DOI: 10.1214/EJP.v14-712

Electronic Journal of Probability

Vol. 14 (2009), Paper no. 84, pages 2418–2437.

Journal URL
http://www.math.washington.edu/~ejpecp/

Generalised stable Fleming-Viot processes as flickering random measures

Matthias Birkner*, Jochen Blath†

Abstract

We study some remarkable path-properties of generalised stable Fleming-Viot processes (including the so-called spatial Neveu superprocess), inspired by the notion of a "wandering random measure" due to Dawson and Hochberg (1982). In particular, we make use of Donnelly and Kurtz' (1999) modified lookdown construction to analyse their longterm scaling properties, exhibiting a rare natural example of a scaling family of probability laws converging in f.d.d. sense, but not weakly w.r.t. any of Skorohod's topologies on path space. This phenomenon can be explicitly described and intuitively understood in terms of "sparks", leading to the concept of a "flickering random measure".

In particular, this completes results of Fleischmann and Wachtel (2006) about the spatial Neveu process and complements results of Dawson and Hochberg (1982) about the classical Fleming Viot process.

Key words: Generalised Fleming-Viot process, flickering random measure, measure-valued diffusion, lookdown construction, wandering random measure, Neveu superprocess, path properties, tightness, Skorohod topology.

AMS 2000 Subject Classification: Primary 60G57; Secondary: 60G17.

Submitted to EJP on September 16, 2008, final version accepted November 2, 2009.

*Institut für Mathematik, Johannes-Gutenberg-Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany. E-mail: birkner@mathematik.uni-mainz.de

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

1 Introduction and statement of the main results

1.1 Classical and generalised 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

$$F(\rho) := \prod_{i=1}^{n} \langle \phi_i, \rho \rangle, \quad \rho \in \mathcal{M}_1(\mathbb{R}^d), \tag{1.1}$$

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

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

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 sense (our description being rather informal). 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' *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.

More general limit population processes describing situations where, from time to time, a single individual produces a non-negligible fraction of the total population have been introduced in [DK99] (see also [BLG03] for a different approach). We follow [BLG03] in calling such processes *generalised Fleming-Viot processes*. The limits of their dual genealogical processes have been classified in [Sa99], [MS01]. See [BB09] for an overview. Generalised Fleming-Viot processes are probability measure valued Markov processes $Y^{\Lambda,\Delta_\alpha}$ whose generator acts on functions F of the form (1.1) with ϕ_i in the domain of Δ_α as

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

where

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

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.

Remark 1.1. *The special form (1.2) which the generalized Fleming-Viot generator takes when acting on functions of type (1.1) highlights its connection to the corresponding dual coalescent processes. Note that (1.2) has been derived e.g. in the proof of Theorem 3 in [BLG03], characterizing the Λ-Fleming-Viot process as solution to a well-posed martingale problem (which implies the strong Markov property); see also [DK99, Thm. 4.3]. For the 'general form' of the generator of the Λ-Fleming-Viot process and its construction as flow of bridges see Section 5.1 and (16) in [BLG03], or, alternatively, the explicit construction via particle systems in [DK99] or [BBM+09], where the latter reference also provides a classical construction via the Hille-Yosida Theorem.*

We endow $\mathcal{M}_1(\mathbb{R}^d)$ with the topology of weak convergence, which we think of being induced by the Prohorov metric $d_{\mathcal{M}_1}$, defined for $\mu, \nu \in \mathcal{M}_1(\mathbb{R}^d)$ by

$$d_{\mathcal{M}_1}(\mu,\nu) := \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B^\varepsilon) + \varepsilon \text{ for all closed } B \subset \mathbb{R}^d\}, \tag{1.4}$$

where B^ε is the usual open ε-enlargement of the set $B \subset \mathbb{R}^d$. It is well known that $d_{\mathcal{M}_1}$ is a complete metric on $\mathcal{M}_1(\mathbb{R}^d)$, cf. e.g., [EK86], Thm. 3.1.7 and Thm. 3.3.1.

By [DK99, Theorem 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], Chapter 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]). Indeed, for $t \geq 0$, while Π_t^Λ has n classes, any k-tuple merges to one at rate $\lambda_{n,k}$. A Λ-Fleming-Viot process is dual to a Λ-coalescent (as shown in [DK99], pp. 195 and [BLG03]), similar to the duality between the classical Fleming-Viot process and Kingman's coalescent established in [DH82]. Note that Kingman's coalescent corresponds to the choice $\Lambda = \delta_0$.

1.2 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 [BBC+05], 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

$$\int_0^t \langle 1, Z_t \rangle^{1-\beta}\, dt, \tag{1.5}$$

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-inhomogeneous 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 1.2 (Normalised spatial Neveu branching process as generalised Fleming-Viot process)**.** *Under the above conditions, we have*

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

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 [BBC+05], it is the only "superprocess" with this property. This observation was the starting point of our investigation.

By considering F as in (1.1) with $n = 1$ resp. $n = 2$, 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 $\Lambda([0,1])$.

Proposition 1.3 (First and second moment measure)**.** *Let* $Y_0 = \mu \in \mathscr{M}_f \setminus \{0\}$. *Then,*

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

and for $t_1 \leq t_2$, *writing* $\rho := \Lambda([0,1])$, *and* $\varphi, \varphi_1, \varphi_2 \in C_b^2$,

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

In particular, the first two moment measures agree with those of the classical Fleming-Viot process, which explains Proposition 3 in [FW06]. Note that in order to establish (1.6) and (1.7), one cannot apply the Laplace-transform method as in [FW06] since the branching property does not hold in general. A proof can be found in Section 3.

Note that a simple explanation can be given in terms of the *dual coalescent process* mentioned before. Indeed, in [BLG03] it is shown that generalised Fleming-Viot processes are *moment dual* to Λ-coalescents (see also [BB09]). Since the first two moments do not involve multiple coalescent events, they cannot "feel" the finer properties of the measure Λ. Of course, for moments greater than two, the moment formulae cannot be expected to agree.

