

On exceptional times for generalized Fleming–Viot processes with mutations

J. Berestycki · L. Döring · L. Mytnik · L. Zambotti

Received: 10 April 2013 / Accepted: 14 February 2014 / Published online: 6 March 2014
© Springer Science+Business Media New York 2014

Abstract If \mathbf{Y} is a standard Fleming–Viot process with constant mutation rate (in the infinitely many sites model) then it is well known that for each $t > 0$ the measure \mathbf{Y}_t is purely atomic with infinitely many atoms. However, Schmuland proved that there is a critical value for the mutation rate under which almost surely there are exceptional times at which the stationary version of \mathbf{Y} is a finite sum of weighted Dirac masses. In the present work we discuss the existence of such exceptional times for the generalized Fleming–Viot processes. In the case of Beta-Fleming–Viot processes with index $\alpha \in]1, 2[$ we show that—irrespectively of the mutation rate and α —the number of atoms is almost surely always infinite. The proof combines a Pitman–Yor type representation with a disintegration formula, Lamperti’s transformation for self-similar processes and covering results for Poisson point processes.

Keywords Fleming–Viot processes · Mutations · Exceptional times · Excursion theory · Jump-type SDE · Self-similarity

J. Berestycki · L. Döring · L. Zambotti (✉)
Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, 4 Place Jussieu,
75252 Paris Cedex 05, France
e-mail: lorenzo.zambotti@upmc.fr

J. Berestycki
e-mail: julien.berestycki@upmc.fr

L. Döring
e-mail: leif.doering@upmc.fr

L. Mytnik
Faculty of Industrial Engineering and Management, Technion Israel Institute of Technology,
32000 Haifa, Israel
e-mail: leonid@ie.technion.ac.il

Mathematics Subject Classification (2000) Primary 60J80 · Secondary 60G18

1 Main result

The measure-valued *Fleming–Viot diffusion processes* were first introduced by Fleming and Viot [20] and have become a cornerstone of mathematical population genetics in the last decades. It is a model which describes the evolution (forward in time) of the genetic composition of a large population. Each individual is characterized by a *genetic type* which is a point in a type-space E . The Fleming–Viot process is a Markov process $(\mathbf{Y}_t)_{t \geq 0}$ on

$$\mathcal{M}_E^1 = \{ \nu : \nu \text{ is a probability measure on } E \}$$

for which we interpret $\mathbf{Y}_t(B)$ as the proportion of the population at time t which carries a genetic type belonging to a Borel set B of types. In particular, the number of (different) types at time t is equal to the number of atoms of \mathbf{Y}_t with the convention that the number of types is infinite if \mathbf{Y}_t has absolutely continuous part.

Fleming–Viot superprocesses can be defined through their infinitesimal generators

$$(\mathcal{L}\phi)(\mu) = \int_E \int_E \mu(dv)(\delta_v(dy) - \mu(dy)) \frac{\delta^2 \phi(\mu)}{\delta \mu(v) \delta \mu(y)} + \int_E \mu(dv) A \left(\frac{\delta \phi(\mu)}{\delta \mu(\cdot)} \right) (v), \tag{1.1}$$

acting on smooth test-functions where $\delta \phi(\mu) / \delta \mu(v) = \lim_{\epsilon \rightarrow 0+} \epsilon^{-1} \{ \phi(\mu + \epsilon \delta_v) - \phi(\mu) \}$ and A is the generator for a Markov process in E which represents the effect of mutations. Here δ_v is the Dirac measure at v . It is well known that the Fleming–Viot superprocess arises as the scaling limit of a Moran-type model for the evolution of a finite discrete population of fixed size if the reproduction mechanism is such that no individual gives birth to a positive proportion of the population in a small number of generations. For a detailed description of Fleming–Viot processes and discussions of variations we refer to the overview article of Ethier and Kurtz [19] and to Etheridge’s lecture notes [17].

The first summand of the generator reflects the genetic resampling mechanism whereas the second summand represents the effect of mutations. Several choices for A have appeared in the literature. In the present work we shall work in the setting of the *infinitely-many-alleles model* where each mutation creates a new type never seen before. Without loss of generality let the type space be $E = [0, 1]$. Then the following choice of A gives an example of an infinite site model with mutations:

$$(Af)(v) = \theta \int_E (f(y) - f(v)) dy, \tag{1.2}$$

for some $\theta > 0$. The choice of the uniform measure dy is arbitrary (we could choose the new type according to any distribution that has a density with respect to the Lebesgue

measure), all that matters is that the newly created type y is different from all other types. With A as in (1.2), mutations arrive at rate θ and create a new type picked at random from E according to the uniform measure, therefore the corresponding process is sometimes called the Fleming–Viot process with neutral mutations.

Let us briefly recall two classical facts concerning the infinite types Fleming–Viot process described above. For any initial condition \mathbf{Y}_0 :

- (i) If there is no mutation, then, for all $t > 0$ fixed, the number of types is almost surely finite.
- (ii) If the mutation parameter θ is strictly positive, then, for all $t > 0$ fixed, the number of types is infinite almost surely.

This can be deduced e.g. from the explicit representation of the transition function given in Ethier and Griffiths in [18].

A beautiful complement to (i) and (ii) was found by Schmuland for exceptional times that are not fixed in advance:

Theorem 1.1 (Schmuland [35]) *For the stationary infinitely-many-alleles model*

$$\mathbb{P}(\exists t > 0 : \#\{\text{types at time } t\} < \infty) = \begin{cases} 1 & \text{if } \theta < 1, \\ 0 & \text{if } \theta \geq 1. \end{cases}$$

Schmuland’s proof of the dichotomy is based on analytic arguments involving the capacity of finite dimensional subspaces of the infinite dimensional state-space. In Sect. 6 we reprove Schmuland’s theorem with a simple proof via excursion theory, that yields the result for arbitrary initial conditions.

In the series of articles [5–7], Bertoin and Le Gall introduced and started the study of Λ -Fleming–Viot processes, a class of stochastic processes which naturally extends the class of standard Fleming–Viot processes. These processes are completely characterized by a finite measure Λ on $]0, 1[$ and a generator A . Similarly to the standard Fleming–Viot process, these processes can be defined through their infinitesimal generator

$$\begin{aligned}
 (\mathcal{L}\phi)(\mu) &= \int_0^1 y^{-2} \Lambda(dy) \int \mu(da) (\phi((1 - y)\mu + y\delta_a) - \phi(\mu)) \\
 &\quad + \int_E \mu(dv) A \left(\frac{\delta\phi(\mu)}{\delta\mu(\cdot)} \right) (v), \tag{1.3}
 \end{aligned}$$

and the sites of atoms are again called types. For $A = 0$, the generator formulation only appeared implicitly in [6] and is explained in more details in Birkner et al. [10] and for A as in (1.2) it can be found in Birkner et al. [9]. The dynamics of a generalized Fleming–Viot process $(\mathbf{Y}_t)_{t \geq 0}$ are as follows: at rate $y^{-2} \Lambda(dy)$ a point a is sampled at time $t > 0$ according to the probability measure $\mathbf{Y}_{t-}(da)$ and a point-mass y is added at position a while scaling the rest of the measure by $(1 - y)$ to keep the total mass at 1. The second term of (1.3) is the same mutation operator as in (1.1). For a detailed description of Λ -Fleming–Viot processes and discussions of variations we refer to the overview article of Birkner and Blath [8].

In the following we are going to focus only on the choice $\Lambda = \text{Beta}(2 - \alpha, \alpha)$, the Beta distribution with density

$$f(u) = C_\alpha u^{1-\alpha} (1 - u)^{\alpha-1} du, \quad C_\alpha = \frac{1}{\Gamma(2 - \alpha)\Gamma(\alpha)},$$

for $\alpha \in]1, 2[$, and mutation operator A as in (1.2). The corresponding Λ -Fleming–Viot process $(\mathbf{Y}_t)_{t \geq 0}$ is called Beta-Fleming–Viot process or (α, θ) -Fleming–Viot process and several results have been established in recent years. The (α, θ) -Fleming–Viot processes converge weakly to the standard Fleming–Viot process as α tends to 2. It was shown in [10] that a Λ -Fleming–Viot process with $A = 0$ is related to measure-valued branching processes in the spirit of Perkin’s disintegration theorem precisely if Λ is a Beta distribution (this relation is recalled and extended in Sect. 2.3 below).

If we chose $\alpha \in]1, 2[$ and \mathbf{Y}_0 uniform on $[0, 1]$, then we find the same properties (i) and (ii) for the one-dimensional marginals \mathbf{Y}_t unchanged with respect to the classical case (1.1), (1.2). In fact, for a general Λ -Fleming–Viot process, (i) is equivalent to the requirement that the associated Λ -coalescent comes down from infinity (see for instance [2]). Here is our main result: contrary to Schmuland’s result, (α, θ) -Fleming–Viot processes with $\alpha \in]1, 2[$ and $\theta > 0$ never have exceptional times:

Theorem 1.2 *Let $(\mathbf{Y}_t)_{t \geq 0}$ be an (α, θ) -Fleming–Viot superprocess with mutation rate $\theta > 0$ and parameter $\alpha \in]1, 2[$. Then for any starting configuration \mathbf{Y}_0*

$$\mathbb{P}(\exists t > 0 : \#\{\text{types at time } t\} < \infty) = 0$$

for any $\theta > 0$.

One way one can get a first rough understanding of why this should be true is by using a heuristic based on the duality between Λ -Fleming–Viot processes and Λ -coalescents. If Λ -Fleming–Viot processes describe how the composition of a population evolve forward in time, Λ -coalescents describe how the ancestral lineages of individuals sampled in the population merge as one goes back in time. The fact that Λ -coalescents describe the genealogies of Λ -Fleming–Viot processes can be seen from Donnelly and Kurtz [14] so-called lookdown construction of Fleming–Viot processes and was also established through a functional duality relation by Bertoin and Le Gall in [5].

The coalescent which corresponds to the classical Fleming–Viot process is the celebrated Kingman’s coalescent. Kingman’s coalescent comes down from infinity at speed $2/t$, i.e. if one initially samples infinitely many individuals in the populations, then the number of active lineages at time t in the past is N_t and $N_t \sim 2/t$ almost surely when $t \rightarrow 0$. It is known (see [6] or more recently [28]) that the process $(N_t, t \geq 0)$ has the same law as the process of the number of atoms of the Fleming–Viot process. For a Beta-coalescent (that is a Λ -coalescent where the measure Λ is the density of a $\text{Beta}(\alpha, 2 - \alpha)$ variable) with parameter $\alpha \in (1, 2)$ we have $N_t \sim c_\alpha t^{-1/(\alpha-1)}$ almost surely as $t \rightarrow 0$ (see [3, Theorem 4]). Therefore Kingman’s coalescents comes down from infinity much quicker than Beta-coalescents. Since the speed at which the generalized Fleming–Viot processes loses types roughly corresponds to the speed at

which the dual coalescent comes down from infinity, it is possible that (α, θ) -Fleming–Viot processes do not lose types fast enough, and hence there are no exceptional times at which the number of types is finite.

2 Auxiliary constructions

To prove Theorem 1.2 we construct two auxiliary objects: a particular measure-valued branching process and a corresponding Pitman–Yor type representation. Those will be used in Sect. 5 to relate the question of exceptional times to covering results for point processes. In this section we give the definitions and state their relations to the Beta-Fleming–Viot processes with mutations. All appearing stochastic processes and random variables will be defined on a common stochastic basis $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{P})$ that is rich enough to carry all Poisson point processes (PPP in short) that appear in the sequel.

2.1 Measure-valued branching processes with immigration

We recall that a continuous state branching process (CSBP in short) with α -stable branching mechanism, $\alpha \in]1, 2]$, is a Markov family $(P_v)_{v \geq 0}$ of probability measures on càdlàg trajectories with values in \mathbb{R}_+ , such that

$$E_v(e^{-\lambda X_t}) = e^{-v u_t(\lambda)}, \quad v \geq 0, \lambda \geq 0, \quad (2.1)$$

where for $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$, $\psi(u) := u^\alpha$, we have the evolution equation

$$u'_t(\lambda) = -\psi(u_t(\lambda)), \quad u_0(\lambda) = \lambda.$$

See e.g. [27] for a good introduction to CSBP. For $\alpha = 2$, $\psi(u) = u^2$ is the branching mechanism for Feller's branching diffusion, where P_v is the law of the unique solution to the SDE

$$X_t = v + \int_0^t \sqrt{2X_s} dB_s, \quad t \geq 0, \quad (2.2)$$

driven by a Brownian motion $(B_t)_{t \geq 0}$. On the other hand, for $\alpha \in]1, 2[$, $\psi(u) = u^\alpha$ gives the so-called α -stable branching processes which can be defined as the unique strong solution of the SDE

$$X_t = v + \int_0^t X_{s-}^{1/\alpha} dL_s, \quad t \geq 0, \quad (2.3)$$

driven by a spectrally positive α -stable Lévy process $(L_t)_{t \geq 0}$, with Lévy measure given by

$$\mathbb{1}_{(x>0)} c_\alpha x^{-1-\alpha} dx, \quad c_\alpha := \frac{\alpha(\alpha - 1)}{\Gamma(2 - \alpha)}.$$

Note that strong existence and uniqueness for (2.3) follows from the fact that the function $x \mapsto x^{1/\alpha}$ is Lipschitz outside zero, and hence strong existence and uniqueness holds for (2.3) until X hits zero. Moreover X , being a non-negative martingale, stays at zero forever after hitting it. For a more extensive discussion on strong solutions for jumps SDEs see [22] and [32].

