

Weierstraß-Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint

ISSN 0946 – 8633

Rescaled stable generalised Fleming-Viot processes: Flickering random measures

Matthias Birkner¹, Jochen Blath²

submitted: 11th September 2007

¹ Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin
Germany
E-mail: birkner@wias-berlin.de

² Institut für Mathematik, Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
Germany
E-mail: blath@math.tu-berlin.de

No. 1252
Berlin 2007



2000 *Mathematics Subject Classification.* Primary: 60G57; Secondary: 60G17.

Key words and phrases. Generalised Fleming-Viot process, measure-valued diffusion, lookdown construction, wandering random measure, path properties, tightness, Skorohod topology.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

Abstract

We show how Donnelly and Kurtz' (modified) lookdown construction for measure-valued processes can be used to analyse the longterm- and scaling properties of spatially stable generalised Λ -Fleming Viot processes, exhibiting a rare "natural" example of a scaling family converging in f.d.d. sense, but not in any of Skorohod's topologies on path space. This completes results of Fleischmann and Wachtel (2004) about the spatial Neveu process and complements results of Dawson and Hochberg (1982) about the classical Fleming Viot process. The lookdown construction provides an elegant machinery and clear intuition to describe the path properties of the family in terms of a "flicker effect", clarifying "what can go wrong."

1 Introduction

1.1 Classical and generalized Fleming-Viot processes

In 1979, Fleming and Viot introduced their now well-known probability-measure-valued stochastic process as a model for the distribution of allelic frequencies in a selectively neutral genetic population with mutation (cf. [FV79]). More formally, they introduced a Markov process $\{Y_t^{\delta_0, \Delta}, t \geq 0\}$, with values in $\mathcal{M}_1(\mathbb{R}^d)$ (denoting the probability measures on \mathbb{R}^d), such that for functions F of the form

$$F(\rho) := \prod_{i=1}^n \langle \phi_i, \rho \rangle, \quad (1.1)$$

where $\phi_i \in C_c^2(\mathbb{R}^d)$ and $\rho \in \mathcal{M}_1(\mathbb{R}^d)$, the generator of $\{Y_t^{\delta_0, \Delta}, t \geq 0\}$ can be written as

$$LF(\rho) = \sum_{i=1}^n \langle \Delta \phi_i, \rho \rangle \prod_{j \neq i} \langle \phi_j, \rho \rangle + \sum_{1 \leq i < j \leq n} \left[\langle \phi_i \phi_j, \rho \rangle - \langle \phi_i, \rho \rangle \langle \phi_j, \rho \rangle \right] \prod_{k \neq i, j} \langle \phi_k, \rho \rangle,$$

with Δ the Laplace operator. The meaning of the superscripts in $\{Y_t^{\delta_0, \Delta}, t \geq 0\}$ will become clear once we identify this process as a special case of a much larger class of processes.

It is well known (cf. [DH82]) that the classical Fleming-Viot process is dual to *Kingman's coalescent* (introduced in [K82]) in the following (our description being rather informal) sense. For $t \geq 0$, if one takes a uniform sample of n individuals from $Y_t^{\delta_0, \Delta}$ and forgets about the respective spatial positions of the n particles, then their genealogical tree backwards in time can be viewed as a realisation of *Kingman's n -coalescent*. That means, at each time $t - s$, where $s \in [0, t]$ (hence *backwards* in time), the ancestral lineages of each particle merge at infinitesimal rate $\binom{k}{2}$, where $k \in \{2, \dots, n\}$ denotes the number of distinct lineages present at time $t - s(-)$. This can be made rigorous, for example, using Donnelly and Kurtz (1996) *lookdown construction* [DK96], and spatial information may also be incorporated, see e.g. [Eth00], Section 1.12.

Since its introduction, the Fleming-Viot process received a great deal of attention from both geneticists and probabilists. One reason is that it is the natural limit of a large class of exchangeable population models with constant size and finite-variance reproduction mechanism, in particular the so-called Moran-model, and can be viewed as the infinite-dimensional analogue of the Wright-Fisher diffusion. See [Eth00] for a good overview.

A corresponding limit population process describing situations where, from time to time, a single individual produces a non-negligible fraction of the total population, has been introduced somewhat implicitly in [DK99], and explicitly in [BLG03]. The limits of the dual genealogical processes have been classified in [Sa99], [MS01].

See [BB07] for an overview. These are probability measure valued Markov processes $Y^{\Lambda, \Delta_\alpha}$ whose generator acts on functions F of the form (1.1) as

$$LF(\rho) = \sum_{i=1}^n \Delta_\alpha \langle \phi_i, \rho \rangle \prod_{j \neq i} \langle \phi_j, \rho \rangle + \sum_{\substack{J \subset \{1, \dots, n\} \\ |J| \geq 2}} \lambda_{n, |J|} \left[\prod_{j \in J} \langle \phi_j, \rho \rangle - \prod_{j \in J} \langle \phi_j, \rho \rangle \right] \prod_{k \notin J} \langle \phi_k, \rho \rangle, \quad (1.2)$$

where

$$\lambda_{n, k} = \int_{[0, 1]} x^{k-2} (1-x)^{n-k} \Lambda(dx), \quad n \geq k \geq 2, \quad (1.3)$$

with Λ a finite measure on $[0, 1]$, and $\Delta_\alpha = -(-\Delta)^{\alpha/2}$ is the fractional Laplacian of index $\alpha \in (0, 2]$, see e.g. [Y65], Chapter IX.11, or [Fe66], Chapter IX.6, i.e. Δ_α is the generator of the semigroup $(P_t^{(\alpha)})_{t \geq 0}$ of the d -dimensional standard symmetric stable process $\{B_t^{(\alpha)}, t \geq 0\}$ of index α . Note that for notational convenience, we denote by $(P_t^{(2)})_{t \geq 0}$ the semigroup of d -dimensional Brownian motion with covariance matrix 2Id at time 1. We endow $\mathcal{M}_1(\mathbb{R}^d)$ with the topology of weak convergence, which we think of being induced the metric (see e.g. [DK96], Remark 2.5)

$$d_{\mathcal{M}_1}(\mu, \nu) := \sum_{k=1}^{\infty} \frac{1}{2^k} |\langle f_k, \mu - \nu \rangle|, \quad \mu, \nu \in \mathcal{M}_1(\mathbb{R}^d), \quad (1.4)$$

where $(f_k) \subset C_c^2(\mathbb{R}^d)$ is dense (w.r.t. the sup-norm of $C_c^2(\mathbb{R}^d)$). By [DK99], Thm. 3.2, the processes $\{Y_t^{\Lambda, \Delta_\alpha}, t \geq 0\}$ take values in $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$, the space of càdlàg paths, endowed with the usual Skorohod (J_1 -)topology (cf. [S56], or [Bi68], Ch. 3).