1.3 Coherent wandering random measures

In the terminology of [DH82], the classical Fleming Viot process $Y^{\delta_0,\Delta}$ is a (compactly) *coherent wandering random measure*, which means that there is a "centring process" $\{x(t), t \geq 0\}$ with values in $\mathbb{R}^d$ and for each $\varepsilon > 0$ a real-valued stationary "radius process" $\{R_\varepsilon(t), t \geq 0\}$ and an a.s. finite T_0, such that

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

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 so-called "level-1 particle" (see Section 2). With this choice, an obvious extension of [DK96], Theorem 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, i.e., almost surely,

$$\operatorname{supp}\big(Y_t^{\Lambda,\Delta_\alpha}\big) \quad \text{is compact for all } t,$$

this will also yield *compact coherence*, i.e. one can choose $\varepsilon = 0$ in (1.8), see [DH82], Theorem 7.2.

It is interesting to see that 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

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

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 pathwise 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 1.2, we recover Proposition 14 of [FS04] on the instant propagation of the spatial Neveu branching process.

Remark 1.4. Observe that for continuous test functions φ with compact support,

$$t^{d/\alpha}\mathbb{E}\Big[\langle \varphi, Y_t^{\Lambda,\Delta_\alpha}\rangle\Big] \;\to\; p_1^{(\alpha)}(0)\int \varphi(x)\,dx \quad \text{as } t \to \infty, \tag{1.9}$$

where $p_t^{(\alpha)}(x)$ is the transition density of $\{B_t^{(\alpha)}, t \geq 0\}$. This is essentially Corollary 6 of [FW06], which was formulated for $\Lambda = U$ only. 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 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 (again, as in [DK96], Theorem 2.9). 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).

1.4 Main results: Longterm-scaling, existence of sparks, and flickering random measures

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: Let $\Lambda \in \mathcal{M}_f([0,1])$ such that $\Lambda([0,1)) > 0$ and define the rescaled process $\{Y_t^{\Lambda,\Delta_\alpha}[k], t \geq 0\}$ via

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

for $\phi \in b\mathscr{B}(\mathbb{R}^d)$ and $t \geq 0$. Let $B^{(\alpha)}$, for $\alpha \in (0,2]$, be the standard symmetric stable process of index α, starting from $B_0^{(\alpha)} = 0$. With these definitions, one readily expects the following convergence of finite-dimensional distributions (f.d.d.) to hold:

Proposition 1.5 (Longterm-Scaling)**.** *For each finite collection of time-points $0 \leq t_1 < \cdots < t_n$, we have*

$$\big(Y_{t_1}^{\Lambda,\Delta_\alpha}[k], \dots, Y_{t_n}^{\Lambda,\Delta_\alpha}[k]\big) \Rightarrow \big(\delta_{B_{t_1}^{(\alpha)}}, \dots, \delta_{B_{t_n}^{(\alpha)}}\big) \quad \text{as} \quad k \to \infty. \tag{1.11}$$

Note that for the classical $\{Y_t^{\delta_0,\Delta}, t \geq 0\}$, this is essentially Theorem 8.1 in [DH82]. Combining Proposition 1.2 and Proposition 1.5, we recover Part (a) of Theorem 1 in [FW06].

In addition to f.d.d.-convergence, Part (b) of Theorem 1 in [FW06] provides weak convergence on $D_{[0,\infty)}(\mathcal{M}_1(\mathbb{R}^d))$ if the underlying motion of the spatial Neveu process (i.e. $\Lambda = U$) is Brownian. However, the question whether this holds in general seems to be inaccessible to the Laplace-transform and moment-based methods of [FW06], and therefore had been left open.

Rather surprisingly, it turns out that pathwise convergence does *not* hold if $\alpha < 2$, and that, with the help of Donnelly and Kurtz' modified lookdown construction ([DK99]), it is possible to understand explicitly "what goes wrong". To this end, we introduce the concept of "sparks" and of a family of "flickering random measures".

Definition 1.6 (Sparks)**.** *Consider a path $\omega = \{\omega_t, t \geq 0\}$ in $D_{[0,\infty)}(\mathcal{M}_1(\mathbb{R}^d))$. We say that ω exhibits an ε-δ-*spark *(on the interval $[0,T]$) if there exist time points $0 < t_1 < t_2 < t_3 \leq T$ such that $t_3 - t_1 \leq \delta$*

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

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

Definition 1.7 (Flickering random measures)**.** *Let $\{Z[k], k \in \mathbb{N}\}$ be a family of measure-valued processes on $D_{[0,\infty)}(\mathcal{M}_1(\mathbb{R}^d))$. If there exists an $\varepsilon > 0$ and a sequence $\delta_k \downarrow 0$, such that*

$$\liminf_{k\to\infty} \mathbb{P}\{Z[k] \text{ exhibits an } \varepsilon\text{-}\delta_k\text{-spark in } [0,T]\} > 0,$$

then we say that $\{Z[k], k \in \mathbb{N}\}$ is a family of "flickering random measures".

The space-time scaling family of many generalised Fleming-Viot processes satisfies this definition:

Lemma 1.8 (Generalised Fleming-Viot processes as flickering random measures)**.** *If* $\alpha < 2$ *and* $\Lambda((0,1)) > 0$*, there exists* $\varepsilon > 0$ *such that*

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

Hence, the scaling family $\{Y^{\Lambda,\Delta_\alpha}[k], k \geq 1\}$ *is a family of* "flickering random measures" *with* $\delta_k = 1/k, k \in N$.

We will see below that the behaviour of $\{Y^{\Lambda,\Delta_\alpha}[k]$, leading to an ε–$1/k$-spark described by condition (1.12) typically 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.