The main tool that we introduce is a particular *measure-valued branching process with interactive immigration* (MBI in short). For a textbook treatment of this subject we refer to Li [31]. Following Dawson and Li [12], we are not going to introduce the MBIs via their infinitesimal generators but as strong solutions of a system of stochastic differential equations instead. On $(\Omega, \mathcal{G}, \mathcal{G}_\square, \mathbb{P})$, let us consider a Poisson point process $\mathcal{N} = (r_i, x_i, y_i)_{i \in I}$ on $(0, \infty) \times (0, \infty) \times (0, \infty)$ adapted to \mathcal{G}_\square and with intensity measure

$$v(dr, dx, dy) := \mathbb{1}_{(r>0)} dr \otimes c_\alpha \mathbb{1}_{(x>0)} x^{-1-\alpha} dx \otimes \mathbb{1}_{(y>0)} dy. \tag{2.4}$$

Throughout the paper we adopt the notation

$$\tilde{\mathcal{N}} := \mathcal{N} - v,$$

i.e. $\tilde{\mathcal{N}}$ is the compensated version of \mathcal{N} . It was shown in [12] that the solution to (2.3) has the same law as the unique strong solution to the SDE

$$X_t = X_0 + \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{(y < X_{r-})} x \tilde{\mathcal{N}}(dr, dx, dy) \tag{2.5}$$

with $X_0 = v$.

Now we are going to switch to the measure-valued setting. The real-valued process X in (2.3), (2.5) describes the evolution of the *total mass* of the CSBP starting at time zero at the mass $X_0 = v$. We are going to consider all initial masses $v \in [0, 1]$ simultaneously, constructing a process $(\mathbf{X}_t)_{t \geq 0}$ taking values in the space $\mathcal{M}_{[0,1]}^F$ of finite measures on $[0, 1]$, endowed with the narrow topology, i.e. the trace of the weak- $*$ topology of $(C[0, 1])^*$. Assume that at time $t = 0$, \mathbf{X}_0 is a finite measure on $[0, 1]$ with cumulative distribution function $(F(v), v \in [0, 1])$, and denote

$$X_t(v) := \mathbf{X}_t([0, v]), \quad t \geq 0, v \in [0, 1].$$

Then the measure-valued branching process $(\mathbf{X}_t)_{t \geq 0}$ can be constructed in such a way that for each v , $(X_t(v))_{t \geq 0}$ solves (2.5) with $X_0 = F(v)$, and with the same driving noise for all $v \in [0, 1]$. In what follows, we deal with a version of (2.5) including an immigration term only depending on the total-mass $X_t(1)$:

$$\begin{cases} X_t(v) = F(v) + \int_{]0,t[\times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{(y < X_{r-}(v))} x \tilde{\mathcal{N}}(dr, dx, dy) + I(v) \int_0^t g(X_s(1)) ds, \\ v \in [0, 1], t \geq 0, \end{cases} \tag{2.6}$$

where $(I(v), v \in [0, 1])$ is the cumulative distribution function of a finite measure on $[0, 1]$ and we assume

(G) $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is monotone non-decreasing, continuous and locally Lipschitz continuous away from zero.

Definition 2.1 An $\mathcal{M}_{[0,1]}^F$ -valued process $(\mathbf{X}_t)_{t \geq 0}$ on $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{P})$ is called a solution to (2.6) if

- it is càdlàg \mathbb{P} -a.s.,
- for all $v \in [0, 1]$, setting $X_t(v) := \mathbf{X}_t([0, v])$, $(X_t(v))_{v \in [0,1], t \geq 0}$ satisfies \mathbb{P} -a.s. (2.6).

Moreover, a solution $(\mathbf{X}_t)_{t \geq 0}$ is strong if it is adapted to the natural filtration \mathcal{F}_t generated by \mathcal{N} . Finally, we say that pathwise uniqueness holds if

$$\mathbb{P}(\mathbf{X}_t^1 = \mathbf{X}_t^2, \forall t \geq 0) = 1,$$

for any two solutions \mathbf{X}^1 and \mathbf{X}^2 on $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{P})$ driven by the same Poisson point process.

Here is a well-posedness result for (2.6):

Theorem 2.2 *Let F and I be as above. For any immigration mechanism g satisfying Assumption (G), there is a strong solution $(\mathbf{X}_t)_{t \geq 0}$ to (2.6) and pathwise uniqueness holds until $T_0 := \inf\{t \geq 0 : \mathbf{X}_t([0, 1]) = 0\}$.*

The proof of Theorem 2.2 relies on ideas from recent articles on pathwise uniqueness for jump-type SDEs such as Fu and Li [22] or Dawson and Li [12]. Our equation (2.6) is more delicate since all coordinate processes depend on the total-mass $X_t(1)$. The uniqueness statement is first deduced for the total-mass $(X_t(1))_{t \geq 0}$ and then for the other coordinates interpreting the total-mass as random environment. To construct a (weak) solution we use a (pathwise) Pitman–Yor type representation as explained in the next section.

2.2 A Pitman–Yor type representation for interactive MBIs

Let us denote by \mathcal{E} the set of càdlàg trajectories $w : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that $w(0) = 0$, w is positive on a bounded interval $]0, \zeta(w)[$ and $w \equiv 0$ on $[\zeta(w), +\infty[$. We recall the construction of the excursion measure of the α -stable CSBP $(P_v)_{v \geq 0}$, also called the *Kuznetsov measure*, see [30, Section 4] or [31, Chapter 8]: For all $t \geq 0$, let $K_t(dx)$ be the unique σ -finite measure on \mathbb{R}_+ such that

$$\int_{\mathbb{R}_+} (1 - e^{-\lambda x}) K_t(dx) = u_t(\lambda) = \left(\lambda^{1-\alpha} + (\alpha - 1)t\right)^{\frac{1}{1-\alpha}}, \quad \lambda \geq 0,$$

where we recall that the function $(u_t(\lambda))_{t \geq 0}$ is the unique solution to the equation

$$u_t(\lambda) + \int_0^t (u_s(\lambda))^\alpha ds = \lambda, \quad t \geq 0, \lambda \geq 0.$$

We also denote by $Q_t(x, dy)$ the Markov transition semigroup of $(P_t)_{t \geq 0}$. Then there exists a unique Markovian σ -finite measure \mathbb{Q} on \mathcal{E} with entrance law $(K_t)_{t \geq 0}$ and transition semigroup $(Q_t)_{t \geq 0}$, i.e. such that for all $0 < t_1 < \dots < t_n, n \in \mathbb{N}$,

$$\begin{aligned} &\mathbb{Q}(w_{t_1} \in dy_1, \dots, w_{t_n} \in dy_n, t_n < \zeta(w)) \\ &= K_{t_1}(dy_1) Q_{t_2-t_1}(y_1, dy_2) \cdots Q_{t_n-t_{n-1}}(y_{n-1}, dy_n). \end{aligned} \tag{2.7}$$

By construction

$$\int_{\mathcal{E}} (1 - e^{-\lambda w_s}) \mathbb{Q}(dw) = u_s(\lambda) = \left(\lambda^{1-\alpha} + (\alpha - 1)s\right)^{\frac{1}{1-\alpha}}, \quad s \geq 0, \lambda \geq 0, \tag{2.8}$$

and under \mathbb{Q} , for all $s > 0$, conditionally on $\sigma(w_r, r \leq s), (w_{t+s})_{t \geq 0}$ has law P_{w_s} . The σ -finite measure \mathbb{Q} is called the *excursion measure* of the CSBP (2.3). By (2.8), it is easy to check that for any $s > 0$

$$\int_{\mathcal{E}} w_s \mathbb{Q}(dw) = \frac{\partial}{\partial \lambda} u_s(\lambda) \Big|_{\lambda=0} = \lim_{\lambda \downarrow 0} \left(1 + \lambda^{\alpha-1}(\alpha - 1)s\right)^{\frac{\alpha}{1-\alpha}} = 1. \tag{2.9}$$

In Duquesne–Le Gall’s setting [15], under the σ -finite measure \mathbb{Q} with infinite total mass, w has the distribution of $(\ell^a(e))_{a \geq 0}$ under $n(de)$, where $n(de)$ is the excursion measure of the height process H and ℓ^a is the local time at level a . For the more general superprocess setting see for instance Dynkin and Kuznetsov [16].

We need now to extend the space of excursions as follows:

$$\mathcal{D} := \{w : \mathbb{R}_+ \mapsto \mathbb{R}_+ : \exists s \geq 0, w \equiv 0 \text{ on } [0, s], w_{-s} \in \mathcal{E}\},$$

i.e. \mathcal{D} is the set of càdlàg trajectories $w : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that w is equal to 0 on $[0, s(w)]$, w is positive on a bounded interval $]s(w), s(w) + \zeta(w)[$ and $w \equiv 0$ on $[s(w) + \zeta(w), +\infty[$. For $s \geq 0$, we denote by $\mathbb{Q}_s(dw)$ the σ -finite measure on \mathcal{D} given by

$$\int_{\mathcal{D}} \Phi(w) \mathbb{Q}_s(dw) := \int_{\mathcal{E}} \Phi(\mathbf{1}_{(\cdot \geq s)} w_{-s}) \mathbb{Q}(dw), \tag{2.10}$$

i.e. \mathbb{Q}_s is the image measure of \mathbb{Q} under the map

$$w \mapsto (\gamma_t := \mathbb{1}_{(t \geq s)} w_{t-s},)_{t \geq 0}. \tag{2.11}$$

Let us consider a Poisson point process $(s_i, u_i, a_i, w^i)_{i \in I}$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{D}$ with intensity measure

$$\Gamma(ds, du, da, dw) := (\delta_0(ds) \otimes \delta_0(du) \otimes F(da) + ds \otimes du \otimes I(da)) \otimes \mathbb{Q}_s(dw) \tag{2.12}$$

where F and I are the cumulative distribution functions appearing in (2.6). An atom (s_i, u_i, a_i, w^i) is a population that has immigrated at time s_i whose size evolution is given by w^i and whose genetic type is given by a_i . The coordinate u_i is used for thinning purposes, to decide whether or not this particular immigration really happened or not.

Theorem 2.3 *Suppose $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ satisfies Assumption (G). Then, for all $v \in [0, 1]$, there is a unique càdlàg process $(Z_t(v), t \geq 0)$ on $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{P})$ satisfying \mathbb{P} -a.s.*

$$\begin{cases} Z_t(v) = \sum_{s_i=0} w_t^i \mathbb{1}_{(a_i \leq v)} + \sum_{s_i>0} w_t^i \mathbb{1}_{(a_i \leq v)} \mathbb{1}(u_i \leq g(Z_{s_i-}(1))), & t > 0, \\ Z_0(v) = F(v). \end{cases} \tag{2.13}$$

Moreover, we can construct on $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{P})$ a PPP \mathcal{N} with intensity ν given by (2.4) such that Z solves (2.6) with respect to \mathcal{N} .

If $I(1) = 1$, then in the special case of branching mechanism $\psi(\lambda) = \lambda^2$ and constant immigration rate $g \equiv \theta$, the total-mass process $X_t = X_t(1)$ for (2.6) also solves

$$\begin{cases} dX_t = \sqrt{2X_t} dB_t + \theta dt, & t \geq 0, \\ X_0 = F(1). \end{cases}$$

for which Pitman and Yor obtained the excursion representation in their seminal paper [34].

Remark 2.4 The recent monograph [31] by Zenghu Li contains a full theory of this kind of Pitman–Yor type representations for measure-valued branching processes, see in particular Chapter 10. We present a different approach below which shows directly how the different Poisson point processes in (2.6) and in (2.13) are related to each other. The most important feature of our construction is that it relates the excursion construction and the SDE construction on a pathwise level.

Observe that an immediate and interesting corollary of Theorem 2.3 is the following:

Corollary 2.5 *Let g be an immigration mechanism satisfying assumption (G) and let $(\mathbf{X}_t)_{t \geq 0}$ be a solution to (2.6). Then almost surely, \mathbf{X}_t is purely atomic for all $t \geq 0$.*

In the proof of our Theorem 1.2 we make use of the fact that the Pitman–Yor type representation is well suited for comparison arguments. If g can be bounded from above or below by a constant, then the righthand side of (2.6) can be compared to an explicit PPP for which general theory can be applied.

2.3 From MBI to Beta-Fleming–Viot processes with mutations

Let us first recall an important characterization started in [6] and completed in [12] which relates Fleming–Viot processes, defined as measure-valued Markov processes by the generator (1.3), and strong solutions to stochastic equations.

Theorem 2.6 (Dawson and Li [12]) *Let Λ be the Beta distribution with parameters $(2 - \alpha, \alpha)$. Suppose $\theta \geq 0$ and \mathcal{M} is a non-compensated Poisson point process on $(0, \infty) \times [0, 1] \times [0, 1]$ with intensity $ds \otimes y^{-2} \Lambda(dy) \otimes du$. Then there is a unique strong solution $(Y_t(v))_{t \geq 0, v \in [0, 1]}$ to*

$$\begin{cases} Y_t(v) = v + \int_{]0, t] \times [0, 1] \times [0, 1]} y [\mathbb{1}_{(u \leq Y_{s-}(v))} - Y_{s-}(v)] \mathcal{M}(ds, dy, du) + \theta \int_0^t [v - Y_s(v)] ds, \\ v \in [0, 1], t \geq 0, \end{cases} \tag{2.14}$$

and the measure-valued process $\mathbf{Y}_t([0, v]) := Y_t(v)$ is an (α, θ) -Fleming–Viot process started at uniformly distributed initial condition.

Existence and uniqueness of solutions for this equation was proved in Theorem 4.4 of [12] while the characterization of the generator of the measure-valued process \mathbf{Y} is the content of their Theorem 4.9.