For a given $\Lambda \in \mathcal{M}_f([0, 1])$, the rates $\lambda_{n, k}$ describe the transitions of an exchangeable partition-valued process $\{\Pi_t^\Lambda, t \geq 0\}$, the so-called Λ -coalescent ([Pi99], [Sa99]). While, for $t \geq 0$, Π_t^Λ has n classes, say, any k -tuple merges to one at rate $\lambda_{n, k}$. Indeed, as shown in [BLG03], a Λ -Fleming-Viot process is dual to a so-called Λ -coalescent, similar to the duality between the standard Fleming-Viot process and Kingman's coalescent established in [DH82]. Note that Kingman's coalescent corresponds to the choice $\Lambda = \delta_0$.

1.2 Relation between generalised Fleming-Viot processes and infinitely divisible superprocesses

Fleischmann and Wachtel ([FW06]) have considered a probability measure valued process $\{Y_t, t \geq 0\}$ obtained by renormalising a spatial version of Neveu's continuous mass branching process $\{X_t, t \geq 0\}$ with underlying α -stable motion (as constructed e.g. in [FS04] via approximation or implicitly in [DK99]) with its total mass, i.e. $\langle \phi, Y_t \rangle = \langle \phi, X_t \rangle / \langle 1, X_t \rangle$, and have investigated its long-time behaviour.

In [BBC05], the relation between stable continuous-mass branching processes $\{Z_t, t \geq 0\}$ and Beta($2 - \beta, \beta$)-Fleming Viot processes, for $\beta \in (0, 2]$, (with a "trivial" spatial motion) has been explored. Informally, $Z_t / \langle 1, Z_t \rangle$, time-changed with the inverse of

$$\int_0^t (Z_t)^{1-\beta} dt, \quad (1.5)$$

is a Beta($2 - \beta, \beta$)-Fleming Viot process. This can be viewed as an extension of Perkins' classical disintegration theorem ([EM91], [Pe91]) to the stable case. It is in principle easy to include a spatial motion component, but note that then the corresponding Fleming-Viot process uses a time-inhomogenous motion, namely an α -stable process time-changed by the inverse of (1.5). However, Neveu's branching mechanism is stable of index $\beta = 1$, so that the time change induced by (1.5) becomes trivial. Thus we obtain

Proposition 1.1.

$$\{X_t / \langle 1, X_t \rangle, t \geq 0\} \stackrel{d}{=} \{Y_t^{U, \Delta_\alpha}, t \geq 0\},$$

where $U = \text{Beta}(1, 1)$ is the uniform distribution on $[0, 1]$.

Note that in particular in this situation, the (randomly) renormalised process $\{X_t / \langle 1, X_t \rangle, t \geq 0\}$ is itself a Markov process. In fact, as observed in [BBC05], it is the only "superprocess" with this property. This observation was the starting point of our investigation.

Remark 1.2 (First two moment measures). By considering F as in (1.1) with $n = 1$ and $n = 2$, it follows from the martingale problem for (1.2) that the first two moments of a generalised Λ -Fleming-Viot process only depend on the underlying motion mechanism and the total mass $\Lambda([0, 1])$, namely

$$\mathbb{E}[\langle \varphi, Y_t^{\Lambda, \Delta_\alpha} \rangle] = \int P_t^{(\alpha)} \varphi(x) \mu(dx), \quad (1.6)$$

and for $t_1 \leq t_2$, writing $\rho := \Lambda([0, 1])$,

$$\begin{aligned} \mathbb{E}[\langle \varphi_1, Y_{t_1}^{\Lambda, \Delta_\alpha} \rangle \langle \varphi_2, Y_{t_2}^{\Lambda, \Delta_\alpha} \rangle] &= \int_0^{t_1} \rho e^{-\rho s} P_s^{(\alpha)} (P_{t_1-s}^{(\alpha)} \varphi_1 P_{t_2-s}^{(\alpha)} \varphi_2)(x) \mu(dx) \\ &\quad + e^{-\rho t_1} \int P_{t_1}^{(\alpha)} \varphi_1(x) \mu(dx) \int P_{t_2}^{(\alpha)} \varphi_2(x) \mu(dx), \end{aligned} \quad (1.7)$$

for $\varphi, \varphi_1, \varphi_2 \in C_c^2$. In particular, they agree with those of the classical Fleming-Viot process, which explains Proposition 3 in [FW06].

Remark 1.3 (Non-compact support property). It is interesting to see that, unlike the classical Fleming-Viot process $Y^{\delta_0, \Delta}$ ([DH82, Thm. 7.1]), generalised Fleming-Viot processes need not have the compact support property, even if the underlying motion is Brownian and the initial state has compact support.

Indeed, if the dual Λ -coalescent Π^Λ does not come down from infinity, i.e. if starting from $\Pi_0^\Lambda = \{\{1\}, \{2\}, \dots\}$, the number of classes $|\Pi_t^\Lambda|$ of Π_t^Λ is (a.s.) infinite for any $t > 0$, then

$$\text{supp}(Y_t^{\Lambda, \Delta}) = \mathbb{R}^d \quad \text{a.s. for any } t.$$

Recall that if the standard Λ -coalescent does not come down from infinity (a necessary and sufficient condition for this can be found in [Sc00]), it either has a positive fraction of singleton classes (so-called “dust”), or countably many families with strictly positive asymptotic mass adding up to one (so called “proper frequencies”), cf. [Pi99], Lemma 25.

Using the path-wise embedding of the standard Λ -coalescent in the Fleming-Viot process provided by the modified lookdown construction (see (2.7) below) we see that in the first case, the positive fraction of singletons contributes an α -heat flow component to $Y_t^{\Lambda, \Delta_\alpha}$, whereas in the latter case there are infinitely many independent families of strictly positive mass, so that by the Borel-Cantelli Lemma any given open ball in \mathbb{R}^d will be charged almost surely.