Technically, we see that Lemma 1.8 shows that the modulus of continuity $w'(\cdot,\delta,T)$ of the processes $Y^{\Lambda,\Delta_\alpha}[k]$, see (3.5) below, does not become small as $\delta \to 0$, contradicting relative compactness of distributions on $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$. 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) lookdown construction.

The situation is different if $\Lambda = c\delta_0$ for some $c > 0$ 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 founds 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".

To complete the picture, we are finally able to provide the full classification of the scaling behaviour of generalised α-stable Fleming-Viot processes.

Theorem 1.9 (Convergence on path space)**.** *If* $\Lambda([0,1)) > 0$*, (1.11) holds weakly on* $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$ *if and only if* $\alpha = 2$.

Remark 1.10 (Other Skorohod topologies)**.** Note that the above-mentioned "afterglow"-phenomenon in the case $\Lambda = c\delta_0$ and $\alpha < 2$ fits well to Skorohod's M_1-topology (see [S56], Definition 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 1.8, 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]. We are not aware of a reasonable topology $\mathscr{T}$ on $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$ such that (the distribution of) $\{Y^{\Lambda,\Delta_\alpha}[k](t)\}$ converges weakly towards (the distribution of) $\{\delta_{B_t^{(\alpha)}}\}$ on $(D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d)),\mathscr{T})$.

Remark 1.11 (The case of star-shaped genealogies)**.** Note that the case $\Lambda = c\delta_1, c > 0$, has been excluded from the setup of this subsection. Here, occasionally (i.e. with rate c), the whole population always jumps to the position of the level-1-particle, producing a star-shaped genealogy. Inbetween, the α-heat flow acts so that the process is continuous between consecutive jumps. Hence, it is clear that $Y^{c\delta_1,\Delta_\alpha}[k]$ converges weakly to the stable unit motion δ_{B^α}, approximated by the rescaled path of the level-1-particle.

2 Donnelly and Kurtz' lookdown 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

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

for some probability measure $\mu \in \mathscr{M}_1(\mathbb{R}^d)$, 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 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. Let $\Delta L_{ij}(t) = L_{ij}(t) - L_{ij}(t-)$. At jump times, types on levels above j are shifted accordingly, in formulas

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

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

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

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

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

$L_J^l(t)$ counts how many times, among the levels in $\{1, \ldots, l\}$, 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

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

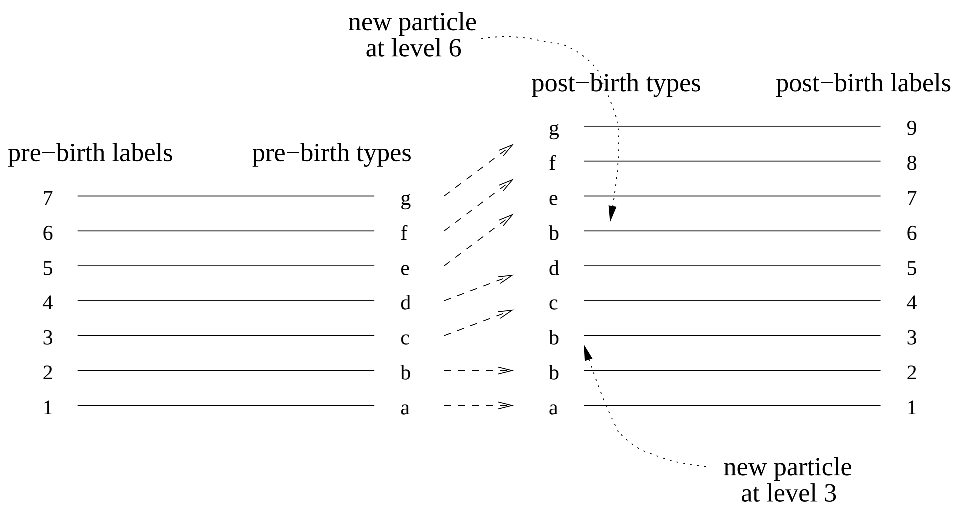

Figure 1: Relabelling after a birth event involving levels 2, 3 and 6.

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 $u_{ij} \le y_i$ 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. $u_{ij} \le y_i$ and $u_{ik} > y_i$ for $k < j$, we put

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

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

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

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

So far, we have only treated the reproductive mechanism of the particle system. In-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

$$Z_t = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \delta_{\xi_t^j}, \quad t \ge 0 \tag{2.6}$$

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

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

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

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

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

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

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

$$A_j^t(s) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{N_i^t(s)=j\}} \tag{2.8}$$

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 Proofs

Fix $\mu \in \mathscr{M}_1(\mathbb{R}^d)$ as the initial condition of the un-scaled 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

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

starting from the image measure of μ under $x \mapsto x/k^{1/\alpha}$, has the same distribution as $\{Y_t^{\Lambda,\Delta_\alpha}[k], t \geq 0\}$ 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

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

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.

3.1 Proof of Proposition 1.5

We have already noted that for $\Lambda = \delta_0$ and $\alpha = 2$, the statement of Proposition 1.5 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].

Instead of following the arguments of [FW06, Lemma 19] (which only make use of the formulas for the first two moments and thus in view of Prop. 1.3 could be adapted as well), we give a proof which is directly based on the lookdown-construction. Indeed, we show that the path of the unit mass $\{\delta_{B_t^{(\alpha)}}, t \geq 0\}$ can be viewed as the trail of the level-1-particle. To this end, we first show convergence in law of the one-dimensional distributions, i.e.

