and p_2 of $\mathbb{U}$ onto $\mathcal{U}$ and $[0, \infty)$ by $p_1(x) := u$, $p_2(x) := \sigma$.

A *chronological tree* $\mathbb{T}$ is a subset of $\mathbb{U}$ such that

(i) $\rho \in \mathbb{T}$,

(ii) the projection $\mathcal{T} := p_1(\mathbb{T})$ of $\mathbb{T}$ onto $\mathcal{U}$ is a discrete tree,

(iii) if $u \in \mathcal{T}$, there exist $0 \leq \beta(u) < \omega(u) \leq \infty$ such that $(u, \sigma) \in \mathbb{T}$ iff $\beta(u) < \sigma \leq \omega(u)$,

(iv) if $u \in \mathcal{T}$ and $j \in \{1, \ldots, K_u\}$, then $\beta(u) < \beta(uj) < \omega(u)$.

(v) for any $u \in \mathcal{T}$ and any $i, j \in \{1, \ldots, K_u\}$, $\beta(ui) = \beta(uj) \implies i = j$.

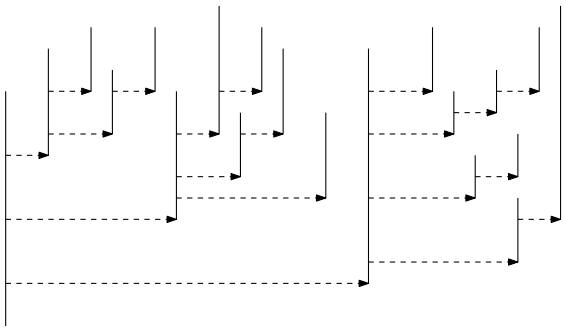


Fig. 2.5: A visual representation of a chronological tree.

Note that while the definition above does not require $\beta(ui) < \beta(uj) \implies i < j$, it can be assumed without loss of generality that this is the case by parts (iii) and (v) of the definition above.

We require a significant amount of content from Lambert [31]. This is used to describe both his contour process and the notion of a splitting tree. Such contour processes are not novel constructions, and appear regularly in the branching process literature.

Let $x, y \in \mathbb{U}$, write $x := (u, \sigma)$, $y := (v, \tau)$ and let $\zeta(u) := \omega(u) - \beta(u)$ denote the *lifespan* of the individual u. Write $u \preceq v$ if u is an *ancestor* of v or $u = v$, i.e if there is $w \in \mathcal{U}$ with $v = uw$. For a chronological tree $\mathbb{T}$, write $x \preceq y$ when $u \preceq v$

and $u \neq v$, $\sigma \leq \beta(uj)$ for the unique j with $uj \preceq v$, or $u = v$ and $\sigma \leq \tau$. This is a notion of ancestry on $\mathbb{T}$.

Similarly the *most recent common ancestor* (mrca) $x \wedge y$ of x and y is the element $z \in \mathbb{T}$ with $z \preceq x$, $z \preceq y$ and $p_2(z)$ maximised.

We use this to define a distance d on $\mathbb{T}$ via $d(x, y) := p_2(x) + p_2(y) - 2p_2(x \wedge y)$, which represents the distance on the chronological tree that must be traversed to go directly from x to y.

Consider $\mathbb{T}$ under the distance d above and let $x \in \mathbb{T}$. Then x is a *branch point* if $\mathbb{T} \setminus \{x\}$ consists of three connected components, a *leaf* if $\mathbb{T} \setminus \{x\}$ consists of a single component and x is not the *root* ρ, and a *simple point* if $\mathbb{T} \setminus \{x\}$ consists of two connected components.

Both $\mathbb{T}$ and $\mathbb{T}'$ denote chronological trees in what follows. If x is a simple point of $\mathbb{T}$, we can graft a copy of $\mathbb{T}'$ onto $x = (u, \sigma)$ as a descendant of individual u. More formally, for any $x \in \mathbb{T}$ and $1 \leq i \leq K_u$, denote by $\widetilde{\mathbb{T}}$ the subset of $\mathbb{U}$ obtained from $\mathbb{T}$ by renaming points $(ukw, \tau) \in \mathbb{T}$ as $(u(k+1)w, \tau)$ for all integers $k \geq i$ and all finite words $w \in \mathcal{U}$ with $ukw \in \mathcal{T}$. Then define $g(\mathbb{T}', \mathbb{T}, x, i) := \widetilde{\mathbb{T}} \cup \{(uiw, \sigma + \tau) : (w, \tau) \in \mathbb{T}'\}$.

Call $\theta(x) := \{z \in \mathbb{T} \colon x \preceq z\}$ the set of *descendants* of x. This can be split into two disjoint subsets $\theta_l(x)$ and $\theta_r(x)$ called the *left-descendance* and *right-descendance* as follows.

If x is not a branch point write $\theta_l(x) := \theta(x)$, $\theta_r(x) := \varnothing$. If $x = (u, \sigma)$ is a branch point, and j is the unique index with $\sigma = \beta(uj)$, write

$$\theta_l(x) := \bigcup_{\epsilon > 0} \theta(u, \sigma + \epsilon), \quad \theta_r(x) := \{x\} \cup \bigcup_{\epsilon > 0} \theta(uj, \sigma + \epsilon).$$

We can also partition the set $\{z \in \mathbb{T} \colon z \preceq x\} \cup \theta(x)$ into two disjoint forests $L(x)$ and $R(x)$ called the *left-hand component* and *right-hand component* by writing

$$L(x) := \bigcup_{y : x \in \theta_r(y)} \theta_l(y), \quad R(x) := \bigcup_{y : x \in \theta_l(y)} \theta_r(y).$$

We say that $\mathbb{T}$ is *finite* if $p_1(\mathbb{T})$ is a finite set. We define a notion of *Lebesgue measure λ on chronological trees.* For any $u \in \mathcal{T}, (a,b] \subseteq (\beta(u), \omega(u)]$, let the *length of a segment* $\lambda(\{u\} \times (a,b])$ be $b-a$, and extend this measure to Borel subsets of $\mathbb{T}$ by a variation of Caratheodory's extension theorem. Note that the generating sets of the Borel σ-algebra are precisely those of the form $\{u\} \times (a,b]$ for any $u \in \mathcal{T}, (a,b] \subseteq (\beta(u), \omega(u)]$. We refer to the *length of a chronological tree* as the quantity $\lambda(\mathbb{T})$.

Define a partial order $\leq$ on $\mathbb{T}$ by $x \leq y$ if and only if $y \preceq x$ or $y \in R(x)$, and for any $x \in \mathbb{T}$, define a Borel set $\mathbb{S}(x)$ by $\mathbb{S}(x) := \{y \in \mathbb{T} \colon y \leq x\}$. [31, p.361] showed that this set is indeed Borel.

Using the above notation, we describe Lambert's construction of an *exploration process* on chronological trees with finite discrete part and finite length. Beginning from $(\varnothing, \omega(\varnothing))$, track down the tree at unit speed, and upon encountering a branch point $x = (u, \sigma)$ with $\sigma = \beta(uj)$ for some $1 \leq j \leq K_u$, jump to the death time $(uj, \omega(uj))$ of the individual j born at the branch point x. Return back to $(u, \beta(uj))$ after exploring the subtree $\theta_r(u, \beta(uj))$, and finish exploring the tree upon reaching ρ. The second projection of the exploration process is called the *jumping chronological contour process* (JCCP), and is illustrated in Figure 2.3.1.

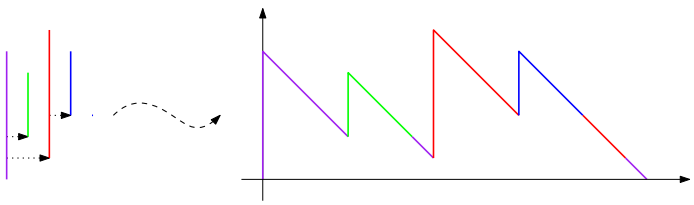


Fig. 2.6: The relationship between a chronological tree and the JCCP.

Here is a more formal definition. Let $\mathbb{T}$ denote a tree with finite length $l := \lambda(\mathbb{T})$, and define a map $\varphi \colon \mathbb{T} \to [0,l]$ by $\varphi(x) := \lambda(\mathbb{S}(x))$. By Theorem 3.1 of [31], the map φ is the unique order preserving, measure preserving bijection from $\mathbb{T}$ to $[0,l]$.

Definition 2.3.1 (Exploration process/JCCP)**.** [31, Definition 3.2] The map $\varphi^{-1} \colon [0,l] \to$

$\mathbb{T}$ is called the *exploration process*, and the *jumping chronological contour process* (JCCP) $X = (X_s)_{0\le s\le l}$ is the second projection of φ^{-1}, i.e. $X_s := p_2(\varphi^{-1}(s))$.

Theorem 2.3.2 ([31], Theorem 3.3)**.** *The exploration process is càdlàg and $s \in [0, l]$ is a jump time of X if and only if $(v, \omega(v)) = \varphi^{-1}(s)$ is a leaf of $\mathbb{T}$. In this case $\varphi^{-1}(s-) = (u, \beta(v))$, where u is v's parent.*

Moreover, the JCCP is also càdlàg and the size of the jump at the birth of an individual $s = \varphi(v, \omega(v))$ is the lifespan $\zeta(v)$ of that individual. In addition, for all $s \in [0, l]$,

$$X_s = -s + \sum_{v\in\mathcal{T}\colon \varphi(v,\omega(v))\le s} \zeta(v).$$

Finally, we describe a class of chronological trees with a random recursive structure. These were introduced by [20]. Start with a segment of length χ, and on this segment choose points according to a Poisson process of rate $\alpha > 0$.

Then to each point graft on another segment with length distributed according to a probability measure Λ on $(0, \infty)$ governing the lifespan, independently of all other segments. Recursively repeat this procedure for each new segment, where the lengths of all segments in the same generation are independent of each other and the lengths of segments in previous generations. Then the limiting object is called a *splitting tree*.

Definition 2.3.3 (Lifespan measure/Splitting tree)**.** [31, p.372] Let Λ be a probability measure on $(0, \infty)$, which we will call a *lifespan measure* in the context of splitting trees below. We define a family of probability measures $(\mathbb{P}_\chi)_{\chi\ge 0}$ by the following procedure. Let $(\beta_i, \zeta_i)_{i\ge 1}$ be the atoms of a Poisson process on $(0, \chi) \times (0, \infty)$ with intensity measure $\alpha Leb \otimes \Lambda$. $(\mathbb{P}_\chi)_{\chi\ge 0}$ is defined to be the unique family of probability measures on chronological trees $\mathbb{T}$ satisfying $\mathbb{T} = \bigcup_{n\ge 1} g(\mathbb{T}_n, \varnothing \times (0, \chi], (\varnothing, \beta_n), n)$, where conditionally on the measure $\sum_{i\ge 1} \delta_{(\beta_i,\zeta_i)}$, the $(\mathbb{T}_n)_{n\ge 1}$ are independent splitting trees, with the property that conditional on $\zeta(n) = \zeta$, $\mathbb{T}_n$ has law $\mathbb{P}_\zeta$. We say that any such random chronological tree $\mathbb{T}$ is a *splitting tree*.

Remark 2.3.4. As noted in [31, p.372], explicit characterisation of $(\mathbb{P}_\chi)_{\chi\ge 0}$ requires

specifying a labelling of the atoms $(\beta_i, \zeta_i)_{i\geq 1}$, and for the sake of completeness we reproduce the details here. Since all lifespans are almost surely finite we can specify that $i < j$ if and only if $\zeta_i > \zeta_j$ or $\zeta_i = \zeta_j$ and $\beta_i < \beta_j$.

Note that Lambert's definition of a splitting tree is slightly more general than is needed in this paper, and includes the cases when $\Lambda(\{\infty\}) > 0$ or $\chi = \infty$, along with $\int_{(0,\infty]}(r \wedge 1)\Lambda(dr) < \infty$. The reader is referred to [31, p.372] for the details.

We conclude this subsection by describing the contour process of a splitting tree, which is in fact a conditioned Lévy process. Let $(\Delta_t)_{t\geq 0}$ be a Poisson point process on $[0,\infty) \times (0,\infty)$ with intensity measure $Leb \otimes \Lambda$, and suppose that $(Y_t)_{t\geq 0}$ is the Lévy process defined by $Y_t := -t + \sum_{s\leq t} \Delta_s$, which has Laplace exponent $\psi(\lambda) := \lambda - \int_0^\infty (1 - e^{-\lambda r})\Lambda(dr)$.

Theorem 2.3.5 ([31], Theorem 4.3)**.** *Suppose $\mathbb{T}$ is a splitting tree with $\chi = \zeta(\varnothing)$ and $\alpha \int_{(0,\infty)} r\Lambda(dr) \leq 1$. Under $\mathbb{P}_\chi$, $\mathbb{T}$ has a JCCP $(X_s)_{0\leq s\leq \lambda(\mathbb{T})}$, which has the law of Y started at χ.*

Remark 2.3.6. Note that in particular we restrict ourselves to the (sub)-critical case $\alpha \int_{(0,\infty)} r\Lambda(dr) \leq 1$ so that the tree is finite almost surely. Lambert describes this result more generally when we are not in the (sub)-critical case by means of conditioning on the Lévy process hitting 0 and being killed there.

2.3.2 The JCCP of the ordered Poissonised up-down CRP

We discuss how to make the ordered Poissonised $(\alpha, 0)$-up-down CRP with $0 < \alpha < 1$ into a splitting tree $\mathbb{T}$, and explain how to derive an (α, θ)-up-down CRP from a modified JCCP for $\theta > 0$. Suppose that our restaurant is initially in the state $(n_1, \ldots, n_k) \in \mathcal{C}$. As an informal description in the case when $\theta = 0$, we view the j-th initial table of initial size n_j as an individual in a population. By convention, the children of the j-th initial table are the new tables inserted directly to the left of this initial table. Note that each initial table of size n_j lives for a time distributed according to $\mathbb{P}_{n_j}(\tau \in \cdot)$, where as in Definition 2.2.9, $(Z_t)_{t\geq 0}$ is a Markov chain

on $\mathbb{N}_0$ which is absorbed at 0, has up rate $n-\alpha$ and down rate n from any state $n \geq 1$, $\mathbb{P}_i$ is the law of this process started from $i \in \mathbb{N}_0$, Q_α is the Q-matrix and $\tau := \inf\{t > 0 : Z_t = 0\}$. All subsequent tables have lifespan distributed as $\mathbb{P}_1(\tau \in \cdot)$ independently of all other lifespans. In particular, the size of the j-th initial table is distributed as a version of Z_t under $\mathbb{P}_{n_j}$, $1 \leq j \leq k$, and all the other table sizes evolve independently as versions of Z_t under $\mathbb{P}_1$.

Let us explain how to view each table and its descendency as a splitting tree $\mathbb{T}^j$. Define $\mathcal{T}_0 := \{\varnothing\}$, and define $\mathcal{T}_n$ inductively as follows. We will use the following notation as adapted from the previous section, where $\mathcal{T} := \bigcup_{n\geq 0} \mathcal{T}_n$. Assume first that we have a single table of initial size n_0.

For the table labelled as $u \in \mathcal{T}$, we will let $\beta(u)$ denote the time at which it came into existence, let $(Z^u_t)_{t\geq 0}$ be the Markov chain governing the evolution of the size of the table labelled u, and observe that $Z^u_0 = 1$ almost surely if $u \neq \varnothing$, but that $Z^\varnothing_0 = n_0$. Set $\zeta(u) := \inf\{t > 0 : Z^u_t = 0\}$ and write $\omega(u) := \beta(u) + \zeta(u)$.

During the lifespan of table $u \in \mathcal{T}_n$, let $\beta(u1) < \beta(u2) < \cdots < \beta(uK_u)$ denote the times at which there is a left insertion directly to the left of u, and label these new tables in order of appearance as $u1, u2, \ldots, uK_u$. Write $\mathcal{T}_{n+1} := \bigcup_{u\in\mathcal{T}_n} \bigcup_{l=1}^{K_u} \{ul\}$, and let $(Z^{ul}_t)_{t\geq 0}$ denote another collection of versions of Z under $\mathbb{P}_1$, independent of each other and of all other Z^u.

These are the basic ingredients needed to construct a chronological tree $\mathbb{T}$ in the notion of the previous section, which is achieved by setting $\mathcal{T} := \bigcup_{n\geq 0} \mathcal{T}_n$, $\mathbb{T} := \big(\bigcup_{u\in\mathcal{T}} \{u\} \times (\beta(u), \omega(u)]\big) \cup \{\rho\}$.

By construction, $\mathbb{T}$ is a splitting tree with lifespan measure $\Lambda(dr) := \mathbb{P}_1(\tau \in dr)$ and initial lifespan distributed according to $\mathbb{P}_{n_0}(\tau \in \cdot)$, where the points of the rate α Poisson processes in this splitting tree correspond to the new tables inserted at rate α.

Now consider the JCCP $(X_s)_{s\geq 0}$ derived from $\mathbb{T}$ as in Definition 2.3.1. Observe that in particular by Theorem 2.3.2, $X_s := -s + \sum_{i=0}^{J_s} \zeta^i$, $s \geq 0$, where $(J_s)_{s\geq 0}$ is a Poisson process of rate α, ζ_0 is a version of τ under $\mathbb{P}_{n_0}$ independent of J and $(\zeta^i)_{i\geq 1}$

are independently distributed as versions of τ under $\mathbb{P}_1$, independently of J and ζ_0.

Note that if we then embed copies $(Z_t^u)_{t\geq 0}$ of the up-down chain $(Z_t)_{t\geq 0}$ into the lifespan of each individual $u \in \mathcal{T}$ and cut through the tree $\mathbb{T}$ at level $y \geq 0$, we would expect to recover a copy of the Poissonised $(\alpha, 0)$-up-down chain started from a single table of initial size n_0. It is natural to expect that doing the same with the JCCP will yield the same result, which is what we will subsequently show in Theorem 2.3.8.

2.3.3 Construction of skewer processes

To begin, observe that the above idea only illustrates a method of recovering the $(\alpha, 0)$-up-down chain started from a specific state $(n_0) \in \mathcal{C}$ corresponding to a single table. It is possible to recover the $(\alpha, 0)$-up-down chain started from a general state $(n_1, \ldots, n_k) \in \mathcal{C}$. However, as the evolution of the size of the initial tables with n_j, $1 \leq j \leq k$ customers and their progenies are independent, we will show the result first for a single initial table in the case when $\theta = 0$ and then generalise it to the case of a general initial composition with $\theta \geq 0$. Much of this material holds in a more general setting, and we investigate the constructions arising from this in Section 2.3.5, choosing at first to focus on detailing the construction of what we call a skewer process and explaining how to recover the up-down Ordered (α, θ)-CRP from this proceedure.

Consider any continuous-time Markov chain Z on $\mathbb{N}_0$, with Q-matrix $Q = (q_{n,n'})_{n,n'\geq 0}$ say, for which 0 is absorbing and $\mathbb{E}_n(\zeta) < \infty$ for all $n \geq 1$, where $\zeta := \inf\{t \geq 0 \colon Z_t = 0\}$ and $\mathbb{E}_n$ denotes the expectation under $\mathbb{P}_n$. Consider any initial distribution $(p_n, n \geq 1)$ on $\mathbb{N}$ such that $\mu = \sum_{n\geq 1} p_n \mathbb{E}_n(\zeta) < \infty$. Let $\alpha \geq 0$, and consider the following independent objects:

- an (n_0, Q)-Markov chain Z^0,
- a sequence Z^i, $i \geq 1$, of Q-Markov chains started from an initial distribution $(p_n, n \geq 1)$,

- a Poisson process $(J_t)_{t\geq 0}$ with intensity α.