Combining this with Proposition 1.1, we recover Proposition 14 of [FS04].

Remark 1.4 (Generalised Λ -Fleming Viot processes as “wandering random measures”). In the terminology of [DH82], the classical Fleming Viot process is a (compactly) coherent wandering random measure, meaning that there is a “centring process” $\{x(t), t \geq 0\}$ with values in \mathbb{R}^d and for each $\varepsilon > 0$ a stationary “radius process” $\{R_\varepsilon(t)\}$ and an a.s. finite T_0 , such that

$$Y_t^{\delta_0, \Delta}(B_{x(t)}(R_\varepsilon(t))) \geq 1 - \varepsilon \quad \text{for } t \geq T_0 \quad \text{a.s.}, \quad (1.8)$$

where $B_x(r)$ is the closed ball of radius r around $x \in \mathbb{R}^d$. One natural choice for $\{x(t), t \geq 0\}$ is the centre of mass process $x(t) = \int x Y_t^{\delta_0, \Delta}(dx)$, see [DH82], Equation 3.10. However, in the context of the lookdown construction, a more convenient choice is $x(t) = \xi_t^1$, the location of the level-1 particle (see Section 2). With this choice, an obvious extension of [DK96], Thm. 2.9, shows that any $Y^{\Lambda, \Delta_\alpha}$ is a coherent wandering random measure. If the process $Y^{\Lambda, \Delta_\alpha}$ has the compact support property, this will also yield compact coherence, i.e. one can choose $\varepsilon = 0$ in (1.8).

In Corollary 6 of [FW06], it is observed that for continuous test functions φ with compact support,

$$t^{d/\alpha} \mathbb{E}[\langle \varphi, Y^{U, \Delta_\alpha} \rangle] \rightarrow p_1^{(\alpha)}(0) \int \varphi(x) dx \quad \text{as } t \rightarrow \infty, \quad (1.9)$$

where $p_t^{(\alpha)}(x)$ is the transition density of $\{B_t^{(\alpha)}, t \geq 0\}$, and in the subsequent Remark 7, Fleischmann and Wachtel ask about convergence of $t^{d/\alpha}\langle \varphi, Y^{U, \Delta_\alpha} \rangle$. With the lookdown construction in mind, (1.9) can be at least qualitatively understood as follows: without loss of generality assume that φ has support in the unit ball, put $C_t := \langle \varphi, Y_t^{\Lambda, \Delta_\alpha} \rangle$. Consider the empirical process $\{Y_t^{\Lambda, \Delta_\alpha}, t \geq 0\}$ together with $\{\xi_t^1, t \geq 0\}$, the position of the level-1 particle. Then $Y_t^{\Lambda, \Delta_\alpha}(\cdot - \xi_t^1)$ converges to some stationary distribution. Thus if ξ_t^1 is “close” to the origin, an event of probability $\approx t^{-d/\alpha}$, C_t is substantial, whereas otherwise it is essentially zero. The terms balance exactly, so that the lefthand side of (1.9) converges, but in fact as $\{B_t^{(\alpha)}, t \geq 0\}$ is not positive recurrent, C_t converges to zero in distribution (and even a.s. if $\alpha < d$, i.e. if ξ_t^1 is transient).

1.3 Statement of the main result

The long-time behaviour of a generalised Fleming-Viot process reflects the interplay between motion and resampling mechanism. If one attempts to capture this via a space-time rescaling, the scaling will be dictated by the underlying (stable) motion process.

Theorem 1.5 (Scaling). *Let $\Lambda \in \mathcal{M}_f([0, 1]) - \{0\}$ and define the rescaled process $\{Y_t^{\Lambda, \Delta_\alpha}[k], t \geq 0\}$ via*

$$\langle \phi, Y_t^{\Lambda, \Delta_\alpha}[k] \rangle := \langle \phi(\cdot/k^{1/\alpha}), Y_{kt}^{\Lambda, \Delta_\alpha} \rangle, \quad (1.10)$$

for $\phi \in b\mathcal{B}(\mathbb{R}^d)$ and $t \geq 0$. Let $B^{(\alpha)}$, for $\alpha \in (0, 2]$, be the standard symmetric stable process of index α , starting from $B_0^{(\alpha)} = 0$. Then,

a)

$$\{Y_t^{\Lambda, \Delta_\alpha}[k], t \geq 0\} \rightarrow \{\delta_{B_t^{(\alpha)}}, t \geq 0\} \quad \text{as } k \rightarrow \infty, \quad (1.11)$$

in the sense of the finite-dimensional distributions (f.d.d.).

b) (1.11) holds weakly on $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$ if and only if $\alpha = 2$.

Remark 1.6. For the classical $\{Y_t^{\delta_0, \Delta}, t \geq 0\}$, this is Theorem 8.1 in [DH82]. Combining Proposition 1.1 and Theorem 1.5, we recover and extend Theorem 1 in [FW06]. This in particular complements Part (b) of Theorem 1 in [FW06] by clarifying that tightness on path space holds *only* in the Brownian case. Our proof as well as our intuition for Part b) rely heavily on Donnelly & Kurtz’ lookdown construction, [DK99], circumventing moment calculations as in [FW06], Sect. 4.4. \square

It is interesting to see why tightness on pathspace can fail. Consider a path $\omega = \{\omega_t, t \geq 0\}$ in $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$. Let us say that ω exhibits an ε - δ -flicker (on the interval $[0, T]$) if there exist time points $0 < t_1 < t_2 < t_3 \leq T$ and $x, y \in \mathbb{R}^d$ such that $t_3 - t_1 \leq \delta$, $|x - y| \geq 2\varepsilon$ and

$$d_{\mathcal{M}_1}(\omega_{t_1}, \omega_{t_3}) \leq \varepsilon, \quad d_{\mathcal{M}_1}(\omega_{t_1}, \omega_{t_2}) \quad \text{and} \quad d_{\mathcal{M}_1}(\omega_{t_2}, \omega_{t_3}) \geq 2\varepsilon, \quad (1.12)$$

where $d_{\mathcal{M}_1}$ denotes the metric (1.4) on $\mathcal{M}_1(\mathbb{R}^d)$.