We next extend a classical relation between Fleming–Viot processes and measure-valued branching processes which is typically known as disintegration formula. Without mutations, for the standard Fleming–Viot process this goes back to Konno and Shiga [26] and it was shown in Birkner et al. [10] that the relation extends to the generalized Λ -Fleming–Viot processes without immigration if and only if Λ is a Beta-measure. Our extension relates (α, θ) -Fleming–Viot processes to (2.6) with immigration mechanism $g(x) = \alpha(\alpha - 1)\Gamma(\alpha)\theta x^{2-\alpha}$ and for $\theta = 0$ gives an SDE formulation of the main result of [10].

Theorem 2.7 *Let $F(v) = I(v) = v$ and let $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ be defined by $g(x) = \alpha(\alpha - 1)\Gamma(\alpha)\theta x^{2-\alpha}$ for some $\alpha \in (1, 2)$. Let then $(\mathbf{X}_t)_{t \geq 0}$ be the unique solution of to (2.6) (in the sense of Definition 2.1) such that*

$$X_t(1) = 0, \quad \forall t \geq T_0 := \inf\{s > 0 : X_s(1) = 0\}.$$

Define

$$S(t) = \alpha(\alpha - 1)\Gamma(\alpha) \int_0^t X_s(1)^{1-\alpha} ds$$

and

$$Y_t(dv) = \frac{X_{S^{-1}(t)}(dv)}{X_{S^{-1}(t)}(1)}, \quad t \geq 0. \tag{2.15}$$

Then $(Y_t)_{t \geq 0}$ is well-defined, i.e. $S^{-1}(t) < T_0$ for all $t \geq 0$, and is an (α, θ) -Fleming–Viot process, i.e. a strong solution to (2.14) with $\Lambda = \text{Beta}(2 - \alpha, \alpha)$.

The proof of the theorem is different from the known result for $\theta = 0$. To prove that $X_{S^{-1}(t)}(1) > 0$ for all $t \geq 0$, Lamperti’s representation for CSBPs was crucially used in [10]. This idea breaks down in our generalized setting since the total-mass process $X_t(1)$ is not a CSBP. Our proof uses instead the fact that for all $\theta \geq 0$ the total-mass process is self-similar and an interesting cancellation effect of Lamperti’s transformation for self-similar Markov processes and the time-change S .

In [1] we study (a generalized version of) the total mass process $(X_t(1), t \geq 0)$ and we show that the extinction time $T_0 = \inf\{t \geq 0 : X_t(1) = 0\}$ is finite almost surely if and only if $\theta < \Gamma(\alpha)$. Otherwise $T_0 = \infty$ almost surely. We will see in the proof of Theorem (2.7) that in both cases

$$\lim_{t \rightarrow \infty} S^{-1}(t) = T_0 \text{ a.s.}$$

Theorem 2.7 thus gives some partial information on the behavior of $(X_t)_{t \geq 0}$ near the extinction time T_0 :

Corollary 2.8 *As $t \rightarrow \infty$ the probability-valued process $(\frac{X_{S^{-1}(t)}(dv)}{X_{S^{-1}(t)}(1)})_{t \geq 0}$ converges weakly to the unique invariant measure of $(Y_t, t \geq 0)$.*

As $t \rightarrow T_0$, almost surely, there exists a (random) sequence of times $t_1 < t_2 < \dots < T_0$ tending to T_0 such that the sets

$$A_i = \text{support of } X_{t_i}$$

are pairwise disjoint.

This corollary is a direct consequence of the result, due to Donnelly and Kurtz [13, 14], that the (α, θ) -Fleming–Viot process (as well as its lookdown particle system) is strongly ergodic and of Theorem 2.7. For the sake of self-containdness, a sketch of the proof is given in Sect. 7 which specialize and explicit the arguments of Donnelly and Kurtz to our case.

3 Proof of Theorems 2.2 and 2.3

Recall that $(s_i, u_i, a_i, w^i)_{i \in I}$ is a Poisson point process on $\mathbb{R}_+^3 \times \mathcal{D}$ with intensity measure Γ given as in (2.12), and that we use the notation (2.11). We are going to show that for all $v \in [0, 1]$ there exists a unique càdlàg process $(Z_t(v), t \geq 0)$ solving

$$\begin{cases} Z_t(v) = \sum_{s_i=0} w_t^i \mathbb{1}_{(a_i \leq v)} + \sum_{s_i>0} w_t^i \mathbb{1}_{(a_i \leq v)} \mathbb{1}_{(u_i \leq g(Z_{s_i-}(1)))}, & t > 0, \\ Z_0 = F(v). \end{cases} \tag{3.1}$$

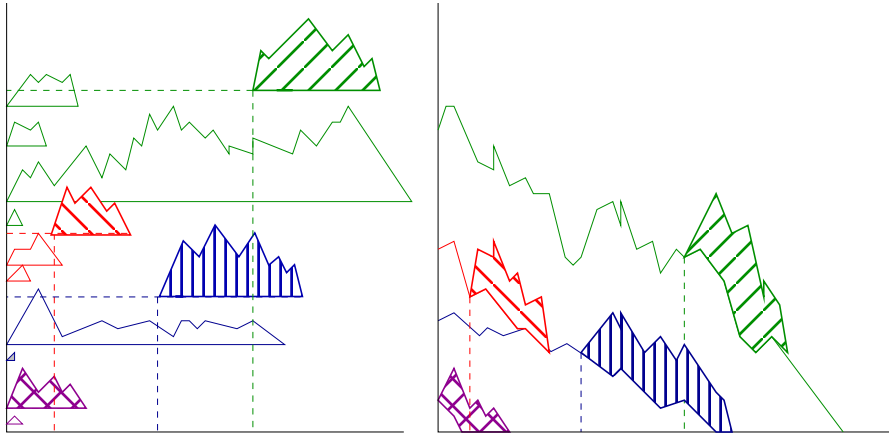


Fig. 1 Definition of \mathcal{N} . On the *left-hand side* we represent the point process (s_i, w^i, a_i) . Observe that $s_4 = 0$ while $s_1, s_2, s_3 > 0$. On the *right-hand side* we show how the w^i are combined to construct the noise \mathcal{N}

Then we are going to construct a PPP \mathcal{N} with intensity $dr \otimes c_\alpha \mathbb{1}(x > 0) x^{-1-\alpha} dx \otimes dy$ such that, for all $v \in [0, 1]$, Z is solution of (2.6).

3.1 The Pitman–Yor type representation with predictable random immigration

We start by replacing the immigration rate $(g(Z_{s-}(1)))_{s>0}$ in the right-hand side of (3.1) with a generic (\mathcal{F}_t) -predictable process $(V_s)_{s \geq 0}$, that we assume to satisfy

$$V_t \geq 0 \text{ and } \int_0^t \mathbb{E}(V_s) ds < +\infty \quad \forall t \geq 0; \tag{3.2}$$

this will be useful when we perform a Picard iteration in the proof of existence of solutions to (2.6) and (3.1). Then we consider

$$\begin{cases} Z_t(v) := \sum_{s_i=0} w_t^i \mathbb{1}(a_i \leq v) + \sum_{s_i>0} w_t^i \mathbb{1}(a_i \leq v) \mathbb{1}(u_i \leq V_{s_i}), & t > 0, v \in [0, 1], \\ Z_0(v) := F(v), & v \in [0, 1]. \end{cases} \tag{3.3}$$

Then we want to show that there is a noise \mathcal{N} on $(\Omega, \mathcal{G}, \mathcal{G}_\square, \mathbb{P})$ such that Z is a solution of an equation of the type (2.6).

3.1.1 Definition of \mathcal{N}

Let us consider a family of independent random variables $(U_{ij})_{i,j \in \mathbb{N}}$ such that U_{ij} is uniform on $[0, 1]$ for all $i, j \in \mathbb{N}$. We also assume that $(U_{ij})_{i,j \in \mathbb{N}}$ is independent

of the PPP (s_i, u_i, a_i, w^i) . Then, for all atoms (s_i, u_i, a_i, w^i) in the above PPP, we define the following point process $\mathcal{N}^i := (r_j^i, x_j^i, y_j^i)_{j \in J^i}$:

- (1) $(r_j^i)_{j \in J^i}$ is the family of jump times of $r \mapsto w_r^i$;
- (2) for each r_j^i we set

$$x_j^i := w_{r_j^i}^i - w_{r_j^i-}^i, \quad y_j^i := w_{r_j^i-}^i \cdot U_{ij}. \tag{3.4}$$

We note that \mathcal{N}^i is not expected to be a Poisson point process. For each $k \in \mathbb{N}$ we set

$$\begin{aligned} L_0^k &:= F(a_k) \text{ and } L_t^k := \sum_{a_i < a_k, u_i \leq V_{s_i}} w_{t-}^i, \quad t > 0, \\ L_t^\infty &:= \sup_k L_t^k, \quad t \geq 0. \end{aligned} \tag{3.5}$$

We consider a PPP $\mathcal{N}^\circ = (r_j^\circ, x_j^\circ, y_j^\circ)_j$ with intensity measure ν given by (2.4) and independent of $((s_i, u_i, a_i, w^i)_i, (U_{ij})_{i,j \in \mathbb{N}}, (V_t)_{t \geq 0})$. We set for any non-negative measurable $f = f(r, x, y)$

$$\begin{aligned} \int f d\mathcal{N} &:= \sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \int f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \\ &+ \int f(r, x, y + L_r^\infty) \mathcal{N}^\circ(dr, dx, dy). \end{aligned} \tag{3.6}$$

The filtration we are going to work with is

$$\begin{aligned} \mathcal{F}_t &:= \sigma((s_i, u_i, a_i, w_r^i, U_{ij}, V_r), (r_i^\circ, x_i^\circ, y_i^\circ)_i : \\ &r \leq t, s_i \leq t, r_i^\circ \leq t, i, j \in \mathbb{N}), \quad t \geq 0. \end{aligned}$$

We are going to prove the following

Proposition 3.1 \mathcal{N} is a PPP with intensity $\nu(dr, dx, dy) = dr \otimes c_\alpha x^{-1-\alpha} dx \otimes dy$.

Proof For $f = f(r, x, y) \geq 0$ we now set

$$I(t) := \sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \int_{]0, t] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy).$$

Since $w_t^i = 0$ if $s_i \geq t$, V is predictable and we can write

$$L_t^k := \sum_{s_i=0} \mathbb{1}_{(a_i < a_k)} w_{t-}^i + \sum_{s_i > 0, u_i \leq V_{s_i}} \mathbb{1}_{(a_i < a_k)} w_{t-}^i,$$

then we obtain that $(L_t^k)_k$ is predictable. Hence, $I(t)$ is \mathcal{F}_t -measurable and for $0 \leq t < T$

$$\begin{aligned} \mathbb{E} (I(T) - I(t) \mid \mathcal{F}_t) &= \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \mid \mathcal{F}_t \right) \\ &= \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k < t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \mid \mathcal{F}_t \right) \\ &\quad + \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \mid \mathcal{F}_t \right) \end{aligned}$$

We will need the following two facts:

- (1) Conditionally on w_t^k and $s_k < t$ the process $w_{\cdot, +t}^k$ has law $P_{w_t^k}$ (this follows for instance from (2.7)).
- (2) Let $(w_t, t \geq 0)$ be a CSBP started from w_0 with law \mathbb{P}_{w_0} . Let $\mathcal{M} = (r_i, x_i, y_i)$ be a point process which is defined from w and a sequence of i.i.d. uniform variables on $[0, 1]$ as \mathcal{N}^k is constructed from w^k and $(U_{ij})_{i,j \in \mathbb{N}}$. Then for any positive function $f = f(r, x, y)$

$$\begin{aligned} &\mathbb{E} \left[\int_{[0, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y) \mathcal{M}(dr, dx, dy) \right] \\ &= \mathbb{E}_{w_0} \left[\int_{[0, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y) \mathbf{1}_{y \leq w_{r-}} \nu(dr, dx, dy) \right]. \end{aligned}$$

Let us start with the case $s_k < t$. Using the above facts we see that

$$\begin{aligned} &\mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k < t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \mid \mathcal{F}_t \right) \\ &= \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k < t)} \mathbb{E} \left[\int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \mid w_t^k, L_r^k \right] \mid \mathcal{F}_t \right) \\ &= \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k < t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbf{1}_{[L_r^k, L_r^k + w_{r-}^k](y)} f(r, x, y) dr \frac{c_\alpha}{x^{1+\alpha}} dx dy \mid \mathcal{F}_t \right) \end{aligned}$$

Let us now consider the case $s_k > t$.

$$\begin{aligned}
 & \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathcal{N}^k(dr, dx, dy) \middle| \mathcal{F}_t \right) \\
 &= \lim_{\epsilon \rightarrow 0} \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathbb{1}_{(s_k + \epsilon < r)} \mathcal{N}^k(dr, dx, dy) \middle| \mathcal{F}_t \right) \\
 &= \lim_{\epsilon \rightarrow 0} \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \mathbb{E} \left[\int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y + L_r^k) \mathbb{1}_{(s_k + \epsilon < r)} \mathcal{N}^k(dr, dx, dy) \middle| w_{s_k + \epsilon}^k, L_r^k \right] \middle| \mathcal{F}_t \right) \\
 &= \lim_{\epsilon \rightarrow 0} \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{[L_r^k, L_r^k + w_{s_k + \epsilon}^k](y)} f(r, x, y) \mathbb{1}_{(s_k + \epsilon < r)} dr \frac{c_\alpha}{x^{1+\alpha}} dx dy \middle| \mathcal{F}_t \right) \\
 &= \mathbb{E} \left(\sum_k \mathbb{1}_{(u_k \leq V_{s_k})} \mathbb{1}_{(s_k \geq t)} \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{[L_r^k, L_r^k + w_{s_k}^k](y)} f(r, x, y) dr \frac{c_\alpha}{x^{1+\alpha}} dx dy \middle| \mathcal{F}_t \right)
 \end{aligned}$$

where we need to introduce the indicator that $r > s_k + \epsilon$ to get a sum of CSBP started from a positive initial mass and thus be in a position to apply the above fact.