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

Since the motion of the level-1-particle $\{\xi_t^1, t \geq 0\}$ is a symmetric α-stable process, i.e. $B_t^{(\alpha)} =^d \xi_t^1$, it suffices by the special form of the limit variable to check that

$$\lim_{k\to\infty} \mathbb{P}\big\{\big|Y_t^{(k)}(B_{\xi_t^1}(\varepsilon)) - \delta_{\xi_t^1}(B_{\xi_t^1}(\varepsilon))\big| < \varepsilon\big\} = \lim_{k\to\infty} \mathbb{P}\big\{Y_t^{(k)}(B_{\xi_t^1}(\varepsilon)) \geq 1-\varepsilon\big\} = 1$$

for each t and ε, where $B_{\xi_t^1}(\varepsilon)$ denotes the ball centred in ξ_t^1 and radius ε. The latter will be implied by

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

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

$$\begin{aligned}\mathbb{E}\big[\langle \Phi_\varepsilon, Y_t^{(k)}\rangle\big] &= \lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N \Phi_\varepsilon(\xi_t^i)\right] \\ &\leq \limsup_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N \Phi_\varepsilon(\xi_t^i) 1_{\{N_i^t(\delta)=1\}}\right] + \mathbb{E}\big[1 - A_1^t(\delta)\big]. \end{aligned} \tag{3.3}$$

The second term in the last line, for each $\delta > 0$, converges to 0 as $k \to \infty$, cf. [Pi99], Prop. 30. Conditioning on $\mathscr{G}_{t-\delta,t}$, the genealogical information as defined in (2.5), we estimate the first term as follows:

$$\begin{aligned}\mathbb{E}\Big[\frac{1}{N}\sum_{i=1}^N \Phi_\varepsilon(\xi_t^i) 1_{\{N_i^t(\delta)=1\}}\Big] &= \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[1_{\{N_i^t(\delta)=1\}} \mathbb{E}\big[\Phi_\varepsilon(\xi_t^i)\big|\mathscr{G}_{t-\delta,t}, \xi_{t-\delta}^1\big]\Big] \\ &\leq \mathbb{E}\left[\int \Phi_\varepsilon(y) p_\delta^{(\alpha)}\big(\xi_{t-\delta}^1, y\big)\, dy\right] \\ &\leq \mathbb{P}\big\{|\xi_t^1 - \xi_{t-\delta}^1| \geq \varepsilon/3\big\} + p_\delta^{(\alpha)}\big(0, B_0(\varepsilon/3)^c\big),\end{aligned}$$

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

For n time points $t_1 < t_2 < \cdots < t_n$ observe that

$$\mathbb{P}\Big\{\exists\, 1 \leq i \leq n : Y_{t_i}^{(k)}\big(B_{\xi_{t_i}^1}(\varepsilon)\big) < 1-\varepsilon\Big\} \leq \sum_{i=1}^n \mathbb{P}\Big\{Y_{t_i}^{(k)}\Big(\big(B_{\xi_{t_i}^1}(\varepsilon)\big)^c\Big) \geq \varepsilon\Big\}$$

which converges to 0 by (3.2) and the Markov inequality.

3.2 Proof of Theorem 1.9

In the case $\alpha = 2$, using Proposition 1.3, tightness on the space $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$ 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 $\alpha < 2$ and $\Lambda((0,1)) > 0$, let us recall the following classical characterisation of relative compactness in $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$, 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 $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$. 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 $\gamma_{\varepsilon,t} \subset \mathscr{M}_1(\mathbb{R}^d)$, such that*

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

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

$$\limsup_{k\to\infty} \mathbb{P}\left\{w'(Y^k, \delta, T) \geq \varepsilon\right\} \leq \varepsilon, \tag{3.4}$$

where

$$w'(y, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} d(y(s), y(t)), \tag{3.5}$$

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

By Lemma 1.8, there is an $\varepsilon > 0$ such that for $k_0 \in \mathbb{N}$ and $\delta > 1/k_0$

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

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

Finally, in the case $\alpha < 2$ and $\Lambda = c\delta_0$ for some $c > 0$, note that, due to the absence of macroscopic birth events, each $Y^{c\delta_0,\Delta_\alpha}[k]$ a.s. has continuous paths (formally, this follows e.g. from Theorem 4.7.2 in [Da93] and the standard disintegration result, see [EM91], [Pe91]). Let μ_k denote the distribution of $Y^{c\delta_0,\Delta_\alpha}[k]$. Since the set of continuous paths $\mathscr{C} := C_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$ is closed in Skorohod's J_1-topology, weak convergence on $D_{[0,\infty)}(\mathscr{M}_1(\mathbb{R}^d))$ to some distribution μ would imply that

$$1 = \limsup_{k\to\infty} \mu_k(\mathscr{C}) \leq \mu(\mathscr{C}),$$

by the Portmanteau Theorem. However, the f.d.d. limit $\{\delta_{B_t^{(\alpha)}}, t \geq 0\}$ has no continuous modification for $\alpha < 2$, and we arrive at a contradiction. □

3.3 Proof of Lemma 1.8

The intuitive mechanism behind a "spark" obtained from the lookdown 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 "spark" 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 quickly, since mass is attracted rapidly towards the position of the level-1 particle. Note that corresponding phenomena will occur on any level $j \geq 2$.

First, we collect some useful notation. Without loss of generality assume $T = 1$. The following choices for the constants δ_2 and ε are justified by Lemma 4.1 and Lemma 4.2 from the Appendix. Indeed, choose $\delta_1 \in (0,1)$ with $\Lambda((\delta_1, 1)) > 0$, and then $\varepsilon = \varepsilon(\delta_1) > 0$ such that for any $y \in \mathbb{R}^d$ and any pair $\mu, \mu' \in \mathscr{M}_1(\mathbb{R}^d)$,

$$\mu(B_y(1)) \geq 1 - \delta_1/2 \quad \text{and} \quad \mu'\big((B_y(2))^c\big) \geq \delta_1 \quad \text{implies} \quad d_{\mathscr{M}_1}(\mu, \mu') > 2\varepsilon \tag{3.6}$$

and choose $\delta_2 = \delta_2(\varepsilon, \delta_1) \in (0, \delta_1]$ such that for any $x, x' \in \mathbb{R}^d$ with $|x - x'| \leq \delta_2$,

$$\mu(B_x(\delta_2)) \geq 1 - \delta_2/2 \quad \text{and} \quad \mu'(B_{x'}(\delta_2)) \geq 1 - \delta_2/2 \quad \text{implies} \quad d_{\mathscr{M}_1}(\mu, \mu') \leq \varepsilon. \tag{3.7}$$