Lemma 1.7. *If $\alpha < 2$ and $\Lambda((0, 1]) > 0$, there exists $\varepsilon > 0$ such that*

$$\liminf_{k \rightarrow \infty} \mathbb{P}\{Y^{\Lambda, \Delta_\alpha}[k] \text{ exhibits an } \varepsilon\text{-}(1/k)\text{-flicker in } [0, T]\} > 0.$$

We will see below that the behaviour described by condition (1.12) arises as follows: At times t_1 and t_3 , $Y^{\Lambda, \Delta_\alpha}[k]$ is (almost) concentrated in a small ball with (random) centre x , say. At time t_2 , suddenly a fraction ε of the total mass appears in a remote ball with centre y , where $|x - y| \geq 1$, and vanishes almost instantaneously, i.e., by time t_3 . Such “sparks” make $Y^{\Lambda, \Delta_\alpha}[k]$ a process of “flickering random measures”. Technically, we see that Lemma 1.7 shows that the modulus of continuity $w'(\cdot, \delta, T)$ of the processes $Y^{\Lambda, \Delta_\alpha}[k]$, see (3.4) below, does not become small as $\delta \rightarrow 0$, contradicting tightness in $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$. Intuitively, at each infinitesimal “spark”, a

limiting process is neither left- nor right-continuous. We will see below how this intuition can be made precise in the framework of the (modified) lockdown construction.

The situation is different if $\Lambda = c\delta_0$ (and $\alpha < 2$). Here, each $Y^{c\delta_0, \Delta_\alpha}[k]$ a.s. has continuous paths, so that any limit in Skorohod's J_1 -topology would necessarily have continuous paths. However, the f.d.d. limit $\{\delta_{B_t^{(\alpha)}}, t \geq 0\}$ has no continuous modification. Intuitively, there is no “flickering”, but an “afterglow” effect: From time to time, a very fertile “infinitesimal” particle jumps some distance, and then finds an extremely large family, so that the population quickly becomes essentially a Dirac measure at this point, while at the same time the rest of the population (continuously) “fades away”. Note that this phenomenon is captured by Skorohod's M_1 -topology ([S56], Def. 2.2.5), which is tailor-made to establish convergence in situations in which a discontinuous process is approximated by a family of continuous processes. However, in the situation of Lemma 1.7, Condition (1.12) implies that the distributions of the processes $Y^{\Lambda, \Delta_\alpha}[k]$ cannot converge with respect to any of the topologies considered in [S56].

2 Donnelly and Kurtz' lockdown construction

2.1 A countable representation for generalised Fleming-Viot processes

We consider a countably infinite system of individuals, each particle being identified by a level $j \in \mathbb{N}$. We equip the levels with types ξ_t^j in \mathbb{R}^d , $j \in \mathbb{N}$. Initially, we require the types $\xi_0 = (\xi_0^j)_{j \in \mathbb{N}}$ to be an i.i.d. vector (in particular exchangeable), so that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \delta_{\xi_0^j} = \mu,$$

for some finite measure $\mu \in \mathcal{M}_1(\mathbb{R}^d)$, which will be the initial condition of the generalised Fleming-Viot process constructed below via (2.6). The point is that the construction will preserve exchangeability.

There are two “sets of ingredients” for the reproduction mechanism of these particles, one corresponding to the “finite variance” part $\Lambda(\{0\})$, and the other to the “extreme reproductive events” described by $\Lambda_0 = \Lambda - \Lambda(\{0\})\delta_0$. Restricted to the first N levels, the dynamics is that of a very particular permutation of a generalised Moran model with the property that always the particle with the highest level is the next to die.

For the first part, let $\{L_{ij}(t), t \geq 0\}$, $1 \leq i < j < \infty$, be independent Poisson processes with rate $\Lambda(\{0\})$. Intuitively, at jump times t of L_{ij} , the particle at level j “looks down” to level i and copies the type from there, corresponding to a single birth event in a(n approximating) Moran model. At jump times, types on levels above j are shifted accordingly, in formulas

$$\xi_t^k = \begin{cases} \xi_{t-}^k, & \text{if } k < j, \\ \xi_{t-}^i, & \text{if } k = j, \\ \xi_{t-}^{k-1}, & \text{if } k > j, \end{cases} \quad (2.1)$$

if $\Delta L_{ij}(t) = 1$. This mechanism is well defined because for each k , there are only finitely many processes L_{ij} , $i < j \leq k$ at whose jump times ξ^k has to be modified.

For the second part, which corresponds to multiple birth events, let n be a Poisson point process on $\mathbb{R}^+ \times (0, 1] \times [0, 1]^{\mathbb{N}}$ with intensity measure $dt \otimes r^{-2}\Lambda_0(dr) \otimes (du)^{\otimes \mathbb{N}}$. Note that for almost all realisations $\{(t_i, y_i, (u_{ij}))\}$ of n , we have

$$\sum_{i: t_i \leq t} y_i^2 < \infty \quad \text{for all } t \geq 0. \quad (2.2)$$

The jump times t_i in our point configuration n correspond to reproduction events. Define for $J \subset \{1, \dots, l\}$ with $|J| \geq 2$,

$$L_J^l(t) := \sum_{i: t_i \leq t} \prod_{j \in J} \mathbf{1}_{u_{ij} \leq y_i} \prod_{j \in \{1, \dots, l\} - J} \mathbf{1}_{u_{ij} > y_i}. \quad (2.3)$$

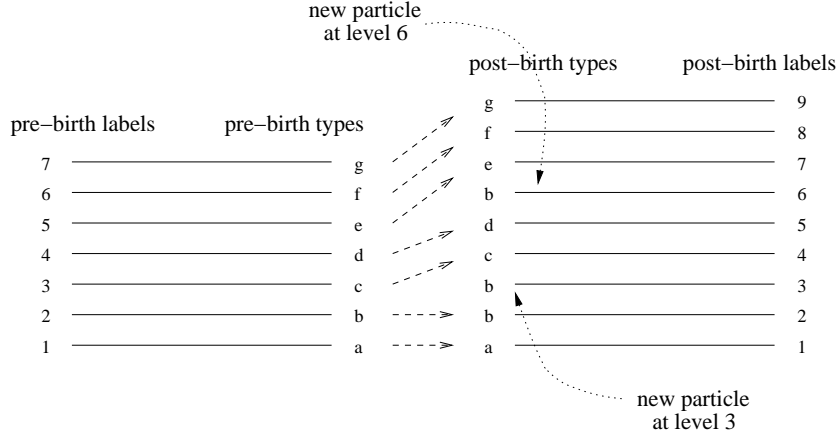


Figure 1: Relabelling after a birth event involving levels 2, 3 and 6.