We conclude that

$$\mathbb{E} (I(T) - I(t) \mid \mathcal{F}_t) = \mathbb{E} \left(\int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{]0, \sup_k L_r^k](y)} f(r, x, y) dr \frac{c_\alpha}{x^{1+\alpha}} dx dy \middle| \mathcal{F}_t \right),$$

Therefore by the Definition (3.6) of \mathcal{N}

$$\begin{aligned}
 & \mathbb{E} \left(\int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y) \mathcal{N}(dr, dx, dy) \middle| \mathcal{F}_t \right) \\
 &= \mathbb{E} \left(\int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} (\mathbb{1}_{]0, L_r^\infty[}(y) + \mathbb{1}_{]L_r^\infty, \infty[}(y)) f(r, x, y) dr \frac{c_\alpha}{x^{1+\alpha}} dx dy \middle| \mathcal{F}_t \right). \\
 &= \int_{]t, T] \times \mathbb{R}_+ \times \mathbb{R}_+} f(r, x, y) dr \frac{c_\alpha}{x^{1+\alpha}} dx dy.
 \end{aligned}$$

By [23, Theorem II.6.2], a point process with deterministic compensator is necessarily a Poisson point process, and therefore the proof is complete. □

Proposition 3.1 tells us how to construct a Poisson noise \mathcal{N} from the (s_i, u_i, a_i, w^i) . Let us now show that Z solves (2.6) with this particular noise.

Proposition 3.2 *Let Z satisfy (3.3). Then for all $v \geq 0$, $(Z(v), \mathcal{N})$ satisfies \mathbb{P} -a.s.*

$$Z_t(v) = F(v) + \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{(y < Z_{r-}(v))} x \tilde{\mathcal{N}}(dr, dx, dy) + I(v) \int_0^t V_s ds, \quad t \geq 0.$$

Proof Using an idea introduced by Dawson and Li [11], we set for $n \in \mathbb{N}^*$

$$Z_t^n(v) := \sum_{a_i \leq v, u_i \leq V_{s_i}} w_t^i \mathbb{1}_{(s_i + \frac{1}{n} \leq t)}. \tag{3.7}$$

Note that $\mathbb{Q}(\{w_{1/n} > 0\}) < +\infty$ for all $n \geq 1$, so that Z_t^n is \mathbb{P} -a.s. given by a finite sum of terms. Moreover, by the properties of PPPs, $((s_i, u_i, a_i, w^i) : w_{1/n}^i > 0)$ is a PPP with intensity $(\delta_0(ds) \otimes \delta_0(du) \otimes F(da) + ds \otimes du \otimes I(da)) \otimes \mathbb{1}_{(w_{1/n} > 0)} \mathbb{Q}(dw)$. Moreover $Z_t^n \uparrow Z_t$ as $n \uparrow +\infty$ for all $t \geq 0$. Now we can write

$$Z_t^n(v) = M_t^n(v) + J_t^n(v),$$

with

$$\begin{aligned} M_t^n(v) &:= \sum_{a_i \leq v, u_i \leq V_{s_i}} (w_t^i - w_{s_i + \frac{1}{n}}^i) \mathbb{1}_{(s_i + \frac{1}{n} \leq t)}, \\ J_t^n(v) &:= \sum_{a_i \leq v, u_i \leq V_{s_i}} w_{s_i + \frac{1}{n}}^i \mathbb{1}_{(s_i + \frac{1}{n} \leq t)}. \end{aligned} \tag{3.8}$$

Let us concentrate on M^n first. We can write, for $s_i + \frac{1}{n} \leq t$,

$$w_t^i - w_{s_i + \frac{1}{n}}^i = \int_{[s_i + \frac{1}{n}, t] \times \mathbb{R}^+ \times \mathbb{R}^+} \mathbb{1}_{(y < w_{r-}^i)} x \tilde{\mathcal{N}}^i(dr, dx, dy)$$

where \mathcal{N}^i is defined in (3.4) and $\tilde{\mathcal{N}}^i(dr, dx, dy)$ is the compensated version of \mathcal{N}^i :

$$\tilde{\mathcal{N}}^i(dr, dx, dy) := \mathcal{N}^i(dr, dx, dy) - \mathbb{1}_{(y < w_{r-}^i)} \nu(dr, dx, dy),$$

with ν defined in (2.4). We set

$$A_{i,n} := \left\{ (y, r) : L_r^i \leq y < L_r^i + w_{r-}^i \mathbb{1}_{(s_i + \frac{1}{n} \leq r)} \right\}, \quad B_n^v := \bigcup_{a_i \leq v, u_i \leq V_{s_i}} A_{i,n}.$$

Since $\mathbb{Q}(\{w_{1/n} > 0\}) < +\infty$, only finitely many $\{A_{i,n}\}_i$ such that $u_i \leq V_{s_i}$ are non-empty \mathbb{P} -a.s and, moreover, the $\{A_{i,n}\}_i$ are disjoint. Then by (3.6)

$$\begin{aligned} & \int_{]0,t] \times \mathbb{R}^+ \times \mathbb{R}^+} \mathbb{1}_{A_{i,n}}(y,r) x \tilde{\mathcal{N}}(dr, dx, dy) \\ &= \mathbb{1}_{(s_i + \frac{1}{n} \leq t)} \int_{[s_i + \frac{1}{n}, t] \times \mathbb{R}^+ \times \mathbb{R}^+} \mathbb{1}_{(y < w_{r-}^i)} x \tilde{\mathcal{N}}^i(dr, dx, dy) \\ &= \left(w_t^i - w_{s_i + \frac{1}{n}}^i \right) \mathbb{1}_{(s_i + \frac{1}{n} \leq t)} \end{aligned}$$

so that

$$\begin{aligned} \int_{]0,t] \times \mathbb{R}^+ \times \mathbb{R}^+} \mathbb{1}_{B_n^v}(y,r) x \tilde{\mathcal{N}}(dr, dx, dy) &= \sum_{a_i \leq v, u_i \leq V_{s_i}} \left(w_t^i - w_{s_i + \frac{1}{n}}^i \right) \mathbb{1}_{(s_i + \frac{1}{n} \leq t)} \\ &= M_t^n(v). \end{aligned}$$

We need first the two following technical lemmas. □

Lemma 3.3 For a (\mathcal{F}_t) -predictable bounded process $f_t : \mathbb{R}_+ \mapsto \mathbb{R}$ we set

$$M_t := \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} f_r(y) x \tilde{\mathcal{N}}(dr, dx, dy), \quad t \geq 0.$$

Then we have

$$\mathbb{E} \left(\sup_{t \in [0,T]} |M_t| \right) \leq C \left(\sqrt{\int_{]0,T] \times \mathbb{R}_+} \mathbb{E}(f_r^2(y)) dr dy} + \int_{]0,T] \times \mathbb{R}_+} \mathbb{E}(|f_r(y)|) dr dy \right).$$

Proof Recall that $\nu_\alpha(dx) = c_\alpha x^{-1-\alpha} dx$. We set

$$\begin{aligned} J_{1,t} &:= \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} f_r(y) \mathbb{1}_{(x \leq 1)} x \tilde{\mathcal{N}}(dr, dx, dy), \quad t \geq 0, \\ J_{2,t} &:= \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} f_r(y) \mathbb{1}_{(x > 1)} x \tilde{\mathcal{N}}(dr, dx, dy), \quad t \geq 0. \end{aligned}$$

Then, by Doob’s inequality,

$$\begin{aligned} \left(\mathbb{E} \left(\sup_{t \in [0,T]} |J_{1,t}| \right) \right)^2 &\leq \mathbb{E} \left(\sup_{t \in [0,T]} |J_{1,t}|^2 \right) \\ &\leq 4 \int_{]0,1]} c_\alpha x^{1-\alpha} dx \int_{]0,T] \times \mathbb{R}_+} \mathbb{E}(f_r^2(y)) dr dy \end{aligned}$$

while

$$\mathbb{E} \left(\sup_{t \in [0, T]} |J_{2,t}| \right) \leq 2 \int_{1, \infty[} c_\alpha x^{-\alpha} dx \int_{[0, T] \times \mathbb{R}_+} \mathbb{E}(|f_r(y)|) dr dy.$$

□

Lemma 3.4 (1) $\lim_{n \rightarrow \infty} \int_{\mathcal{E}} (z_{\frac{1}{n}})^2 \mathbf{1}_{(z_{\frac{1}{n}} \leq 1)} \mathbb{Q}(dz) = 0.$

(2) $\lim_{n \rightarrow \infty} \int_{\mathcal{E}} z_{\frac{1}{n}} \mathbf{1}_{(z_{\frac{1}{n}} \geq 1)} \mathbb{Q}(dz) = 0.$

Proof First recall from (2.9) that $\int_{\mathcal{E}} z_{\frac{1}{n}} \mathbb{Q}(dz) = 1$ for all n . The proof of (1) is based on the estimate $\frac{1}{e}x \leq 1 - e^{-x}$ for $x \in [0, 1]$ which follows from differentiating both sides. Of course, the inequality also implies that

$$x^2 \mathbf{1}_{(x \leq 1)} \leq ex(1 - e^{-x}), \quad x \geq 0.$$

We apply this estimate to the excursion measure:

$$\begin{aligned} \int_{\mathcal{E}} (z_{\frac{1}{n}})^2 \mathbf{1}_{(z_{\frac{1}{n}} \leq 1)} \mathbb{Q}(dz) &\leq e \int_{\mathcal{E}} z_{\frac{1}{n}} (1 - e^{-z_{\frac{1}{n}}}) \mathbb{Q}(dz) \\ &= e \left(\int_{\mathcal{E}} z_{\frac{1}{n}} \mathbb{Q}(dz) - \int_{\mathcal{E}} z_{\frac{1}{n}} e^{-z_{\frac{1}{n}}} \mathbb{Q}(dz) \right). \end{aligned} \tag{3.9}$$

Next, by (2.8),

$$\int_{\mathcal{E}} z_{\frac{1}{n}} e^{-z_{\frac{1}{n}}} \mathbb{Q}(dz) = \frac{d}{d\lambda} u_{1/n}(\lambda) \Big|_{\lambda=1} = \frac{1}{(1 + (\alpha - 1)\frac{1}{n})^{\alpha/(\alpha-1)}} \xrightarrow{n \rightarrow \infty} 1,$$

so that (3.9) combined with $\int_{\mathcal{E}} z_{\frac{1}{n}} \mathbb{Q}(dz) = 1$ proves (1). For (2) we use that $x \mathbf{1}_{(x > 1)} \leq \frac{e}{(e-1)}x(1 - e^{-x})$ to get

$$\int_{\mathcal{E}} z_{\frac{1}{n}} \mathbf{1}_{(z_{\frac{1}{n}} > 1)} \mathbb{Q}(dz) \leq \frac{e}{e-1} \int_{\mathcal{E}} z_{\frac{1}{n}} (1 - e^{-z_{\frac{1}{n}}}) \mathbb{Q}(dz)$$

which goes to zero as argued above. □

Lemma 3.5 For all $v \geq 0$ and $T \geq 0$ we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\sup_{t \in [0, T]} |Z_t(v) - Z_t^n(v)| \right) = 0,$$

$(Z_t(v))_{t \geq 0}$ is \mathbb{P} -a.s. càdlàg and \mathbb{P} -a.s.

$$Z_t(v) = F(v) + \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{(y < Z_{r-}(v))} x \tilde{\mathcal{N}}(dr, dx, dy) + I(v) \int_0^t V_s ds.$$

Proof We have obtained above the representation

$$Z_t^n(v) = \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{B_n^v}(y, r) x \tilde{\mathcal{N}}(dr, dx, dy) + J_t^n(v). \tag{3.10}$$

First, let us note that

$$B_n^v \subset B^v := \bigcup_{a_i \leq v, u_i \leq V_{s_i}} A_i, \quad A_i := \left\{ (y, r) : L_r^i \leq y < L_r^i + w_{r-}^i \right\},$$

and moreover

$$B^v \setminus B_n^v = \bigcup_{a_i \leq v} \left\{ (y, r) : L_r^i \leq y < L_r^i + w_{r-}^i \mathbb{1}_{(s_i + \frac{1}{n} > r)} \right\}$$

and the latter union is disjoint. If we set

$$M_t(v) := \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{B^v}(y, r) x \tilde{\mathcal{N}}(dr, dx, dy),$$

then

$$M_t(v) - M_t^n(v) = \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{B^v \setminus B_n^v}(y, r) x \tilde{\mathcal{N}}(dr, dx, dy)$$

and by Lemma 3.3

$$\begin{aligned} \frac{1}{C} \mathbb{E} \left(\sup_{t \in [0, T]} |M_t - M_t^n| \right) &\leq \sqrt{\int_0^T \mathbb{E}(\mathbb{1}_{B^v \setminus B_n^v}(y, r)) dr dy} + \int_0^T \mathbb{E}(\mathbb{1}_{B^v \setminus B_n^v}(y, r)) dr dy \\ &= \sqrt{\mathbb{E} \left(\sum_{a_i \leq v, u_i \leq V_{s_i}} \int_{s_i \wedge T}^{(s_i + \frac{1}{n}) \wedge T} w_r^i dr \right)} + \mathbb{E} \left(\sum_{a_i \leq v, u_i \leq V_{s_i}} \int_{s_i \wedge T}^{(s_i + \frac{1}{n}) \wedge T} w_r^i dr \right). \end{aligned} \tag{3.11}$$