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

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

(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

$$\begin{aligned} H_{s,t} := & \left\{L_{12}(t) - L_{12}(s) = 0\right\} \\ & \bigcap \quad \left\{n\big((s,t] \times \{(x,(u_m)) \in (0,1] \times [0,1]^{\mathbb{N}} : u_1, u_2 \leq x\}\big) = 0\right\} \end{aligned} \tag{3.8}$$

be the event that in the time interval $(s, t]$, no lookdown event involving both levels 1 and 2 occurs. Furthermore, note that since symmetric α-stable processes do not have fixed times of discontinuity,

$$\lim_{k \to \infty} \mathbb{P}\left\{ \sup_{0 \leq t \leq 1/k} |B_t^{(\alpha)}| \leq \frac{\delta_2}{2} \right\} = 1. \tag{3.9}$$

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

- Within the time-interval $(a_i, b_i]$, consider the event $\mathscr{A}_i^{(k)}$ 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 in space, more precisely, recalling (2.8),

$$\begin{aligned} \mathscr{A}_i^{(k)} := & \left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} 1_{\{N_j^{b_i}(1/(4k))=1\}} 1_{\{|\xi^j_{b_i} - \xi^1_{b_i - \sigma_j^{b_i}}| \leq \delta_2/2\}} \geq 1 - \frac{\delta_2}{2} \right\} \\ & \bigcap \quad \left\{ \sigma_2^{b_i} < \frac{1}{4k} \right\} \quad \bigcap \quad \left\{ |\xi^2_{b_i} - \xi^2_{b_i - \sigma_2^{b_i}}| \leq \frac{1}{2} \right\}. \end{aligned}$$

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

$$\mathscr{B}_i^{(k)} := H_{b_i, c_i} \quad \bigcap \quad \left\{ |\xi^2_{c_i} - \xi^2_{b_i}| > 4 \right\}.$$

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

$$\mathscr{C}_i^{(k)} := H_{c_i,d_i} \bigcap \left\{ \sup_{t\in(c_i,d_i]} |\xi_t^2 - \xi_{c_i}^2| < 1 \right\} \\ \bigcap \left\{ n\big([c_i, d_i] \times \{(x,(u_m)) \in (0,1] \times [0,1]^{\mathbb{N}} : x > \delta_1,\, u_2 < x \le u_1\}\big) \ge 1 \right\}.$$

- Finally, let $\mathscr{D}_i^{(k)}$ 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,

$$\mathscr{D}_i^{(k)} := \left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} 1_{\{N_j^{a_{i+1}}(1/(4k))=1\}} 1_{\{|\xi^j_{a_{i+1}} - \xi^1_{a_{i+1}-\sigma_j^{a_{i+1}}}| \le \delta_2/2\}} \ge 1 - \frac{\delta_2}{2} \right\}.$$

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

$$\left(\xi^j_{b_i} - \xi^1_{b_i - \sigma_j^{b_i}}\right) 1_{\{\sigma_j^{b_i} \le 1/(4k)\}}, \quad \left(\xi^j_{a_{i+1}} - \xi^1_{a_{i+1} - \sigma_j^{a_{i+1}}}\right) 1_{\{\sigma_j^{a_{i+1}} \le 1/(4k)\}},$$

for $j = 2, 3, \dots$, and

$$\left(\xi_t^2 - \xi_{b_i}^2\right) 1_{H_{b_i,d_i}},\ b_i \le t \le d_i.$$

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

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

Define

$$\mathscr{O}_i^{(k)} := \left\{ \sup_{t\in(a_i,a_{i+1}]} |\xi_t^1 - \xi_{a_i}^1| \le \delta_2/2 \right\}.$$

On the event

$$\mathscr{E}_i^{(k)} := \mathscr{O}_i^{(k)} \cap \mathscr{A}_i^{(k)} \cap \mathscr{B}_i^{(k)} \cap \mathscr{C}_i^{(k)} \cap \mathscr{D}_i^{(k)}, \tag{3.10}$$

we see from (3.6), (3.7) and the definitions of $\mathscr{A}_i^{(k)}, \mathscr{B}_i^{(k)}, \mathscr{C}_i^{(k)}, \mathscr{D}_i^{(k)}$ that there is a (random) time $\tau \in (c_i, d_i]$ such that $Y^{\Lambda,\Delta_\alpha}[k]$ exhibits an ε-$(1/k)$-spark in $(a_i, a_{i+1}]$.

Indeed, $\mathscr{O}_i^{(k)}$ guarantees that the level-1-particle did not move more than $\delta_2/2$ units away during $(a_i, a_{i+1}]$ from its initial location at time a_i, and this combined with the first set in the definition of $\mathscr{A}_i^{(k)}$ guarantees that

$$Y_{b_i}^{\Lambda,\Delta_\alpha}[k](B_{\xi^1_{b_i}}(\delta_2)) \ge 1 - \frac{\delta_2}{2}.$$

In a similar fashion, $\mathscr{O}_i^{(k)}$ and $\mathscr{D}_i^{(k)}$ guarantee that

$$Y_{a_{i+1}}^{\Lambda,\Delta_\alpha}[k](B_{\xi^1_{a_{i+1}}}(\delta_2)) \ge 1 - \frac{\delta_2}{2}.$$