$L_J^l(t)$ counts how many times, among the levels in $\{1, \dots, l\}$, exactly those in J were involved in a birth event up to time t . Note that for any configuration n satisfying (2.2), since $|J| \geq 2$, we have

$$\mathbb{E}[L_J^l(t) \mid n|_{[0,t] \times (0,1]}] = \sum_{i: t_i \leq t} y_i^{|J|} (1 - y_i)^{l-|J|} \leq \sum_{i: t_i \leq t} y_i^2 < \infty,$$

so that $L_J^l(t)$ is a.s. finite.

Intuitively, at a jump t_i , each level performs a uniform coin toss, and all the levels j with $u_{ij} \leq y_i$ participate in this birth event. Each participating level adopts the type of the smallest level involved. All the other individuals are shifted upwards accordingly, keeping their original order with respect to their levels (see Figure 1). More formally, if $t = t_i$ is a jump time and j is the smallest level involved, i.e. $u_{ij} \leq y_i$ and $u_{ik} > y_i$ for $k < j$, we put

$$\xi_t^k = \begin{cases} \xi_{t-}^k, & \text{for } k \leq j, \\ \xi_{t-}^j, & \text{for } k > j \text{ with } u_{ik} \leq y_i, \\ \xi_{t-}^{k-J_t^k}, & \text{otherwise,} \end{cases} \quad (2.4)$$

where $J_{t_i}^k = \#\{m < k : U_{im} \leq y_i\} - 1$. Let us define $\mathcal{G} = (\mathcal{G}_{u,v})_{u < v}$, where for $u \leq v$

$$\begin{aligned} \mathcal{G}_{u,v} = & \sigma(L_{ij}(t) - L_{ij}(s), u < s \leq t \leq v, i, j \in \mathbb{N}) \\ & \vee \sigma(n([t, s] \times A \times B), u < s \leq t \leq u, A \subset (0, 1], B \subset [0, 1]^\mathbb{N}) \end{aligned} \quad (2.5)$$

is the σ -algebra describing all “genealogical events” between times u and v .

So far, we have treated the reproductive mechanism of the particle system. Between reproduction events, all the levels follow independent α -stable motions. For a rigorous formulation, all three mechanisms together can be cast into a suitable countable system of stochastic differential equations driven by Poisson processes and α -stable processes, see [DK99], Section 6.

Then, for each $t > 0$, $(\xi_t^1, \xi_t^2, \dots)$ is an exchangeable random vector and

$$Z_t = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \delta_{\xi_t^j}, \quad t \geq 0 \quad (2.6)$$

exists almost surely on $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$, and $\{Z_t, t \geq 0\}$ is the Markov process with generator (1.2) and initial condition $Z_0 = \mu$, see [DK99], Thm. 3.2.

2.2 Pathwise embedding of Λ -coalescents in generalised Λ -Fleming-Viot processes

Note that for each $t > 0$ and $s \leq t$, the modified lookdown construction gives rise to the ancestral partition of the levels at time t with respect to the ancestors at time s before t by describing

$$N_i^t(s) = \text{level of level } i\text{'s ancestor at time } t - s.$$

For fixed t , the vector-valued process $\{N_i^t(s) : i \in \mathbb{N}\}_{0 \leq s \leq t}$ satisfies an “obvious” system of Poisson-process driven stochastic differential equations, see [DK99], p. 195, (note that we have indulged in a time re-parametrisation), and the partition-valued process defined by

$$\{\{i : N_i^t(s) = j\}, j = 1, 2, \dots\} \quad (2.7)$$

is a standard Λ -coalescent with time interval $[0, t]$. This implies in particular by Kingman’s theory of exchangeable partitions, [K82], see e.g. [Pi06] for an introduction, that the empirical family sizes

$$A_j^t(s) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{N_i^t(s)=j\}}$$

exist a.s. in $[0, 1]$ for each j and $s \leq t$, describing the relative frequency at time t of descendants of the particle at level j at time $t - s$.

3 Proof of Theorem 1.5

Fix $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ as the initial condition of the unscaled process $Y^{\Lambda, \Delta_\alpha}$. We begin with the useful observation that, due to the scaling properties of the underlying motion process, for each k , the process $\{Y_t^{(k)}, t \geq 0\}$, defined by

$$Y_t^{(k)} = Y_t^{k\Lambda, \Delta_\alpha}, \quad t \geq 0, \quad (3.1)$$

(and starting from the image measure of μ under $x \mapsto x/k^{1/\alpha}$), has the same distribution as $\{Y_t^{\Lambda, \Delta_\alpha}[k]\}$ defined in (1.10). It will be convenient to work in the following with a version of $Y^{(k)}$ which is obtained from a lookdown construction with “parameter” $k\Lambda$, in particular, we have

$$Y_t^{(k)} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \delta_{\xi_t^i}, \quad t \geq 0.$$

Note that the family ξ^i , $i \in \mathbb{N}$, used to construct $Y^{(k)}$ depends (implicitly) on k , but for the sake of readability, we suppress this in our notation.

Proof of Part a)

We have already noted that for $\Lambda = \delta_0$ and $\alpha = 2$, this is Theorem 8.1 in [DH82], and that, for $\Lambda = U = \text{Beta}(1, 1)$, the uniform distribution on $[0, 1]$, this is essentially Theorem 1 in [FW06], see Remark 1.6. Using Remark 1.2, the proof of Fleischmann and Wachtel can easily be adapted, as it relies only on the first two moments.