Then we get

$$\begin{aligned}
 & \mathbb{E} \left(\sum_{a_i \leq v, u_i \leq V_{s_i}} \int_{s_i \wedge T}^{(s_i + \frac{1}{n}) \wedge T} w_r^i dr \right) \\
 &= \mathbb{E} \left(\sum_{a_i \leq v} \mathbb{1}_{(s_i=0)} \int_{s_i \wedge T}^{(s_i + \frac{1}{n}) \wedge T} w_r^i dr \right) + \mathbb{E} \left(\sum_{a_i \leq v, u_i \leq V_{s_i}} \mathbb{1}_{(s_i > 0)} \int_{s_i \wedge T}^{(s_i + \frac{1}{n}) \wedge T} w_r^i dr \right) \\
 &= \int_0^{T \wedge \frac{1}{n}} F(v) \int_{\mathcal{E}} w_r \mathbb{Q}(dw) dr + \int_0^T \mathbb{E}(V_s I(v)) \int_s^{(s + \frac{1}{n}) \wedge T} \int_{\mathcal{E}} w_{r-s} \mathbb{Q}(dw) dr ds \\
 &= F(v) \left(T \wedge \frac{1}{n} \right) + \int_0^T \mathbb{E}(V_s I(v)) \left(T \wedge \left(s + \frac{1}{n} \right) - s \right) ds,
 \end{aligned}$$

where the last equality follows by (2.9). By our assumptions on V the right hand side in the above display converges to 0, as $n \rightarrow \infty$. Hence (3.11) also converges to 0, as $n \rightarrow \infty$. Let us now deal with $(J_t^n)_{\geq 0}$. Note that we can write

$$J_{t+\frac{1}{n}}^n(v) = \sum_{a_i \leq v, u_i \leq V_{s_i}} w_{s_i+\frac{1}{n}}^i \mathbb{1}_{(s_i \leq t)} = A_t^n + \sum_{a_i \leq v} w_{\frac{1}{n}}^i \mathbb{1}_{(s_i=0)} + I(v) \int_0^t V_s ds,$$

where

$$A_t^n := \sum_{0 < s_i \leq t} w_{s_i+\frac{1}{n}}^i \mathbb{1}_{(a_i \leq v, u_i \leq V_{s_i})} - I(v) \int_0^t V_s ds,$$

and $(A_t^n)_{t \geq 0}$ is a martingale such that $A_0^n = 0$. We have by an analog of Lemma 3.3 and its proof

$$\mathbb{E} \left(\sup_{t \in [0, T]} |A_t^n| \right) \leq 2 \sqrt{K_V \int_{\mathcal{E}} (z_{\frac{1}{n}})^2 \mathbb{1}_{(z_{\frac{1}{n}} \leq 1)} \mathbb{Q}(dz)} + 2K_V \int_{\mathcal{E}} z_{\frac{1}{n}} \mathbb{1}_{(z_{\frac{1}{n}} > 1)} \mathbb{Q}(dz),$$

where $K_V := I(v) \int_0^T \mathbb{E}(V_s) ds$. The righthand side tends to zero as $n \rightarrow \infty$ by Lemma

3.4. Analogously

$$\mathbb{E} \left(\left| \sum_{a_i \leq v} w_{\frac{1}{n}}^i \mathbb{1}_{(s_i=0)} - F(v) \right| \right)$$

$$\leq 2 \sqrt{F(1) \int_{\mathcal{E}} (z_{\frac{1}{n}})^2 \mathbb{1}_{(z_{\frac{1}{n}} \leq 1)} \mathbb{Q}(dz) + 2F(1) \int_{\mathcal{E}} z_{\frac{1}{n}} \mathbb{1}_{(z_{\frac{1}{n}} > 1)} \mathbb{Q}(dz),}$$

which again tends to 0 as $n \rightarrow \infty$ by Lemma 3.4. Therefore

$$\mathbb{E} \left[\sup_{t \in [0, T]} |Z_t(v) - Z_t^n(v)| \right] \rightarrow 0.$$

and, passing to a subsequence, we see that a.s.

$$\sup_{t \in [0, T]} |Z_t(v) - Z_t^{n_k}(v)| \rightarrow 0 \tag{3.12}$$

(observe that in fact we don't need to take a subsequence since Z_t^n is monotone non-decreasing in n).

In particular, a.s. $(Z_t(v), t \geq 0)$ is càdlàg and we obtain

$$Z_t(v) = F(v) + \int_{]0, t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{B^v}(y, r) x \tilde{\mathcal{N}}(dr, dx, dy) + I(v) \int_0^t V_s ds.$$

It remains to prove that a.s. $B^v = \{(y, r) : y < Z_{r-}(v)\}$. By definition a.s.

$$Z_{r-}(v) = \sum_{a_i \leq v, u_i \leq V_{s_i}} w_{r-}^i, \quad r \geq 0.$$

If $a_i \leq v$ and $u_i \leq V_{s_i}$, then $L_r^i + w_{r-}^i \leq Z_{r-}(v)$, so that $B^v \subset \{(y, r) : y < Z_{r-}(v)\}$. On the other hand, if $y < Z_{r-}(v)$, then there is one j such that

$$\sum_{a_i < a_j, u_i \leq V_{s_i}} w_{r-}^i = L_r^j \leq y < L_r^j + w_{r-}^j.$$

Therefore we have obtained the desired results. □

The proof of Proposition 3.2 is complete.

3.2 Proof of Theorem 2.3

With a localisation argument we can suppose that g is globally Lipschitz. Let us first show uniqueness of solutions to (3.1). Let $v = 1$. If $(Z_t^i, t \geq 0)$ for $i = 1, 2$ is a càdlàg process satisfying (3.1) with $v = 1$, then taking the difference we obtain

$$\begin{aligned} \mathbb{E}(|Z_t^1 - Z_t^2|) &\leq I(1) \int_0^t ds \int_{\mathcal{E}} z_{t-s} \mathbb{Q}(dz) \mathbb{E}(|g(Z_s^1) - g(Z_s^2)|) \\ &= I(1) \int_0^t ds \mathbb{E}(|g(Z_s^1) - g(Z_s^2)|), \end{aligned}$$

where the second equality follows by (2.9). By the Lipschitz-continuity of g and the Gronwall Lemma we obtain $Z^1 = Z^2$ a.s., i.e. uniqueness of solutions to (3.1).

The next step is to use an iterative Picard scheme in order to construct a solution of (3.1) (and thus of (2.6)). Let $v := 1$, and let us set $Z_t^0 := 0$ and for all $n \geq 0$

$$Z_t^{n+1} := \sum_{s_i=0} w_t^i \mathbb{1}_{(u_i \leq 1)} + \sum_{s_i>0} w_t^i \mathbb{1}_{(u_i \leq g(Z_{s_i-}^n))}, \quad t \geq 0.$$

By recurrence and monotonicity of g , $Z_t^{n+1} \geq Z_t^n$ and therefore a.s. there exists the limit $Z_t := \lim_n Z_t^n$.

To show that Z is actually the solution of (3.1) we show first that it is càdlàg (by proving that the convergence holds in a norm that makes the space of càdlàg processes on $[0, T]$ complete) and then by proving that (3.1) holds almost surely for each fixed $t \geq 0$. Let us first show that Z^n is a Cauchy sequence for the norm $\|Z\| = \mathbb{E}(\sup_{t \in [0, T]} |Z_t|)$ for which first we set

$$Z_t^{n,k} := Z_t^{n+k+1} - Z_t^{n+1} = \sum_{s_i>0} w_t^i \mathbb{1}_{(g(Z_{s_i-}^n) < u_i \leq g(Z_{s_i-}^{n+k}))}.$$

By an analog of Proposition 3.2 we can construct a PPP $\mathcal{N}^{n,k}$ with the intensity measure $\mathbb{1}_{(r>0)} dr \otimes c_\alpha \mathbb{1}_{(x>0)} x^{-1-\alpha} dx \otimes \mathbb{1}_{(y>0)} dy$ such that for all $t \geq 0$

$$Z_t^{n,k} = \int_0^t \mathbb{1}_{(y < Z_{r-}^{n,k})} x \tilde{\mathcal{N}}^{n,k}(dr, dx, dy) + I(1) \int_0^t [g(Z_s^{n+k}) - g(Z_s^n)] ds.$$

Then by the Lipschitz-continuity of g with the Lipschitz constant L , and by Lemma 3.3

$$\mathbb{E} \left(\sup_{t \in [0, T]} |Z_t^{n,k}| \right) \leq \sqrt{\int_0^T \mathbb{E} (Z_s^{n,k}) ds} + \int_0^T \mathbb{E} (Z_s^{n,k}) ds + I(1)L \int_0^T \mathbb{E} (Z_s^{n-1,k}) ds.$$

We show now that the right hand side in the latter formula vanishes as $n \rightarrow +\infty$ uniformly in k . Indeed

$$\mathbb{E}(Z_t^0) = 0, \quad 0 \leq \mathbb{E}(Z_t^{n+1}) = F(1) + \int_0^t \mathbb{E}(g(Z_s^n)) ds \leq C + L \int_0^t \mathbb{E}(Z_s^n) ds.$$

Then by recurrence $\mathbb{E}(Z_t^{n+1}) \leq Ce^{tL}$ and by monotone convergence we obtain that $\mathbb{E}(Z_t^{n+1}) \uparrow \mathbb{E}(Z_t) \leq Ce^{tL}$. By dominated convergence it follows that

$$\int_0^T \mathbb{E}(Z_s^n) ds \rightarrow \int_0^T \mathbb{E}(Z_s) ds,$$

i.e. the sequence $\int_0^T \mathbb{E}(Z_s^n) ds$ is Cauchy and we conclude that $Z_n \rightarrow Z$ in the sense of the above norm and therefore Z is almost surely càdlàg. The above argument also show that

$$Z_t = \sum_{s_i=0} w_t^i \mathbb{1}_{(u_i \leq 1)} + \sum_{s_i>0} w_t^i \mathbb{1}_{(u_i \leq g(Z_{s_i-}))},$$

holds almost surely for each fixed t and therefore for all $t \geq 0$, i.e. Z is a solution of (3.1) for $v = 1$. Setting $V_s := g(Z_{s-}(1))$ and applying Proposition 3.2, we obtain (3.1) and the proof of Theorem 2.3 is complete.

3.3 Proof of Theorem 2.2

Let us start from existence of a weak solution to (2.6); by Theorem 2.3 we can build a process $(Z_t(v), t \geq 0, v \in [0, 1])$ and a Poisson point process $\mathcal{N}(dr, dx, dy)$ such that (3.1) and (2.6) hold. Now, we set

$$\mathbf{X}_t := \sum_{s_i=0} w_t^i \delta_{u_i} + \sum_{s_i>0} w_t^i \mathbb{1}_{(g(Z_{s_i-}(1))>0)} \delta_{u_i/g(Z_{s_i-}(1))},$$

where δ_a denote the Dirac mass at a ; by construction it is clear that $X_t(v) := \mathbf{X}_t([0, v])$, for all $v \in [0, 1]$, is a solution to (2.6). It remains to prove that $(\mathbf{X}_t)_{t \geq 0}$ is càdlàg in the space of finite measures on the space $[0, 1]$. By Lemma 3.5, for all $v \in [0, 1]$, $(X_t(v))_{t \geq 0}$ is càdlàg; by countable additivity, a.s. $(X_t(v))_{t \geq 0}$ is càdlàg for all $v \in \mathbb{Q} \cap [0, 1]$; then, by the compactness of $[0, 1]$, it is easy to see that $(\mathbf{X}_t)_{t \geq 0}$ is càdlàg: for instance, a.s. any limit point of \mathbf{X}_{t_n} , for $t_n \geq t$ and $t_n \rightarrow t$, is equal on each interval $]a, b]$, $a, b \in \mathbb{Q} \cap [0, 1]$, to $X_t(b) - X_t(a) = \mathbf{X}_t(]a, b])$. Therefore, we have proved that $(\mathbf{X}_t)_{t \geq 0}$ is a solution to (2.6) in the sense of Definition 2.1.

It remains to prove pathwise uniqueness. Let $(\mathbf{X}_t^i)_{t \geq 0}, i = 1, 2$, be two solutions to (2.6) driven by the same Poisson noise \mathcal{N} and let us set $X_t^i(v) := \mathbf{X}_t^i([0, v])$, $v \in [0, 1]$. Let us first consider the case $v = 1$: then $(X_t^i(1), t \geq 0), i = 1, 2$, solves a particular case of the equation considered by Dawson and Li [12, (2.1)]; therefore, by [12, Theorem 2.5], $\mathbb{P}(X_t^1(1) = X_t^2(1), \forall t \geq 0) = 1$.

Let us now consider $0 \leq v < 1$; in this case the equation satisfied by $(X_t^i(v), t \geq 0)$ depends on $(X_t^i(1), t \geq 0)$ and therefore the uniqueness result by Dawson and Li does not apply directly. Instead, we consider the difference $D_t := X_t^1(v) - X_t^2(v)$ so that the drift terms cancel since $X^1(1) = X^2(1)$. Hence, $(D_t, t \geq 0)$ can be treated as if g were identically equal to 0. The same proof as in [12] shows that $\mathbb{P}(X_t^1(v) = X_t^2(v), \forall t \geq 0) = 1$. Finally, since a.s. the two finite measures \mathbf{X}_t^1 and \mathbf{X}_t^2 are equal on each interval $]a, b]$, $a, b \in \mathbb{Q} \cap [0, 1]$, they coincide. Therefore, pathwise uniqueness holds for (2.6).

Finally, in order to obtain existence of a strong solution, we apply the classical Yamada-Watanabe argument, for instance in the general form proved by Kurtz [25, Theorem 3.14].