Hence, an application of (3.7) using the observation that $|\xi^1_{b_i} - \xi^1_{a_{i+1}}| \leq \delta_2/2$ (due to $\mathscr{O}_i^{(k)}$) gives

$$d_{\mathscr{M}_1}(Y_{b_i}^{\Lambda,\Delta_\alpha}[k], Y_{a_{i+1}}^{\Lambda,\Delta_\alpha}[k]) \leq \varepsilon. \tag{3.11}$$

Second, if we choose τ to be the first time point in $(c_i, d_i]$ when a lookdown event conforming to the requirements of the third component of the definition of $\mathscr{C}_i^{(k)}$ takes place, we have by $\mathscr{O}_i^{(k)}$, $\mathscr{B}_i^{(k)}$ and $\mathscr{C}_i^{(k)}$ that the level-2-particle is sufficiently far away from the location of the level-1-particle in $(b_i, c_i]$, stays away more than 2 units during $(c_i, d_i]$ and at some time $\tau \in (c_i, d_i]$ it accumulates a proportion of mass at least δ_1. Hence,

$$Y_\tau^{\Lambda,\Delta_\alpha}[k]\big((B_{\xi^1_\tau}(2))^c\big) \geq \delta_1,$$

and we know from above that under the assumptions as before (again, given $\mathscr{O}_i^{(k)}$),

$$Y_{b_i}^{\Lambda,\Delta_\alpha}[k](B_{\xi^1_\tau}(1)) \geq 1 - \frac{\delta_2}{2} \geq 1 - \frac{\delta_1}{2}$$

by our choice of $\delta_1 \geq \delta_2$. Using (3.6) with $y = \xi^1_\tau$ we arrive at

$$d_{\mathscr{M}_1}(Y_{b_i}^{\Lambda,\Delta_\alpha}[k], Y_\tau^{\Lambda,\Delta_\alpha}[k]) > 2\varepsilon. \tag{3.12}$$

Finally, $\mathscr{D}_i^{(k)}$ guarantees that most of the mass returns to the close vicinity of the level-1-particle at time a_{i+1}, which is, due to $\mathscr{B}_i^{(k)}$, $\mathscr{C}_i^{(k)}$ and $\mathscr{O}_i^{(k)}$, more than 2 units away from the position of the level-2-particle at time τ, but still close to its own position at time τ. Hence, as before,

$$Y_\tau^{\Lambda,\Delta_\alpha}[k]\big((B_{\xi^1_\tau}(2))^c\big) \geq \delta_1,$$

and

$$Y_{a_{i+1}}^{\Lambda,\Delta_\alpha}[k](B_{\xi^1_\tau}(1)) \geq 1 - \delta_1/2.$$

Using (3.6) again, we finally arrive at

$$d_{\mathscr{M}_1}(Y_\tau^{\Lambda,\Delta_\alpha}[k], Y_{a_{i+1}}^{\Lambda,\Delta_\alpha}[k]) > 2\varepsilon. \tag{3.13}$$

Recalling Definition 1.6, we see that (3.11), (3.12) and (3.13) guarantee the existence of an ε-$(1/k)$-spark in $(a_i, a_{i+1}]$.

It remains to show that

$$\inf_{k\in\mathbb{N}} \mathbb{P}\Big(\bigcup_{i=0}^{k-1} \mathscr{E}_i^{(k)}\Big) > 0, \tag{3.14}$$

which yields the claim. In order to verify (3.14), note that by (3.9), for all $i < k$,

$$\lim_{k\to\infty} \mathbb{P}\{\mathscr{O}_i^{(k)}\} = \lim_{k\to\infty} \mathbb{P}\Big\{ \sup_{0\leq t\leq 1/k} |B_t^{(\alpha)}| \leq \delta_2/2 \Big\} = 1$$

and that for fixed k, the family of events

$$\{\mathscr{E}_i^{(k)} : 1 \leq i \leq k\}$$

is independent (by the usual independence properties of the driving Poisson processes and the increments of the motion processes for each level). Moreover,

$$\forall\, k, i < k : \; \mathbb{P}\big(\mathscr{E}_i^{(k)}\big) \geq \frac{C}{k} \tag{3.15}$$

for some $C = C(\alpha, \Lambda, \delta, (\delta_m)) > 0$, which together with the independence properties implies (3.14). To see (3.15), note that the 'crucial' component in $\mathscr{E}_i^{(k)}$ is $\mathscr{B}_i^{(k)}$, which satisfies by the scaling properties and tail behaviour of symmetric α-stable processes, see e.g. [Be96, Chapter VIII],

$$\begin{aligned}
\mathbb{P}\big(\mathscr{B}_i^{(k)}\big) &= \mathbb{P}\Big(\big\{|\xi^2_{c_i} - \xi^2_{b_i}| > 4\big\} \cap H_{b_i,c_i}\Big) \\
&= \mathbb{P}\Big(\big\{|\xi^2_{c_i} - \xi^2_{b_i}| > 4\big\} \Big| H_{b_i,c_i}\Big)\mathbb{P}\big(H_{b_i,c_i}\big) \\
&= \mathbb{P}\big\{|B^{(\alpha)}_{1/(4k)}| > 4\big\}\mathbb{P}\big(H_{b_i,c_i}\big) \\
&= \mathbb{P}\big\{|B^{(\alpha)}_{1}| > 4k^{1/\alpha}\big\}\mathbb{P}\big(H_{b_i,c_i}\big) \sim \text{Const.} \times \frac{1}{k}.
\end{aligned}$$

The probability of the genealogical components of $\mathscr{E}_i^{(k)}$ does not converge to zero, since smaller time intervals (of length $1/k$) are compensated by higher resampling intensities due to the time rescaling (proportional to k) in each step. Finally, the probability of the remaining requirements (upper bounds on the distance that the level-1-particle or the level-2-particle are allowed to travel) do not converge to zero either by (3.9) or the fact that these upper bounds are fixed values, independent of k.