Alternatively, since the motion of the level-1 particle $\{\xi_t^1, t \geq 0\}$ is a symmetric α -stable process, it suffices to check that

$$\lim_{k \rightarrow \infty} \mathbb{P}\{Y_t^{(k)}(B_{\xi_t^1}(\varepsilon)) \geq 1 - \varepsilon\} = 1.$$

for each t and ε , which will be implied by

$$\lim_{k \rightarrow \infty} \mathbb{E}[Y_t^{(k)}(B_{\xi_t^1}(\varepsilon)^c)] = 0 \quad \text{for each } \varepsilon > 0. \quad (3.2)$$

In order to check this, let Φ_ε be a ‘‘mollified’’ (continuous) indicator of $B_\varepsilon(\xi_t^1)^c$, and note, by dominated convergence, that for any $\delta > 0$

$$\begin{aligned}\mathbb{E}[\langle \Phi_\varepsilon, Y_t^{(k)} \rangle] &= \lim_{N \rightarrow \infty} \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \Phi_\varepsilon(\xi_t^i) \right] \\ &\leq \limsup_{N \rightarrow \infty} \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \Phi_\varepsilon(\xi_t^i) \mathbf{1}_{\{N_i^t(\delta)=1\}} \right] + \mathbb{E}[1 - A_1^t(\delta)].\end{aligned}$$

The second term in the last line, for each $\delta > 0$, converges to 0 as $k \rightarrow \infty$, cf. [Pi99], Prop. 30. For the first term note that, where $\mathcal{G}_{t-\delta, t}$ describes the genealogical information as defined in (2.5),

$$\begin{aligned}\mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \Phi_\varepsilon(\xi_t^i) \mathbf{1}_{\{N_i^t(\delta)=1\}} \right] &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\mathbf{1}_{\{N_i^t(\delta)=1\}} \mathbb{E}[\Phi_\varepsilon(\xi_t^i) | \mathcal{G}_{t-\delta, t}] \right] \\ &\leq \mathbb{E} \left[\int \Phi_\varepsilon(y) p_\delta^{(\alpha)}(\xi_{t-\delta}^1, y) dy \right] \\ &\leq \mathbb{P}\{|\xi_t^1 - \xi_{t-\delta}^1| \geq \varepsilon/2\} + p_\delta^{(\alpha)}(0, B_0(\varepsilon/2)^c),\end{aligned}$$

which for fixed ε tends to 0 as $\delta \rightarrow 0$.

3.1 Proof of Part b)

In the case $\alpha = 2$, using Remark 1.2, tightness on the space $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$ can be proved by inspection, literally tracing through the corresponding arguments of [FW06], Lemma 20 and 21 (note that even though Equations (133)–(137) in [FW06] estimate a fourth moment, this refers only to an increment of a d -dimensional Brownian motion).

For the case $\alpha < 2$, let us recall the following classical characterisation of relative compactness in $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$, cf. e.g. [Bi68], Theorem 15.2.

Theorem 3.1 (Relative compactness on path space). *Let $\{Y^k\}$ be a sequence of processes taking values in $D_{[0, \infty)}(\mathcal{M}_1(\mathbb{R}^d))$. Then $\{Y^k\}$ is relatively compact if and only if the following two conditions hold.*

- For every $\varepsilon > 0$ and every (rational) $t \geq 0$, there exists a compact set $\gamma_{\varepsilon, t} \subset \mathcal{M}_1(\mathbb{R}^d)$, such that

$$\liminf_{k \rightarrow \infty} \mathbb{P} \{ Y_t^k \in \gamma_{\varepsilon, t} \} \geq 1 - \varepsilon.$$

- For every $\varepsilon > 0$ and $T > 0$, there exists $\delta > 0$, such that

$$\limsup_{k \rightarrow \infty} \mathbb{P} \{ w'(Y^k, \delta, T) \geq \varepsilon \} \leq \varepsilon, \quad (3.3)$$

where

$$w'(y, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s, t \in [t_{i-1}, t_i)} d(y(s), y(t)), \quad (3.4)$$

and $\{t_i\}$ ranges over all finite partitions of $[0, T]$ with $t_i - t_{i-1} > \delta$ for all i .

Then we obtain from Lemma 1.7 a $\varepsilon > 0$ such that for $k_0 \in \mathbb{N}$ and $\delta > 1/k_0$

$$\mathbb{P} \{ w'(Y^{\Lambda, \delta_\alpha}[k], \delta, T) \geq \varepsilon \} \geq \mathbb{P} \{ Y^{\Lambda, \Delta_\alpha}[k] \text{ exhibits an } \varepsilon\text{-}(1/k)\text{-flicker on } [0, T] \}$$

is bounded away from 0 uniformly in $k \geq k_0$. □

3.2 Proof of Lemma 1.7 The intuitive mechanism behind a “flicker” obtained from the lockdown construction is as follows: Typically when k is large, most of the total mass of $Y^{(k)}$ as defined in (3.1) will be in the immediate vicinity of the location of the level-1 particle. A “flicker” arises if the level-2 particle jumps to a remote position and shortly afterwards participates in an extreme reproduction event involving a positive fraction of the current population, but not the level-1 particle. In this situation, a new atom appears in the support of $Y^{(k)}$, which is then removed very quickly, since mass is attracted rapidly towards the position of the level-1 particle. Note that corresponding phenomena will occur on any level $j \geq 2$.

A technical obstacle to turn this intuition into a rigorous proof stems from the fact that the metric d on $\mathcal{M}_1(\mathbb{R}^d)$, inducing the weak topology, is insensitive to such flickers if they occur far away. Hence, in what follows, we require the level-1 particle to stay within a fixed ball around the origin. This forces us to disentangle $\sigma\{\xi_t^1, t \geq 0\}$ and the information about the genealogy and the increments of the other particles relative to the position of ξ^1 at the time of their respective most recent common ancestor.

First, we collect some useful notation. Without loss of generality assume $T = 1$, choose $\delta \in (0, 1]$ with $\Lambda((\delta, 1]) > 0$ and $\varepsilon > 0$ such that for any $\mu, \mu' \in \mathcal{M}_1(\mathbb{R}^d)$,

$$\mu(B_0(1)) \geq 1 - \delta/2 \text{ and } \mu'(B_0(2)^c) \geq \delta \text{ implies } d_{\mathcal{M}_1}(\mu, \mu') > \varepsilon. \quad (3.5)$$