4 Proof of Theorem 2.7

We consider the immigration rate function $g(x) = \alpha(\alpha - 1)\Gamma(\alpha)\theta x^{2-\alpha}$, $x \geq 0$. Now g is not Lipschitz-continuous, so that Theorem 2.3 does not apply directly. However, by considering $g^n(x) = \alpha(\alpha - 1)\Gamma(\alpha)\theta(x \vee n^{-1})^{2-\alpha}$, we obtain a monotone non-decreasing and Lipschitz continuous function for which Theorem 2.3 yields existence and uniqueness of a solution $(X_t^n(v), t \geq 0, v \geq 0)$ to (2.6). We now define $T^0 := 0$, $T^n := \inf\{t > 0 : X_t^n(1) = n^{-1}\}$ and

$$X_t(v) := \sum_{n \geq 1} X_t^n(v) \mathbb{1}_{(T^{n-1} \leq t < T^n)}.$$

Since $(X_t(1))_{t \geq 0}$ has no downward jumps, it follows that $T_0 := \sup_n T^n$ is equal to $\inf\{s > 0 : X_s(1) = 0\}$, and moreover $X_t(1) = 0$ for all $t \geq T_0$. By pathwise uniqueness, if $n \geq m$ then $X_t^n(v) = X_t^m(v)$ on $\{t \leq T^m\}$, and therefore $(X_t(v), t \geq 0, v \geq 0)$ is a solution to (2.6) for $g(x) = \alpha(\alpha - 1)\Gamma(\alpha)\theta x^{2-\alpha}$ with the desired properties. Pathwise uniqueness follows from the same localisation argument.

To prove that the right-hand side of (2.15) is well-defined, i.e. the denominator is always strictly positive, we are going to apply Lamperti’s representation for self-similar Markov process. A positive self-similar Markov process of index w is a strong Markov family $(\mathbb{P}^x)_{x>0}$ with coordinate process denoted by $(U_t)_{t \geq 0}$ in the Skorohod space of càdlàg functions with values in $[0, +\infty[$, satisfying

$$\text{the law of } (cU_{c^{-1/w}t})_{t \geq 0} \text{ under } \mathbb{P}^x \text{ is given by } \mathbb{P}^{cx} \tag{4.1}$$

for all $c > 0$. Lamperti has shown in [29] that this property is equivalent to the existence of a Lévy process ξ such that, under \mathbb{P}^x , the process $(U_{t \wedge T_0})_{t \geq 0}$ has the same law as $(x \exp(\xi_{A^{-1}(tx^{-1/w})}))_{t \geq 0}$, where

$$A^{-1}(t) := \inf\{s \geq 0 : A_s > t\} \quad \text{and} \quad A(t) := \int_0^t \exp\left(\frac{1}{w} \xi_s\right) ds.$$

We now use Lamperti’s representation to find a surprisingly simple argument for the well-posedness of (2.15).

Lemma 4.1 *The right-hand side of (2.15) is well-defined for all $v \in [0, 1]$ and $t \geq 0$.*

Proof In Lemma 1 of [1] it was shown that, if L is a spectrally positive α -stable Lévy process as in (2.3), solutions to the SDE

$$X_t = X_0 + \int_0^t X_{s-}^{1/\alpha} dL_s + \alpha(\alpha - 1)\Gamma(\alpha)\theta \int_0^t X_s^{2-\alpha} ds \tag{4.2}$$

trapped at zero induce a positive self-similar Markov process of index $1/(\alpha - 1)$. The corresponding Lévy process ξ has been calculated explicitly in [1, Lemma 2.2], but for the proof here we only need that ξ has infinite lifetime and additionally a remarkable cancellation effect between the time-changes. Since, by Lemma 1 of Fournier [21], the unique solution to the SDE (4.2) for $X_0 = 1$ coincides in law with the unique solution to

$$X_t = 1 + \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}_+} \mathbb{1}_{(y < X_{s-})} x \tilde{N}(ds, dx, dy) + \alpha(\alpha - 1)\Gamma(\alpha)\theta \int_0^t X_s^{2-\alpha} ds,$$

we see that the total-mass process $(X_t(1))_{t \geq 0}$ and $(\exp(\xi_{A^{-1}(t)}))_{t \geq 0}$ are equal in law up to first hitting 0. Applying the Lamperti transformation for $t < T_0$ yields

$$\begin{aligned} \bar{S}(t) &:= \int_0^t X_s(1)^{1-\alpha} ds \\ &= \int_0^t \exp((1 - \alpha)\xi_{A^{-1}(s)}) ds \\ &= \int_0^{A^{-1}(t)} \exp((1 - \alpha)\xi_s) \exp((\alpha - 1)\xi_s) ds \\ &= A^{-1}(t) \end{aligned}$$

so that \bar{S} and A are reciprocal for $t < T_0$. Plugging this identity into the Lamperti transformation yields

$$0 = X_{T_0}(1) = \lim_{t \uparrow T_0} X_t(1) = \lim_{t \uparrow T_0} \exp(\xi_{A^{-1}(t)}) = \lim_{t \uparrow T_0} \exp(\xi_{\bar{S}(t)}). \tag{4.3}$$

For the second equality we used left-continuity of $X(1)$ at T_0 which is due to Sect. 3 of [29] because the Lévy process ξ does not jump to $-\infty$. Using that $\xi_t > -\infty$ for

any $t \in [0, \infty)$, from (4.3) we see that \bar{S} explodes at T_0 , that is $\bar{S}(T_0) = \infty$. Since S and \bar{S} only differ by the factor $\alpha(\alpha - 1)\Gamma(\alpha)$, it also holds that $S(T_0) = \infty$ so that $X_{S^{-1}(t)}(1) > 0$ for all $t \geq 0$. \square

We can now show how to construct on a pathwise level the Beta-Fleming–Viot processes with mutations the measure-valued branching process.

Proof of Theorem 2.7 Suppose \mathcal{N} is the PPP with compensator measure ν that drives the strong solution of (2.6) with atoms

$(r_i, x_i, y_i)_{i \in I} \in (0, \infty) \times (0, \infty) \times (0, \infty)$. Then we define a new point process on $(0, \infty) \times (0, \infty) \times (0, \infty)$ by

$$\mathcal{M}(ds, dz, du) := \sum_{i \in I} \delta_{\left\{S(r_i), \frac{x_i}{X_{r_i-(1)+x_i}\mathbf{1}_{(y_i \leq X_{r_i-(1)})}}, \frac{y_i}{X_{r_i-(1)}}\right\}}(ds, dz, du).$$

If we can show that the restriction $\mathcal{M}|_1$ of \mathcal{M} to $(0, \infty) \times (0, 1) \times (0, 1)$ is a PPP with intensity measure $\mathcal{M}'_1(ds, dz, du) = ds \otimes C_\alpha z^{-2} z^{1-\alpha} (1 - z)^{\alpha-1} dz \otimes du$ and furthermore that

$$\left(R_t(v) := \frac{X_{S^{-1}(t)}(v)}{X_{S^{-1}(t)}(1)}\right)_{t \geq 0, v \in [0,1]}$$

is a solution to (2.14) with respect to $\mathcal{M}|_1$, then the claim follows from the pathwise uniqueness of (2.14).

Step 1: We have

$$\begin{aligned} R_t(v) &= \frac{X_{S^{-1}(t)}(v)}{X_{S^{-1}(t)}(1)} = \\ &= \frac{X_0(v)}{X_0(1)} + \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \left[\frac{X_{r-}(v) + x\mathbf{1}_{(y \leq X_{r-}(v))}}{X_{r-}(1) + x\mathbf{1}_{(y \leq X_{r-}(1))}} - \frac{X_{r-}(v)}{X_{r-}(1)} \right] (\mathcal{N} - \nu)(dr, dx, dy) \\ &\quad + \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \left[\frac{X_{r-}(v) + x\mathbf{1}_{(y \leq X_{r-}(v))}}{X_{r-}(1) + x\mathbf{1}_{(y \leq X_{r-}(1))}} - \frac{X_{r-}(v)}{X_{r-}(1)} - \frac{x\mathbf{1}_{(y \leq X_{r-}(v))}}{X_{r-}(1)} \right. \\ &\quad \left. + \frac{x\mathbf{1}_{(y \leq X_{r-}(1))} X_{r-}(v)}{X_{r-}(1)^2} \right] \nu(dr, dx, dy) \\ &\quad + \alpha(\alpha - 1)\Gamma(\alpha)\theta v \int_0^{S^{-1}(t)} \frac{1}{X_r(1)} X_r(1)^{2-\alpha} dr - \theta \int_0^{S^{-1}(t)} \frac{X_r(v)}{X_r(1)^2} X_r(1)^{2-\alpha} dr \\ &= v + \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \left[\frac{X_{r-}(v) + x\mathbf{1}_{(y \leq X_{r-}(v))}}{X_{r-}(1) + x\mathbf{1}_{(y \leq X_{r-}(1))}} - \frac{X_{r-}(v)}{X_{r-}(1)} \right] \mathcal{N}(dr, dx, dy) \\ &\quad + \alpha(\alpha - 1)\Gamma(\alpha)\theta v \int_0^{S^{-1}(t)} X_r(1)^{1-\alpha} dr - \theta \int_0^{S^{-1}(t)} \frac{X_r(v)}{X_r(1)} X_r(1)^{1-\alpha} dr. \end{aligned}$$

To verify the third equality, first note that due to Lemma II.2.18 of [24] the compensation can be split from the martingale part and then can be canceled by the compensator integral since integrating-out the y -variable yields

$$\int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \left[-\frac{x \mathbf{1}_{(y \leq X_{r-}(v))}}{X_{r-}(1)} + \frac{x \mathbf{1}_{(y \leq X_{r-}(1))} X_{r-}(v)}{X_{r-}(1)^2} \right] c_\alpha x^{-1-\alpha} dr dx dy = 0.$$

To replace the jumps governed by the PPP \mathcal{N} by jumps governed by \mathcal{M} note that by the definition of \mathcal{M} we find, for measurable non-negative test-functions h for which the first integral is defined, the almost sure transfer identity

$$\begin{aligned} & \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty h \left(S(r), \frac{x}{X_{r-}(1) + x \mathbf{1}_{(y \leq X_{r-}(1))}}, \frac{y}{X_{r-}(1)} \right) \mathcal{N}(dr, dx, dy) \\ &= \int_0^t \int_0^1 \int_0^\infty h(s, z, u) \mathcal{M}(ds, dz, du) \end{aligned} \tag{4.4}$$

or in an equivalent but more suitable form

$$\begin{aligned} & \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty h \left(r, \frac{x}{X_{S^{-1}(r-)}(1) + x \mathbf{1}_{(y \leq X_{S^{-1}(r-)}(1))}}, \frac{y}{X_{S^{-1}(r-)}(1)} \right) \mathcal{N}(dr, dx, dy) \\ &= \int_0^t \int_0^1 \int_0^\infty h(S^{-1}(s), z, u) \mathcal{M}(ds, dz, du). \end{aligned} \tag{4.5}$$

Since the integrals are non-compensated we actually defined \mathcal{M} in such a way that the integrals produce exactly the same jumps.

Let us now rewrite the equation found for R in such a way that (4.5) can be applied:

$$\begin{aligned} R_t(v) &= v + \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \left[\frac{x \mathbf{1}_{(y \leq X_{r-}(v))} X_{r-}(1) - X_{r-}(v) x \mathbf{1}_{(y \leq X_{r-}(1))}}{(X_{r-}(1) + x \mathbf{1}_{(y \leq X_{r-}(1))}) X_{r-}(1)} \right] \\ &\quad \times \mathcal{N}(dr, dx, dy) + \alpha(\alpha - 1) \Gamma(\alpha) \theta \int_0^{S^{-1}(t)} \\ &\quad \times \left[v X_r(1)^{1-\alpha} - \frac{X_r(v)}{X_r(1)} X_r(1)^{1-\alpha} \right] dr \\ &= v + \int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \frac{x}{X_{r-}(1) + x \mathbf{1}_{(y \leq X_{r-}(1))}} \end{aligned}$$

$$\begin{aligned} & \times \left[\mathbf{1}_{(y \leq X_{r-}(v))} - \frac{X_{r-}(v)}{X_{r-}(1)} \mathbf{1}_{(y \leq X_{r-}(1))} \right] \mathcal{N}(dr, dx, dy) \\ & + \alpha(\alpha - 1)\Gamma(\alpha)\theta \int_0^{S^{-1}(t)} \left[v X_r(1)^{1-\alpha} - \frac{X_r(v)}{X_r(1)} X_r(1)^{1-\alpha} \right] dr. \end{aligned}$$

The stochastic integral driven by \mathcal{N} can now be replaced by a stochastic integral driven by \mathcal{M} via (4.5):

$$\begin{aligned} R_t(v) &= v + \int_0^t \int_0^1 \int_0^\infty z \left[\mathbf{1}_{(u X_{S^{-1}(s)-}(1) \leq X_{S^{-1}(s)-}(v))} - \right. \\ & \quad \left. - R_{S^{-1}(s)-}(v) \mathbf{1}_{(u X_{S^{-1}(s)-}(1) \leq X_{S^{-1}(s)-}(1))} \right] \mathcal{M}(ds, dz, du) \\ & \quad + \theta \int_0^t [v - R_s(v)] ds \\ &= v + \int_0^t \int_0^1 \int_0^\infty z \left[\mathbf{1}_{(u \leq R_{s-}(v))} - R_{s-}(v) \mathbf{1}_{(u \leq 1)} \right] \mathcal{M}(ds, dz, du) \\ & \quad + \theta \int_0^t [v - R_s(v)] ds. \end{aligned}$$

By monotonicity in v , $R_t(v) \leq 1$ so that the du -integral in fact only runs up to 1 and the second indicator can be skipped:

$$\begin{aligned} R_t(v) &= v + \int_0^t \int_0^1 \int_0^1 z \left[\mathbf{1}_{(u \leq R_{s-}(v))} - R_{s-}(v) \right] \mathcal{M}_1(ds, dz, du) \\ & \quad + \theta \int_0^t [v - R_s(v)] ds. \end{aligned}$$

This is precisely the equation we wanted to derive.