□

3.4 Proof of Proposition 1.3

The proof is a straightforward modification of arguments in [Eth00], Prop. 2.27. Indeed, it is easy to read off the function-valued dual process from (1.2), which is an obvious extension of [Da93], Section 5.6 (see [BBM+09, Section 5.2] for an explicit description in the context of Ξ-Fleming-Viot processes, which contains Λ-Fleming-Viot processes as a special case).

For the first moment, the function valued dual consists of one component only, to which the α-heat flow is applied. This gives, for $t > 0$,

$$\mathbb{E}\big[\langle \varphi, Y_t^{\Lambda,\Delta_\alpha}\rangle\big] = \int P_t^{(\alpha)}\varphi(x)\,\mu(dx). \tag{3.16}$$

For the second moment, the function valued dual consists of two components, and the formula follows from considering the exponential jump time at which both components merge, yielding, for $t > 0$,

$$\begin{aligned}
&\mathbb{E}\big[\langle \varphi_1, Y_t^{\Lambda,\Delta_\alpha}\rangle\langle \varphi_2, Y_t^{\Lambda,\Delta_\alpha}\rangle\big] \\
&\qquad = \int_0^t \int \rho e^{-\rho s} P_s^{(\alpha)}\big(P_{t-s}^{(\alpha)}\varphi_1 P_{t-s}^{(\alpha)}\varphi_2\big)(x)\,\mu(dx)\,ds \\
&\qquad\quad + e^{-\rho_t}\int P_t^{(\alpha)}\varphi_1(x)\,\mu(dx)\int P_t^{(\alpha)}\varphi_2(x)\,\mu(dx).
\end{aligned} \tag{3.17}$$

Note that this can easily be extended to distinct time points t_1, t_2 using the tower property and the strong Markov property. Indeed, w.l.g. assume $t_1 < t_2$ and write $t := t_1, t + s = t_2$ for $s := t_2 - t_1$. Then, using (3.16),

$$\begin{aligned}\mathbb{E}\big[\langle\varphi_1, Y_{t_1}^{\Lambda,\Delta_\alpha}\rangle\langle\varphi_2, Y_{t_2}^{\Lambda,\Delta_\alpha}\rangle\big] &= \mathbb{E}\Big[\langle\varphi_1, Y_t^{\Lambda,\Delta_\alpha}\rangle \mathbb{E}_{Y_t^{\Lambda,\Delta_\alpha}}\big[\langle\varphi_2, \tilde{Y}_s^{\Lambda,\Delta_\alpha}\rangle\big]\Big] \\ &= \mathbb{E}\Big[\langle\varphi_1, Y_t^{\Lambda,\Delta_\alpha}\rangle \int P_s^{(\alpha)}\varphi(x)\, Y_t^{\Lambda,\Delta_\alpha}(dx)\Big],\end{aligned}$$

where $\tilde{Y}_s^{\Lambda,\Delta_\alpha}$ is an independent copy of $Y_s^{\Lambda,\Delta_\alpha}$. Denoting $\tilde{\varphi}_2 := P_s^{(\alpha)}\varphi(x)$ and observing that $\varphi_2 \in C_b^2$ implies $\tilde{\varphi}_2 \in C_b^2$, we obtain from (3.17)

$$\begin{aligned}&\mathbb{E}\big[\langle\varphi_1, Y_{t_1}^{\Lambda,\Delta_\alpha}\rangle\langle\varphi_2, Y_{t_2}^{\Lambda,\Delta_\alpha}\rangle\big] \\ &\quad= \mathbb{E}\big[\langle\varphi_1, Y_t^{\Lambda,\Delta_\alpha}\rangle\langle\tilde{\varphi}_2, Y_t^{\Lambda,\Delta_\alpha}\rangle\big] \\ &\quad= \int_0^t \int \rho e^{-\rho u} P_u^{(\alpha)}\big(P_{t-u}^{(\alpha)}\varphi_1 P_{t-u}^{(\alpha)} P_s^{(\alpha)}\varphi_2\big)(x)\,\mu(dx)\,du \\ &\qquad + e^{-\rho_{t_1}} \int P_t^{(\alpha)}\varphi_1(x)\,\mu(dx) \int P_t^{(\alpha)} P_s^{(\alpha)}\varphi_2(x)\,\mu(dx) \\ &\quad= \int_0^{t_1} \int \rho e^{-\rho u} P_u^{(\alpha)}\big(P_{t_1-u}^{(\alpha)}\varphi_1 P_{t_2-u}^{(\alpha)}\varphi_2\big)(x)\,\mu(dx)\,du \\ &\qquad + e^{-\rho_{t_1}} \int P_{t_1}^{(\alpha)}\varphi_1(x)\,\mu(dx) \int P_{t_2}^{(\alpha)}\varphi_2(x)\,\mu(dx),\end{aligned}$$

as required. □

4 Appendix

Consider the metric space of probability measures $(\mathscr{M}_1(\mathbb{R}^d), d_{\mathscr{M}_1})$ on $\mathbb{R}^d$ and recall that, for $\mu, \nu \in \mathscr{M}_1(\mathbb{R}^d)$,

$$d_{\mathscr{M}_1}(\mu,\nu) := \inf\{\varepsilon > 0 : \mu(B) \le \nu(B^\varepsilon) + \varepsilon \text{ for all closed sets } B \subset \mathbb{R}^d\}, \tag{4.1}$$

where B^ε is the usual ε-enlargement of the set $B \subset \mathbb{R}^d$. Let $B_y(r)$ be the d-dimensional closed ball centered in y with radius r. We collect two technical lemmas.