To see how the above generalises from the setting of the $(\alpha, 0)$-up-down CRP, simply take $p_1 = 1$, $p_n = 0$ for $n \geq 2$, let $Q = Q_\alpha$ and note that the values of α and μ are consistent with Proposition 2.2.11.

Set $\zeta^i := \inf\{t \geq 0 : Z^i_t = 0\}$ for the extinction time of Z^i, $i \geq 0$. Denote by $(C^i)_{i\geq 1}$, the arrival times of J and set $C^0 := 0$. Write $X_t := -t + \sum_{i=0}^{J_t} \zeta^i$, $B^i := X_{C^i-}$, $D^i := B^i + \zeta^i$, $i \geq 1$, for the birth and death heights of the chains in each jump of X, and set $B^0 := 0, D^0 := \zeta^0$, where $X_{a-} := \lim_{s\uparrow a} X_s$. Set

- $T := \inf\{s \geq 0 : X_s = 0\}$,
- $I^y_1 := \inf\{i \geq 0 \colon B^i \leq y < D^i\}$ and
  $I^y_{l+1} := \inf\{i \geq I^y_l + 1 \colon B^i \leq y < D^i\}$ for $y \geq 0$, $l \geq 1$,
- $K^y := \sup\{0\} \cup \{l \geq 1 \colon C^{I^y_l} \leq T\}$ and
  $N^y_l := Z^{I^y_l}(y - B^{I^y_l})$, for $y \geq 0$, $1 \leq l \leq K^y$.

To clarify the meaning of the notation, T is the time it takes for the Lévy process excursion X to hit 0, the indices I^y_l track in left to right order the jumps of X which cross level y, K^y counts the number of jumps that X makes over y before hitting 0, and N^y_l is the height of the l-th up-down chain embedded in X crossing level y at such a level, upon viewing X as jumping instantaneously from 0 to X_0.

Definition 2.3.7 (Skewer process)**.** Using the notation above, define the *skewer process* $(\mathcal{S}^y)_{y\geq 0}$ *of* X *and* $(Z^i)_{i\geq 0}$ by $\mathcal{S}^y := (N^y_{K^y}, \ldots, N^y_2, N^y_1)$.

The skewer process introduced here is a discrete analogue of a process introduced in [14] known as the *type*-1 *evolution*, and the $(\alpha, \theta) = (\frac{1}{2}, 0)$ case of the discrete process is itself mentioned in [14], albeit with the left to right ordering reversed. This is to account for the fact that insertions at rate α are located to the left of current tables, while the Lévy process $(X_t)_{t\geq 0}$ is a contour process of planar splitting trees with new branches drawn to the right. However, the analogue of X in the type-1

evolution of [14] is instead a $\frac{3}{2}$-stable Lévy process without negative jumps, and the $\mathbb{N}_0$-valued up-down chains are replaced by $BES^2(-1)$ excursions. We discuss connections to BES^2 processes in Section 4.

Intuitively, the skewer at level y is obtained by embedding up-down chains into the jumps of the Lévy process, recording the heights of these up-down chains at level y and ordering these from right to left. For now, we investigate the behaviour of the skewer process in the special case of the up-down Chinese restaurant process with $p_1 = 1, p_n = 0$ for $n \geq 2$, $Q = Q_\alpha$.

Theorem 2.3.8. *Suppose that $p_1 = 1$, $p_n = 0, n \geq 2$, $Q = Q_\alpha$ and $Z_0^0 = n_0$. Then the skewer process $(\mathcal{S}^y)_{y\geq 0}$ is Markovian and evolves as an ordered Poissonised $(\alpha, 0)$-up-down CRP started from $(n_0) \in \mathcal{C}$.*

Observe that this extraction of a Markov process from the level sets of another process is similar to the Ray-Knight theorem that identifies the local time process of a suitably stopped Brownian motion as a Bessel-squared process, see also [44].

Before proving this result, we will describe how to extend to the (α, θ)-case started from a general state $(n_1, \ldots, n_k) \in \mathcal{C}_n$. First we must describe how to start from the general state $(n_1, \ldots, n_k) \in \mathcal{C}_n$.

Define the concatenation $uv \in \mathcal{C}$ of $u, v \in \mathcal{C}$ with $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_m)$ by $uv := (u_1, \ldots, u_n, v_1, \ldots, v_m)$. Now, let $(\mathcal{S}^{j,y})_{y\geq 0}$, $1 \leq j \leq k$ be independent skewers, with $Z_0^0 = n_j$ for $\mathcal{S}^j$, and define $\mathcal{S}^y := \mathcal{S}^{k,y} \cdots \mathcal{S}^{2,y}\mathcal{S}^{1,y}$.

Corollary 2.3.9. *Suppose that $(\mathcal{S}^{j,y})_{y\geq 0}$ are skewer processes with independent versions of $X, (Z^i)_{i\geq 0}$, each with $p_1 = 1$, $p_n = 0, n \geq 2$, $Q = Q_\alpha$, $Z_0^0 = n_j$ for the j-th skewer process, analogously to Theorem 2.3.8, and also that $\mathcal{S}^{j,0} = (n_j) \in \mathcal{C}$. Then the skewer process $(\mathcal{S}^y)_{y\geq 0}$ evolves as an ordered $(\alpha, 0)$-up down CRP chain started from $(n_1, \ldots, n_k) \in \mathcal{C}_n$.*

Note that we have to order the skewers corresponding to each excursion from right to left. This is needed in order to ensure that our insertion rules are consistent with the ordering in Chinese Restaurant Process of Definition 2.2.2. Our convention

is such that the regenerativity property of the Discrete Ordered CRP holds and can be applied. An inconsistency with [14] who instead insert single tables of the Ordered CRP to the right of each existing table at rate α and to the left of all tables at rate θ is therefore natural.

Remark 2.3.10. It is also possible to derive the $(\alpha, 0)$-up-down CRP from a single path of the Lévy process. Consider the Lévy process $(X_s)_{s\geq 0}$ up to time T^k, defined to be the time of the k-th down-crossing over 0. Use the same notation as previously, but with T replaced by T^k in the definition of K^y. Then conditional on the event $N_1^0 = n_k, N_2^0 = n_{k-1} \dots N_k^0 = n_1$, and assuming that $p_1 = 1$, $p_n = 0, n \geq 2$, $Q = Q_\alpha$ as before, the skewer $\mathcal{S}^y$ evolves as an up-down Chinese restaurant started from state $(n_1, n_2, \dots, n_k) \in \mathcal{C}$. A proof involves showing that the excursions above level 0 are independent and is then a trivial enough corollary of Theorem 2.3.8 that the proof is omitted.

To account for the right insertions of rate $\theta > 0$, suppose that we have the following objects, independent of each other and of the previous objects:

- a Poisson process $\widetilde{J}$ with intensity $\theta > 0$,
- an IID sequence $(\tilde{Z}^{-i})_{i\geq 1}$, of Q-Markov chains started from the initial distribution $p' = (p'_n, n \geq 1)$,
- the associated sequence $\widetilde{\zeta}_{-i} := \inf\{t \geq 0 : \tilde{Z}_t^{-i} = 0\}$ of jumps distributed as $\mathbb{P}_{p'}(\zeta \in \cdot)$, $i \geq 1$,
- an IID sequence X^{-i} of excursions distributed as X, started from levels $\widetilde{\zeta}_{-i}$, $i \geq 1$.

Let $\xi^{-i} := \inf\{s \geq 0 : X_s^{-i} = 0\}$ be a sequence of hitting times of 0 by the excursions $(X^{-i})_{i\geq 1}$, let $(C^{-i})_{i\geq 1}$ denote the arrival times of $\widetilde{J}$ where $0 < C^{-1} < C^{-2} < \cdots$ and for $s < 0$, define $(\widetilde{X}_s)_{s<0}$ by $\widetilde{X}_s := \inf\{t \geq 0 : t + \sum_{i=1}^{\widetilde{J}_t} \xi^{-i} \geq |s|\}$. Observe that $(\widetilde{X}_t)_{t<0}$ has intervals $[L_{-i}, R_{-i}], i \geq 1$ of length ξ^{-i} along which it is

constant, with $R_{-(i+1)} < L_{-i} < R_{-i}$ at heights C^{-i}. In particular $C^{-i} = \widetilde{X}_{R_{-i}}$. Along each interval of constancy we add an excursion X^{-i} of X. Define for $s < 0$, $(X_s)_{s<0}$ by

$$X_s := \begin{cases} C^{-i} + X^{-i}_{s-L^{-i}} \text{ if there exists } i \text{ such that } s \in [L_{-i}, R_{-i}], \\ \widetilde{X}_s \text{ otherwise.} \end{cases}$$

Let $\cdots < S^{-3} < S^{-2} < S^{-1} < 0$ be the jump times of $(X_s)_{s\leq 0}$. Write $\zeta^{-i} := X_{S^{-i}} - X_{S^{-i}-}$ for the height of the $(-i)$-th jump. Analogously to before, let $y \geq 0$, $i \geq 1$. Recalling that $B^0 := 0$,

$$\begin{aligned} &B^{-i} = X_{S^{-i}-}, && \text{and} \quad D^{-i} = X_{S^{-i}}, \\ &I^y_{-1} := \inf\{l > 0 \colon B^{-l} \leq y < D^{-l}\} && \text{and} \quad I^y_{-(i+1)} := \inf\{l > I^y_{-i}, : B^{-l} \leq y < D^{-l}\}, \\ &\widetilde{K}^y := \inf\{i \geq 0 : I^y_{-(i+1)} = \infty\} && \text{and} \quad N^{-i,y} := Z^{-I^y_{-i}}(y - B^{-I^y_{-i}}),\ -\widetilde{K}^y \leq -i \leq -1. \end{aligned}$$

Define the *skewer* $(\widetilde{\mathcal{S}}^y)_{y\geq 0}$ by $\widetilde{\mathcal{S}}^y := \mathcal{S}^y(N^{-1,y}, N^{-2,y}, \ldots, N^{-\widetilde{K}^y,y})$.

Theorem 2.3.11. *Suppose that $(\mathcal{S}^{j,y})_{y\geq 0}$ are skewer processes with independent versions of $X, (Z^i)_{i\geq 0}$, each with $p_1 = 1$, $p_n = 0, n \geq 2$, $Q = Q_\alpha$, as in Corollary 2.3.9, that $p' = p$ and that $\mathcal{S}^{j,0} = (n_j) \in \mathcal{C}$. Then the skewer process $(\widetilde{\mathcal{S}}^y)_{y\geq 0}$ evolves as an ordered (α, θ)-up down CRP chain started from $(n_1, \ldots, n_k) \in \mathcal{C}_n$.*

2.3.4 The skewer is a Poissonised up-down CRP

Proof of Theorem 2.3.8. Let $(X_s)_{s\geq 0}$ be a copy of the Lévy process constructed from $(Z^i)_{i\geq 0}$ and $(J_s)_{s\geq 0}$ as before via $X_t := -t + \sum_{i=0}^{J_t} \zeta^i$. Denote the distribution of $(X_t)_{0\leq t\leq T}$ by P_{n_0}.

Set $\zeta_\varnothing := X_0$. Let us show that there is a stopped Poisson process of rate α corresponding to the birth times of the tables inserted directly to the left of the initial table. Note that this stopped Poisson process is distinct from, and not obtained by stopping the process J in the definition of $(X_t)_{t\geq 0}$ above. It arises by considering the levels at which the running minimum of the Lévy process is constant, noting that

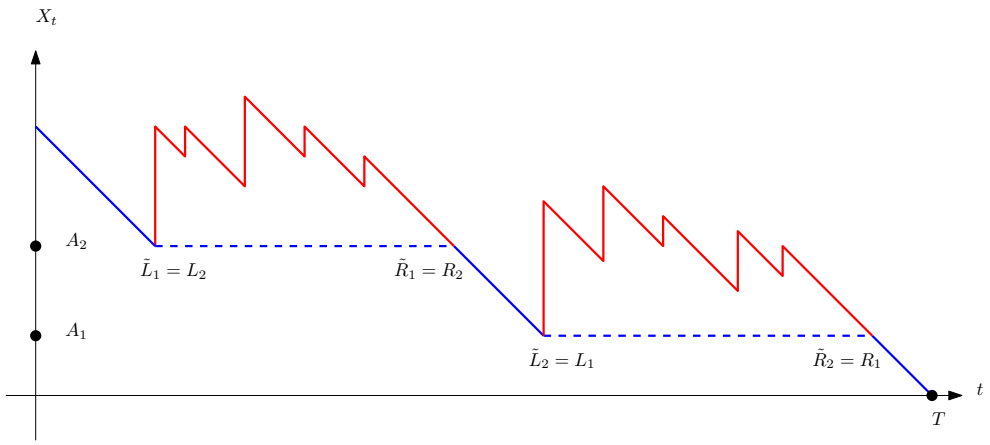


Fig. 2.7: An excursion of the JCCP with notation used in the proof of Theorem 2.3.8 below.

in this proof we work on a probability space for which the Lévy process continues even after reaching 0 for the first time. Define

$$\tilde{L}_1 := \inf\{s \geq 0 \colon X_s \neq X_{s-}\}, \qquad \tilde{R}_1 := \inf\{s \geq \tilde{L}_1 \colon X_s = X_{\tilde{L}_1-}\},$$
$$\tilde{L}_j := \inf\{s \geq \tilde{R}_{j-1} \colon X_s \neq X_{s-}\}, \quad \tilde{R}_j := \inf\{s \geq \tilde{L}_j \colon X_s = X_{\tilde{L}_j-}\} \qquad \text{for } j \geq 2.$$

Set $K_\varnothing := \sup\{j \geq 0 \colon \tilde{R}_j < T\}$ where $\tilde{R}_0 := 0$, and define the first generation of a genealogy by $(L_j, R_j) := (\tilde{L}_{K_\varnothing - j+1}, \tilde{R}_{K_\varnothing - j+1})$, where the individuals in the genealogy are tables and the children of each table are those inserted directly to the left of it at rate α. Write $\tilde{A}_j := X_0 - X_{\tilde{R}_j}$, for $j \geq 1$ and $A_j := X_{R_j}$ for any $1 \leq j \leq K_\varnothing$, where $\tilde{A}_0 := 0$.

We claim that the $(\tilde{A}_j)_{j\geq 1}$ are the points of a rate α Poisson process independent of the associated IID sequence of excursions $(X_{t+\tilde{L}_j} - X_{\tilde{R}_j})_{0\leq t\leq \tilde{R}_j - \tilde{L}_j}$, $j \geq 1$. To show this, observe that $\tilde{A}_1 = X_0 - X_{\tilde{R}_1} = X_0 - X_{\tilde{L}_1-}$. This is by definition equal to $\tilde{L}_1$ which is exponentially distributed of rate α.

Now for each $j \geq 1$, apply the strong Markov property to X at times $\tilde{L}_j$ and $\tilde{R}_j$. Then it follows that the inter-arrival times $(\tilde{A}_j - \tilde{A}_{j-1})_{j\geq 1}$ are all IID $Exp(\alpha)$ random variables independent of the IID sequence of excursions $(X_{t+\tilde{L}_j} - X_{\tilde{R}_j})_{0\leq t\leq \tilde{R}_j - \tilde{L}_j}$, $j \geq 1$, as claimed. In particular, the $\tilde{A}_j$ are the points of a $PP(\alpha)$ as claimed.

Note that since X_0 and $(X_t - X_0)_{t\geq 0}$ are independent, it follows that X_0 is independent of the $(\tilde{A}_j)_{j\geq 1}$ as this infinite sequence can be defined solely in terms of $(X_t - X_0)_{t\geq 0}$. By time-reversal, it follows that

(S) under P_{n_0}, we have $Z^0 \sim \mathbb{P}_{n_0}$ and the $(A_j)_{1\leq j\leq K_\varnothing}$ are the jump times of a $PP(\alpha)$ restricted to $[0, \zeta_\varnothing]$, and the excursions $(X_{t+L_j} - X_{R_j})_{0\leq t\leq R_j - L_j}$, $1 \leq j \leq K_\varnothing$, are IID conditionally given $K_\varnothing$, with distribution P_1.

We will now verify the jump-chain-holding-time structure of the skewer process by an induction on the steps of the jump chain. Note that the total number of steps M of the jump chain is finite almost surely.

Consider the inductive hypothesis that the first $m \wedge M$ steps satisfy the jump-chain-holding-time structure of the Poissonised ordered $(\alpha, 0)$-up-down CRP, and given those jump chain and holding time variables with sum T_m, and skewer $\mathcal{S}^{T_m} = (n_1, \ldots, n_k)$, the excursions above T_m are independent excursions with distributions P_{n_j}, $1 \leq j \leq n_k$.

Assuming the inductive hypothesis for some $m \geq 0$ and $(n_1, \ldots, n_k) \in \mathcal{C}$, we note that if $T_m = \infty$ or if $T_m < \infty$ with $(n_1, \ldots, n_k) = \varnothing$ the induction proceeds trivially, while if $T_m < \infty$ with $(n_1, \ldots, n_k) \neq \varnothing$, there are independent exponential clocks

(i) of rate $n_j - \alpha$, from Z^0 of (S) under P_{n_j}, triggering a transition to $(n_1, \ldots, n_{j-1}, n_j + 1, n_{j+1}, \ldots, n_k)$, for each $j \in [k]$,

(ii) of rate n_j, from Z^0 of (S) under P_{n_j}, triggering a transition to $(n_1, \ldots, n_{j-1}, n_j - 1, n_{j+1}, \ldots, n_k)$, or to $(n_1, \ldots, n_{j-1}, n_{j+1}, \ldots, n_k)$ if $n_j = 1$, for each $j \in [k]$,

(iii) of rate α, from the $PP(\alpha)$ of (S) under P_{n_j}, triggering a transition to $(n_1, \ldots, n_{j-1}, 1, n_j, \ldots, n_k)$, for each $j \in [k]$.

Whichever is the smallest, denoted by E_{m+1}, the lack of memory property of the other exponential clocks makes the system start afresh, with either $\mathcal{S}^{T_{m+1}} = \varnothing$ or

independent excursions above $T_{m+1} = T_m + E_{m+1}$. Specifically, in (iii), we note that the post-E_{m+1} $PP(\alpha)$ is still $PP(\alpha)$, and is independent of the first excursion as required for the induction to proceed. □

To prove Theorem 2.3.11, observe that $(X_s)_{s\leq 0}$ and $(X^j_s)_{0\leq s\leq T_j}$ are independent by construction. There is a fourth class of new transitions from $(X_s)_{s\leq 0}$, which are those from state $(n_1, \ldots, n_k) \in \mathcal{C}_n$ into state $(n_1, \ldots, n_k, 1)$ at rate θ corresponding to the insertion of a rightmost table at rate θ. All other transitions from $(X_s)_{s\leq 0}$ are still of the correct rate by construction, and essentially the same argument as at that used at the end of the last proof will suffice to show that we have an (α, θ)-up-down chain as required. Note that the probabilities for each of the three types of transitions are now $(2n_j - \alpha)/(2n + \theta)$, $\alpha/(2n + \theta)$ and $\theta/(2n + \theta)$ respectively.