Step 2: The proof is complete if we can show that the restriction \mathcal{M}_1 of \mathcal{M} to $(0, \infty) \times [0, 1] \times [0, 1]$ is a PPP with intensity $\mathcal{M}'(ds, dz, du) = ds \otimes C_\alpha z^{-1-\alpha} (1 - z)^{\alpha-1} dz \otimes du$. For this sake, we choose a non-negative measurable predictable function $W : \Omega \times (0, \infty) \times (0, 1) \times (0, 1) \rightarrow \mathbb{R}$ bounded in the second and third variable and compactly supported in the first, plug-in the definition of \mathcal{M}_1 and use the compensator measure ν of \mathcal{N} to obtain via (4.4)

$$\begin{aligned} & \mathbb{E} \left(\int_0^t \int_0^1 \int_0^1 W(s, z, u) \mathcal{M}_1(ds, dz, du) \right) \\ &= \mathbb{E} \left(\int_0^t \int_0^1 \int_0^\infty \mathbf{1}_{(u \leq 1)} W(s, z, u) \mathcal{M}(ds, dz, du) \right) \\ &= \mathbb{E} \left(\int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \mathbf{1}_{\left(\frac{y}{X_{r-(1)}} \leq 1\right)} \right. \\ & \quad \left. \times W \left(S(r), \frac{x}{X_{r-(1)} + x \mathbf{1}_{(y \leq X_{r-(1)})}}, \frac{y}{X_{r-(1)}} \right) \mathcal{N}(dr, dx, dy) \right) \end{aligned}$$

which, by predictable projection and change of variables, equals

$$\begin{aligned} & \mathbb{E} \left(\int_0^{S^{-1}(t)} \int_0^\infty \int_0^\infty \mathbf{1}_{\left(\frac{y}{X_r(1)} \leq 1\right)} \right. \\ & \quad \left. \times W \left(S(r), \frac{x}{X_r(1) + x \mathbf{1}_{\left(\frac{y}{X_r(1)} \leq 1\right)}} , \frac{y}{X_r(1)} \right) c_\alpha x^{-1-\alpha} dr dx dy \right). \end{aligned}$$

Now we substitute the three variables r, x, y (in this order), using $C_\alpha = \frac{1}{\alpha(\alpha-1)\Gamma(\alpha)} c_\alpha$ for the substitution of r and the identity

$$\int_0^\infty g \left(\frac{x}{a+x} \right) x^{-1-\alpha} dx = a^{-\alpha} \int_0^1 g(z) z^{-1-\alpha} (1-z)^{\alpha-1} dz$$

for the substitution of x to obtain

$$\begin{aligned} & \mathbb{E} \left(\int_0^t \int_0^1 \int_0^1 W(s, z, u) \mathcal{M}_1(ds, dz, du) \right) \\ &= \mathbb{E} \left(\int_0^t \int_0^1 \int_0^1 W(s, z, u) C_\alpha z^{-1-\alpha} (1-z)^{\alpha-1} ds dz du \right). \end{aligned}$$

It now follows from Theorems II.4.8 of [24] and the definitions of c_α, C_α that \mathcal{M}_1 is a PPP with intensity $ds \otimes C_\alpha z^{-2} z^{1-\alpha} (1-z)^{\alpha-1} dz \otimes du$. □ □

5 Proof of Theorem 1.2

Let us briefly outline the strategy for the proof: In order to show that the measure-valued process \mathbf{Y}, \mathbb{P} -a.s., does not possess times t for which \mathbf{Y}_t has finitely many atoms,

by Theorem 2.7 it suffices to show that \mathbb{P} -a.s. the same is true for the measure-valued branching process \mathbf{X} . In order to achieve this, it suffices to deduce the same property for the Pitman-Yor type representation up to extinction, i.e. we need to show that

$$\mathbb{P}(\#\{v \in]0, 1] : Z_t(v) - Z_t(v-) > 0\} = \infty, \forall t \in]0, T_0[) = 1. \tag{5.1}$$

The upshot of working with Z instead of \mathbf{Y} is that things are easier due to a comparison property that is not available for \mathbf{Y} . More precisely, we are going to prove that with probability 1, the number of *immigrated types* alive is infinite at all times, therefore proving that the result in Theorem 1.2 is indeed independent of the starting configuration \mathbf{Y}_0 .

We start the proof with a technical result on the covering of a half line by the shadows of a Poisson point process defined on some probability space $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$. Suppose $(s_i, h_i)_{i \in I}$ are the points of a Poisson point process Π on $(0, \infty) \times (0, \infty)$ with intensity $dt \otimes \Pi'(dh)$. For a point (s_i, h_i) we define the shadow on the half line \mathbb{R}^+ by $(s_i, s_i + h_i)$ which is precisely the line segment covered by the shadow of the line segment connecting $(s_i, 0)$ and (s_i, h_i) with light shining in a 45 degrees angle from the above left-hand side. Shepp proved that the half line \mathbb{R}^+ is almost surely fully covered by the shadows induced by the points $(s_i, h_i)_{i \in I}$ if and only if

$$\int_0^1 \exp\left(\int_t^1 (h-t) \Pi'(dh)\right) dt = \infty. \tag{5.2}$$

The reader is referred to the last remark of [36]. For our purposes we need the following variant:

Lemma 5.1 *Suppose Π is a PPP with intensity $dt \otimes \Pi'(dh)$ and Shepp’s condition (5.2) holds, then*

$$\mathbb{P}(\#\{s_i \leq t : (s_i, h_i) \in \Pi \text{ and } s_i + h_i > t\} = \infty, \forall t > 0) = 1,$$

i.e. almost surely every point of \mathbb{R}^+ is covered by the shadows of infinitely many line segments.

Proof The proof is an iterated use of Shepp’s result for the sequence of restricted Poisson point processes Π_k obtained by removing all the atoms (s_i, h_i) with $h_i > \frac{1}{k}$ from Π , i.e. restricting the intensity measure to $[0, \frac{1}{k}]$. Since Shepp’s criterion (5.2) only involves the intensity measure around zero, the shadows of all point processes Π_k cover the half line. Consequently, if there is some $t > 0$ such that t is only covered by the shadows of finitely many points $(s_i, h_i) \in \Pi$, then t is not covered by the shadows generated by $\Pi_{k'}$ for some k' large enough. But this is a contradiction to Shepp’s result applied to $\Pi_{k'}$. \square

Now we want to apply Shepp’s result to the Pitman-Yor type representation. We want to prove that (5.1) holds for any $\theta > 0$. Let us set for all $\epsilon > 0$

$$T_\epsilon := \inf\{t > 0 : Z_t(1) \leq \epsilon\}.$$

Then it is clearly enough to prove that for all $\epsilon > 0$

$$\mathbb{P}(\#\{v \in]0, 1] : Z_t(v) - Z_t(v-) > 0\} = \infty, \forall t \in]0, T_\epsilon[) = 1.$$

In order to connect the covering lemma with the question of exceptional times, we use the comparison property of the Pitman-Yor representation to reduce the problem to the process Z^ϵ explicitly defined by

$$Z_t^\epsilon(v) = \sum_{s_i > 0} w_t^i \mathbb{1}_{(u_i \leq v \alpha(\alpha-1)\Gamma(\alpha)\theta\epsilon^{2-\alpha})}, \quad v \in [0, 1], t \geq 0. \tag{5.3}$$

Setting

$$N_t := \#\{v \in]0, 1] : Z_t(v) - Z_t(v-) > 0\},$$

$$N_t^\epsilon := \#\{v \in]0, 1] : Z_t^\epsilon(v) - Z_t^\epsilon(v-) > 0\},$$

it is obvious by the definition of Z and Z^ϵ that

$$\mathbb{P}(N_t \geq N_t^\epsilon, \forall t \in]0, T_\epsilon[) = 1. \tag{5.4}$$

We are now prepared to prove our main result.

Proof of Theorem 1.2 Due to (5.4) we only need to show that almost surely $v \mapsto Z_t^\epsilon(v)$ has infinitely many jumps for all $t > 0$ and arbitrary $\epsilon > 0$. To verify the latter, Lemma 5.1 will be applied to a PPP defined in the sequel. If Π denotes the Poisson point process with atoms $(s_i, w^i, u_i)_{i \in I}$ from which $Z_t^\epsilon(v)$ is defined, then we define a new Poisson point process Π_l via the atoms

$$(s_i, h_i, u_i)_{i \in I} := (s_i, \ell(w^i), u_i)_{i \in I},$$

where $\ell(w) := \inf\{t > 0 : w_t = 0\}$ denotes the length of the trajectory w . In order to apply Lemma 5.1 we need the intensity of Π_l . Using the definition of \mathbb{Q} and the Laplace transform duality (2.8) with the explicit form

$$\int (1 - e^{-\lambda w_t}) \mathbb{Q}(dw) = (\lambda^{1-\alpha} + (\alpha - 1)t)^{\frac{1}{1-\alpha}},$$

we find the distribution

$$\begin{aligned} \mathbb{Q}(\ell(w) > h) &= \mathbb{Q}(w_h > 0) = \lim_{\lambda \rightarrow +\infty} \mathbb{Q}(1 - e^{-\lambda w_h}) \\ &= \lim_{\lambda \rightarrow +\infty} u_h(\lambda) = ((\alpha - 1)h)^{\frac{1}{1-\alpha}}. \end{aligned} \tag{5.5}$$

Differentiating in h shows that Π_t is a Poisson point process on $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ with intensity measure

$$\Pi'_t(dt, dh, du) = dt \otimes ((\alpha - 1)h)^{\frac{\alpha}{(1-\alpha)}} dh \otimes du.$$

Plugging-in the new definitions leads to

$$\begin{aligned} N_t^\epsilon &= (\text{number of non-zero summands of } Z_t^\epsilon(1))_{t \geq 0} \\ &= (\#\{s_i \leq t : (s_i, w_i, u_i) \in \Pi \text{ and } w_{t-s_i}^i \mathbb{1}_{(u_i \leq \alpha(\alpha-1)\Gamma(\alpha)\theta\epsilon^{2-\alpha})} > 0\})_{t \geq 0} \\ &= (\#\{s_i \leq t : (s_i, w_i, u_i) \in \Pi \text{ and } \ell(w_i) > t - s_i, u_i \leq \alpha(\alpha-1)\Gamma(\alpha)\theta\epsilon^{2-\alpha}\})_{t \geq 0} \\ &= (\#\{s_i \leq t : (s_i, h_i, u_i) \in \Pi_t \text{ and } s_i + h_i > t, u_i \leq \alpha(\alpha-1)\Gamma(\alpha)\theta\epsilon^{2-\alpha}\})_{t \geq 0}. \end{aligned} \tag{5.6}$$

There is one more simplification that we can do. Let us define $\Pi_{t,\epsilon}$ as a Poisson point process on $(0, \infty) \times (0, \infty)$ with intensity measure

$$\Pi'_{t,\epsilon}(dt, dh) = \alpha(\alpha - 1)\Gamma(\alpha)\theta\epsilon^{2-\alpha} dt \otimes (\alpha - 1)^{\alpha/(1-\alpha)} h^{\frac{\alpha}{(1-\alpha)}} dh, \tag{5.7}$$

then by the properties of Poisson point processes we have the equality in law

$$\{(s_i, h_i) : (s_i, h_i, u_i) \in \Pi_t \text{ and } u_i \leq \alpha(\alpha - 1)\Gamma(\alpha)\theta\epsilon^{2-\alpha}\} \stackrel{(d)}{=} \Pi_{t,\epsilon}.$$

Then (5.6) yields

$$(N_t^\epsilon)_{t \geq 0} \stackrel{(d)}{=} (\#\{s_i \leq t : (s_i, h_i) \in \Pi_{t,\epsilon} \text{ and } s_i + h_i > t\})_{t \geq 0}.$$

Now we are precisely in the setting of Shepp’s covering results and the theorem follows from Lemma 5.1 if (5.2) holds. Shepp’s condition can be checked easily for $\Pi_{t,\epsilon}$ for (5.7) independently of θ and ϵ . □

6 A Proof of Schmuland’s Theorem

In this section we sketch how our lines of arguments can be adopted for the continuous case corresponding to $\alpha = 2$. The proofs go along the same lines (reduction to a measure-valued branching process and then to an excursion representation for which the covering result can be applied) but are much simpler due to a constant immigration structure. The crucial difference, leading to the possibility of exceptional times, occurs in the final step via Shepp’s covering results.

Proof of Schmuland’s Theorem 1.1 We start with the continuous analogue to Theorem 2.2. Suppose W is a white-noise on $(0, \infty) \times (0, \infty)$, then one can show via the standard Yamada-Watanabe argument that there is a unique strong solution to

$$\begin{cases} X_t(v) = v + \sqrt{2} \int_{]0,t] \times \mathbb{R}_+} \mathbb{1}_{(u \leq X_s(v))} W(ds, du) + \theta vt, \\ v \in [0, 1], t \geq 0. \end{cases} \tag{6.1}$$

In fact, since the immigration mechanism g is constant, pathwise uniqueness holds. For every $v \in [0, 1]$, $(X_t(v))_{t \geq 0}$ satisfies

$$X_t(v) = v + \int_0^t \sqrt{2X_s(v)} dB_s + v\theta t$$

for a Brownian motion B . Recalling (2.2), we see that (6.1) is a measure-valued process with branching mechanism $\psi(u) = u^2$ and constant-rate immigration.