For $k \in \mathbb{N}$, we split the time interval $[0, 1]$ into k disjoint intervals $(a_i, a_{i+1}]$, where $a_i = i/k$, $i = 0, \dots, k-1$. Moreover, we define $b_i = a_i + 1/(4k)$, $c_i = a_i + 2/(4k)$, $d_i = a_i + 3/(4k)$. Let

$$\sigma_j^t := \inf\{s > 0 : N_j^t(s) = 1\}$$

(with the usual convention $\inf \emptyset = +\infty$) be the backwards time to the most recent common ancestor of the particles at level j and at level 1 at time t , and let

$$H_{s,t} := \{L_{12}(t) - L_{12}(s) = 0\} \cap \left\{ n((s, t] \times \{(x, (u_m)) \in (0, 1] \times [0, 1]^{\mathbb{N}} : u_1, u_2 \leq x\}) = 0 \right\} \quad (3.6)$$

be the event that in the time interval $(s, t]$, no lockdown event involving both levels 1 and 2 occurs. Furthermore, let (δ_k) be such that

$$\lim_{k \rightarrow \infty} \delta_k = 0 \text{ and } \lim_{k \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq 1/k} |B_t^{(\alpha)}| \leq \delta_k \right\} = 1. \quad (3.7)$$

In order to cook up a “flicker” within $(a_i, a_{i+1}]$, we collect the following “ingredients”:

- Within the time-interval $(a_i, b_i]$, consider the event $\mathcal{A}_i^{(k)}$ that at time b_i most of the population (including the level-2 particle) is sufficiently closely related to the level-1 particle and has not moved too far away, more precisely

$$\begin{aligned} \mathcal{A}_i^{(k)} := & \left\{ A_1^{b_i}(1/(4k)) \geq 1 - \frac{\delta}{4} \right\} \cap \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{N_j^{b_i}(1/(4k))=1\}} \mathbf{1}_{\{|\xi_{b_i}^j - \xi_{b_i - \sigma_j^{b_i}}^1| \leq \delta_k\}} \geq 1 - \delta/2 \right\} \\ & \cap \left\{ \sigma_2^{b_i} < 1/(4k) \right\} \cap \left\{ |\xi_{b_i}^2 - \xi_{b_i - \sigma_2^{b_i}}^2| \leq 1/2 \right\}. \end{aligned}$$

- Within the time-interval $(b_i, c_i]$, the event $\mathcal{B}_i^{(k)}$ requires that the level-2 particle jumps to a sufficiently remote position and there is no subsequent lockdown-event involving level-1 and level-2, more precisely,

$$\mathcal{B}_i^{(k)} := H_{b_i, c_i} \cap \left\{ |\xi_{c_i}^2 - \xi_{b_i}^2| > 4 \right\}.$$

- Within the time-interval $(c_i, d_i]$, the event $\mathcal{C}_i^{(k)}$ requires that the level-2 particle does not travel very far, and that there is a lockdown event involving a sufficiently large fraction of the population, but *not* the level-1 particle:

$$\begin{aligned} \mathcal{C}_i^{(k)} := & H_{c_i, d_i} \cap \left\{ \sup_{t \in (c_i, d_i]} |\xi_t^2 - \xi_{c_i}^2| < 1 \right\} \\ & \cap \left\{ n([c_i, d_i] \times \{(x, (u_m)) \in (0, 1] \times [0, 1]^{\mathbb{N}} : x > \delta, u_2 < x \leq u_1\}) \geq 1 \right\}. \end{aligned}$$

- Finally, let $\mathcal{D}_i^{(k)}$ be the event that most of the mass returns to the location of the level-1 particle, and stays there, (which essentially is the same behaviour as within $(a_i, b_i]$), namely,

$$\mathcal{D}_i^{(k)} := \left\{ A_1^{a_{i+1}}(1/(4k)) \geq 1 - \frac{\delta}{4} \right\} \\ \cap \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{N_j^{a_{i+1}}(1/(4k))=1\}} \mathbf{1}_{\{|\xi_{a_{i+1}}^j - \xi_{a_{i+1}-\sigma_j^{a_{i+1}}}^1| \leq \delta_k\}} \geq 1 - \delta/2 \right\}.$$

Now let us introduce a family of σ -algebras containing our ingredients: Recall $\mathcal{G}_{u,v}$ from (2.5) and let $\mathcal{H}_i^{(k)}$ be the σ -algebra generated by $\mathcal{G}_{a_i, a_{i+1}}$ and the random variables

$$\left(\xi_{b_i}^j - \xi_{b_i - \sigma_j^{b_i}}^1 \right) \mathbf{1}_{\{\sigma_j^{b_i} \leq 1/(4k)\}}, \quad \left(\xi_{a_{i+1}}^j - \xi_{a_{i+1} - \sigma_j^{a_{i+1}}}^1 \right) \mathbf{1}_{\{\sigma_j^{a_{i+1}} \leq 1/(4k)\}}, \quad j = 2, 3, \dots, \quad \text{and} \\ \left(\xi_t^2 - \xi_{b_i}^2 \right) \mathbf{1}_{H_{b_i, d_i}}, \quad b_i \leq t \leq d_i.$$

Note that for fixed k , the family $\mathcal{H}_i^{(k)}$, $i = 0, 1, \dots, k-1$ is independent and independent of $\sigma\{\xi_t^1, t \geq 0\}$, and

$$\mathcal{A}_i^{(k)}, \mathcal{B}_i^{(k)}, \mathcal{C}_i^{(k)}, \mathcal{D}_i^{(k)} \in \mathcal{H}_i^{(k)}, \quad i = 0, 1, \dots, k-1.$$

On the event

$$\mathcal{E}_i^{(k)} := \left\{ \sup_{t \in (a_i, a_{i+1}]} |\xi_t^1 - \xi_{a_i}^1| \leq \delta_k \right\} \cap \mathcal{A}_i^{(k)} \cap \mathcal{B}_i^{(k)} \cap \mathcal{C}_i^{(k)} \cap \mathcal{D}_i^{(k)}, \quad (3.8)$$

we see from (3.5) that there is a (random) time $\tau \in (c_i, d_i]$ such that

$$d_{\mathcal{M}_1}(Y_{b_i}^{\Lambda, \Delta_\alpha}[k], Y_{a_{i+1}}^{\Lambda, \Delta_\alpha}[k]) \leq \varepsilon, \quad d_{\mathcal{M}_1}(Y_{b_i}^{\Lambda, \Delta_\alpha}[k], Y_\tau^{\Lambda, \Delta_\alpha}[k]) \quad \text{and} \quad d_{\mathcal{M}_1}(Y_\tau^{\Lambda, \Delta_\alpha}[k], Y_{a_{i+1}}^{\Lambda, \Delta_\alpha}[k]) \geq 2\varepsilon, \quad (3.9)$$