2.3.5 Generalised Chinese Restaurants

Note that much of the proof of Theorem 2.3.11 above does not rely on the specific transition rates. Under certain conditions, the skewer processes are part of a new class of processes which we refer to as generalised (α, θ)-up-down Chinese Restaurant Processes. To the best of our knowledge this is a novel construction which does not appear in any existing literature. We describe how these processes evolve, and then state a set of conditions under which they coincide with the skewer process constructions of 2.3.3.

Definition 2.3.12 (Generalised (α, θ)-up-down Chinese Restaurant Process)**.** Let $\alpha, \theta \geq 0$, let Z and $\tilde{Z}$ be any pair of continuous-time Q-Markov chains on $\mathbb{N}_0$ absorbed at 0 with initial distributions $p = (p_n, n \geq 1)$ and $p' = (p'_n, n \geq 1)$ and let $\alpha, \theta \geq 0$. The *generalised* (α, θ)*-up-down Chinese Restaurant Process* is a continuous-time Markov process where from state $(n_1, \ldots n_k) \in \mathcal{C}_n$ there are independent exponential clocks

- of rate $q_{n_i,l}$, triggering a transition into $(n_1, \ldots, n_{i-1}, l, n_{i+1}, \ldots, n_k)$ if $l \geq 1$ or into $(n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_k)$ if $l = 0$, for each $1 \leq i \leq k$,

- of rate α triggering a transition into $(n_1, \ldots, n_{i-1}, P, n_{i+1}, \ldots, n_k)$ for each $1 \leq i \leq k$, where $P \sim (p_n)_{n\geq 1}$,
- of rate θ triggering a transition into $(n_1, \ldots, n_{i-1}, n_i n_{i+1}, \ldots, n_k, P')$, where $P' \sim (p'_n)_{n\geq 1}$.

To understand more clearly how this process evolves, it should be thought of as a continuous time up-down ordered (α, θ)-CRP, but with the following changes made to its evolution.

- The chains that govern the evolution of table sizes are replaced by more general chains on $\mathbb{N}_0$, which are absorbed at 0. They correspond to multiple customers arriving or leaving at once.
- When a new table arrives to the left of an existing one at rate α, its initial size follows the $\mathbb{N}$-valued distribution p.
- When a new table arrives to the right of all existing ones at rate θ, its initial size follows the $\mathbb{N}$-valued distribution p'.

Theorem 2.3.13. *Let Z and $\tilde{Z}$ be any pair of continuous-time Q-Markov chains on $\mathbb{N}_0$ absorbed at 0 with initial distributions $p = (p_n, n \geq 1)$ and $p' = (p'_n, n \geq 1)$ respectively, and define $\zeta := \inf\{t \geq 0 \colon Z_t = 0\}$. If $\mathbb{E}_n(\zeta) < \infty$ for every $n \geq 1$ and $\mu := \sum_{n=1}^{\infty} p_n \mathbb{E}_n(\zeta) \leq \frac{1}{\alpha}$, then the skewer process $(\widetilde{\mathcal{S}}^y)_{y\geq 0}$ is still Markov.*

Note that the assumption $\mu \leq \frac{1}{\alpha}$ of (sub)-criticality is required for the excursions of the Lévy process to be almost surely finite. If this does not hold, then it is not necessarily possible to apply the Strong Markov property at time $\tilde{R}_1$ and the proof of Theorem 2.3.11 breaks down.

As is reasonable to expect, the generalised (α, θ)-up-down CRP arises as the distribution of a skewer process under the conditions of Theorem 2.3.13. The proof follows by using essentially the same argument as with Theorem 2.3.8, and is therefore omitted.

Theorem 2.3.14. *Following the same construction as described before Theorem 2.3.11 and under the assumptions of Theorem 2.3.13 above, the skewer process* $(\tilde{S}^y)_{y\geq 0}$ *is a Generalised* (α, θ)*-up-down Chinese Restaurant process.*

2.4 Scaling limits and $(1+\alpha)$-stable Lévy processes

In this section, we describe a scaling limit relating a version of the chain governing table sizes in the Poissonised up-down CRP and the Bessel-squared process, as well as for the Lévy process controlling the JCCP and the $(1+\alpha)$-stable Lévy process. We will also show how the stable Lévy process arises naturally from Bessel-squared processes. Before discussing these new results in Section 4.2, it will be necessary to cover some preliminary material in Section 4.1 below. The reader acquainted with this material may wish to skip ahead to Section 4.2, referring back to Section 4.1 as required.

The reader is invited to recall the definition of the Skorokhod topology in Section 4.1.1 below, which roughly follows Billingsley [6]. Similarly, we will also recall the definition of the Bessel-squared process from Revuz and Yor [44], and Göing-Jaeschke and Yor [22] in Section 4.1.2, and will discuss some material on semimartingales and diffusions in Section 4.1.3 following Jacod and Shiryaev [24]. A reader familiar with these topics is advised only to read over Lemma 2.4.4 and Theorem 2.4.15, as they will be used in the proof of Theorem 2.4.16 in Section 4.2.

2.4.1 Relevant material from the literature

2.4.1.1 The Skorokhod topology

In order to describe convergence of stochastic processes, we will work with the Skorokhod topology on the space of càdlàg paths. Note that in particular, we will view our stochastic processes as random variables taking values in a space of càdlàg functions.

The Skorokhod topology allows us to view functions as converging not only in

space, but also in time. We begin by defining Skorokhod convergence for functions defined on $[0,t]$.

Definition 2.4.1 (Skorokhod topology on $\mathbb{D}([0,t],E)$)**.** Let E be a metric space and let $\mathbb{D}([0,t],E)$ denote the space of càdlàg functions from $[0,t]$ to E. The *Skorokhod topology on* $\mathbb{D}([0,t],E)$ is the topology induced by the metric

$$d_t(f,g) := \inf_{\lambda\in\mathbb{L}} \Big(||\lambda||_t^\circ \vee ||f - g\circ\lambda||_t \Big),$$

where $\mathbb{L}$ is the set of all strictly increasing bijections of $[0,t]$ into itself, $||\cdot||^\circ$ is the function on $\mathbb{L}$ with $||\lambda||_t^\circ := \sup_{r,s:\, 0\le r<s\le t} \big|\ln\big(\frac{\lambda(s)-\lambda(r)}{s-r}\big)\big|$, and $||\cdot||_t$ is the supremum norm on $\mathbb{D}([0,t],E)$.

Now we extend to functions defined on $[0,\infty)$. It ought to be a natural definition to simply say that $f_n \to f$ in the Skorokhod topology on $\mathbb{D}([0,\infty),E)$ if and only if for every $t>0$, $d_t(f_n,f)\to 0$ upon restricting f_n and f to $[0,t]$ as necessary. However, this does not work well as a definition, since we should have $\mathbf{1}_{[0,1-\frac{1}{n})} \to \mathbf{1}_{[0,1)}$ in $\mathbb{D}([0,\infty),\mathbb{R})$, yet $d_1(\mathbf{1}_{[0,1-\frac{1}{n})}, \mathbf{1}_{[0,1)}) = 1$. The issue is that 1 is a discontinuity point of $\mathbf{1}_{[0,1)}$, but we can fix this issue by smoothing out our approximating functions.

Definition 2.4.2 (Skorokhod topology on $\mathbb{D}([0,\infty),E)$)**.** Let (E,ρ) be a metric space, let $\mathbb{D}([0,\infty),E)$ denote the space of càdlàg functions from $[0,\infty)$ to E and view $\mathbb{D}([0,t],E)$ as a subspace by restricting these functions to $[0,t]$. The *Skorokhod topology on* $\mathbb{D}(E,[0,\infty))$ is the topology induced by the metric

$$d(f,g) := \sum_{k=1}^{\infty} \frac{1\wedge d_k(f^{(k)},g^{(k)})}{2^k},$$

$$\text{where} \qquad f^{(k)}(t) := f(t)\mathbf{1}_{\{t<k-1\}} + (k-t)f(t)\mathbf{1}_{\{k-1\le t\le k\}}.$$

The following properties of the Skorokhod topology are used later.

Lemma 2.4.3. *The following results hold.*

(a) *[6, Theorem 16.2]* $d(f_n, f) \to 0$ *in* $\mathbb{D}([0,\infty), E)$ *if and only if* $d_t(f_n, f) \to 0$ *for every continuity point* t *of* f.

(b) *[6, Lemma 16.1, Theorem 16.2] If* $f_n \to f$ *in* $\mathbb{D}([0,\infty), E)$ *and* f *is continuous, then* $f_n \overset{unif}{\to} f$ *on* $[0,t]$ *for any* $t > 0$.

(c) *[6, Theorem 16.3] If* E *is a Polish space, so is the space* $\mathbb{D}([0,\infty), E)$. *In particular,* $\mathbb{D}([0,\infty), \mathbb{R})$ *is a Polish space.*

(d) *If* $f_n \overset{unif}{\to} f$, *then* $f_n \to f$ *in* $\mathbb{D}([0,\infty), \mathbb{R})$.

To show that convergence in distribution of hitting times holds, we can use the following lemma.

Lemma 2.4.4. *Define a pair of functions* $\tau\colon \mathbb{D}([0,\infty), \mathbb{R}) \to [0,\infty]$ *and* $\tau^<\colon \mathbb{D}([0,\infty), \mathbb{R}) \to [0,\infty]$ *by*

$$\tau(f) := \inf\{t \geq 0\colon f(t) \leq 0\}, \quad \tau^<(f) := \inf\{t \geq 0 : f(t) < 0\},$$

and write $C([0,\infty), \mathbb{R}) := \{f \in \mathbb{D}([0,\infty), \mathbb{R})\colon f$ *is continuous on* $[0,\infty)\}$,

$$\mathbb{D}^*([0,\infty), \mathbb{R}) := \{f \in \mathbb{D}([0,\infty), \mathbb{R})\colon \forall t \geq \tau(f), f(t) \leq 0,\ \tau^<(f) = \tau(f) < \infty\}.$$

If $f_n \to f$ *in* $\mathbb{D}([0,\infty), \mathbb{R})$ *and* $f \in C([0,\infty), \mathbb{R}) \cap \mathbb{D}^*([0,\infty), \mathbb{R})$, *then* $\tau(f_n) \to \tau(f)$.

Proof. Suppose that $f_n \in \mathbb{D}([0,\infty), \mathbb{R})$ converges under the Skorokhod topology to some function $f \in \mathbb{D}^*([0,\infty), \mathbb{R}) \cap C([0,\infty), \mathbb{R})$. For any $t \geq 0$, $d_t(f_n, f) \to 0$ by Lemma 2.4.3 (a), and as $f \in C([0,\infty), \mathbb{R})$, $f_n \overset{\text{unif}}{\to} f$ on $[0,t]$ by Lemma 2.4.3 (b).

Let $\varepsilon > 0$ and observe that $\tau(f) < \infty$. Then it follows from uniform convergence on $[0, \tau(f) - \varepsilon]$, the definition of τ and the fact that $\inf_{0 \leq t \leq \tau(f)-\varepsilon} f(t) > 0$, that there exists $N_1 \in \mathbb{N}$ such that for any $n > N_1, f_n > 0$ on $[0, \tau(f) - \varepsilon]$. This yields $\tau(f_n) > \tau(f) - \varepsilon$ for every $n > N_1$.

To show that $\tau(f_n) < \tau(f) + \varepsilon$ for sufficiently large n, recall that as $f \in \mathbb{D}^*([0,\infty), \mathbb{R})$, $\tau(f) = \tau^<(f)$. By assumption, there is some $\delta \in (0, \varepsilon)$ with $f(\tau^<(f) + \delta) < 0$ by continuity of f. Now $f_n \overset{\text{unif}}{\to} f$ on $[0, \tau^<(f) + \delta]$, and so there exists

$N_2 \in \mathbb{N}$ such that for any $n > N_2$, $f_n(\tau(f) + \delta) < 0$. It follows that $\tau(f_n) \leq \tau(f) + \delta \leq \tau(f) + \varepsilon$. Setting $N := \max(N_1, N_2)$, we have shown that for all $n > N$, $|\tau(f) - \tau(f_n)| < \varepsilon$ as required. □

2.4.1.2 Bessel-squared processes

It will be shown that scaling limits of versions of the chain governing the evolution both the mass chain and table sizes converge to Bessel-squared processes, so we describe them here.

Definition 2.4.5 (Bessel-squared process)**.** [22, 44] Let $\delta \in \mathbb{R}$ and suppose that $(W_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}$. The unique strong solution of the SDE $dY_t = \delta dt + 2\sqrt{|Y_t|}dW_t$ given the initial state Y_0 is called a *Bessel-squared process with dimension* δ. If $Y_0 = y$, its distribution on $C([0,\infty), \mathbb{R})$ is denoted by $BES^2_y(\delta)$.

Remark 2.4.6. As noted by [22], the SDE defining the $BES^2(\delta)$ does indeed have a unique strong solution. The drift parameter δ is referred to as the dimension since it can be shown for any integer $n \in \mathbb{N}$, that if $(W_t)_{t\geq 0}$ is an n-dimensional Brownian motion, $||W_t||^2$ is a $BES^2(n)$ process.

We state some properties of Bessel-squared processes needed later.

Proposition 2.4.7 (Pathwise properties of the Bessel-squared process)**.** *Let $\delta \in \mathbb{R}$, $y \geq 0$, and suppose that Y is a $BES^2_y(\delta)$ process.*

(a) [22, p.330] $-Y$ is a $BES^2_{-y}(-\delta)$ process.

(b) [22, Corollary 1] If $\delta \geq 2$ and $y > 0$, then Y does not hit 0 almost surely, and if $\delta < 2$ then Y hits 0 almost surely.

(c) [44, XI. Proposition 1.5] If $\delta \geq 0$ then $Y_t \geq 0$ $\forall t \geq 0$, with 0 an absorbing state if and only if $\delta = 0$. If $\delta \in (0, 2)$, then 0 is instantaneously reflecting. When $\delta \geq 2$, 0 is inaccessible.

(d) [44, p.442] If $\delta > 2$ then $Y_t \to \infty$ as $t \to \infty$ almost surely, and if $\delta < -2$ then $Y_t \to -\infty$ as $t \to \infty$ almost surely.

Remark 2.4.8. Note that in part (d), [44] only gives the result for $\delta > 2$. The result for $\delta < -2$ follows from this and part (a).

Corollary 2.4.9. *When $\delta < 0$ and $y \geq 0$, almost all the paths of the $BES^2_y(\delta)$ lie in $C([0,\infty),\mathbb{R}) \cap \mathbb{D}^*([0,\infty),\mathbb{R})$.*

Proof. Let f be a path of the $BES^2_y(\delta)$. Proposition 2.4.7 (a) tells us that $-f$ is a $BES^2_{-y}(-\delta)$ path, and so by Proposition 2.4.7 (b)

- f hits 0 almost surely if $\delta \in (-2, 0)$,
- f does not hit 0 almost surely if $\delta \leq -2$.

By Proposition 2.4.7 (c),

- $f \leq 0$ on $[\tau(f), \infty)$ almost if $\delta \in (-2, 0)$,
- $f < 0$ on $(\tau(f), \infty)$ almost surely if $\delta \leq -2$.

It is clear by almost sure continuity of f that $f(\tau(f)) = 0$, since $f > 0$ on $[0, \tau(f))$. To show that $\tau^<(f) = \tau(f)$, suppose for a contradiction that this was not the case. Since $f \leq 0$ on $[\tau(f), \infty)$, $\tau^<(f) > \tau(f)$ would imply that $f = 0$ on $[\tau(f), \tau^<(f))$. This contradicts part (c) of Proposition 2.4.7, and so $\tau^<(f) = \tau(f)$. Hence $f \in \mathbb{D}^*([0,\infty),\mathbb{R})$ almost surely. □

Remark 2.4.10. Observe by Proposition 2.4.7 (c) that when $\delta \geq 0$, $\tau^<(f) = \infty$. In particular Corollary 2.4.9 fails for $\delta \geq 0$.

2.4.1.3 Semi-martingales and diffusions

To describe some of the technicalities needed for a proof later, we follow [24]. In what follows, $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P})$ is a filtered probability space and $\mathbf{F} = (\mathcal{F}_t)_{t\geq 0}$.

Definition 2.4.11 (Continuous semi-martingale/characteristics)**.** A *continuous semi-martingale* $(X_t)_{t\geq 0}$ is a real-valued stochastic process on $(\Omega, \mathbf{F}, \mathcal{F}, \mathbb{P})$ of the form $X_t = X_0 + M_t + B_t$, where X_0 is $\mathcal{F}_0$-measurable, M is a continuous local $\mathbf{F}$-martingale started from 0 and B_t is a continuous $\mathbf{F}$-adapted process of finite variation. The *quadratic variation* $C_t = \langle M \rangle_t$ of X, is the unique continuous increasing $\mathbf{F}$-adapted process such that $M_t^2 - \langle M \rangle_t$ is a local martingale, and we refer to the pair $(B_t, C_t)_{t\geq 0}$ as the *characteristics* of X.

If η is any Borel probability measure, we say that the *martingale problem* $s(\sigma(X_0), X|\eta; B, C)$ has a solution (P, X) if P is a probability measure on $(\Omega, \mathcal{F})$ with $P(X_0 \in A) = \eta(A)$ for all Borel sets A, and X_t is a continuous semi-martingale on $(\Omega, \mathbf{F}, \mathcal{F}, P)$ with characteristics $(B_t)_{t\geq 0}$ and $(C_t)_{t\geq 0}$.

Definition 2.4.12 (Homogeneous diffusion)**.** A *homogeneous diffusion* is a semi-martingale X on $(\Omega, \mathbf{F}, \mathcal{F}, \mathbb{P})$ for which there exist Borel functions $b\colon \mathbb{R} \to \mathbb{R}$ and $c\colon \mathbb{R} \to [0, \infty)$ called the *drift coefficient* and *diffusion coefficient* with that $B_t = \int_0^t b(X_s)ds$ and $C_t = \int_0^t c(X_s)ds$.

Remark 2.4.13. Note that in particular, the $BES^2(\delta)$ process is a homogeneous diffusion with drift coefficient $b(x) = \delta$ and diffusion coefficient $c(x) = 4|x|$. This can be seen from the SDE characterisation in Definition 2.4.5.

We will need two results from [24].

Theorem 2.4.14. *[24, Theorem III 2.26] Consider a homogeneous diffusion X and let our notation be as above. The martingale problem $s(\sigma(X_0), X|\eta; B, C)$ has a unique solution if and only if the SDE $dY_t = b(Y_t)dt + \sqrt{c(Y_t)}dW_t$, $Y_0 = \xi$ has a unique weak solution, for $(W_t)_{t\geq 0}$ an $\mathbf{F}$-Brownian motion and ξ an $\mathcal{F}_0$-measurable random variable with distribution η.*

Theorem 2.4.15. *[24, Theorem IX 4.21] Suppose that for any $x \in \mathbb{R}$, the martingale problem $s(\sigma(X_0), X|\delta_x; B, C)$ has a unique solution P_x, and that $x \mapsto P_x(A)$ is Borel measurable for all $A \in \mathcal{F}$. Let $X^{(n)}$ be a Markov process started from an initial*

distribution η^n, where $X^{(n)}$ has generator $\mathcal{L}^n$ given by $\mathcal{L}^n f(x) = \int_{y\in\mathbb{R}} \big(f(x+y) - f(x)\big)K^n(x,dy)$, for K^n a finite kernel on $\mathbb{R}$. Suppose further that for any $x \in \mathbb{R}$, the quantities $b^n(x)$ and $c^n(x)$ given by

$$b^n(x) := \int_{y\in\mathbb{R}} yK^n(x,dy) \quad \text{and} \quad c^n(x) := \int_{y\in\mathbb{R}} y^2K^n(x,dy)$$

are well-defined and finite, and that the drift and diffusion coefficients $b(x)$ and $c(x)$ of X satisfy

(a) $b^n \to b$ and $c^n \to c$ locally uniformly,

(b) $\sup_{\{x:\,|x|\le a\}} \int_{y\in\mathbb{R}} |y|^2 \mathbf{1}_{\{|y|>\varepsilon\}} K^n(x,dy) \to 0$ as $n \to \infty$ for any $\varepsilon > 0$ and any $a > 0$,

(c) $\eta^n \to \eta$ weakly on $\mathbb{R}$, where η is the initial distribution of X.