The Pitman–Yor type representation corresponding to Theorem 2.3 looks as follows: in the setting of Sect. 2.2, we consider a Poisson point process $(s_i, u_i, w^i)_i$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{D}$ with intensity measure $(\delta_0(ds) \otimes F(du) + ds \otimes I(du)) \otimes \mathbb{Q}_s(dw)$, where the excursion measure \mathbb{Q} is defined via the law of the CSBP (2.2) with branching mechanism $\psi(\lambda) = \lambda^2$. Then the analog of Theorem 2.3 is the following:

$$\begin{cases} Z_t(v) = \sum_{s_i=0} w_t^i \mathbb{1}_{(u_i \leq v)} + \sum_{0 < s_i \leq t} w_{t-s_i}^i \mathbb{1}_{(u_i \leq v\theta)}, \\ v \in [0, 1], t \geq 0, \end{cases} \tag{6.2}$$

can be shown to solve (6.1); this result, for fixed v , goes back to Pitman and Yor [34]. The calculation (5.5), now using that $u_t(\lambda) = (\lambda^{-1} + t)^{-1}$ is the unique non-negative solution to

$$\partial_t u_t(\lambda) = -(u_t(\lambda))^2, u_0(\lambda) = \lambda,$$

yields $\mathbb{Q}(\ell(w) \in dh) = \frac{1}{h^2} dh$.

For the analogue for Theorem 2.7 we define now the process

$$R_t(v) = \frac{X_{S^{-1}(t)}(v)}{X_{S^{-1}(t)}(1)},$$

with $S(t) = \int_0^t X_s(1)^{-1} ds$. It then follows again from the self-similarity that R is well-defined and from Itô’s formula that R is a standard Fleming–Viot process on $[0, 1]$. The arguments here involve a continuous SDE which has been studied in [12]:

$$\begin{cases} Y_t(v) = v + \int_0^t \int_0^1 [\mathbb{1}_{(u \leq Y_s(v))} - Y_s(v)] W(ds, du) + \theta \int_0^t [v - Y_s(v)] ds, \\ v \in [0, 1], t \geq 0, \end{cases} \tag{6.3}$$

where W is a white-noise on $(0, \infty) \times (0, 1)$. It was shown in Theorem 4.9 of [12] that the measure-valued process \mathbf{Y} associated with $(Y_t(v), t \geq 0, v \in [0, 1])$ solves

the martingale problem for the infinitely many sites model with mutations, i.e. \mathbf{Y} has generator (1.1) with the choice (1.2) for A .

Finally, in order to prove Schmuland’s Theorem 1.1 on exceptional times it suffices to prove the same result for (6.2). We proceed again via Shepp’s covering arguments as we did in Sect. 5. The crucial difference is that the immigration is already constant θ so that (5.3) becomes superfluous. The role of the Poisson point process $\Pi_{l,\epsilon}$ is played by $\Pi_{\theta,l}$ with intensity measure

$$\Pi'_{\theta,l}(dt, dh) = dt \otimes \frac{\theta}{h^2} dh.$$

Plugging into Shepp’s criterion (5.2), by Lemma 5.1 and

$$\int_0^1 \exp\left(-\theta \log(t)\right) dt = \int_0^1 t^{-\theta} dt \tag{6.4}$$

we find that there are no exceptional times if $\theta \geq 1$. Conversely, let us assume $\theta < 1$. Recalling that for $\theta = 0$ the Fleming–Viot process has almost surely finitely many atoms for all $t > 0$, we see that the first term in (6.2) almost surely has finitely non-zero summands for all $t > 0$. Hence, it suffices to show the existence of exceptional times for which the second term in (6.2) vanishes. Arguing as before, this question is reduced to Shepp’s covering result applied to $\Pi_{\theta,l}$: (6.4) combined with (5.2) leads to the result. □

7 Proof of Corollary 2.8

The fact that the (α, θ) -Fleming–Viot process $(\mathbf{Y}_t, t \geq 0)$ converges in distribution to its unique invariant distribution and that this invariant distribution is not trivial (i.e. it charges measures with at least two atoms) is proven by Donnelly and Kurtz in [14] at the end of Sect. 5.1 and [13] Sect. 4.1. Here we re-sketch their argument that relies on the so-called *lookdown* construction of $(\mathbf{Y}_t, t \geq 0)$ which was introduced in the same papers. Let us very briefly describe how the lookdown construction works (for more details we refer to [9, 14]).

The idea is to construct a sequence of processes $(\xi_i(t), t \geq 0), i = 1, 2, \dots$ which take their values in the type-space E (here $E = [0, 1]$). We say that $\xi_i(t)$ is the *type* of the *level* i at time t . The types evolve by two mechanisms :

- *lookdown events*: with rate $x^{-2} \Lambda(dx)$ a proportion x of lineages are selected by i.i.d. Bernoulli trials. Call i_1, i_2, \dots the selected levels at a given event at time t . Then, $\forall k > 1, \xi_{i_k}(t) = \xi_{i_1}(t-)$, that is the levels all adopt the type of the smallest participating level. The type $\xi_{i_k}(t-)$ which was occupying level i_k before the event is pushed up to the next available level.
- *mutation events*: On each level i there is an independent Poisson point process $(t_j^{(i)}, j \geq 1)$ of rate θ of mutation events. At a mutation event $t_j^{(i)}$ the type $\xi_i(t_j^{(i)}-)$

is replaced by a new independent variable uniformly distributed on $[0, 1]$ and the previous type is pushed up by one level (as well as all the types above him).

The point is then that

$$\Xi_t := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{\xi_i(t)}$$

exists simultaneously for all $t \geq 0$ almost surely and that $(\Xi_t, t \geq 0) = (\mathbf{Y}_t, t \geq 0)$ in distribution.

Fix $n \in \mathbb{N}$, and define a process $(\pi_t, t \geq 0)$ with values in the partitions of $\{1, 2, \dots\}$ by saying that $i \sim j$ for $\pi_t^{(n)}$ if and only $\xi_i(t) = \xi_j(t)$. It is well known that this is an *exchangeable* process. Recall from Corollary 2.5 that for each $t \geq 0$ fixed, Ξ_t is almost surely purely atomic. Alternatively this can be seen from the lookdown construction since at a fixed time $t > 0$, the level one has been looked down upon by infinitely many level above since the last mutation event on level one. We can thus write

$$\Xi_t = \sum a_i(t) \delta_{x_i(t)},$$

where the a_i are enumerated in decreasing order. It is also known that the sequences $(a_i(t), i \geq 1)$ of atom masses and $(x_i(t), i \geq 1)$ of atom locations are independent. The $a_i(t)$ are the asymptotic frequencies of the blocks of $\pi(t)$ which are thus in one-to-one correspondence with the atoms of Ξ_t . Furthermore the sequence $(x_i(t), i \geq 1)$ converges in distribution to a sequence of i.i.d random variables with common distribution I because all the types that were present initially have been replaced by immigrated types after some time. To see this note that after the first mutation on level 1, the type $\xi_1(0)$ is pushed up to infinity in a finite time which is stochastically dominated by the fixation time of the type at level 1 in a Beta-Fleming–Viot without mutation. This also proves the second point of the corollary.

For each $n \geq 1$, let us consider $\pi^{(n)}(t) = \pi_{[n]}(t)$ the restriction to $\{1, \dots, n\}$ of $\pi(t)$. Then, for all $n \geq 1$, the process $(\pi_t^{(n)}, t \geq 0)$ is an irreducible Markov process on a finite state-space and thus converges to its unique invariant distribution. This now implies that $(\pi(t), t \geq 0)$ must also converges to its invariant distribution. By Kingman continuity Theorem (see [33, Theorem 36] or [4, Theorem 1.2]) this implies that the ordered sequence of the atom masses $(a_i(t))$ converges in distribution as $t \rightarrow \infty$. Because $(x_i(t), i \geq 1)$ also converges in distribution this implies that Ξ_t itself converges in distribution to its invariant measure. (Alternatively, this second part of the Corollary could be deduced from Theorem 2 in [28])

Furthermore it is also clear that the invariant distribution of $(\pi_t^{(n)}, t \geq 0)$ must charge configurations with at least two non-singleton blocks. Since π is an exchangeable process, so is its invariant distribution. Exchangeable partitions have only two types of blocks: singletons and blocks with positive asymptotic frequency so this proves that the invariant distribution of π charges partition with at least two blocks of positive asymptotic frequency.

Acknowledgments The authors are very grateful for stimulating discussions with Zenghu Li and Marc Yor and to an anonymous referee whose insightful comments greatly helped us improve this manuscript. L. Döring was supported by the Fondation Science Mathématiques de Paris, L. Mytnik is partly supported by the Israel Science Foundation. J. Berestycki, L. Mytnik and L. Zambotti thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for a very pleasant stay during which part of this work was produced. L. Mytnik thanks the Laboratoire de Probabilités et Modèles Aléatoires for the opportunity to visit it and carry out part of this research there

References

- Berestycki, J., Döring, L., Mytnik, L., Zambotti, L.: Hitting properties and non-uniqueness for SDEs driven by stable processes, Preprint <http://arxiv.org/abs/1111.4388>
- Berestycki, J., Berestycki nd, N., Limic, V.: A small-time coupling between Λ -coalescents and branching processes, Preprint <http://arxiv.org/abs/1101.1875>
- Berestycki, J., Berestycki, N., Schweinsberg, J.: Beta-coalescents and continuous stable random trees. *Ann. Probab.* **35**, 1835–1887 (2007)
- Berestycki, N.: Recent progress in coalescent theory. *Ens. Mat.* **16**, 1–193 (2009)
- Bertoin, J., Le Gall, J.F.: Stochastic flows associated to coalescent processes. *Probab. Theory Relat. F.* **126**(2), 261–288 (2003)
- Bertoin, J., Le Gall, J.F.: Stochastic flows associated to coalescent processes. II. Stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **41**(3), 307–333 (2005)
- Bertoin, J., Le Gall, J.F.: Stochastic flows associated to coalescent processes. III. Limit theorems. III. *J. Math.* **50**(1–4), 147–181 (2006)
- Birkner, M., Blath, J.: *Measure-Valued Diffusions, General Coalescents and Population Genetic Inference Trends in Stochastic Analysis*. Cambridge University Press, Cambridge (2009)
- 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)
- Birkner, M., Blath, J., Capaldo, M., Etheridge, A., Möhle, M., Schweinsberg, J., Wakolbinger, A.: Alpha-stable branching and beta-coalescents. *Electron. J. Probab.* **10**(9), 303–325 (2005)
- Dawson, D.A., Li, Z.: Construction of immigration superprocesses with dependent spatial motion from one-dimensional excursions. *Probab. Theory Relat. F.* **127**, 37–61 (2003)
- Dawson, D.A., Li, Z.: Stochastic equations, flows and measure-valued processes. *Ann. Probab.* **40**(2), 813–857 (2012)
- Donnelly, Kurtz, T.: A countable representation of the Fleming–Viot measure-valued diffusion. *Ann. Probab.* **24**, 698–742 (1996)
- Donnelly, P., Kurtz, T.: Particle representations for measure-valued population models. *Ann. Probab.* **27**, 166–205 (1999)
- Duquesne, T., Le Gall, J.F.: Random trees, Levy processes and spatial branching processes. *Asterisque* **281**, 1–147 (2002)
- Dynkin, E.B., Kuznetsov, S.E.: \mathbb{N} -measures for branching Markov exit systems and their applications to differential equations. *Prob. Theory Relat. F.* **130**, 135–150 (2004)
- Etheridge, A.M.: Some mathematical models from population genetics. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009. *Lecture Notes in Mathematics*, 2012. Springer, Heidelberg, viii+119 pp (2011)
- Ethier, S.N., Griffiths, R.C.: The transition function of a Fleming–Viot process. *Ann. Prob.* **21**, 1571–1590 (1993)
- Ethier, S.N., Kurtz, T.G.: Fleming–Viot processes in population genetics. *SIAM J. Control Optim.* **31**, 345–386 (1993)
- Fleming, W.H., Viot, M.: Some measure-valued Markov processes in population genetics. *Indiana Univ. Math. J.* **28**, 817–843 (1979)
- Fournier, N.: On pathwise uniqueness for stochastic differential equations driven by stable Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **49**(1), 138–159 (2013)
- Fu, Z., Li, Z.: Stochastic equations of non-negative processes with jumps. *Stoch. Process. Appl.* **120**, 306–330 (2010)

23. Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library, North-Holland Publishing Co, Amsterdam (1989)
24. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin (2003)
25. Kurtz, T.G.: The Yamada–Watanabe–Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.* **12**, 951–965 (2007)
26. Konno, N., Shiga, T.: Stochastic differential equations for some measure-valued diffusions. *Probab. Theory Relat. F.* **79**, 201–225 (1988)
27. Kyprianou, A.E.: *Fluctuations of Lévy Processes with Applications, Introductory Lectures*. Universitext, 2nd edn, p. XVIII. Springer-Verlag, Berlin (2014)
28. Labbé, C.: *Flots stochastiques et representation lookdown*. PhD Thesis (2013). <http://tel.archives-ouvertes.fr/tel-00874551>
29. Lamperti, J.: Semi-stable Markov processes. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **22**, 205–225 (1972)
30. Li, Z.: Skew convolution semigroups and related immigration processes. *Theory Probab. Appl.* **46**(2), 274–296 (2002)
31. Li, Z.: *Measure-Valued Branching Markov Processes, Probability and Its Applications*. Springer, Berlin (2010)
32. Li, Z., Mytnik, L.: Strong solutions for stochastic differential equations with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **47**(4), 1055–1067 (2011)
33. Pitman, J.: Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902 (1999)
34. Pitman, J., Yor, M.: A decomposition of Bessel bridges. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **59**, 425–457 (1982)
35. Schmuland, B.: A result on the infinitely many neutral alleles diffusion model. *J. Appl. Probab.* **28**(2), 253–267 (1991)
36. Shepp, L.A.: Covering the line with random intervals. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **23**, 163–170 (1972a)