Lemma 4.1. *Let $\delta \in (0,1)$. There exists $\varepsilon = \varepsilon(\delta) > 0$ such that for any two measures $\mu, \mu' \in \mathscr{M}_1(\mathbb{R}^d)$*

$$\mu(B_y(1)) \ge 1 - \frac{\delta}{2} \quad \textit{and} \quad \mu'\big((B_y(2))^c\big) \ge \delta \quad \textit{for some } y \in \mathbb{R}^d \tag{4.2}$$

implies

$$d_{\mathscr{M}_1}(\mu,\mu') > 2\varepsilon.$$

Proof. Assume (4.2) holds for μ, μ'. Then $\mu'(B_y(1)^{2\varepsilon}) < 1 - \delta$ for any $\varepsilon \in (0, 1/2)$. By definition, $d_{\mathscr{M}_1}(\mu,\mu') > 2\varepsilon$ is implied by

$$\mu(B_y(1)) > \mu'(B_y(1)^{2\varepsilon}) + 2\varepsilon$$

which holds if $\varepsilon < \delta/4$. □

Lemma 4.2. *Pick some $\varepsilon \in (0, 1/2)$. Then, there exists a constant $\tilde{\delta}(\varepsilon) > 0$, such that for each $0 < \delta < \tilde{\delta}$ and each x, x' with $|x - x'| \le \delta$, we have that for any two measures $\mu, \mu' \in \mathscr{M}_1(\mathbb{R}^d)$*

$$\mu(B_x(\delta)) \ge 1 - \frac{\delta}{2} \quad \textit{and} \quad \mu'(B_{x'}(\delta)) \ge 1 - \frac{\delta}{2} \tag{4.3}$$

implies

$$d_{\mathscr{M}_1}(\mu, \mu') \le \varepsilon.$$

Proof. Let $A \subset \mathbb{R}^d$ be closed. Let $\delta < \tilde{\delta} := \varepsilon/4$. We consider two cases. First, assume $A \cap B_x(\delta) \neq \emptyset$. Then, $B_x(\delta), B_{x'}(\delta) \subset A^\varepsilon$. Hence, $\mu'(A^\varepsilon) + \varepsilon \ge 1 - \delta/2 + \varepsilon \ge 1$ and the result holds.

In the second case, i.e. $A \cap B_x(\delta) = \emptyset$, we have $\mu(A) < \delta/2 < \varepsilon$, and the result holds, too. □

Acknowledgement We would like to thank an anonymous referee for her/his careful reading and helpful comments.

References

[Be96] BERTOIN, J.: *Lévy Processes*, Wiley (1996).

[BLG03] BERTOIN, J.; LE GALL, J. F.: Stochastic flows associated to coalescent processes, *Probab. Theory Related Fields* **126** (2003), no. 2, 261–288.MR1990057

[Bi68] BILLINGSLEY, P.: *Convergence of probability measures*, Wiley (1968).

[BB09] BIRKNER, M.; BLATH, J.: Measure-valued diffusions, general coalescents and population genetic inference, *Trends in Stochastic Analysis*, LMS **353**, Cambridge University Press, 329–363 (2009).MR0233396

[BBC+05] 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).

[BBM+09] BIRKNER, M.; BLATH, J.; MÖHLE, M.; STEINRÜCKEN, M.; TAMS, J.: A modified lookdown construction for the Xi-Fleming-Viot process with mutation and populations with recurrent bottlenecks, *ALEA Lat. Am. J. Probab. Math. Stat.* **6**, 25–61 (2009).MR2485878

[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.MR1242575

[DH82] DAWSON, D.; HOCHBERG, K.: Wandering random measures in the Fleming-Viot Model. *Ann. Probab.* **10**, no. 3, 554–580, (1982).MR0659528

[DK96] DONNELLY, P.; KURTZ, T.: A countable representation of the Fleming-Viot measure-valued diffusion, *Ann. Appl. Probab.* **24**, 698–742, (1996).MR1404525

[DK99] DONNELLY, P.; KURTZ, T.: Particle representations for measure-valued population models, *Ann. Probab.* **27**, no. 1, 166–205, (1999).MR1681126

[EKR91] EL KAROUI, N.; RŒLLY, S.: Proprietés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchemant a valeurs mesures, *Stoch. Proc. Appl.* **38**, 239–266, (1991).MR1119983

[Eth00] ETHERIDGE, A.: *An Introduction to Superprocesses.* AMS University Lecture Series, Vol. 20, (2000).MR1779100

[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).MR0838085

[Fe66] FELLER, W.: *Introduction to Probability Theory and its Applications,* Volume II, Wiley, (1966).MR0210154

[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).MR2086012

[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).MR2238611

[FV79] FLEMING, W.; VIOT, M.: Some measure-valued markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843, (1979).MR0542340

[K82] KINGMAN, J. F. C.: The coalescent. *Stoch. Proc. Appl.* **13**, 235–248, (1982).MR0671034

[MS01] MÖHLE, M.; SAGITOV, S.: A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562, (2001).MR1880231

[Pe91] PERKINS, E.: Conditional Dawson-Watanabe processes and Fleming-Viot processes, Seminar in Stochastic Processes, Birkhäuser, pp 142–155, (1991).MR1172149

[Pi99] PITMAN, J.: Coalescents with multiple collisions, *Ann. Probab.* **27**, no. 4, 1870–1902, (1999).MR1742892

[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). MR2245368

[Sa99] SAGITOV, S.: The general coalescent with asynchronous mergers of ancestral lines, *J. Appl. Probab.* **26**, 1116-1125 (1999).MR1742154

[Sc00] SCHWEINSBERG, J.: A necessary and sufficient condition for the Λ-coalescent to come down from infinity. *Electron. Comm. Probab.* **5**, 1–11, (2000).MR1736720

[S56] SKOROHOD, A.V.: Limit theorems for stochastic processes. *Theory Probab. Appl.* **1**, 261–290, (1956). MR0084897

[Y65] YOSIDA, K.: *Functional Analysis*, Grundlehren der mathematischen Wissenschaften **123**, Springer–Verlag, (1965).