Then $X^{(n)}$ converges in distribution to X as $n \to \infty$ under the Skorokhod topology on $\mathbb{D}([0,\infty),\mathbb{R})$.

2.4.2 Results on convergence to scaling limits

As in the introduction, we let $0 \le \alpha \le 1$ and let $\theta \ge 0$. Recall from the remarks before Proposition 2.2.11 that the mass chain is the number of customers in a Poissonised Chinese Restaurant, and that it is an up-down Markov chain on $\mathbb{N}_0$ with up rate $n+\theta$ from state $n \ge 0$ and down rate n from state $n \ge 1$. Similarly, the number of customers at each table has up rate $n-\alpha$ and down rate n from state $n \ge 1$ and is absorbed at state 0.

Theorem 2.4.16. *Let $\big(Y_z^\delta(t)\big)_{t\ge0}$ be a modified $BES_z^2(\delta)$ process absorbed at 0, where $z > 0$. If $(Z_t^n)_{t\ge0}$ is a version of the chain governing the table size of the Poissonised up-down chain with $\lfloor nz \rfloor$ initial customers, and M_t^n is a version of*

the mass chain started from $\lfloor ny \rfloor, y > 0$, *then for* $\zeta_n' := \inf\{t > 0 : \frac{1}{n} Z^n_{2nt} = 0\}$ *and* $\zeta_z^\delta := \inf\{t > 0 : Y_z^\delta \le 0\}$,

$$\Big(\frac{1}{n}(Z^n_{2nt})_{t\ge 0}, \zeta_n'\Big) \xrightarrow{D} \Big(\big(Y_z^{-2\alpha}(t)\big)_{t\ge 0}, \zeta_z^{-2\alpha}\Big) \text{ in } \mathbb{D}([0,\infty),\mathbb{R}) \times [0,\infty),$$
$$\text{and } \frac{1}{n}(M^n_{2nt})_{t\ge 0} \xrightarrow{D} \big(Y_y^{2\theta}(t)\big)_{t\ge 0} \text{ in } \mathbb{D}([0,\infty),\mathbb{R}).$$

Theorem 2.4.17 (Skorokhod's representation theorem)**.** *Let* S *be a separable metric space, and suppose that* $X, X_1, X_2, \dots$ *are* S*-valued random variables. If* $X_n \xrightarrow{D} X$, *then there is a probability space on which are defined* S*-valued random variables* $Y \overset{D}{=} X$ *and* $Y_n \overset{D}{=} X_n$ *with* $Y_n \xrightarrow{a.s} Y$.

Proof of Theorem 2.4.16. Let $x \in \mathbb{R}$. Observe by Theorem 2.4.14 that the martingale problem associated to the SDE characterising the $BES^2(\delta)$ diffusion has a unique solution P_x. Showing that $x \mapsto P_x(A)$ is Borel measurable is fairly elementary, and so we leave this to the reader. Consider an extension of the Markov chain $Z_n(t)$ to $\mathbb{Z}$ that transitions from 0 into -1 at rate α, and from state $-n$, $n \ge 1$, to $-n-1$ at rate n and into $-n+1$ at rate $n-\alpha$. This chain is an instance of a Markov chain on $\mathbb{R}$ with generator $\mathcal{L}$ given by

$$\mathcal{L}f(x) = \big[(|x| \vee \alpha) - \alpha\big](f(x+1) - f(x)) + \big[|x| \vee \alpha\big]\big(f(x-1) - f(x)\big), \quad x \in \mathbb{R}.$$

Let $X^{(n)}(t)$ be any version of $\frac{1}{n} Z^n_{2nt}$ with Z^n extended to $\mathbb{R}$ as above. Then $X^{(n)}$ has generator $\mathcal{L}^n$ and increment kernel $K^n(x, dy)$ given by
$\mathcal{L}^n f(x) = \int_{y\in\mathbb{R}} \big[f(x+y) - f(x)\big] K^n(x, dy)$, where $\forall x \in \mathbb{R}$,

$$K^n(x, dy) = n\bigg(\big[(2n|x| \vee 2\alpha) - 2\alpha\big]\, \delta_{\frac{1}{n}}(dy) + \big[2n|x| \vee 2\alpha\big]\, \delta_{-\frac{1}{n}}(dy)\bigg).$$

Now $$\int_{y\in\mathbb{R}} y K^n(x, dy) = (2n|x| \vee 2\alpha) - 2\alpha - (2n|x| \vee 2\alpha) = -2\alpha,$$

and $$\int_{y\in\mathbb{R}} y^2 K^n(x, dy) = \frac{2n|x| \vee 2\alpha - 2\alpha}{n} + \frac{2n|x| \vee 2\alpha}{n} = 4(|x| \vee \frac{\alpha}{n}) - \frac{2\alpha}{n} \to 4|x|.$$

This convergence is uniform in $x \in \mathbb{R}$ under the supremum norm $||\cdot||$ since

$$\left|\left| 4(|x| \vee \frac{\alpha}{n}) - \frac{2\alpha}{n} - 4|x| \right|\right| = \left|\left| 4|x| + \frac{2\alpha}{n} - 4(|x| \vee \frac{\alpha}{n}) \right|\right| = \frac{2\alpha}{n} \to 0.$$

Moreover, for any $a \geq 0$, $\varepsilon > 0$ and n sufficiently large,

$$\sup_{\{x:|x|\leq a\}} \int_{y\in\mathbb{R}} |y|^2 \mathbf{1}_{\{|y|>\varepsilon\}} K^n(x, dy) = 0.$$

It is clear that as $n \to \infty$, $\frac{\lfloor nz \rfloor}{n} \to z$. This shows that all the assumptions of Theorem 2.4.15 are satisfied.

It follows that the law of $X^{(n)}(t)$ converges weakly in the Skorokhod topology on $\mathbb{D}([0,\infty), \mathbb{R})$ to that of the diffusion process with drift coefficient -2α and diffusion coefficient $4|x|$, namely a $BES^2(-2\alpha)$ process. The convergence of the scaled mass chain to the $BES^2(2\theta)$ diffusion is proved in a very similar way, where the extension of M_t^n to $\mathbb{R}$ has generator given by

$$\mathcal{L}f(x) := \big[|x| + \theta\big]\big(f(x+1) - f(x)\big) + |x|\big(f(x-1) - f(x)\big)$$

instead. We apply Lemma 2.4.4 to show joint convergence of distribution of the scaled process and the hitting times for $0 < \alpha \leq 1$. Note that the paths of the $BES^2(-2\alpha)$ process lie in $\mathbb{D}^*([0,\infty), \mathbb{R}) \cap C([0,\infty), \mathbb{R})$ by Corollary 2.4.9 since $-2\alpha \in [-2, 0)$. By applying Theorem 2.4.17 on the separable space $\mathbb{D}([0,\infty), \mathbb{R})$, there are versions $V^{(n)}$ and V' of $X^{(n)}$ and a $BES_z^2(-2\alpha)$ process $Z_z^{-2\alpha}$ with $V^{(n)} \overset{a.s}{\to} V'$, such that V' takes values in $\mathbb{D}^*([0,\infty), \mathbb{R}) \cap C([0,\infty), \mathbb{R})$ almost surely. By Lemma 2.4.4, $\tau(V^{(n)}) \overset{a.s}{\to} \tau(V')$. Therefore $(V^{(n)}, \tau(V^{(n)})) \overset{a.s}{\to} (V', \tau(V'))$, and in particular $(X^{(n)}, \zeta_n') \overset{D}{\to} (V_z^{-2\alpha}, \zeta_z^{-2\alpha})$ as required.

If instead $\alpha = 0$, then we cannot apply Corollary 2.4.9. Observe that a version of the chain describing each table's evolution is a simple birth-death process. By a standard reference such as [23, Corollary 6.11.12], the hitting time ξ_m of 0 by a simple birth-death process started from height $m \in \mathbb{N}_0$ satisfies $\mathbb{P}(\xi_m \leq t) = (\frac{t}{t+1})^m$,

so that $\mathbb{P}(\zeta'_n \leq t) = (\frac{2nt}{2nt+1})^{\lfloor nz \rfloor}$. As $n \to \infty$, this converges to $e^{-z/2t}$.

By [22, p.319], it is known that $\zeta^\delta_z \stackrel{D}{=} \frac{z^2}{2G}$ for $G \sim Gamma(\frac{2-\delta}{2}, 1)$ when $\delta < 2$. Hence $\mathbb{P}(\zeta^0_z \leq t) = \mathbb{P}(Exp(1) \geq z/2t) = e^{-z/2t}$. Thus we still have $\tau(X^n) \stackrel{D}{\to} \tau(X)$, in the case when $\alpha = 0$.

We need to strengthen this to joint convergence in distribution. Since the distributions of pairs $(X^n, \tau(X^n))$, $n \geq 1$ are tight, it suffices to show that any subsequential distributional limit for the pairs is such that the limiting time is the extinction time of the limiting process. Specifically, if $(X^{n_k}, \tau(X^{n_k}))$ converges, we may assume convergence holds almost surely, by application of Skorokhod's representation theorem 2.4.17, and the limit $(\overline{X}, \overline{\tau})$ is such that $\overline{\tau} \stackrel{D}{=} \tau(\overline{X})$, as the marginal distributions converge. We will use deterministic arguments as in the proof of Lemma 2.4.4 to show that $\overline{\tau} \geq \tau(\overline{X})$ almost surely, and since $\overline{\tau} \stackrel{D}{=} \tau(\overline{X})$, this will yield $\overline{\tau} = \tau(\overline{X})$ almost surely, as required. Let $\varepsilon > 0$. Since $X_{n_k} \to \overline{X}$ almost surely and $\tau(\overline{X}) < \infty$ almost surely, we have $\inf_{0 \leq t \leq \tau(\overline{X})-\varepsilon} X^{n_k}(t) > 0$ for sufficiently large k, i.e $\tau(X^{n_k}) \geq \tau(\overline{X}) - \varepsilon$. Since $\tau(X^{n_k}) \to \overline{\tau}$ almost surely, this yields $\overline{\tau} \geq \tau(\overline{X}) - \varepsilon$, and so $\overline{\tau} \geq \tau(\overline{X})$ almost surely as ε was arbitrary.

□

Let us now show convergence of the process $(X_t)_{t \geq 0}$ defining the JCCP to a $(1+\alpha)$-stable Lévy process, in the special case of an $(\alpha, 0)$-up-down Chinese restaurant with $0 < \alpha < 1$. By convention, we choose to work with the Lévy process $X' = (X'_t)_{t \geq 0}$ defined by $X'_t := X_t - X_0$. Write $X_s = X_0 - s + \sum_{i=1}^{J_s} \zeta_i$, where as before $(J_s)_{s \geq 0}$ is a Poisson process of rate α and the ζ_i are IID distributed as τ under $\mathbb{P}_1$, independently of $(J_s)_{s \geq 0}$, with X_0 distributed as τ under $\mathbb{P}_{n_0}$ for some $n_0 \geq 1$. The following result is due to [46, Theorem 2.7], and we use the formulation from [24, Corollary VII 3.6].

Lemma 2.4.18. *Suppose that X^n and X are càdlàg processes with stationary independent increments. Then $X^n \stackrel{D}{\to} X$ on $\mathbb{D}([0,\infty)) \times \mathbb{R}$ if and only if $X^n_1 \stackrel{D}{\to} X_1$. Moreover, whenever this holds and F_{X^n} and F_X are the Lévy measures of X^n and*

X respectively, then for every $g \in C_2([0,\infty),\mathbb{R})$, $F_{X^n}(g) \to F_X(g)$, where

$C_2([0,\infty),\mathbb{R}) := \{g : [0,\infty) \to \mathbb{R} : g$ *is continuous, bounded and* $\exists \varepsilon > 0$ *with* $g = 0$ *on* $[0,\varepsilon)\}$.

Theorem 2.4.19. *The Lévy process $(X'_t)_{t\geq 0}$ has Laplace exponent $\varphi(\lambda) = \frac{\lambda^{1+\alpha}e^{-\lambda}}{\Gamma(1+\alpha,\lambda)}$, and as $n \to \infty$, $\left(\frac{X'_{2n^{1+\alpha}t}}{2n}\right)_{t\geq 0}$ converges in distribution on $\mathbb{D}([0,\infty),\mathbb{R})$ to the $(1+\alpha)$-stable Lévy process $(\overline{X}_t)_{t\geq 0}$ with Laplace exponent $\psi(\lambda) = \frac{\lambda^{1+\alpha}}{2^\alpha\Gamma(1+\alpha)}$ and no negative jumps.*

Proof. By conditioning on the value of J_t, applying independence of the ζ_i and using Proposition 2.2.11,

$$\begin{aligned}\mathbb{E}(e^{-\lambda X'_t}) &= e^{\lambda t}\mathbb{E}\bigg(\exp\big(-\lambda\sum_{i=1}^{J_t}\zeta_i\big)\bigg) = e^{\lambda t}\sum_{k=0}^{\infty}\mathbb{E}_1(e^{-\lambda\zeta})^k\mathbb{P}(J_t = k)\\ &= e^{\lambda t}\sum_{k=0}^{\infty}\Big(1 - \frac{\lambda}{\alpha} + \frac{\lambda^{1+\alpha}}{\alpha e^{\lambda}\Gamma(1+\alpha,\lambda)}\Big)^k\frac{e^{-\alpha t}(\alpha t)^k}{k!}\\ &= e^{(\lambda-\alpha)t}\sum_{k=0}^{\infty}\Big(\alpha t - \lambda t + \frac{\lambda^{1+\alpha}t}{e^{\lambda}\Gamma(1+\alpha,\lambda)}\Big)^k\frac{1}{k!} = \exp\left(\frac{\lambda^{1+\alpha}t}{e^{\lambda}\Gamma(1+\alpha,\lambda)}\right).\end{aligned}$$

Hence X'_t has Laplace exponent given by

$$\varphi(\lambda) = \frac{\ln\mathbb{E}(e^{-\lambda X'_t})}{t} = \frac{\lambda^{1+\alpha}e^{-\lambda}}{\Gamma(1+\alpha,\lambda)}.$$

To show convergence of the scaled JCCP, note that as $n \to \infty$, the Laplace exponent of $\frac{1}{2n}X'_{2n^{1+\alpha}t}$ satisfies

$$\frac{\ln\mathbb{E}\Big(\exp\big(-\lambda X'_{2n^{1+\alpha}t}/2n\big)\Big)}{t} = \frac{(\lambda/2n)^{1+\alpha}e^{-\lambda/2n}}{\Gamma(1+\alpha,\lambda/2n)}\cdot 2n^{1+\alpha} = \frac{\lambda^{1+\alpha}e^{-\lambda/2n}}{2^\alpha\Gamma(1+\alpha,\lambda/2n)} \to \frac{\lambda^{1+\alpha}}{2^\alpha\Gamma(1+\alpha)}.$$

By the convergence Theorem for Laplace transforms, we have $\frac{1}{2n}X'_{2n^{1+\alpha}} \xrightarrow{D} \overline{X}_1$ as $n \to \infty$. By independence and stationarity of increments, it follows by Lemma 2.4.18 that as $n \to \infty$, $\left(\frac{X'_{2n^{1+\alpha}t}}{2n}\right)_{t\geq 0} \xrightarrow{D} (\overline{X}_t)_{t\geq 0}$ under the Skorokhod topology on

$\mathbb{D}([0,\infty),\mathbb{R})$. □

Proposition 2.4.20. *Let $\alpha \in (0,1)$, and suppose that Q_y^δ is the law of a $BES_y^2(\delta)$ process V_t^δ started from $y>0$. If $\zeta := \inf\{t>0 : V_t^\delta = 0\}$ denotes the absorption time of 0 by V^δ, then as $y \downarrow 0$, $y^{-(1+\alpha)}Q_y^{-2\alpha}(\zeta > s)$ converges to $\widetilde{\Pi}(s,\infty)$, where*

$$\widetilde{\Pi}(ds) = \frac{s^{-(2+\alpha)}}{2^{1+\alpha}\Gamma(1+\alpha)}ds$$

is the Lévy measure of a $(1+\alpha)$-stable Lévy process with no negative jumps.

Proof. Recall from [22, p.319] that under $Q_y^{-2\alpha}$, $\zeta = \frac{y}{2G}$ in distribution, where $G \sim \text{Gamma}(1+\alpha,1)$ has shape parameter $1+\alpha$ and rate parameter 1. Observe that

$$Q_y^{-2\alpha}(\zeta > s) = \mathbb{P}(G < \frac{z}{2s}) = \int_0^{\frac{y}{2s}} \frac{1}{\Gamma(1+\alpha)} e^{-u}u^\alpha du.$$

Substituting $x = \frac{2su}{y}$, we see that

$$y^{-(1+\alpha)}Q_y^{-2\alpha}(\zeta > s) = \int_0^1 \frac{1}{\Gamma(1+\alpha)2^{1+\alpha}} e^{-\frac{y}{2s}x}x^\alpha s^{-(1+\alpha)}dx.$$

By letting $y \downarrow 0$ and using dominated convergence, the right hand expression converges to

$$\int_0^1 \frac{1}{\Gamma(1+\alpha)2^{1+\alpha}}x^\alpha s^{-(1+\alpha)}dx = \frac{(2s)^{-(1+\alpha)}}{\Gamma(2+\alpha)}.$$

Writing $\Pi(s,\infty) = \int_s^\infty \Pi(du)$, we observe that indeed

$$y^{-(1+\alpha)}Q_y^{-2\alpha}(\zeta > s) \to \Pi(s,\infty) \text{ as } y \downarrow 0.$$

□

Remark 2.4.21. *$\widetilde{\Pi}$ is the Lévy measure of a stable process with Laplace exponent $c\lambda^{1+\alpha}/2^\alpha\Gamma(1+\alpha)$, and no negative jumps where $c = \frac{\Gamma(1-\alpha)}{2\alpha(1+\alpha)}$.*

Proposition 2.4.22. *For every $\varepsilon > 0$, it is the case that as $n \to \infty$,*

$$2\alpha n^{1+\alpha}\mathbb{P}_1\Big(\frac{\tau(Z)}{2n} > \varepsilon\Big) \to \Pi(\varepsilon, \infty)$$
$$\text{and} \qquad \mathbb{P}_1\Big(\frac{\tau(Z)}{2n} \in \cdot \Big| \frac{\tau(Z)}{2n} > \varepsilon\Big) \to \frac{\Pi_{|(\varepsilon,\infty)}(\cdot)}{\Pi(\varepsilon, \infty)},$$

where $\Pi(dx) = \frac{1}{c}\widetilde{\Pi}(dx)$ is the Lévy measure with Laplace exponent $\psi(\lambda) = \frac{\lambda^{1+\alpha}}{2^\alpha\Gamma(1+\alpha)}$ from Theorem 2.4.19, τ is the function from Lemma 2.4.4 defined by $\tau(f) := \inf\{t \geq 0\colon f(t) \leq 0\}$ for $f \in \mathbb{D}([0,\infty), \mathbb{R})$ and $c = \frac{\Gamma(1-\alpha)}{2\alpha(1+\alpha)}$ as in remark 2.4.21 above.