i.e. $Y^{\Lambda, \Delta_\alpha}[k]$ exhibits an ε - $(1/k)$ -flicker in $(a_i, a_{i+1}]$. It is easy to see that

$$\inf_{k \in \mathbb{N}} \mathbb{P} \left(\bigcup_{i=0}^{k-1} \mathcal{E}_i^{(k)} \mid \left\{ \sup_{0 \leq t \leq 1} |\xi_t^1| \leq 1/2 \right\} \right) > 0, \quad (3.10)$$

which yields the claim. In order to verify (3.10), note that

$$\forall k, i < k : \mathbb{P}(\mathcal{A}_i^{(k)} \cap \mathcal{B}_i^{(k)} \cap \mathcal{C}_i^{(k)} \cap \mathcal{D}_i^{(k)}) \geq C/k$$

for some $C = C(\alpha, \Lambda, \delta, (\delta_m)) > 0$, which basically comes from the fact that

$$\mathbb{P}\{|B_{1/(4k)}^{(\alpha)}| > 4\} \sim \text{Const.} \times \frac{1}{k}.$$

Furthermore, let

$$I_k := \left\{ i \in \{0, 1, \dots, k-1\} : \sup_{t \in (a_i, a_{i+1}]} |\xi_t^1 - \xi_{a_i}^1| \leq \delta_k \right\}$$

and observe that for each k , I_k is independent of $\bigvee_{i=0}^{k-1} \mathcal{H}_i^{(k)}$ and we have

$$\inf_{k \in \mathbb{N}} \mathbb{P} \left(\{|I_k| \geq k/2\} \mid \left\{ \sup_{0 \leq t \leq 1} |\xi_t^1| \leq 1/2 \right\} \right) > 0.$$

References

- [BLG03] BERTOIN, J.; LE GALL, J. F.: Stochastic flows associated to coalescent processes, *Probab. Theory Related Fields* **126** (2003), no. 2, 261–288.
- [Bi68] BILLINGSLEY, P.: *Convergence of probability measures*, Wiley (1968).
- [BB07] BIRKNER, M.; BLATH, J.: Measure-valued diffusions, general coalescents and population genetic inference, to appear in: *Trends in Stochastic Analysis*, Cambridge University Press, (2007).
- [BBC05] BIRKNER, M.; BLATH, J.; CAPALDO, M.; ETHERIDGE, A.; MÖHLE, M.; SCHWEINSBERG, J.; WAKOLBINGER, A.: α -stable branching and β -coalescents, *Electronic Journal of Probability* **10**, Paper no. 9, 303–325, (2005).
- [Da93] DAWSON, D.: Measure-valued Markov processes, École d’été de Probabilités de Saint Flour XXI, Lecture Notes in Mathematics **1541** pp 1–260, Springer-Verlag, 1993.
- [DH82] DAWSON, D.; HOCHBERG, K.: Wandering random measures in the Fleming-Viot Model. *Ann. Probab.* **10**, no. 3, 554–580, (1982).
- [DK96] DONNELLY, P.; KURTZ, T.: A countable representation of the Fleming-Viot measure-valued diffusion, *Ann. Appl. Probab.* **24**, 698–742, (1996).
- [DK99] DONNELLY, P.; KURTZ, T.: Particle representations for measure-valued population models, *Ann. Probab.* **27**, no. 1, 166–205, (1999).
- [EKR91] EL KAROUI, N.; RÖLLY, S.: Propriétés de martingales, explosion et représentation de Lévy-Khintchine d’une classe de processus de branchement a valeurs mesures, *Stoch. Proc. Appl.* **38**, 239–266, (1991).
- [Eth00] ETHERIDGE, A.: *An Introduction to Superprocesses*. AMS University Lecture Series, Vol. 20, (2000).
- [EM91] ETHERIDGE, A.; MARCH, P.: A note on superprocesses, *Probab. Theory Rel. Fields* **89**, 141–148, (1991).
- [EK86] ETHIER, S.; KURTZ, T.: *Markov Processes: Characterization and Convergence*, Wiley, New York, (1986).
- [Fe66] FELLER, W.: *Introduction to Probability Theory and its Applications*, Volume II, Wiley, (1966).
- [FS04] FLEISCHMANN, K.; STURM, A.: A super-stable motion with infinite mean branching, *Ann. Inst. H. Poincaré Probab. Statist.* **40**, no. 5, 513–537, (2004).
- [FW06] FLEISCHMANN, K.; WACHTEL, V.: Large scale localization of a spatial version of Neveu’s branching process. *Stochastic Process. Appl.* **116**, no. 7, 983–1011, (2006).
- [FV79] FLEMING, W.; VIOT, M.: Some measure-valued markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843, (1979).
- [K82] KINGMAN, J. F. C.: The coalescent. *Stoch. Proc. Appl.* **13**, 235–248, (1982).
- [MS01] MÖHLE, M.; SAGITOV, S.: A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562, (2001).
- [Pe91] PERKINS, E.: Conditional Dawson-Watanabe processes and Fleming-Viot processes, Seminar in Stochastic Processes, Birkhäuser, pp 142–155, (1991).
- [Pi99] PITMAN, J.: Coalescents with multiple collisions, *Ann. Probab.* **27**, no. 4, 1870–1902, (1999).
- [Pi06] PITMAN, J.: *Combinatorial stochastic processes. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. Lecture Notes in Mathematics*, **1875**. Springer-Verlag, (2006).
- [Pr56] PROHOROV, Y.V.: Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1**, 157–214, (1956).
- [Sa99] SAGITOV, S.: The general coalescent with asynchronous mergers of ancestral lines, *J. Appl. Probab.* **26**, 1116–1125 (1999).

- [Sc00] SCHWEINSBERG, J.: A necessary and sufficient condition for the Λ -coalescent to come down from infinity. *Electron. Comm. Probab.* **5**, 1–11, (2000).
- [S56] SKOROHOD, A.V.: Limit theorems for stochastic processes. *Theory Probab. Appl.* **1**, 261–290, (1956).
- [Y65] YOSIDA, K.: *Functional Analysis*, Grundlehren der mathematischen Wissenschaften **123**, Springer-Verlag, (1965).