Proof. Write F_{X^n} and $F_{\overline{X}}$ for the Lévy measures of X^n and $\overline{X}$ respectively. Setting $X^n_t := \frac{X'_{2n^{1+\alpha}t}}{2n}$ and noting that $X^n_0 = 0$, a combination of Lemma 2.4.18 and Theorem 2.4.19 show that for every $g \in C_2([0,\infty), \mathbb{R})$, $F_{X^n}(g) \to F_{\overline{X}}(g)$. Being a compound Poisson process with drift, one observes by a standard reference such as [30, Section 1.2.2] that X'_t has Lévy measure $\alpha\mathbb{P}(\zeta \in dx)$, so that

$$F_{X^n}(dx) = 2\alpha n^{1+\alpha}\mathbb{P}\Big(\frac{\tau(Z)}{2n} \in dx\Big).$$

Observe that this quantity converges as $n \to \infty$ to the Lévy measure of the $(1+\alpha)$-stable Lévy process with Laplace exponent $\psi(\lambda) = \frac{\lambda^{1+\alpha}}{2^\alpha\Gamma(1+\alpha)}$, and the Lévy measure is the quantity $\Pi(ds)$. The second result is a corollary of the first. □

Remark 2.4.23. *More generally, $2\alpha n^{1+\alpha}\mathbb{P}_1\Big(\big(\frac{Z_{2nt}}{n}\big)_{t\geq 0} \in \cdot\Big) \to \Theta(\cdot)$, where Θ is a $BES^2(-2\alpha)$ excursion measure with $\Theta(\tau \in \cdot) = \Pi(\cdot)$. See [15, 12, 13] for the direct study of limiting structures consisting of $(1+\alpha)$-stable Lévy processes with $BES^2(-2\alpha)$ excursions embedded into the jumps.*

3. MIXING TIMES AND THE ORDERED KINGMAN LATTICE

Abstract

Fulman 2009 shows that the framework of commutation relations can be used to study the behaviour of down-up Markov chains, and he applies this to the unordered up-down (α, θ)-CRP from Pitman 2006. We demonstrate that these techniques still yield valid results when applied to the ordered down-up (α, θ)-CRP of Pitman and Winkel 2009, and generalise a theorem of Fulman 2009 on the explicit form and asymptotics of the maximal separation distance from the case $\theta = 1$ to general the case $\theta > 0$.

3.1 Introduction

In this paper, we adapt and generalise results of Fulman [17] on the maximal separation distance after r steps of a down-up unordered (α, θ)-Chinese restaurant Markov chain on the Kingman lattice to the ordered Chinese restaurant, and furthermore extend his results on asymptotics and exact forms of the maximal separation distance from the case $\theta = 1$ to the case $\theta > 0$.

Let K be the space $\{(\lambda_1, \lambda_2, \ldots, \lambda_l, 0, 0, \ldots) : \lambda_i \in \mathbb{N} \text{ if } 1 \leq i \leq l\}$ known as the *Ordered Kingman lattice* with subspace $K_n := \{\lambda \in K : \sum_{i=1}^{l} \lambda_i = n\}$, where $K_0 := \{(0, 0, 0 \ldots)\}$. These spaces can trivially be identified with the set $\mathcal{C}$ of compositions and subset $\mathcal{C}_n$ of compositions of n; details are the same as those found in [38, p.15]. Note that in Fulman, the Kingman lattice has decreasing elements, this corresponds to the standard two-parameter Dubins-Pitman CRP [38] with the

structure of [40], we refer to this as the *Unordered Kingman Lattice*.

The Ordered Chinese Restaurant process, or OCRP(α, θ) is a K-valued Markov chain which is only capable of making the following transitions from state $(\lambda_1, \ldots, \lambda_l, 0, \ldots) \in K_n$, where $0 \leq \alpha \leq 1$ and $\theta \geq 0$.

- With probability $\frac{\lambda_i - \alpha}{n+\theta}$, a new customer joins the i-th table, leading to a transition into state $(\lambda_1, \ldots, \lambda_{i-1}, \lambda_i + 1, \lambda_{i+1}, \ldots, \lambda_l, 0, \ldots)$, $1 \leq i \leq l$.

- With probability $\frac{\alpha}{n+\theta}$ a new customer opens a new table inserted directly to the left of the i-th table, leading to state $(\lambda_1, \ldots, \lambda_{i-1}, 1, \lambda_i, \lambda_{i+1}, \ldots, \lambda_l, 0, 0, \ldots)$, $1 \leq i \leq l$.

- With probability $\frac{\theta}{n+\theta}$ a new customer opens a new table inserted in the rightmost position, leading to state $(\lambda_1, \ldots, \lambda_l, 1, 0, 0, \ldots)$.

We also consider reverse transitions as follows: each existing customer leaves with equal probability, choosing the i-th non-empty table with probability $\frac{\lambda_i}{n}$ for $1 \leq i \leq l$, and transitioning leading to either $(\lambda_1, \ldots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \ldots, \lambda_l, 0, \ldots)$ if $\lambda_i \geq 2$ or $(\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_l, 0, \ldots)$ if $\lambda_i = 1$.

The first three bullet points transition us from K_n into K_{n+1} and we call any one of them an *up-step*, while the reverse transitions from K_n into K_{n-1} are called a *down-step*; the *down-up chain* is a Markov chain on K_n that consists of a downward followed by an upward step, similarly the *up-down chain* consists of an upward followed by a downward step.

If P is the transition matrix of a Markov chain X on a finite set I with stationary distribution π, then define the *separation distance* $s(r, x)$ of X started from $x \in I$ and *maximal separation distance* $s^*(r)$ by

$$s(r, x) := \max_{y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right),$$
$$s^*(r) := \max_{x \in I} s(r, x) = \max_{x,y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right).$$

These definitions are a measure of how far from stationarity our Markov chain is, with the separation distance allowing the chain to start from a general state $x \in I$ and the maximal separation distance considering any starting location. Note that in the case of a transitive Markov chain, it is the case that $s(r,x) = s^*(r)$ for any $x \in I$, $r \in \mathbb{N}$, as noted in [3, p.79]. Note that from Corollaries 2.2.6 and 2.2.7, we are able to use the fact that an up-down ordered (α,θ)-CRP chain has a stationary distribution along with the consistency condition to conclude that the maximal separation distance is well-defined. We recall from Corollary 2.2.7 that the stationary distribution π has the form

$$\pi(n_1,\ldots,n_k) = \left(\prod_{i=1}^{k-1}\binom{N_i-1}{n_i-1}\Big(\frac{N_{i+1}\alpha+n_i\theta}{n_i}\Big)\right)\frac{\prod_{i=1}^{k}(1-\alpha)_{(n_i-1)\uparrow}}{(\theta+1)_{(n-1)\uparrow}}, \text{ if } \alpha \neq 1,$$

$$\pi(n_1,\ldots,n_k) = \prod_{i=1}^{k}\delta_{n_i 1} \text{ if } \alpha = 1, \text{ where } \delta_{ij} \text{ is the Kronecker delta,}$$

$(n_1,\ldots,n_k) \in \mathcal{C}_n$ is a composition of n, and $N_i := \sum_{j=i}^{k} n_j$ for $i \in [k]$.

We can study the transitions of these Markov chains as follows via a pair of linear maps. For any set S, $\mathbb{C}S$ denotes the complex vector space with basis S. Consider two linear maps $D_n := \mathbb{C}K_n \to \mathbb{C}K_{n-1}$ and $U_n := \mathbb{C}K_n \to \mathbb{C}K_{n+1}$ on $\mathbb{C}K_n$ that are uniquely specified by how they act on the basis elements $\lambda \in K_n$ and capture the transitions of the up-down and down-up (α,θ)-CRP chains in the following way.

Let $p_D(\lambda,\rho)$ and $p_U(\lambda,\tau)$ be the respective probabilities of transitioning from λ to ρ or τ in upward or downward steps respectively, where we assume that $\rho \in K_{n-1}, \lambda \in K_n$ and $\tau \in K_{n+1}$ and define

$$D_n(\lambda) := \sum_{\rho \in K_{n-1}} p_D(\lambda,\rho)\rho \quad \text{and} \quad U_n(\lambda) := \sum_{\tau \in K_{n+1}} p_U(\lambda,\tau)\tau.$$

These maps are such that the coefficients are suitable downward and upward transition probabilities when they act on basis elements, and moreover the coefficients of basis vectors $\mu \in K_n$ in $U_{n-1}D_n(\lambda)$ or $D_{n+1}U_n(\lambda)$ are the probabilities of

transitioning from λ to μ after a down-up or up-down step respectively, where we note that the quantities $U_{n-1}D_n(\lambda)$ and $D_{n+1}U_n(\lambda)$ are a linear combination of the elements of the basis vectors $\mu \in K_n$.

Our first main result confirms that the maps satisfy the following commutation relation, so that Fulman's [17] theory applies and is adaptable. We will summarise the relevant portions of his work as required, so as to avoid significant confusion to a reader unfamiliar with it. In the case of the unordered Kingman Lattice, the result below is [17, Proposition 7.1].

Theorem 3.1.1. *Suppose that $\theta > 0$, $0 \leq \alpha < 1$. For $I_n : \mathbb{C}K_n \to \mathbb{C}K_n$ the identity map, the functions U_n and D_n satisfy the relation*

$$D_{n+1}U_n = a_n U_{n-1}D_n + b_n I_n \qquad (+)$$

with $a_n = \frac{n(\theta+n-1)}{(n+1)(\theta+n)}$ and $b_n = 1 - a_n$, where K_n is the ordered Kingman Lattice.

Our second main result is the following statement about the maximal separation distance. If we consider the unordered Kingman lattice when $\theta = 1$, this is [17, Theorem 7.5]. We generalise this his result to the case when $\theta > 0$, and the lattice may be either one of the ordered and unordered Kingman lattices.

Theorem 3.1.2. *If $s^*(r)$ is the maximal separation distance after $r \geq 1$ iterations of the down-up chain on the unordered Kingman lattice or ordered Kingman lattice and $0 \leq \alpha < 1$, then if $\theta > 0$*

$$s^*(r) = \sum_{i=2}^{n} \left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^r \frac{(-1)^i(i-1)(\theta+2i-1)}{\theta+i-1} \binom{n}{i} \frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)}.$$

Moreover for any fixed $c > 0$,

$$\lim_{n\to\infty} s^*(cn(\theta+n-1)) = \sum_{i=2}^{\infty} e^{-ci(\theta+i-1)} \frac{(-1)^i(i-1)(\theta+2i-1)}{i!(\theta+i-1)} \prod_{j=1}^{i}(\theta+j).$$

This paper has the following structure. In Section 3.2 we will recall relevant

theory from [17], and in 3.3 we will show how it is adapted to the ordered case, proving Theorems 3.1.1 and 3.1.2 in particular.

3.2 Main results from Fulman

3.2.1 Markov chains on branching diagrams

We summarise the main results and notation from Fulman's paper [19] that are of use to us later on.

Let I be a finite set, and P the transition matrix of a Markov chain X on I. Assume that the Markov chain is reversible, so that there is a distribution π on I such that $\forall x, y \in I$, $\pi(x)K(x,y) = \pi(y)K(y,x)$, which is therefore a stationary distribution. Consider the inner product

$$\langle f, g \rangle := \sum_{x \in I} f(x)g(x)\pi(x),$$

for $f, g \in L^2(\pi)$, defined via $L^2(\pi) := \{f : I \to \mathbb{R}, \text{ such that } \langle f, f \rangle \text{ exists in } \mathbb{R}^+\}$, the set of functions from I to $\mathbb{R}$ as I is finite. Now $Pf(x) := \sum_{y \in I} K(x,y)f(y)$ is self-adjoint and the eigenvalues β_i satisfy $-1 \leq \beta_{|I|-1} \leq \cdots \leq \beta_1 \leq \beta_0 = 1$. The chain is called *ergodic* if $|\beta_1| < 1$ and $|\beta_{|I|-1}| < 1$.

When P and Q are probability measures on I, define the *Total Variation Distance* by

$$||P - Q|| := \frac{1}{2} \sum_{x \in I} |P_x - Q_x| = \max_{A \subseteq I} |P_A - Q_A|.$$

When P_x^r is viewed as the probability measure for the distribution of the Markov chain started from $x \in I$ after r steps then the following result from [9, Equation 2.2] formulated as in [19, Lemma 2.1] holds.

Lemma 3.2.1. *$4||P_x^r - \pi||^2 \leq \sum_{i=1}^{|I|-1} \beta_i^{2r} |f_i(x)|^2$, for f_i the orthonormal eigenvectors with $(Pf_i)(x) = \beta_i f_i(x)$ $\forall x \in I$, where the β_i and $f_i(x)$ are all real.*

This result is very powerful, but is often hard to use since the eigenvalues are

generally unknown. Bounds using $\max(|\beta_1|, |\beta_{|I|-1}|)$ are typically used despite yielding weaker results.

Define the *separation distance* $s(r, x)$ of X started from $x \in I$ and *maximal separation distance* $s^*(r)$ by

$$s(r, x) := \max_{y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right),$$
$$s^*(r) := \max_{x \in I} s(r, x) = \max_{x,y \in I} \left(1 - \frac{P^r_{x,y}}{\pi(y)}\right).$$

Indeed, the following hold:

- $\max_{x \in I} ||P^r_x - \pi|| \leq s^*(r)$,
- $\forall r_1 \geq r_2,\ s^*(r_1) \leq s^*(r_2)$,
- $\forall r_1 \geq r_2,\ s^*(r_1 + r_2) \leq s^*(r_1)s^*(r_2)$.

To see why these results hold, note that by the definition of the total variation distance with P_x and Q_x the atoms of P and Q on I,

$$||P - Q|| = \sum_{x \in I : Q_x \geq P_x} (Q_x - P_x) \leq \sum_{x \in I : Q_x \geq P_x} Q_x \max_{y \in I} \left(1 - \frac{P_y}{Q_y}\right) \leq \max_{y \in I} \left(1 - \frac{P_y}{Q_y}\right).$$

Replacing P with the r-step transition measure started from x, Q with the stationary distribution π and taking the maximum over all $x \in I$, the first bullet point follows.

The third bullet point is just [3, Lemma 3.7], and since $s^*(r) \leq 1$ by construction, the second bullet point follows from the third upon noting that whenever $r_1 \geq r_2$,

$$s^*(r_1) \leq s^*(r_1 - r_2)s^*(r_2) \leq s^*(r_2).$$

We also note that both the total variation distance $||P - Q||$ and the quantity $s(P, Q) := \max_{y \in I} \left(1 - \frac{P_y}{Q_y}\right)$ can be viewed as a way of measuring the distance between two probability measures, although the quantity $s(P, Q)$ is not a metric, unlike the total variation distance $||P - Q||$.

For any $\varepsilon > 0$, let n^*_ε be the smallest integer with $s^*(n^*_\varepsilon) \leq \varepsilon$, which is called a *mixing time*. It is a way of measuring the time taken for the distance between an r-step transition matrix and stationary distribution to be small enough that the former is a good approximation to the latter. [1] shows that the mixing time $n^*_{1/4}$ of a certain kind of simple up-down Markov chain on uniform trees with n leaves and $n-2$ internal branchpoint vertices is of order n^3 and conjectured there to be of order n^2, which connects to Conjecture 1.2.5 of the introduction, and provides informal evidence that the time scaling chosen is correct.

Let T be a random variable such that for any $r \geq 0$, $s^*(r) = \mathbb{P}(T > r)$, which exists by [3, Proposition 3.2]. The following lemma is essentially [34, Cor 2.2.9].

Lemma 3.2.2. *For T an $\mathbb{N}_0$-valued random variable as above, $\frac{1}{2}\mathbb{E}(T) \leq n^*_{1/2} \leq 2\mathbb{E}(T)$.*

The following two results are due to Fulman.

Proposition 3.2.3 ([18], Proposition 5.1)**.** *Let P be the transition matrix of a reversible ergodic Markov chain on a set I. Let $1, \beta_1, \ldots, \beta_d$ be the eigenvalues of P, listed without repetition in the case that they are not distinct, and choose $x, y \in I$ such that it is possible to transition from x to y in d steps but not fewer, so that necessarily $|I| \geq d+1$. Then for any $r \geq 0$,*

$$1 - \frac{P^r_{x,y}}{\pi(y)} = \sum_{i=1}^{d} \beta_i^r \left(\prod_{j \neq i} \frac{1-\beta_j}{\beta_i - \beta_j} \right).$$

We use the convention that a $Geo(p)$ random variable has probability mass function $(p_n)_{n \geq 1}$ given by then $p_n = p(1-p)^{n-1}$ for $n \geq 1$ and $p \in (0, 1]$, and correspondingly write $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Proposition 3.2.4 ([17], Proposition 2.4)**.** *Suppose that $T = \sum_{i=1}^{d} Y_i$, where the Y_i are independent $Geo(1-\beta_i)$ with $\beta_i \in [0, 1)$ $\forall i \in \{1, 2, \ldots, d\}$. If the β_i are distinct then for any $r \in \mathbb{N}$,*

$$\mathbb{P}(T > r) = \sum_{i=1}^{d} \beta_i^r \Big(\prod_{j \neq i} \frac{1-\beta_j}{\beta_i - \beta_j} \Big).$$

Remark 3.2.5. *As is noted in remarks made by [17], we cannot in general assume that $s^*(r) = \mathbb{P}(T > r)$ for any reversible Markov chain. It is however the case that the result holds for Markov chains within Fulman's framework, provided that the commutation relations of Theorem 3.1.1 hold.*

In order to study the down-up chain from Section 3.1, we will need to use some standard definitions. The formulations used are as in [17].

Definition 3.2.6. A *partially ordered set* S, or *poset* is a set with a relation $\leq$ such that

1) $\forall t \in S$, $t \leq t$,

2) if $s \leq t$ and $t \leq s$ then $s = t$,

3) if $s \leq t$ and $t \leq u$ then $s \leq u$.

An *oriented path* is a set $C \subseteq S$ with the property that $s \leq t$ if and only if $s \leq t$ in S and $\forall s, t \in C$, $s \leq t$ or $t \leq s$. An oriented path is *maximal* if it is not contained in any other oriented path. We say that the poset S is *graded* if the length of every maximal oriented path from x to y depends only on x and y, and $[s,t] := \{u \in S : s \leq u \leq t\}$ is finite for any $s, t \in S$ then S is said to be *locally finite*.

Remark 3.2.7. *Note that what we refer to as an oriented path would more typically be called a chain as is the case in [19]. We choose to use alternative terminology to avoid there being any ambiguity caused by the term* Markov chain.

Suppose that S is a locally finite poset, and that there exists $\hat{0} \in S$ with the property that for every $x \in S$, $x \geq \hat{0}$, where $\hat{0}$ is sometimes referred to as the *smallest element* of the poset. Let $\mathbb{C}S$ denote the complex vector space with basis S, and for

$x \in S$, let the *rank* be the maximal l such that there are $x_0, x_1, \dots x_{l-1} \in S$ with $x_0 < x_1 < \cdots < x_{l-1} < x_l = x$. $\mathbb{C}S_n$ is the subspace of $\mathbb{C}S$ spanned by the rank n elements.

We write $x \nearrow y$ if y *covers* x *in* S, that is if $x < y$ and there is no $u \in S$ with $x < u < y$, or equivalently if $[x, y] = \{x, y\}$.

Definition 3.2.8. A *branching diagram* is an oriented graph $\Gamma = \bigcup_{n \geq 0} \Gamma_n$ with edge set $E(\Gamma)$ and vertex set $V(\Gamma)$ such that

i) $\Gamma_0 = \{\varnothing\}$,

ii) If $xy \in E(\Gamma)$ and $x \in V(\Gamma_n)$ then $y \in V(\Gamma_{n+1})$,

iii) $\forall x \in V(\Gamma), \exists y \in V(\Gamma)$ such that $xy \in E(\Gamma)$,

iv) Γ_n is finite for any $n \geq 0$.

Note that a branching diagram Γ gives rise to a graded poset on $V(\Gamma)$, with $x \leq y$ if there is a sequence of vertices leading from x to y which increase in distance from $\varnothing$ at each step. More formally, if $x, y \in V(\Gamma)$ are such that $x \in V(\Gamma_n)$ and $y \in V(\Gamma_{n+m})$ with $m, n \in \mathbb{N}_0$, we say that $x \leq y$ if there is a sequence $x = x_0, x_1, \dots, x_m = y$ with $x_l \in V(\Gamma_{n+l})$ and $x_{l-1}x_l \in E(\Gamma)$.

For any $\lambda, \Lambda \in V(\Gamma), \lambda \nearrow \Lambda$ means that there is an edge in Γ from λ to Λ. Suppose we are given a *multiplicity function* $\kappa(\lambda, \Lambda)$ on the set of (λ, Λ) with $\lambda \nearrow \Lambda$, which is a positive integer valued function that can be thought of as the number of directed edges from λ to Λ if Γ was a multigraph. The weight of a path is the product of the multiplicities of the edges, viewing an edge $\lambda\Lambda \in E(\Gamma)$ as corresponding to a pair (λ, Λ) of vertices. Define the *dimension* d_Λ of a vertex $\Lambda \neq \varnothing$ to be the sum of the weights over all minimal length paths from $\varnothing$ to Λ and let $d_\varnothing := 1$.

A sequence $(M_n)_{n \in \mathbb{N}_0}$ of probability distributions, with M_n distributed on $V(\Gamma_n)$ is called *coherent* if

$$M_n(\lambda) = \sum_{\Lambda : \lambda \nearrow \Lambda} \frac{d_\lambda \kappa(\lambda, \Lambda)}{d_\Lambda} M_{n+1}(\Lambda).$$

The *up-chain* is a Markov chain on Γ which transitions from $\lambda \in \Gamma_n$ to $\tau \in \Gamma_{n+1}$, $n \geq 0$ with probability

$$\frac{d_\lambda M_{n+1}(\tau)\kappa(\lambda,\tau)}{d_\tau M_n(\lambda)}.$$

The *down-chain* is a Markov chain on Γ which transitions from $\lambda \in \Gamma_n$ to $\rho \in \Gamma_{n-1}$, $n \geq 1$ with probability

$$\frac{d_\rho \kappa(\rho,\lambda)}{d_\lambda},$$

and is absorbed at $\varnothing \in \Gamma_0$.

The *up-down chain* is a Markov chain on Γ_n, $n \geq 0$, that first moves up one state and then down one state, with explicit transition probability from $\lambda \in \Gamma_n$ to $\mu \in \Gamma_n$ of

$$\frac{d_\lambda d_\mu}{M_n(\lambda)} \sum_{\tau \in \Gamma_{n+1}} \frac{M_{n+1}(\tau)\kappa(\lambda,\tau)\kappa(\mu,\tau)}{d_\tau^2}.$$

The *down-up chain* is a Markov chain on Γ_n, $n \geq 1$, that first moves down one state and then up one state, with explicit transition probability from $\lambda \in \Gamma_n$ to $\mu \in \Gamma_n$, $n \geq 1$ of

$$\frac{M_n(\mu)}{d_\lambda d_\mu} \sum_{\rho \in \Gamma_{n-1}} \frac{d_\rho^2 \kappa(\rho,\lambda)\kappa(\rho,\mu)}{M_{n-1}(\rho)},$$

and is absorbed when considered as a chain only on Γ_0.

It is an elementary argument to verify that the up-down chain has M_n as its stationary distribution. Indeed, it is reversible with respect to M_n, and the same is true of the down-up chain.

We will consider the *Unordered Kingman lattice*. This is a branching diagram where Γ_n consists of the partitions of n, which are sequences $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l, 0, 0, \ldots)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_l > 0$, $\lambda_i \in \mathbb{N}$ for every $1 \leq i \leq l$ and $\sum_{i=1}^{l} \lambda_i = n$. We identify these sequences with collections of boxes in the same way as in a standard reference such as [33], a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l, 0, 0, \ldots)$ is identified with a collection of boxes, where the i-th row has λ_i boxes. See Figure 3.1 for an illustration

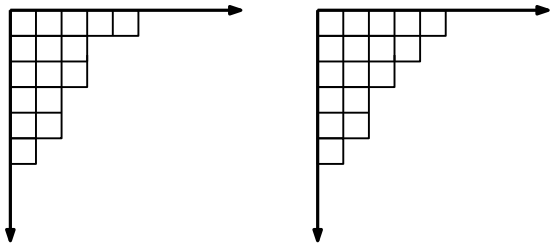


Fig. 3.1: The unordered partitions $\lambda = (5, 3, 3, 2, 2, 1, 0, \ldots)$ and $\Lambda = (5, 4, 3, 2, 2, 1, 0, \ldots)$ represented as collections of boxes. Observe that $\kappa(\lambda, \Lambda) = 2$ here.

of this, and Figure 3.2 for a visualisation of the branching diagram up to level 4.

That is, for any $\lambda \in \Gamma_n$ and $\Lambda \in \Gamma_{n+1}$, we have $\lambda \nearrow \Lambda$ if Λ can be obtained from λ by adding on a box to the end of one of the rows or directly below all boxes in the first column.

We could consider the case when $\kappa(\lambda, \Lambda) = 1$ on all edges, but we choose not to study this case otherwise known as the *Young lattice*. Instead, suppose that it is possible to obtain λ from Λ by removing a box from a row of length $j \geq 1$ and let $\kappa(\lambda, \Lambda)$ be the number of these length j rows in Λ; the structure here is known as the *Kingman lattice*, and is referred to as unordered for the remainder of this document to avoid ambiguity. In this case, $d_\lambda = \frac{n!}{\lambda_1!\lambda_2!\cdots\lambda_l!}$.

The *Pitman distributions* are a sequence $(M_n)_{n\geq 0}$ of coherent probability distributions on Γ_n, where for $\theta > 0$, $0 \leq \alpha < 1$,

$$M_n(\lambda) = \frac{\theta(\theta+\alpha)\cdots(\theta+(l(\lambda)-1)\alpha)}{\theta(\theta+1)\cdots(\theta+n-1)} \frac{n!}{\prod_{k=1}^{\infty} m_k(\lambda)! \prod_{i=1}^{l(\lambda)} \lambda_i!} \prod_{i:\, \lambda_i \geq 2} (\lambda_j - 1 - \alpha),$$

with $m_k(\lambda)$ the number of size k parts of λ, that is $m_k(\lambda) := |\{i : \lambda_i = k\}|$ and $l(\lambda)$ denotes the number of non-zero blocks in $\lambda \in K_n$, i.e $l(\lambda) := \max\{i : \lambda_i \neq 0\}$, with $l(0, 0, 0, \ldots) := 0$. As noted in [18, Chapter 3, Example 2], the Pitman distributions correspond to the κ given above for the unordered Kingman lattice. We show briefly here that the Pitman distributions are coherent. Observe that coherency is equivalent to the statement that

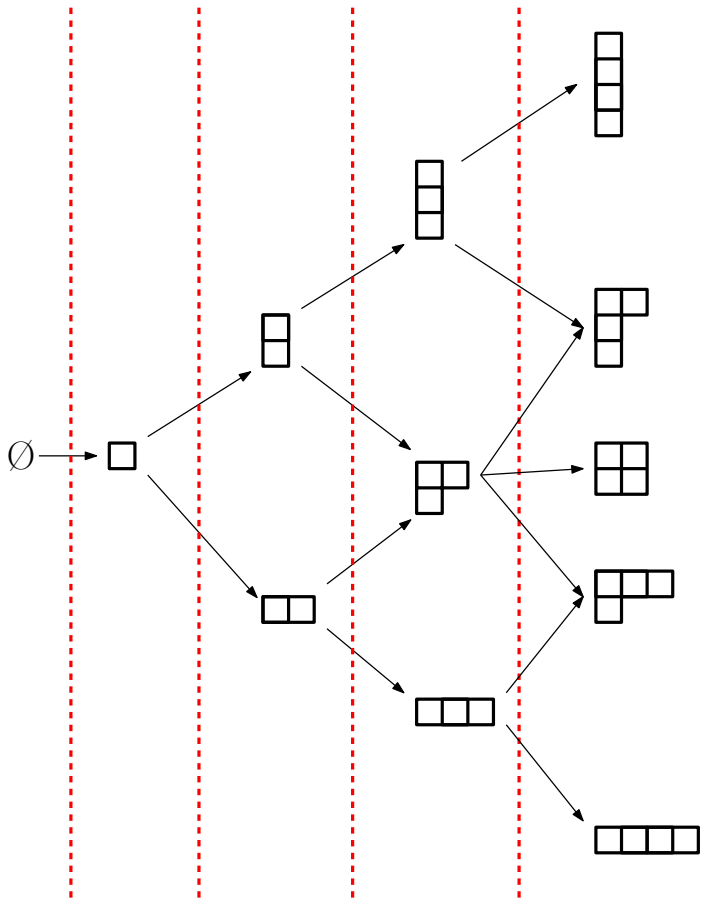


Fig. 3.2: The first few levels of unordered partitions with possible upward transitions according to the rules of an ordered CRP with $(\alpha, \theta) \neq (0, 0)$.

$$\frac{M_n(\lambda)}{d_\lambda} = \sum_{\Lambda:\lambda\nearrow\Lambda} \kappa(\lambda,\Lambda)\frac{M_{n+1}(\lambda)}{d_\Lambda} \qquad (\#)$$

If λ is obtained from $\Lambda \in K_{n+1}$ by removing a box from a row of length j then we have

$$\frac{\kappa(\lambda,\Lambda)}{\prod_{k=1}^{\infty} m_k(\Lambda)!} = \frac{m_j(\Lambda)}{\prod_{k=1}^{\infty} m_k(\Lambda)!} = \frac{m_{j-1}(\lambda)}{\prod_{k=1}^{\infty} m_k(\lambda)!}$$

with the convention that $m_0(\lambda) = 1$. Then, the RHS of $(\#)$ is a sum over $j \geq 1$ where for each j the term in the sum simplifies to the LHS, up to factors summing up to

$$\frac{1}{n+\theta}\left(\theta + l(\lambda)\alpha + \sum_{j\geq 2} m_{j-1}(\lambda)(j-1-\alpha)\right) = 1,$$

observing that $\sum_{k\geq 1} k m_k(\lambda) = n$ and $\sum_{k\geq 1} m_k(\lambda) = l(\lambda)$.

3.2.2 Commutation relations

For a linear map $A : \mathbb{C}S \to \mathbb{C}S$ and $n \geq 0$, the restriction to $\mathbb{C}S_n$ is denoted by A_n, and AB_n means $A(B_n) = (AB)_n$, and also $ABv = (AB)(v) = A(B(v))$; I is the identity map on $\mathbb{C}S$. Let $U_n : \mathbb{C}S_n \to \mathbb{C}S_{n+1}$ for $n \geq 0$ and $D_n : \mathbb{C}S_n \to \mathbb{C}S_{n-1}$ for $n \geq 1$ be linear maps satisfying

$$D_{n+1}U_n = a_n U_{n-1} D_n + b_n I_n \text{ for some } a_n, b_n \in \mathbb{R}, \qquad (*)$$

these formulae are known as the *commutation relations*. Since it is worth ensuring consistency in the notation, we can construct a linear map $U : \mathbb{C}S \to \mathbb{C}S$ by specifying that if $\lambda \in \mathbb{C}S_n$ then $U(\lambda) = U_n(\lambda)$ and linearly extending this to $\mathbb{C}S$, noting that the union of a the sets of basis vectors for $\mathbb{C}S_n$ over all $n \geq 1$ forms a basis for $\mathbb{C}S$. A near identical construction can be used to build a linear map $D : \mathbb{C}S \to \mathbb{C}S$. We list a sequence of relevant results from [18].

Theorem 3.2.9 ([18], Theorem 4.1)**.** *Suppose that the commutation relations* $(*)$ *hold with* $a_n > 0\ \forall n \geq 1$. *Let* $p_j := |\{x \in S : rank(x) = j\}|$. *Then* UD_n *has*

eigenvalues

$$\begin{cases} 0 \text{ of multiplicity } p_n - p_{n-1}, \\ \sum_{j=i}^{n-1} b_j \prod_{k=j+1}^{n-1} a_k \text{ of multiplicity } p_i - p_{i-1},\ 0 \le i \le n-1. \end{cases}$$

If $b_i = 1 - a_i$ for all $i \in \mathbb{N}_0$, then the eigenvalues simplify to

$$\begin{cases} 0 \text{ of multiplicity } p_n - p_{n-1}, \\ 1 - \prod_{k=i}^{n-1} a_k \text{ of multiplicity } p_i - p_{i-1},\ 0 \le i \le n-1. \end{cases}$$

Lemma 3.2.10 ([18], Lemma 4.2)**.** *If (*) holds with $b_0 = 1$ and $a_n, b_n > 0\ \forall n \ge 1$, then the U_n are injective and the D_{n+1} are surjective.*

Theorem 3.2.11 ([18], Theorem 4.3)**.** *If (*) holds with $b_0 = 1$ and $a_n, b_n > 0 \forall n \ge 1$ and $E_n(\lambda)$ is the λ-eigenspace of UD_n, then*

$$E_n(0) = Ker(D_n) \quad \text{and} \quad E_n\Big(\sum_{j=i}^{n-1} b_j \prod_{k=j+1}^{n-1} a_k\Big) = U^{n-i} E_i(0).$$

Proposition 3.2.12 ([18], Proposition 4.5)**.** *Suppose that (*) holds, then*

$$(UD)_n^r = \sum_{k=0}^{n} A_n(r,k)(U^k D^k)_n,$$

where the $A_n(r,k)$ are determined by the recurrence

$$A_n(r,k) = A_n(r-1,k-1) \prod_{j=n-k+1}^{n-1} a_j + A_n(r-1,k) \sum_{j=n-k}^{n-1} b_j \prod_{l=j+1}^{n-1} a_l$$

with initial conditions $A_n(0,0) = 1$ and $A_n(0,m) = 0$ for $m \neq 0$.

If $0 \le a_i \le 1$ and $b_i = 1 - a_i$ for all i, the recurrence becomes

$$A_n(r,k) = A_n(r-1,k-1) \prod_{j=n-k+1}^{n-1} a_j + A_n(r-1,k)\Big(1 - \prod_{j=n-k}^{n-1} a_j\Big)$$

and all $A_n(r,k)$ are then non-negative.

3.2.3 Results on the unordered down-up Chinese restaurant

Observe that the down chain removes a box from a row of length j with probability $\frac{jm_j(\lambda)}{n}$, and the up chain adds a box onto a row of length $k \geq 1$ with probability $\frac{(k-\alpha)m_k(\lambda)}{\theta+n}$ or a length 0 row with probability $\frac{\theta+\alpha l(\lambda)}{\theta+n}$.

Let S be the poset of partitions with the same partial order as the unordered Kingman's lattice (that is $\lambda \leq \Lambda$ if and only if $\lambda = \Lambda$ or $\exists\, \lambda_1, \ldots, \lambda_n \in P$ with $\lambda_1 = \lambda, \lambda_n = \Lambda,\ \lambda_i \nearrow \lambda_{i+1}$ if $1 \leq i \leq n-1$). We will define operators $D, U : \mathbb{C}S \to \mathbb{C}S$ as follows.

The coefficient of ρ in $D_n(\lambda)$ is defined as the probability of transitioning to ρ from λ, and the coefficient of τ in $U_n(\lambda)$ is defined as the probability of transitioning to τ from λ. The down-up chain on the unordered Kingman lattice arising from Pitman's distribution corresponds to the linear map $U_{n-1}D_n$, where the probability of transitioning from λ to μ is the coefficient of μ in $U_{n-1}D_n(\lambda)$.

Explicitly, we let

$$p_D(\lambda,\rho) := \frac{jm_j(\lambda)}{n} \text{ if } \rho \text{ can be obtained from } \lambda \text{ by removing a box from a row of length } j,$$

$$p_U(\lambda,\tau) = \begin{cases} \frac{(k-\alpha)m_k(\lambda)}{\theta+n} \text{ if } \tau \text{ is obtained from } \lambda \text{ by adding a box to a length } k \text{ row,} \\ \frac{\theta+l(\lambda)\alpha}{n+\theta} \text{ if } \tau \text{ is obtained from } \lambda \text{ by adding a box to the last row.} \end{cases}$$

Then $D_n(\lambda) := \sum_{\rho \in K_{n-1}} p_D(\lambda,\rho)\rho$ and $U_n(\lambda) := \sum_{\tau \in K_{n+1}} p_U(\lambda,\tau)\tau$.

The remaining collection of results in this subsection are proved in [17], and so the proofs are not given here.

Proposition 3.2.13 ([17], Proposition 7.1)**.** *Consider the down-up chain on the unordered Kingman lattice with $\theta > 0$, $0 \leq \alpha < 1$. Letting $a_n = \frac{n(\theta+n-1)}{(n+1)(n+\theta)}$,*

$$D_{n+1}U_n = a_n U_{n-1}D_n + (1-a_n)I_n \quad \textit{for } n \geq 1.$$

Corollary 3.2.14 ([17], Corollary 7.2). *If $p(j)$ denotes the number of integer partitions of j, then the eigenvalues of UD_n are 1 with multiplicity 1 and $1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}$ with multiplicity $p(i) - p(i-1)$ for $2 \leq i \leq n$.*

Proposition 3.2.15 ([17], Proposition 7.3). *Let π be the Pitman distribution with parameters $\theta > 0$, $0 \leq \alpha < 1$ and let $r \geq 0$. Then among all partitions μ, λ of n, $\frac{p_{UD}^{(r)}(\mu,\lambda)}{\pi(\lambda)}$ is minimised by $\mu = (n)$ and $\lambda = (1, 1, \ldots, 1)$ or $\lambda = (n)$ and $\mu = (1, 1, \ldots, 1)$, where $p_{UD}^{(r)}(\mu, \lambda)$ probability r step transition probability of the down-up chain from μ to λ.*

Corollary 3.2.16 ([17], Corollary 7.4). *For the down-up chain with $\theta > 0$ and $0 \leq \alpha < 1$ on the n-th level of the unordered Kingman lattice, $s^*(r) = \mathbb{P}(T > r)$, where T is a sum of independent $Geo\left(\frac{i(\theta+i-1)}{n(\theta+n-1)}\right)$ random variables for $2 \leq i \leq n$.*

Theorem 3.2.17 ([17], Theorem 7.5). *If $\theta = 1$ and $0 \leq \alpha < 1$, then in the unordered case*

$$s^*(r) = 2\sum_{i=2}^{n}(-1)^i(i^2-1)\frac{(n!)^2}{(n-i)!(n+i)!}\left(1 - \frac{i^2}{n^2}\right)^n.$$

Moreover for any fixed $c > 0$,

$$\lim_{n\to\infty} s^*(cn^2) = 2\sum_{i=2}^{\infty}(-1)^i(i^2-1)e^{-ci^2}.$$

Corollary 3.2.18 ([17], Corollary 7.6). *With T as in Corollary 3.2.16 above, we have*

$$\mathbb{E}(T) = \sum_{i=2}^{n}\frac{n^2}{i^2} \sim n^2\left(\frac{\pi^2}{6} - 1\right) \text{ when } \theta = 1. \text{ If } \theta \neq 1, \text{ then}$$

$$\frac{n(\theta+n-1)}{\theta-1}\log\left(\frac{(n+1)(\theta+1)}{2(n+\theta)}\right) \leq \mathbb{E}(T) \leq \frac{n(\theta+n-1)}{\theta-1}\log\left(\frac{n\theta}{n+\theta-1}\right).$$

Proposition 3.2.19 ([17], Proposition 7.7). *Let $s^*_{UD_n}(r)$ be the maximal separation distance after r iterations of the down-up chain on the unordered Kingman lattice, and let $s^*_{DU_n}(r)$ be the corresponding quantity for the up-down chain. Then for any $n \geq 1$, $r \geq 0$,*

$$s^*_{DU_n}(r) = s^*_{UD_{n+1}}(r+1).$$

3.3 New results for the ordered lattice

We now explain how to extend Fulman's work as described above, to include the case of the ordered down-up chain and also prove Theorem 3.1.2, extending Fulman's result to the case when $\theta > 0$. Consider the *ordered Kingman lattice* $K := \{(\lambda_1, \lambda_2, \ldots, \lambda_l, 0, 0, \ldots) : \lambda_i \in \mathbb{N}, 1 \leq i \leq l\}$ and introduce the subspace $K_n := \{\lambda \in K : \sum_{i=1}^{l(\lambda)} \lambda_i = n\}$ where $l(\lambda) = \max\{i : \lambda_i > 0\}$ is the number of non-zero entries in λ. If there is no ambiguity, we write elements of K without an infinite sequence of zeros. The spaces K_n with $n \geq 0$ have upward transition functions U_n corresponding to insertion according to the rules of an ordered (α, θ)-CRP as introduced by Pitman and Winkel [40], and downward transition functions D_n which correspond to uniform deletion as before. Explicitly, we identify these ordered partitions with diagrams of boxes in much the same way as in Figure 3.1, and move up from $\lambda \in K_n$ by inserting a new box

- onto the i-th row with probability $\frac{\lambda_i - \alpha}{n+\theta}$,
- directly above the i-th row with probability $\frac{\alpha}{n+\theta}$, moving all rows beneath it down,
- directly below the final non-empty row with probability $\frac{\theta}{n+\theta}$.

We move down from K_n by choosing a box uniformly at random and removing the rightmost box in its row, deleting the row it is in if this is necessary to avoid an empty row, and in this case moving any rows below it up. Note that unlike in the previous section, it is no longer necessary that the number of boxes in each row be non-increasing.

An up-down chain related to this process has been studied in Chapter 2.

We will adapt Fulman's [17] results as recalled in the previous section. We begin by proving Theorem 3.1.1 from the introduction.

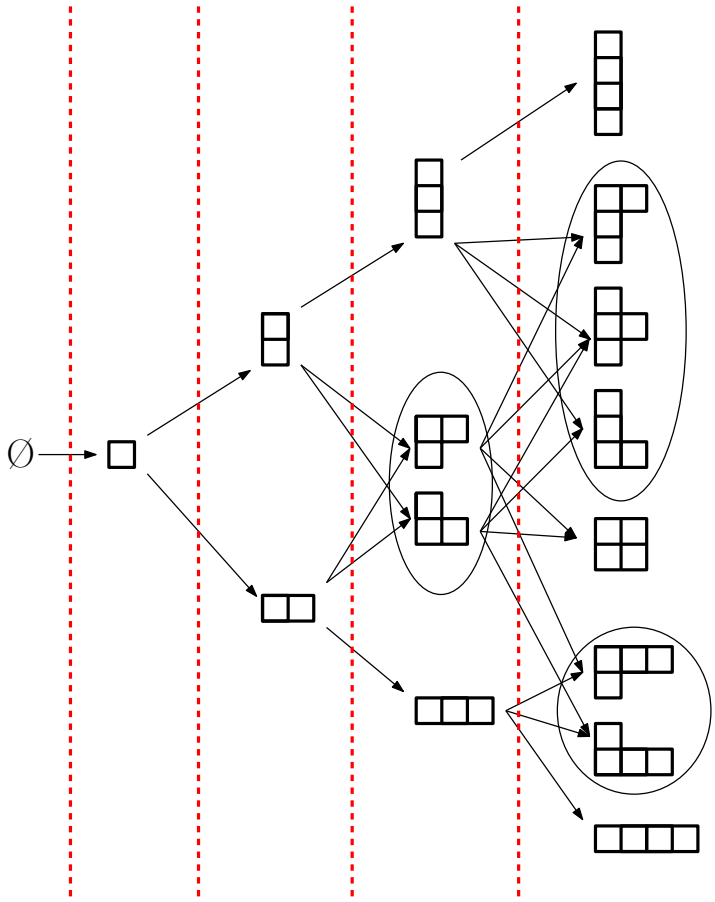


Fig. 3.3: The first few levels of the ordered Kingman lattice, which also show the possible upward transitions according to the rules of an ordered CRP when $(\alpha, \theta) \neq (0, 0)$ and the partitions that would be identical in the unordered case.

Theorem 3.3.1 (Restatement of Theorem 3.1.1)**.** *Suppose that $\theta > 0$, $0 \leq \alpha < 1$. For $I_n : \mathbb{C}K_n \to \mathbb{C}K_n$ the identity map, the functions U_n and D_n satisfy the relation*

$$D_{n+1}U_n = a_n U_{n-1} D_n + b_n I_n \qquad (+)$$

with $a_n = \frac{n(\theta+n-1)}{(n+1)(\theta+n)}$ and $b_n = 1 - a_n$.

Proof of Theorem 3.1.1. We first show that if $D_{n+1}U_n = a_n U_{n-1} D_n + b_n I_n$ for $n \geq 2$, then $a_n = \frac{n(\theta+n-1)}{(n+1)(\theta+n)}$ and $b_n = 1 - a_n$ as before.

Let us write $x_n = (n)$, $y_n = (1, n-1)$ and $z_n = (n-1, 1)$. We have

$$\begin{aligned} DU_n(x_n) &= D_{n+1}\left(\frac{n-\alpha}{n+\theta}x_{n+1} + \frac{\alpha}{n+\theta}y_{n+1} + \frac{\theta}{n+\theta}z_{n+1}\right) \\ &= \frac{n-\alpha}{n+\theta}x_n + \frac{1}{n+\theta}\left[\frac{\alpha}{n+1}x_n + \frac{\alpha n}{n+1}y_n + \frac{\theta}{n+1}x_n + \frac{\theta n}{n+1}z_n\right] \\ &= \frac{1}{n+\theta}\left[\left(n - \alpha + \frac{\alpha+\theta}{n+1}\right)x_n + \frac{\alpha n}{n+1}y_n + \frac{\theta n}{n+1}z_n\right]. \end{aligned}$$

Similarly,

$$UD_n(x_n) = U_{n-1}(x_{n-1}) = \frac{n-1-\alpha}{n-1+\theta}x_n + \frac{\alpha}{n-1+\theta}y_n + \frac{\theta}{n-1+\theta}z_n.$$

Substituting these into equation $(+)$, we have

$$\frac{1}{(n+\theta)(n+1)}[(n^2+n-n\alpha+\theta)x_n+\alpha n y_n+\theta n z_n] = \frac{a_n}{n-1+\theta}[(n-1-\alpha)x_n+\alpha y_n+\theta z_n]+b_n x_n.$$

As $\theta > 0$, we see by comparing the coefficients of z_n that $\frac{\theta n}{(n+\theta)(n+1)} = \frac{\theta a_n}{n-1+\theta}$, so that $a_n = \frac{n(n-1+\theta)}{(n+\theta)(n+1)}$ as before. Comparison of the coefficients of x_n yields $b_n = \frac{\theta+2n}{(n+\theta)(n+1)}$, and so $a_n + b_n = 1$ as before.

Following the proof of [17] let $\lambda \in K_n$. We would like to identify the coefficient of any $\mu \in K_n$ in $(n+1)(\theta+n)DU_n(\lambda) - n(\theta+n-1)UD_n(\lambda)$ and show that for any $\mu \neq \lambda$ it is 0, which is equivalent to showing that $(+)$ holds. While [17] also treats

the case $\mu = \lambda$, this is not strictly necessary as both the up-down and down-up chains are Markovian and it suffices to check the entries of the transition matrix off the diagonal.

Observe that the possible μ that can be reached in one up-down transition are precisely those that can be reached in one down-up transition; and clearly $(+)$ holds for the $\mu \in K_n$ not of this form.

Recall Figure 3.3, and use it as an illustration of how to represent any $\lambda \in K_n$ via a collection of boxes. With such a representation, we refer to the imagined space between the $(i-1)$-st and i-th rows in λ from top to bottom as the i-th gap, understanding that the empty regions above the 1-st row and below the $l(\lambda)$-th row are considered the 1-st and $(1+l(\lambda))$-th gaps here. $G_\lambda(i)$ counts the number of rows consisting of single boxes around the i-th gap, using the convention that $G_\lambda(i) = 0$ if $i < 1$ or $i > 1 + l(\lambda)$.

We split the μ reachable from λ in one down-up step according to the λ_i which was reduced by one in the down step.

Case 1: $\lambda_i > 2$

We have the following subcases arising from the possible up steps:

i) Add onto row j with $j \neq i$,

ii) Insert directly above row j,

iii) Insert directly below all rows.

Observing that $G_\lambda(i) = G_{(\lambda_1,\ldots,\lambda_{i-1},\lambda_i-1,\lambda_{i+1}\ldots,\lambda_l)}(i)$ in this case, the expressions for $p_{UD}(\lambda,\mu)$ in each subcase are $\frac{\lambda_i}{n} \cdot \frac{\lambda_j-\alpha}{n-1+\theta}$, $\frac{\lambda_i}{n} \cdot \frac{\alpha(G_\lambda(j)+1)}{n-1+\theta}$ and $\frac{\lambda_i}{n} \cdot \frac{\alpha G_\lambda(1+l(\lambda))+\theta}{n-1+\theta}$, and for $p_{DU}(\lambda,\mu)$ they are $\frac{\lambda_j-\alpha}{n+\theta} \cdot \frac{\lambda_i}{n+1}$, $\frac{\alpha(G_\lambda(j)+1)}{n+\theta} \cdot \frac{\lambda_i}{n+1}$ and $\frac{\alpha G_\lambda(1+l(\lambda))+\theta}{n+\theta} \cdot \frac{\lambda_i}{n+1}$ and so clearly $(+)$ holds.

Case 2: $\lambda_i = 2$

We consider the following three subcases:

i) Add onto row j with $j \neq i$,

ii) Insert directly above or below a row not in the same sequence of single boxes as in the ith and $(i+1)$-st gaps,

iii) Insert directly above or below a row in the same sequence of single boxes as in the i-th and $(i+1)$-st gaps.

It is fairly clear that subcases i) and ii) work as in case 1. For subcase iii), the deletion from $\lambda_i = 2$ means that by the convention of $G_\lambda(i) = 0$ if $i < 1$ or $i > 1 + l(\lambda)$, $G_{(\lambda_1,\ldots,\lambda_{i-1},\lambda_i-1,\lambda_{i+1}\ldots,\lambda_l)}(i) = G_\lambda(i) + G_\lambda(i+1) + 2$ and so we then have that $p_{UD}(\lambda,\mu)$ is $\frac{\lambda_i}{n} \cdot \frac{\alpha(G_\lambda(i)+1+G_\lambda(i+1)+1)}{n-1+\theta}$ or $\frac{\lambda_i}{n} \cdot \frac{\alpha(G_\lambda(i)+1+G_\lambda(i+1))+\theta}{n-1+\theta}$ depending on if $i = l(\lambda)+1$ or not, and $p_{DU}(\lambda,\mu)$ is $\frac{\alpha(G_\lambda(i)+1)+\alpha(G_\lambda(i+1)+1)}{n+\theta} \cdot \frac{\lambda_i}{n+1}$ or $\frac{\alpha(G_\lambda(i)+1)+\alpha G_\lambda(i+1)+\theta}{n+\theta} \cdot \frac{\lambda_i}{n+1}$. Again $(+)$ holds.

Case 3: $\lambda_i = 1$

Here, we note that we do not have to consider the case of inserting directly above or below a row in the same sequence of single boxes as the i-th and $(i+1)$-st gaps of $(\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1}\ldots,\lambda_l)$, as this would result in us reaching a state $\mu = \lambda$, which we excluded from consideration. The two subcases are as follows.

i) Insert directly above, below or onto a row not in the same sequence of single box rows as the i-th and $(i+1)$-st gaps of $(\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1}\ldots,\lambda_l)$.

ii) Insert onto one of rows in the same sequence of single box rows as the i-th row.

Subcase i) is more or less entirely as in cases 1 and 2 before, although we replace all occurrences of $\lambda_i = 1$ in the explicit expressions for $UD_n(\lambda)$ and $DU_n(\lambda)$ with $G_\lambda(i)$ since there are $G_\lambda(i)$ possible places where insertion leads to the same final partition μ.

In subcase ii), we have $p_{UD}(\lambda,\mu) = \frac{G_\lambda(i)}{n} \cdot \frac{1-\alpha}{n-1+\theta}$ and have $p_{DU}(\lambda,\mu) = \frac{1-\alpha}{n+\theta} \cdot \frac{G_\mu(i)+1}{n+1} + \frac{1-\alpha}{n+\theta} \cdot \frac{G_\mu(i+1)+1}{n+1}$ and $G_\lambda(i) = G_\mu(i) + G_\mu(i+1) + 2$, so $(+)$ still holds. $\square$

We note that by and large, nothing else from here on is really that different to [17]. All that changes in Proposition 3.2.14 is the multiplicities of the eigenvalues; the proof is the same as Fulman's after noting that $p_0 = 1$ and $p_n = 2^{n-1}$ for $n \geq 1$ here.

Proposition 3.3.2. *The eigenvalues of UD_n are 1 and $1 - \frac{i(i+\theta-1)}{n(n+\theta-1)}$, for $2 \leq i \leq n$. Moreover, these have multiplicities 1 and 2^{i-2} respectively.*

Proposition 3.3.3. *Let $\theta > 0$, $\leq \alpha < 1$ and let $r \geq 0$. Then among all partitions μ, λ of n, $\frac{p_{UD}^{(r)}(\mu,\lambda)}{\pi(\lambda)}$ is minimised by $\mu = (n)$ and $\lambda = (1, 1, \ldots, 1)$ or $\lambda = (n)$ and $\mu = (1, 1, \ldots, 1)$, where π is the stationary distribution from Proposition 2.2.6 given by*

$$\pi(\lambda) = \left(\prod_{i=1}^{k-1} \binom{\Lambda_i - 1}{\lambda_i - 1} \left(\frac{\Lambda_{i+1}\alpha + \lambda_i \theta}{n_i} \right) \right) \frac{\prod_{i=1}^{k} (1-\alpha)_{(\lambda_i - 1)\uparrow}}{(\theta+1)_{(n-1)\uparrow}},$$

where $\Lambda_i = \sum_{j=i}^{l} \lambda_j$ for $1 \leq i \leq l$, $x_{0\uparrow} := 1$ and $x_{y\uparrow} := (x + (y-1))x_{(y-1)\uparrow}$, for $x \in \mathbb{R}$, $y \in \mathbb{N}$.

This is proved in exactly the same way as in [17]. We also note that Corollary 3.2.16, Corollary 3.2.18 and Proposition 3.2.19 from [17] still hold here and are proved in exactly the same way as before, so let us conclude by proving Theorem 3.1.2, an extension of [17, Proposition 7.5] to the case when $\theta > 0$.

The proof will require the following lemmas.

Lemma 3.3.4. *If $\exists i_c \in \mathbb{N}$ such that the sequences $(a_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}}$ with $a_{n,i} \geq 0$ are decreasing in i for $i \geq i_c$ and $a_{n,i} \to b_i$ as $n \to \infty$, then provided that $\sum_{i=i_c}^{\infty} b_i < \infty$,*

$$\sum_{i=i_c}^{n} (-1)^i a_{n,i} \to \sum_{i=i_c}^{\infty} (-1)^i b_i$$

as $n \to \infty$.

Proof. Since the $a_{n,i} \geq 0$ are decreasing in $i \geq i_c$, for any $m, n \in \mathbb{N}$,

$$\sum_{i=i_c}^{2m-1}(-1)^i a_{n,i} \leq \sum_{i=i_c}^{n}(-1)^i a_{n,i} \leq \sum_{i=i_c}^{2m}(-1)^i a_{n,i}.$$

Observe that

$$\limsup_{n\to\infty}\sum_{i=i_c}^{n}(-1)^i a_{n,i} \leq \inf_{m\geq 1}\sum_{i=i_c}^{2m}(-1)^i b_i = \sum_{i=i_c}^{\infty}(-1)^i b_i$$
$$\text{and } \liminf_{n\to\infty}\sum_{i=i_c}^{n}(-1)^i a_{n,i} \geq \sup_{m\geq 1}\sum_{i=i_c}^{2m-1}(-1)^i b_i = \sum_{i=i_c}^{\infty}(-1)^i b_i.$$

Since the limsup and liminf of $\sum_{i=i_c}^{n}(-1)^i a_{n,i}$ are equal, it follows that

$$\lim_{n\to\infty}\sum_{i=i_c}^{n}(-1)^i a_{n,i} = \sum_{i=i_c}^{\infty}(-1)^i b_i.$$

□

Theorem 3.3.5 (Restatement of Theorem 3.1.2)**.** *If $s^*(r)$ is the maximal separation distance after $r \geq 1$ iterations of the down-up chain on the Kingman lattice or ordered Kingman lattice and $0 \leq \alpha < 1$, then if $\theta > 0$,*

$$s^*(r) = \sum_{i=2}^{n}\left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^r \frac{(-1)^i(i-1)(\theta+2i-1)}{\theta+i-1}\binom{n}{i}\frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)}.$$

Moreover for any fixed $c > 0$,

$$\lim_{n\to\infty} s^*(cn(\theta+n-1)) = \sum_{i=2}^{\infty} e^{-ci(\theta+i-1)}\frac{(-1)^i(i-1)(\theta+2i-1)}{i!(\theta+i-1)}\prod_{j=1}^{i}(\theta+j).$$

Lemma 3.3.6. *For fixed $\theta > 0$, there exist some $A > 0$ and $I \in \mathbb{N}$ with the property that if $i \geq I$, and $An(\theta+n-1) \leq i^2(\theta+i-1)^2$ then $n < \frac{2i^2+2i-\theta}{\theta+2}$.*

Proof. Clearly there are natural numbers i_1, i_2 and i_3 such that

- If $i \geq i_1$, $\theta + i - 1 \leq \sqrt{2}i$.
- If $i \geq i_2$, $2i^2 + 2i - \theta > 0$.

- If $i \geq i_3$, $2i^4 < 4i^4 + 8i^3 + 4i^2 - 4i^2\theta - 4i\theta + \theta^2 + (\theta - 1)(\theta + 2)(2i^2 + 2i - \theta)$.

Let $I := \max\{i_1, i_2, i_3\}$, and assume that $i \geq I$ for the remainder of this proof.

We make the claim that when b, $x \geq 0$, $x^2 + x(\theta - 1) < b$ if and only if $x < a$, where $a > 0$ is the unique positive root of $b = x^2 + x(\theta - 1)$, the proof of this statement is completely elementary. Then as $i \geq i_2$, $n < \frac{2i^2+2i-\theta}{\theta+2}$ is equivalent to

$$n^2 + n(\theta - 1) < \frac{4i^4 + 8i^3 + 4i^2 - 4i^2\theta - 4i\theta + \theta^2}{(\theta + 2)^2} + \frac{(\theta - 1)(2i^2 + 2i - \theta)}{\theta + 2}$$

Let $A > (\theta + 2)^2$, then $An(\theta + n - 1) \leq i^2(\theta + i - 1)^2 \leq 2i^4$ as $i \geq i_1$ and so $n(\theta + n - 1) \leq \frac{2i^4}{A} < \frac{2}{(\theta+2)^2} i^4$. As $i \geq i_3$, the result follows.

□

Proof of Theorem 3.1.2. By the definition of the maximal separation distance and Propositions 3.3.3 and 3.2.3,

$$\begin{aligned} s^*(r) &= \max_{\lambda,\mu} \left(1 - \frac{p_{UD}^{(r)}(\mu, \lambda)}{\pi(\lambda)}\right) \\ &= \sum_{i=2}^{r} \left(1 - \frac{i(\theta + i - 1)}{n(\theta + n - 1)}\right)^r \prod_{j=2,\, j\neq i}^{n} \left(\frac{j(\theta + j - 1)}{j(\theta + j - 1) - i(\theta + i - 1)}\right), \end{aligned} \tag{1}$$

since the eigenvalues of $(U_{n-1}D_n)^r$ are 1 and $1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}$, $2 \leq i \leq n$. Factorising the denominator $j(\theta + j - 1) - i(\theta + i - 1)$,

$$\prod_{j=2,\, j\neq i}^{n} \left(\frac{j(\theta + j - 1)}{j(\theta + j - 1) - i(\theta + i - 1)}\right) = \frac{\prod_{j=2,\, j\neq i}^{n} [j(\theta + j - 1)]}{\prod_{j=2,\, j\neq i}^{n} [(j - i)(i + j + \theta - 1)]}. \tag{2}$$

The numerator of (2) is

$$\frac{n!}{i} \frac{(\theta + 1)(\theta + 2) \ldots (\theta + n - 1)}{\theta + i - 1}.$$

The denominator of (2) is $(-1)^{i-2}(i - 2)! \prod_{j=i+1}^{2i-2} (\theta + j) \times (n - i)! \prod_{j=2i}^{i+n-1} [\theta + j]$.

If we substitute this into (1), we get

$$s^*(r) = \sum_{i=2}^{n} \left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^r \frac{n!}{i} \cdot \frac{\prod_{j=1}^{n-1}(\theta+j)}{\theta+i-1} \times \frac{1}{(-1)^i (i-2)!(n-i)! \prod_{j=i+1}^{2i-2}(\theta+j) \prod_{j=2i}^{i+n-1}(\theta+j)}$$

and this simplifies to the required expression.

For the second part, we wish to apply Lemma 3.3.4 and must show that when

$$a_{n,i} = \left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^{cn(\theta+n-1)} \frac{(i-1)(\theta+2i-1)}{\theta+i-1} \binom{n}{i} \frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)}$$

that these terms are decreasing for some $i \geq i_c$, as clearly

$$\left(1 - \frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^{cn(\theta+n-1)} \frac{(i-1)(\theta+2i-1)}{i!(\theta+i-1)} \to e^{-ci(\theta+i-1)} \frac{(i-1)(\theta+2i-1)}{i!(\theta+i-1)}.$$

Now

$$\begin{aligned} \frac{n!}{(n-i)!} \frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)} &= \prod_{j=1}^{i}(n+1-j) \frac{\prod_{j=1}^{i}(\theta+j)}{\prod_{j=n}^{i+n-1}(\theta+j)} \\ &= \prod_{j=1}^{i}(\theta+j) \prod_{j=1}^{i} \frac{n+1-j}{\theta+n+j-1} \to \prod_{j=1}^{i}(\theta+j). \end{aligned}$$

By the algebra of limits, it follows that

$$a_{n,i} \to b_i := e^{-ci(\theta+i-1)} \frac{(i-1)(\theta+2i-1)}{i!(\theta+i-1)} \prod_{j=1}^{i}(\theta+j)$$

and so the formula for $\lim_{n\to\infty} s^*(cn(\theta+n-1))$ will follow from Lemma 3.3.4 upon showing that $\exists i_c \in \mathbb{N}$ such that for $i \geq i_c$ and all n, the $a_{n,i}$ are decreasing, as it is elementary to show that $\sum_{i=i_c}^{\infty} b_i < \infty$ here.

To do this, we first show that there are constants $A > 0$ and $i_1 \in \mathbb{N}$ for which the $a_{n,i}$ decrease if $i \geq i_1$ and $An(\theta+n-1) \leq i^2(\theta+i-1)^2$, if n is sufficiently large.

Write

$$a_{n,i} = \frac{(i-1)(\theta+2i-1)}{i!(\theta+i-1)}\left(1-\frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^{cn(\theta+n-1)}\frac{n!}{(n-i)!}\frac{\prod_{j=1}^{n-1}(\theta+j)}{\prod_{j=i+1}^{i+n-1}(\theta+j)}.$$

and note that since $\left(1-\frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^{cn(\theta+n-1)}$ is decreasing then the claim about $a_{n,i}$ is implied by the claim that

$$c_{n,i} := \frac{(i^2-1)(\theta+2i-1)}{(\theta+i-1)}\cdot\frac{n!}{(n-i)!}\cdot\frac{\prod_{j=1}^{n-1}(\theta+j)}{(i+1)!\prod_{j=i+1}^{i+n-1}(\theta+j)}$$

is decreasing. Via elementary simplifications,

$$\begin{aligned}\frac{c_{n,i+1}}{c_{n,i}} &= \frac{i(\theta+2i+1)(\theta+i-1)}{(i^2-1)(\theta+2i-1)(\theta+i)}\cdot\frac{(n-i)(\theta+i+1)}{(\theta+i+n)}\\ &= \left[\frac{i(\theta+i-1)}{\theta+2i-1}\div\frac{(i+1)(\theta+i)}{\theta+2i+1}\right]\frac{(n-i)(\theta+i+1)}{(i-1)(\theta+i+n)}.\end{aligned}$$

There exists some constant $i_1 \in \mathbb{N}$ for which $d_i := \frac{i(\theta+i-1)}{\theta+2i-1}$ is increasing when $i \geq i_1$ as it behaves asymptotically like $\frac{i}{2}$, and so showing that $\frac{(n-i)(\theta+i+1)}{(i-1)(\theta+i+n)} < 1$ is sufficient to check that the $c_{n,i}$ are decreasing. Now this is equivalent to $n < \frac{2i^2+2i\theta-\theta}{\theta+2}$, and by Lemma 3.3.6 this holds if there exist some $A > 0$ and $i_2 \in \mathbb{N}$ for which $An(\theta+n-1) \leq i^2(\theta+i-1)^2$ when $i \geq i_2$.

To deal with the $i \in \mathbb{N}$ for which $i^2(\theta+i-1)^2 \leq An(\theta+n-1)$, write

$$\begin{aligned}a_{n,i} &= q(i)\left(1-\frac{i(\theta+i-1)}{n(\theta+n-1)}\right)^{cn(\theta+n-1)}\tilde{c}_{n,i},\\ \text{where } \tilde{c}_{n,i} &= \frac{n!\prod_{j=1}^{n-1}(\theta+j)}{(n-i)!(i+\lfloor\theta\rfloor+1)!\prod_{j=i+1}^{i+n-1}(\theta+j)},\\ \text{and } q(i) &= \frac{(i-1)(\theta+2i+1)\prod_{j=1}^{\lfloor\theta\rfloor+1}(j+i)}{\theta+i-1}.\end{aligned}$$

Now

$$\frac{\tilde{c}_{n,i+1}}{\tilde{c}_{n,i}} = \frac{(n-i)!(i+\lfloor\theta\rfloor+1)!}{(n-i-1)!(i+\lfloor\theta\rfloor+2)!}\cdot\frac{\theta+i+1}{\theta+i+n} = \frac{(n-i)(\theta+i+1)}{(i+\lfloor\theta\rfloor+2)(\theta+i+n)},$$

so that $\tilde{c}_{n,i}$ is easily seen to be decreasing if and only if

$$n(\theta - \lfloor\theta\rfloor - 1) < i(\theta + i + 1) + (\theta + i)(i + \theta + 2),$$

which is trivially true from the definition of the floor function $\lfloor\cdot\rfloor$.

Thus we show that the $a_{n,i}$ are decreasing by showing that

$$\widetilde{d}_{n,i} := q(i)\left(1 - \frac{i(\theta + i - 1)}{n(\theta + n - 1)}\right)^{cn(\theta+n-1)}$$

is a sequence decreasing in i. Observe that

$$\frac{\widetilde{d}_{n,i}}{\widetilde{d}_{n,i+1}} = \frac{q(i)}{q(i+1)} \exp\left(cn(\theta+n-1)\left[\ln\left(1-\frac{i(\theta+i-1)}{n(\theta+n-1)}\right)-\ln\left(1-\frac{(i+1)(\theta+i)}{n(\theta+n-1)}\right)\right]\right).$$

Use the inequalities $\leq -x - x^2 \leq \ln(1-x) \leq -x$ for $0 \leq x \leq \frac{1}{2}$ on the natural logarithms in the exponential, which is seen to be valid upon observing that as $i^2(\theta + i - 1)^2 \leq An(\theta + n - 1)$, $(i + 1)(\theta + i) \leq \frac{1}{2}n(\theta + n - 1)$ when $i \geq A$ and n is large enough. Then

$$\begin{aligned}
&cn(\theta + n - 1)\left[\ln\left(1 - \frac{i(\theta + i - 1)}{n(\theta + n - 1)}\right) - \ln\left(1 - \frac{(i + 1)(\theta + i)}{n(\theta + n - 1)}\right)\right] \\
\geq &- ci(\theta + i - 1) - \frac{ci^2(\theta + i - 1)^2}{n(\theta + n - 1)} + c(i + 1)(\theta + i) \\
= &\, c\left[\theta + 2i - \frac{i^2(\theta + i - 1)^2}{n(\theta + n - 1)}\right] \geq ci
\end{aligned}$$

when $i \geq A$, as then $i^2(\theta+i-1)^2 \leq An(\theta+n-1) \leq i(\theta+n-1)$ and so $\frac{i^2(\theta+i-1)^2}{n(\theta+n-1)} \leq i + \theta$. Since $\frac{q(i)}{q(i+1)}$ is a rational fraction converging to 1 as $i \to \infty$ which does not depend on n and in particular eventually exceeds $\frac{1}{2}$, it follows that $\exists i_3 \in \mathbb{N}$ with $\frac{\widetilde{d}_{n,i}}{\widetilde{d}_{n,i+1}} \geq \frac{1}{2}e^{ci}$, for $i \geq i_3$ and so $\exists i_4 \in \mathbb{N}$ for which $\widetilde{d}_{n,i+1} < \widetilde{d}_{n,i}$ when $i \geq i_4$. In particular upon choosing $i_c \in \mathbb{N}$ with $i_c \geq \max\{i_1, i_2, i_3, i_4, A\}$, we observe that the conditions needed for Lemma 3.3.4 hold and therefore so does the limiting expression for $s^*(cn(\theta + n - 1))$. □

BIBLIOGRAPHY

[1] D. Aldous. Mixing Time for a Markov Chain on Cladograms. *Combinatorics, Probability and Computing*, 9:191–204, 2000.

[2] D. Aldous. Wright-Fisher diffusions with negative mutation rate! *https://www.stat.berkeley.edu/ aldous/Research/OP/fw.html*, 1999.

[3] D. Aldous and P. Diaconis. Strong Uniform Times and Finite Random Walks. *Advances in Applied Mathematics*, 8:69–97, 1987.

[4] S. Asmussen, F. Avram and M. R. Pistorius. Russian and American put options under exponential phase-type Lévy models. *Stochastic Processes and their Applications*, 109(1):79–111, 2004.

[5] J. Berestycki, J. Fontbona and M. C. Knight. Ray–Knight representation of flows of branching processes with competition by pruning of Lévy trees. *Probability Theory and Related Fields*, 172(3-4):725–788, 2018.

[6] P. Billingsley. *Convergence of Probability Measures*. John Wiley and Sons, New York, 2nd edition 1999.

[7] A. Borodin and G. Olshanski. Infinite-dimensional diffusions as limits of random walks on partitions. *Probability Theory and Related Fields*, 144:281–318, 2009.

[8] M. A. Carlton. A family of densities derived from the three-parameter Dirichlet process. *Journal of Applied Probability*, 39(4):764-774, 2002.

[9] P. Diaconis and P. Hanlon. Eigen-analysis for Some Examples of the Metropolis Algorithim. *Contemporary Mathematics*, 138:99–117, 1992.

[10] S. Feng and W. Sun. Some diffusion processes associated with two parameter Poisson-Dirichlet distribution and Dirichlet process. *Probability Theory and Related Fields*, 148(3-4):501-525, 2010.

[11] N. Forman, S. Pal, D. Rizzolo, Q. Shi and M. Winkel. A two-parameter family of measure-valued diffusions with Poisson–Dirichlet stationary distributions. Work in progress. 2019.

[12] N. Forman, S. Pal, D. Rizzolo and M. Winkel. Diffusions on a space of interval partitions: construction from marked Lévy processes. *ArXiv e-prints, arXiv:1909.02584*, 2019.

[13] N. Forman, S. Pal, D. Rizzolo and M. Winkel. Diffusions on a space of interval partitions: Poisson–Dirichlet stationary distributions. work in progress, revising parts of arxiv: 1609.06706v2. *ArXiv e-prints, arXiv:1909.07626*, 2019.

[14] N. Forman, S. Pal, D. Rizzolo and M. Winkel. Diffusions on a space of interval partitions with Poisson-Dirichlet stationary distributions. *ArXiv e-prints, 1609.06706v2*, 2017.

[15] N. Forman, S. Pal, D. Rizzolo and M. Winkel. Uniform control of local times of spectrally positive stable processes. *The Annals of Applied Probability*, 28(4):2592–2634, 2018.

[16] N. Forman, S. Pal, D. Rizzolo and M. Winkel. Unpublished work, private correspondence.

[17] J. Fulman. Commutation relations and Markov chains. *Probability Theory and Related Fields*, 144:99–136, 2009.

[18] J. Fulman. Separation cutoffs for random walk on irreducible representations. *Annals of Combinatorics*, 14:319–337, 2010.

[19] J. Fulman. Stein's method and Plancherel measure of the symmetric group. *Transactions of the American mathematical society*, 357(2):555–570, 2005.

[20] J. Geiger and G. Kersting. Depth-first search of random trees and Poisson point processes. *IMA Volumes in Mathematics and its Applications*, 84:111–126, 1997.

[21] A. Gnedin and J. Pitman. Regenerative composition structures. *The Annals of Probability*, 33:445–479, 2005.

[22] A. Göing-Jaeschke and M. Yor. A survey and some generalisations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[23] G. Grimmett and D Stirzaker. *Probability and Random Processes*. Oxford University Press, Oxford, 3rd edition 2001.

[24] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of Grundlehren der Mathematischen Wissenschaften Springer-Verlag, Berlin, 2nd edition, 2003.

[25] P. Jagers. A general stochastic model for population development. *Scandinavian Actuarial Journal*, 1969(1-2):84–103, 1969.

[26] P. Jagers. *Branching processes with biological applications*. John Wiley and Sons, New York, 1975.

[27] L. James. Poisson Calculus for spatial neutral to the right processes. *The Annals of Statistics*, 34(1):416–440, 2006.

[28] P. Joyce and S. Tavaré. Cycles, permutations and the structure of the Yule process with immigration. *Stochastic Processes and their Applications*, 25:309–314, 1987.

[29] F. B. Knight. Random walks and a sojourn density process of Brownian motion. *Transactions of the American Mathematical Society*, 109(1):56–86, 1963.

[30] A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer-Verlag, Berlin, 1st edition 2006.

[31] A. Lambert. The contour of splitting trees is a Lévy process. *The Annals of Probability*, 38(1):348–395, 2010.

[32] L. Lorentzen and H. Waadeland. *Continued fractions, 2nd edition, Volume 1: Convergence Theory.* Atlantis Press, 2008.

[33] I. Macdonald. *Symmetric functions and Hall Polynomials.* Oxford University Press, Oxford, 2nd edition, 1995.

[34] I. Pak. *Random walks on groups: Strong uniform time approach.* PhD thesis, Havard, 1997.

[35] S. Pal. On the Aldous diffusion on Continuum Trees. I. *ArXiv e-prints, 1104.4184*, 2011.

[36] S. Pal. Wright-Fisher diffusion with Negative Mutation Rates. *The Annals of Probability*, 41(2):503–526, 2013.

[37] L. Petrov. Two-parameter family of diffusion processes in the Kingman simplex. *Functional Analysis and its applications*, Vol 43(4):279–296, 2009.

[38] J. Pitman. *Combinatorial Stochastic Processes.* Springer-Verlag, Berlin, 2006.

[39] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probability theory and related fields*,102:145–158, 1995.

[40] J. Pitman and M. Winkel. Regenerative tree growth: Binary self-similar continuum random trees and Poisson-Dirichlet compositions. *The Annals of Probability*, 37:1999–2041, 2009.

[41] J. Pitman and Y. Yakubovich. Ordered and size-biased frequencies in gem and Gibbs' models for species sampling. *The Annals of Applied Probability*, 28(3):1793–1820, 2018.

[42] J. Pitman and M. Yor. A decomposition of Bessel bridges. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 59:425–457, 1982.

[43] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO - Informatique Théorique et Applications*, 19(2):179–195, 1985.

[44] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, 1999.

[45] D. Rogers and M. Winkel. Ray-Knight representation of up-down Chinese Restaurant Processes. Work in progress.

[46] A. V. Skorokhod. Limit Theorems for Stochastic Processes with Independent Increments. *Theory of Probability and its Applications*, II(2):138–171, 1957.

[47] Y. W. Teh. A hierarchical Bayesian language model based on Pitman-Yor processes. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 985-992. Association for Computational Linguistics 2006.